Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2024, 2024:129
https://doi.org/10.28919/cmbn/8920
ISSN: 2052-2541

ROBUST SPATIAL DURBIN APPROACH IN MODELLING THE SPREADING OF TUBERCULOSIS IN INDONESIA

FERRA YANUAR[1,*], DILLA PERMATA SARI[1], YUDIANTRI ASDI[1], FENNI KURNIA MUTIYA[2]

[1]Department of Mathematics and Data Science, Andalas University, Padang, Indonesia

[2]Department of Statistics, Padang State University, Padang, Indonesia

Copyright © 2024 the author(s).This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: Spatial regression analysis is a method used for data that has spatial effects. One method in spatial regression analysis is the Spatial Durbin Model (SDM) which shows spatial effects on both dependent and independent variables. In the spatial regression model, inaccuracy can occur in predicting the model due to spatial outliers. To overcome outliers in the SDM model, a robust method is needed, namely the Robust Spatial Durbin Model (RSDM). This study was conducted to model the factors that influence the spread of tuberculosis (TB) cases in Indonesia and determine the best method. The results obtained are that the RSDM model is better with a larger Adjusted R^2 value and a smaller Mean Squared Error (MSE) value than SDM. The factors suspected of influencing the spread of TB cases in Indonesia are the percentage of households with access to proper sanitation, the number of HIV cases and population density.

Keywords: spatial Durbin model (SDM); robust spatial Durbin model (RSDM); tuberculosis (TB); Indonesia.

2020 AMS Subject Classification: 62F15.

---

*Corresponding author
E-mail address: ferrayanuar@sci.unand.ac.id
Received September 24, 2024

1. INTRODUCTION

Tuberculosis (TB) is an infectious disease caused by the bacteria Mycobacterium Tuberculosis. The bacteria spread from TB sufferers through the air. Geographically, the largest TB cases are in Southeast Asia at 46%. The number of new cases diagnosed with TB in 2022 globally is 7.5 million people with an estimated death toll of around 1.30 million. TB is still being a health problem in the world today. In 2022, Indonesia was ranked second at 10% as the largest contributor to the burden of TB cases in the world after India [1]. Based on the Indonesian Ministry of Health, the number of TB cases in 2022 increased compared to the previous year.

TB is an infectious disease that allows it to spread from one region to another adjacent region. Therefore, a method is needed to determine the factors that influence the spread of TB cases, namely by multiple regression analysis. However, this method will be less appropriate for use on data containing spatial information. If there is a spatial effect on the research data due to the influence between one region and another adjacent region, then the appropriate method to use is spatial regression analysis [2], [3]. One method in spatial regression analysis is the Spatial Durbin Model (SDM) [4].

SDM is a regression method that shows spatial effects on both dependent and independent variables. However, in some spatial data, outliers are often found which cause the parameter estimates in the model to be biased. To overcome data contaminated by outliers, a robust modeling method is needed against outliers, namely the Robust Spatial Durbin Model (RSDM) [5]. RSDM was applied to the life expectancy rate of Central Java Province using the M-estimator, revealing that RSDM is the best model for explaining the life expectancy rate in Central Java [6]. Mukrom, et al. [7] also applied RSDM in the case of life expectancy in Central Java province. Meanwhile, Khofifah [8] investigated the modeling of the open unemployment rate in West Java Province using RSDM and M-estimator. Syam et al. [9] applied RSDM with MM estimator to model tuberculosis data. This study will carry out an analysis of RSDM in other fields, namely modeling the spread of TB cases in Indonesia using M-estimator.

2. Data and Methods

2.1. Data

Data used in this study are secondary data sourced from web of Central Bureau of Statistics Indonesia https://www.bps.go.id/id in 2023. The response variable used is the number of tuberculosis cases in each province in Indonesia. In this study, the predictor variables are percentage of population who smoke, percentage of households with access to proper sanitation. number of HIV cases, and population density. The selection of predictor variables was based on information obtained from a literature review of several research studies related to tuberculosis cases. This study uses 27 provinces in Indonesia which neighbor to each other. Eight provinces are excluded from analysis since their position are cycle with sea not land.

2.2. Spatial Regression

Spatial regression is a development of the multiple linear regression method which is based on the location effect (spatial effect) on the data being analyzed. Therefore, spatial regression is a statistical method used to determine the relationship between dependent variables and independent variables by considering the spatial effect. The general form of the spatial regression model can be written as follows [10]:

$$\boldsymbol{y} = \rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \text{ with } \quad \boldsymbol{u} = \lambda\boldsymbol{W}\boldsymbol{u} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma_\varepsilon^2\boldsymbol{I_n}), \tag{1}$$

where $\boldsymbol{y}$ is the response variable vector, ρ is the spatial lag parameter coefficient of the dependent variable, $\boldsymbol{W}$ is the spatial weight matrix, $\boldsymbol{X}$ is the independent variable matrix, $\boldsymbol{\beta}$ is the regression parameter coefficient vector, $\boldsymbol{u}$ is the residual vector that has spatial effects, λ is the residual spatial parameter coefficient, and $\boldsymbol{\varepsilon}$ is the residual vector [12]. The spatial weight matrix, $\boldsymbol{W}$**,** is a symmetric matrix of size $n \times n$ that describes the closeness of the relationship between one region and another. The spatial weight matrix used in this study is the queen contiguity. The spatial weight matrix can be obtained in two ways, namely the standardized weight matrix ($\boldsymbol{W}$) and the unstandardized weight matrix ($\boldsymbol{W}^*$). The elements of the weight matrix $\boldsymbol{W}$ are w_{ij} with i is the row and j is the column, with $j = 1,2,\dots,n$ [11]. In obtaining the value of w_{ij} is formulated as follows:

$$w_{ij} = \frac{c_{ij}}{\sum_{j=1}^{n} c_{ij}} \quad . \tag{2}$$

Value for $c_{ij} = 1$ if i is neighbor with j, and $c_{ij} = 0$ both regions are not neighbor [13].

2.3. Spatial Dependence Test

Spatial dependence shows that observations in one area depend on observations in another area. The test used to identify the presence of spatial dependencies is the Moran Index. Moran Index is defined as follows:

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i-\bar{y})(y_j-\bar{y})}{S^2 \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}, \tag{3}$$

with *i* and *j* are the number of data (*n*), for $i \neq j$, y_i is the response variable for all observations I, $\bar{y}$ is the mean of *y*, S^2 is sample variance. If *I* is positive, adjacent areas have similar values, and the data pattern tends to be clustered. If *I* is negative, it means that adjacent areas have different values, and the data pattern tends to spread out. If *I* is 0, it means no spatial dependence detected [14]. The existence of spatial dependence on the dependent variable is also checked using Lagrange Multiplier lag $(LM\ lag)$ with the general form as follows [4]:

$$LM\ lag = \frac{\left(\frac{\boldsymbol{\varepsilon}^T \boldsymbol{W y}}{s^2}\right)^2}{\boldsymbol{D}}, \tag{4}$$

where $D = \left(\frac{(\boldsymbol{W X\beta})^T\ (\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}^T\ \boldsymbol{X})^{-1}\boldsymbol{X}^T\ )\boldsymbol{W X\beta}}{s^2}\right) + tr(\boldsymbol{W}^T\boldsymbol{W} + \boldsymbol{W}^2)$ and $s^2 = \frac{\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}}{n}$. If $LM\ lag >$ $\chi^2_{\alpha(1)}$ or *p-value* $> \alpha$ it means spatial dependence due among dependent variables.

The existence of spatial dependence on the error is checked using Lagrange Multiplier error (LM error) with the following form:

$$LM\ error = \frac{\left(\frac{\boldsymbol{\varepsilon}^T \boldsymbol{W y}}{s^2}\right)^2}{tr(\boldsymbol{W}^T\boldsymbol{W}+\boldsymbol{W}^2)}. \tag{5}$$

If $LM\ error$ more than $\chi^2_{\alpha(1)}$ or *p-value* more than α, it means spatial dependence on error due in the hypothesis model [10]. If both spatial dependence on dependent and on error presence and/or presence outliers in the proposed model, RSDM is then applied to estimate the parameter model.

2.4. Spatial Durbin Model (SDM)

Spatial Durbin Model is a spatial regression method that has a spatial lag on the response variable

(y) and predictor variables $(\boldsymbol{X})$ [15]. Hypothesis model of SDM is presented as follows [12]:

$$\boldsymbol{y} = \rho \boldsymbol{W}\boldsymbol{y} + \alpha \boldsymbol{1_n} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}, \text{ with } \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I_n}) \tag{6}$$

Can be written as follows:

$$\boldsymbol{y} = \rho \boldsymbol{W}\boldsymbol{y} + \boldsymbol{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}. \tag{7}$$

Where α is a constant parameter, $\boldsymbol{\theta}$ is the spatial lag parameter vector of predictor variable of size $k \times 1$, $\boldsymbol{Z} = [\boldsymbol{1_n} \quad \boldsymbol{X} \quad \boldsymbol{WX}]$, $\boldsymbol{\delta}$=$[\alpha \quad \boldsymbol{\beta} \quad \boldsymbol{\theta}]^T$.

2.5. Spatial Outlier

Outliers are observations that show significant differences with other observations in a data set collection [9]. One graphical method for detecting spatial outliers is Moran's Scatterplot, which is a normal plot graph with attribute values $\left(Z[f(i)] = \frac{f(i) - \mu_f}{\sigma_f}\right)$, namely the average value of the neighbors of the normalized attribute values. Data plots in the upper left and lower right quadrants can be said to be spatial outliers. Spatial outlier detection can also be identified mathematically with the following formula:

$$\left(Z[f(i)] \times \left(\textstyle\sum_j \boldsymbol{W_{ij}} \boldsymbol{Z}[f(j)]\right)\right). \tag{8}$$

With $Z[f(j)]$ is the normalized transpose of the attribute values, μ_f and σ_f are the mean and standard deviation of the function $f(i)$. If the value obtained is less than zero, then it can be said to be a spatial outlier [16].

2.6. Robust Spatial Durbin Model (RSDM) Using M-Estimator

Robust Spatial Durbin Model (RSDM) is applied if outlier present in error of SDM. The initial estimation of the RSDM parameters was carried out using Ordinary Least Square (OLS) by minimizing the sum of squares of the residual SDM [17]:

$$\boldsymbol{\varepsilon} = (\boldsymbol{I_n} - \rho \boldsymbol{W})\, \boldsymbol{y} - \boldsymbol{Z}\boldsymbol{\delta} \tag{9}$$

Then, based on equation above, it can result the Sum of Square Error (SSE):

$$SSE = [\boldsymbol{y}^T (\boldsymbol{I} - \rho \boldsymbol{W})^T (\boldsymbol{I} - \rho \boldsymbol{W})\boldsymbol{y} - 2\boldsymbol{\delta}^T \boldsymbol{Z}^T (\boldsymbol{I} - \rho \boldsymbol{W})\boldsymbol{y} + \boldsymbol{\delta}^T \boldsymbol{Z}^T \boldsymbol{Z}\boldsymbol{\delta}] \tag{10}$$

Using OLS, formula to estimate $\boldsymbol{\delta}$ is as follows:

$$\widehat{\boldsymbol{\delta}}_{\boldsymbol{OLS}} = (\boldsymbol{Z^T Z})^{-1} \boldsymbol{Z^T} (\boldsymbol{I} - \rho \boldsymbol{W})\boldsymbol{y}. \tag{11}$$

Then, we will construct the estimate for δ using M estimator. The estimation process begins with minimizing the following objective function $\rho(.)$:

$$\underset{\delta}{min}\,\rho(u_i) = \underset{\delta}{min}\,\rho\left(\frac{\boldsymbol{\varepsilon}}{s}\right) = \underset{\delta}{min}\,\rho\left(\frac{(\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y} - \boldsymbol{Z}\hat{\boldsymbol{\delta}}_{\boldsymbol{OLS}}}{s}\right), \tag{12}$$

where s is robust estimate scale obtained from:

$$s = \frac{MAD}{0{,}6745} = \frac{median|\boldsymbol{\varepsilon} - median(\boldsymbol{\varepsilon})|}{0{,}6745}. \tag{13}$$

By using Iteratively Reweighted Least Square (IRLS) estimation method, it can be obtained the estimated δ as following:

$$\hat{\boldsymbol{\delta}}_{\boldsymbol{OLS}} = (\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{B}\boldsymbol{Z})^{-\boldsymbol{1}}\,\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{B}(\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y}, \tag{14}$$

where $\boldsymbol{B}$ is diagonal matrix obtained by using Tukey Bisquare weighting:

$$b_i = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 & , |u_i| \le c \\ 0 & , |u_i| \le c \end{cases}$$

with $c = 4{,}685$. For $(m+1)th$ iteration with diagonal element in matrix $B^{(m)}$ is $b_i^{(m)}$, estimation for $\delta^{(m+1)}$ is:

$$\hat{\boldsymbol{\delta}}^{(\boldsymbol{m}+\boldsymbol{1})} = (\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{B}^{\boldsymbol{m}}\boldsymbol{Z})^{-\boldsymbol{1}}\,\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{B}^{\boldsymbol{m}}(\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y}, \tag{15}$$

where $\boldsymbol{\delta} = [\alpha \quad \boldsymbol{\beta} \quad \boldsymbol{\theta}]^T$. Iteration is continued until a convergent value of $\hat{\delta}$ is obtained, namely when the difference in the values of $\hat{\delta}^{(m+1)}$ and $\hat{\delta}^{(m)}$ approaches 0 with *m* being the number of iterations.

2.7. Best Model Selection

The selection of the best model is based on the value of the coefficient of determination (R^2) and Mean Squared Error (MSE). The coefficient of determination is a measure used to determine the ability of a model to explain the variation of the dependent variable. Some researchers prefer to use the adjusted coefficient of determination (Adjusted R^2) with the following formula [18].

$$R_{adj}^2 = 1 - \frac{JKS/(n-k-1)}{JKT/(n-1)} \tag{16}$$

In regression analysis, MSE refers to the unbiased estimate of the residual variance. The MSE calculation is done using the following formula:

$$MSE = \frac{SSE}{n-k-1} \tag{17}$$

A good model is a model that has a larger R_{adj}^2 value and a smaller MSE value [7].

3. MAIN RESULTS

According to Central Bureau of Statistics [19] the population of Indonesia in 2022 is 275,773.8 people which is in the fourth position as the country with the largest population in the world. This can facilitate the spread of infectious diseases, one of them is tuberculosis. The following is a graph showing the number of tuberculosis case notifications per 100,000 population on each province in Indonesia.

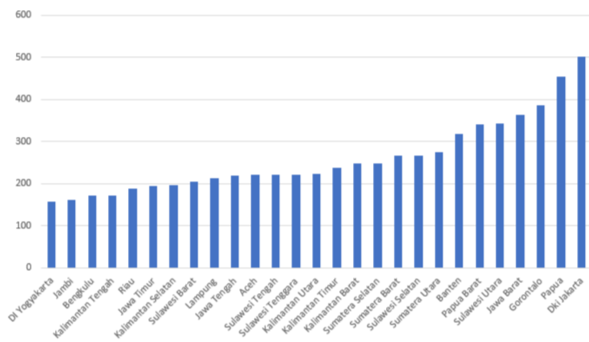

Figure 1. The number of TB case notifications per 100.000 population

Based on the graph, it can be seen that the highest number of TB case notifications per 100,000 population is in DKI Jakarta province and the lowest is in Yogyakarta province. The following is a thematic map to see the pattern of TB case distribution in Indonesia based on data on TB case notification numbers per 100,000 population in Indonesia.

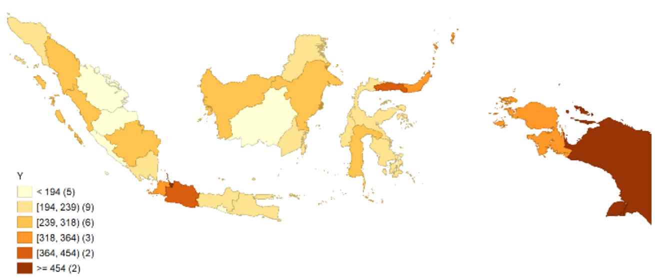

Figure 2. Map of TB case distribution in Indonesia

From the thematic map, it can be seen that several adjacent areas have the same color. This indicates that there is a possibility of a spatial effect on the data on the spread of TB in Indonesia.

3.1. Spatial Dependence

In the preliminary analysis using the least squares method, the estimated model obtained did not meet the residual assumption, namely the data was not normally distributed. Thus, the estimated model produced by the least squares method cannot be accepted. Based on the TB distribution map image in Indonesia, it is feared that there is a spatial effect problem in the data. For this reason, tests were carried out to determine whether the TB distribution data in Indonesia could be modeled using spatial analysis or not. Following are any test should be done to make sure that spatial analysis suitable to be implemented.

a. Moran Index Test

Using equation (4) above, the following Moran Index test is carried out to detect whether there is an auto-correlation problem between areas. The results of the test with the Moran Index are presented in Table 1. Based on Table 1, it can be seen that the variables Y, X_1 and X_3 have the p-value is lower than significant level ($\alpha = 0.05$) which means that H_0 is rejected. This shows that there is spatial auto-correlation between areas in the dependent variable and the independent variable. The dependent variable and at least one of the independent variables indicate spatial auto-correlation between areas, the SDM model can be applied in research.

Table 1. Moran Index Test

Variable	Moran Index Value	Z_{count}	*p-value*	Decision
Y	0.552	3.25	0.0006	Reject H_0
X_1	0.332	2.04	0.0208	Reject H_0
X_2	0.197	1.30	0.0973	Not Reject H_0
X_3	0.611	3.58	0.0002	Reject H_0
X_4	0.104	0.78	0.2165	Not Reject H_0

b. Lagrange Multiplier Test

Lagrange multiplier test consist of two types, there are LM lag and LM error. Following Table 2 presents the results from both tests. Based on both tests, it proves that LM value is higher than $\chi^2_{\alpha(1)} = 3.841$ and p-value from both tests are lower that significant level ($\alpha = 0.05$). The decision of that is reject the hypothesis and it can be concluded that dependence spatial present on dependent variable and error. Therefore, the SDM model can be used in this research.

Table 2. Lagrange Multiplier Test

LM Test	LM Value	*p-value*	Decision
LM lag	8.0567	0.0045	Reject H_0
LM error	4.2508	0.0392	Reject H_0

3.2. Spatial Durbin Model (SDM)

Based on both spatial dependence test, it can be concluded that SDM model can be applied in this data. Estimation process in SDM is done by using MLE method. The result of estimation process to this data are presented in Table 3.

Table 3. Estimation of SDM

Parameter	Coefficient Mean	Standard Error	*p-value*
ρ	0.4483*	0.1446	0.0019
α	0.5679*	0.1981	0.0042
β_1	-0.0443	0.1105	0.6883
β_2	-0.3923*	0.1656	0.0179
β_3	0.2517*	0.1265	0.0466
β_4	0.7852*	0.1681	0.0000
θ_1	0.0149	0.1396	0.9152
θ_2	-0.1638	0.1894	0.3870
θ_3	-0.4905*	0.1680	0.0035
θ_4	0.4724	0.3332	0.1563

*significant at level $\alpha = 0.05$

It can be seen in Table 3 that not all variables entered into the model have a significant effect on the response. Therefore, the model was improved by removing insignificant variables and entering only significant variables into the model. The results of the parameter estimation for Stage 2 are presented in Table 4. The proposed spatial regression model of the spread of TB based on SDM method is as follows:

$$\hat{y} = 0.5609Wy + 0.3849 - 0.3737X_2 + 0.2693X_3 + 0.7356X_4 - 0.7356WX_3.$$

Table 4. Estimated Parameter Using SDM Stage 2

Parameter	Estimate Mean	Std. Error	p-value
ρ	0.5609*	0.1080	0.0000
α	0.3849*	0.1154	0.0009
β_2	-0.3737*	0.1465	0.0107
β_3	0.2693*	0.1221	0.0274
β_4	0.7356*	0.1432	0.0000
θ_3	-0.7356*	0.1374	0.0063

*significant at level $\alpha = 0.05$

3.2. Spatial Outlier Detection

Spatial outliers from the residuals of the proposed model will be detected using Moran's Scatterplot. Figure 3 presents the Moran's scatterplot of the residuals of the model. From the graph, it is known that there are outliers in the 6th, 13th, and 24th data. Therefore, the proposed model based on SDM could not be accepted. A method that is more resistant to the presence of these outliers is needed.

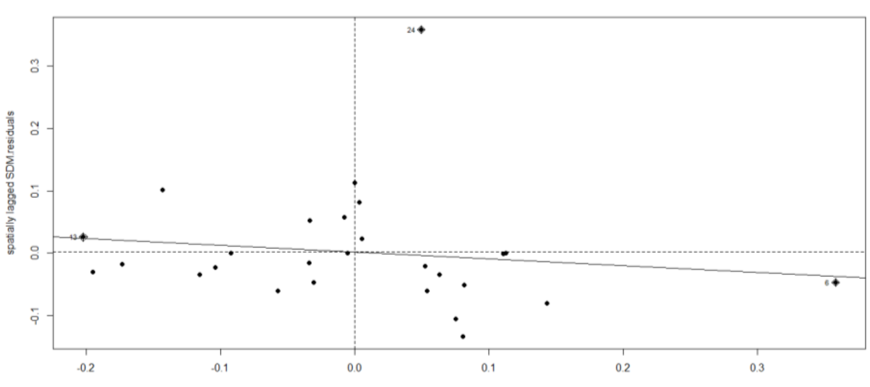

Figure 3. Outlier Detection on Proposed SDM Model.

3.3. Robust Spatial Durbin Model (RSDM) Using M-Estimator

The estimated value of RSDM parameters is obtained based on the results of the iteration process using the IRLS method. The modeling results with this method are presented in Table 5.

Table 5. Iteration Process to Estimate Parameter Based on RSDM

Parameter	1th Iteration	2nd Iteration	3th Iteration	4th Iteration
α	0.3817184	0.3803433	0.3800011	0.3799287
β_2	-0.3825581	-0.3828311	-0.382663	-0.3826032
β_3	0.2765417	0.2775277	0.2776683	0.2776908
β_4	0.7323289	0.7319185	0.7318258	0.7318061
θ_4	-0.3621845	-0.3599960	0.3597005	-0.3596668

Based on Table 5, it is known that the iteration stops at the 4th iteration where the difference between the values $\hat{\delta}^{(4)}$ and $\hat{\delta}^{(3)}$ is less than 0.0001. The estimated values using the RSDM method based on the final results of the iteration process are presented in Table 6. The proposed model based on RSDM approach is as following:

$$\hat{y} = 0.5609Wy + 0.3799 - 0.3826X_2 + 0.2777X_3 + 0.7318X_4 - 0.3597WX_3$$

Table 6. Estimated Parameter Using RSDM.

Parameter	Estimate Mean	Std. Error	p-value
ρ	0.5609*	0.1215	0.0000
α	0.3799*	0.1194	0.0015
β_2	-0.3826*	0.1461	0.0088
β_3	0.2777*	0.1139	0.0147
β_4	0.7318*	0.1328	0.0000
θ_4	-0.3597*	0.1338	0.0072

3.4. Best Model Selection

A measure of model goodness is determined based on the R^2_{adj} value and MSE value generated from the SDM model and RSDM. The following presents a comparison between the two models to determine the best regression model. From Table 7, it is known that the RSDM model is better than the SDM model, as indicated by a larger value of R^2_{adj} and a smaller MSE.

Table 7. Comparison of Model Goodness Measures Using SDM and RSDM

Model	R^2_{adj} Value	MSE
SDM	73.06%	0.0138
RSDM	80.90%	0.0113

4. CONCLUSIONS

This study aims to construct a model of the spread of TB disease in Indonesia. In the preliminary analysis, it was found that the data were not normally distributed and there was a significant spatial influence on the response. Therefore, the analysis of the SDM and RSDM models was used to produce the model. This study verifies that the RSDM model is able to produce a model with a larger value of R^2_{adj} and a smaller MSE value. The estimated model of the spread of TB cases in Indonesia produced by the RSDM model with the M estimator is as follows:

$$\hat{y} = 0,5609Wy + 0,3799 - 0,3799X_2 + 0,2777X_3 + 0,7318X_4 - 0,3597WX_3.$$

Factors that significantly influence the number of TB case notifications in Indonesia are the variables of the percentage of households with access to proper sanitation (X_2), number of HIV cases (X_3), population density and spatial lag of the number of HIV cases (WX_3).

ACKNOWLEDGMENT

This research was funded by LPPM, Andalas University of Indonesia under *Penelitian Skripsi Sarjana* (PSS) fund, with Contract Number 155/UN16.19/PT.01.03/PSS/2024.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] WHO, Global tuberculosis report 2023, World Heath Organization, Geneva, 2023.

[2] Y. Zhang, J. Jiang, Y. Feng, Penalized quantile regression for spatial panel data with fixed effects, Commun. Stat. - Theory Methods 52 (2023), 1287–1299. https://doi.org/10.1080/03610926.2021.1934028.

[3] J. Zhang, Q. Lu, L. Guan, et al. Analysis of factors influencing energy efficiency based on spatial quantile autoregression: evidence from the panel data in China, Energies 14 (2021), 504. https://doi.org/10.3390/en14020504.

[4] L. Anselin, Spatial econometrics: methods and models, Kluwer Academic Publishers, Dordrecht, 1988.

[5] R.D. Bekti, W.R. Rahmadhani, N. Noeryanti, Robust spatial Durbin model for modeling open unemployment rates in Central Java Province, JNANALOKA 4 (2023), 91–103. https://doi.org/10.36802/jnanaloka.2022.v4-no2-91-103.

[6] A.R. Hakim, B. Warsito, H. Yasin, Live expectancy modelling using spatial durbin robust model, J. Phys.: Conf. Ser. 1655 (2020), 012098. https://doi.org/10.1088/1742-6596/1655/1/012098.

[7] M.H. Mukrom, H. Yasin, A.R. Hakim, Pemodelan angka harapan hidup Provinsi Jawa Tengah menggunakan robust spatial Durbin model, J. Gaussian 10 (2021), 44–54. https://doi.org/10.14710/j.gauss.v10i1.30935.

[8] H.N. Khofifah, Robust spatial Durbin model (RSDM) untuk pemodelan tingkat pengangguran terbuka (TPT) di Provinsi Jawa Barat, J. Riset Stat. 1 (2022), 135–142. https://doi.org/10.29313/jrs.v1i2.522.

[9] U.A. Syam, S. Siswanto, N. Sunusi, Robust spatial Durbin modelling on tuberculosis data using the MM-estimator method, Stat. Transit. New Ser. 25 (2024), 23–38. https://doi.org/10.59170/stattrans-2024-013.

[10] W. Sanusi, H. Ihsan, H. Syam, Model regresi spasial dan aplikasinya dalam menganalisis angka putus sekolah usia wajib belajar di Provinsi Sulawesi Selatan, J. Math. Comput. Stat. 1 (2018), 183–192.

[11] H. Yasin, B. Warsito, A.R. Hakim, Regresi spasial (Aplikasi dengan R), Wade Group, Pekalongan, 2020.

[12] J. LeSage, R.K. Pace, Introduction to spatial econometrics, Taylor & Francis, 2009.

[13] A. Mahmud, E. Pasaribu, Permodelan spasial pada analisis faktor yang mempengaruhi tingkat pengangguran terbuka Provinsi Bangka Belitung tahun 2018, Eng. Math. Comput. Sci. J. 3 (2021), 47–58. https://doi.org/10.21512/emacsjournal.v3i2.7034.

[14] F. Yanuar, T. Abrari, I.R. Hg, The construction of unemployment rate model using SAR, quantile regression, and

SARQR model, Pak. J. Stat. Oper. Res. 19 (2023), 447–458. https://doi.org/10.18187/pjsor.v19i3.4241.

[15] L. Anselin, Spatial econometrics: methods and models, Springer, Dordrecht, 1988. https://doi.org/10.1007/978-94-015-7799-1.

[16] S. Shekhar, C.T. Lu, P. Zhang, A unified approach to detecting spatial outliers, GeoInformatica 7 (2003), 139–166. https://doi.org/10.1023/A:1023455925009.

[17] A.R. Hakim, B. Warsito, H. Yasin, Live expectancy modelling using spatial Durbin robust model, J. Phys.: Conf. Ser. 1655 (2020), 012098. https://doi.org/10.1088/1742-6596/1655/1/012098.

[18] R.E. Walpole, R.H. Myers, S.L. Myers, et al. Probability & statistics for engineers & scientists, Pearson, New York, 2011.

[19] Badan Pusat Statistik, Statistik Indonesia 2023, Badan Pusat Statistik Indonesia, Jakarta, 2023.