Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2026, 2026:25
https://doi.org/10.28919/cmbn/9813
ISSN: 2052-2541

ROBUST SPATIAL DURBIN MODEL FOR DENGUE FEVER CASES IN INDONESIA

FENNI KURNIA MUTIYA[1,*], FERRA YANUAR[2], TESSY OKTAVIA MUKHTI[1], FITRI MUDIA SARI[1], HENDRY FRANANDA[3]

[1]Department of Statistics, Universitas Negeri Padang, Padang, Indonesia

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

[3]Department of Geography, Universitas Negeri Padang, Padang, Indonesia

Copyright © 2026 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 employed for data exhibiting spatial dependencies. The Spatial Durbin Model (SDM) is one such method that accounts for spatial effects in both the dependent and independent variables. However, the presence of spatial outliers can compromise the accuracy of SDM predictions. To address this issue, a robust method is needed, namely the Robust Spatial Durbin Model (RSDM). This study applies the RSDM to model the factors influencing the spread of dengue haemorrhagic fever cases across Indonesia and to identify the superior modeling approach. The results indicate that the RSDM outperforms the standard SDM, evidenced by a higher Adjusted R^2 and a lower Mean Squared Error (MSE). Key factors identified as significantly influencing dengue haemorrhagic fever cases are population density, the number of doctors in healthcare facilities, and the percentage of the population covered by health insurance.

Keywords: dengue haemorrhagic fever cases; spatial durbin model; m-estimator; robust spatial durbin model.

2020 AMS Subject Classification: 62H11, 62P10.

---

[*]Corresponding author
E-mail address: fennikurnia@unp.ac.id
Received February 06, 2026

1. INTRODUCTION

Dengue Hemorrhagic Fever (DHF) is one of the infectious diseases caused by a virus and transmitted by vectors. The virus that causes this disease is Dengue. The vectors of this disease come from the Aedes aegypti and Aedes albopictus mosquito species [1]. Aedes mosquitoes prefer stagnant or water-holding places such as ditches, vases or plant pots, pet drinking places, swimming pools, or trash bins as breeding sites. The characteristics of the vectors determine the distribution and timing of infection occurrence [2]. Aedes mosquitoes generally inhabit areas with a tropical climate, high rainfall, and hot and humid temperatures [3].

The incidence of dengue has grown dramatically worldwide in recent decades, with the number of cases reported to WHO increasing from 505,430 cases in 2000 to 14.6 million in 2024 [4]. Most cases are asymptomatic or mild and self-managed, and hence the actual numbers of dengue cases are under-reported. The disease is now endemic in more than 100 countries. In 2024, more cases of dengue were recorded than ever before in a 12-month period, affecting over 100 countries on all continents. During 2024, ongoing transmission, combined with an unexpected spike in dengue cases, resulted in a historic high of over 14.6 million cases and more than 12,000 dengue-related deaths reported. The Region of the Americas contributed a significant proportion of the global burden, with over 13 million cases reported to WHO. Based on the Indonesian Ministry of Health, in 2024 there were 257,271 DHF cases with a total of 1,461 deaths [5]. This number represents a significant increase compared to 2023, which had 114,720 cases and 894 deaths.

DHF is a communicable disease whose transmission is inherently spatial, spreading between contiguous geographic areas [6]. Consequently, analyzing the factors driving its spread necessitates methods that account for this spatial dependence [7] [8]. While standard multiple regression can identify influencing factors, it is inadequate for data with inherent spatial autocorrelation [9]. When such spatial effects are present due to inter-regional influences, spatial regression analysis becomes essential. The Spatial Durbin Model (SDM) is one robust method within this framework for this purpose [10], [11], [12].

SDM is a regression method that captures spatial effects from both dependent and independent variables. However, spatial data often contain outliers, which can bias parameter estimates in the model. To address data contaminated with outliers, a robust modeling method is needed against outliers, namely the Robust Spatial Durbin Model (RSDM). 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 [13]. Yanuar et al [12] applied RSDM with M estimator in modelling the spreading of tuberculosis in Indonesia. Hermalia et al. [14] applied SDM and Spatial Autoregressive to model the open unemployment rate in West Java. Wibowo et al. [15] investigated the modeling of crime in East Java Province using RSDM and M-estimator. This study carried out an analysis of RSDM in other fields, namely modeling the DHF cases in Indonesia using M-estimator.

2. DATA AND METHODS

2.1. Data

Data used in this study are secondary data sourced from the Indonesia Health Survey https://kemkes.go.id/id/profil-kesehatan-indonesia-2024 in 2024. The response variable used is the number of dengue fever cases in 38 provinces of Indonesia, but there was no dengue fever case in South Papua. In this study, the predictor variables are population density, percentage of poor population, the number of doctors in healthcare facilities, percentage of health insurance, and percentage of households with access to proper sanitation. The predictor variables were derived from Central Bureau of Statistics Indonesia in 2024. The detail descriptive statistics about the response and predictor variables are presented in Table 1.

Table. 1. Descriptive Statistics of Variables

Variable	Mean	Minimum	Q_1	Median	Q_3	Maximum
Number of DHF case (Y)	6953.27	249	1914	3232	6263	61423
Population density (X_1)	698.4865	10	42	103	268	16155
Percentage of poor population (X_2)	10.43	3.8	5.7	10	12.52	29.66
The number of doctors in healthcare facilities (X_3)	3123.432	187	760	1602	3366	17449
Percentage of health insurance (X_4)Percentage of the population covered by health insurance (X_4)	76.6068	58.06	69.58	75	83.55	97.28
Percentage of households with access to proper sanitation (X_5)	81.6932	12.61	80.51	84	87.31	96.83

2.2. Spatial Regression

As an advancement of multiple linear regression, spatial regression integrates the influence of location (spatial effect) into the data analysis [16]. It is, therefore, a statistical method designed to identify relationships between dependent and independent variables by explicitly considering spatial dependencies. The general formulation of this model can be expressed as [17]:

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

In this model, the response variable vector is $\boldsymbol{y}$, 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.

2.3. Spatial Weighting Matrix

To model how districts influence each other spatially, a queen contiguity matrix was constructed [18]. This framework identifies neighboring districts through shared borders or vertices, incorporating both lateral and diagonal connections. Row standardization was applied, normalizing the weights of each district's neighbors to sum to one. This step was crucial for maintaining regional comparability and facilitating the interpretation of spatial coefficients in our model. The matrix $\boldsymbol{W}$**,** known as the spatial weight matrix, is an $n \times n$ a symmetric structure that models the closeness of neighboring regions [19]. The spatial weight matrix can be constructed in two forms: a standardized weight matrix ($\boldsymbol{W}$) and an unstandardized weight matrix ($\boldsymbol{W}^*$). The components of the weight matrix $\boldsymbol{W}$ are w_{ij} with I is the row and j is the column, with $j = 1,2, \ldots, n$. In obtaining the value of w_{ij} is formulated as follows:

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

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

2.4. Spatial Dependence

Spatial dependence indicates the existence of assemblies between the locations of the research objects [20]. It can be detected using Moran Index test. Moran's Index can be formulated as follows[21]:

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

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. The existence of spatial dependence on the dependent variable is also checked using Lagrange Multiplier lag $(LM\ lag)$ with the general form as follows [22]:

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

where $D = \left(\frac{(\boldsymbol{W}\,\boldsymbol{X}\boldsymbol{\beta})^T\,(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}^T\,\boldsymbol{X})^{-1}\boldsymbol{X}^T\,)\boldsymbol{W}\,\boldsymbol{X}\boldsymbol{\beta}}{s^2}\right) + tr(\boldsymbol{W}^T\boldsymbol{W} + \boldsymbol{W}^2)$ 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:

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

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. The Robust Spatial Durbin Model is applied to estimate the model parameters when spatial dependence exists in the dependent variable and the errors, and/or when outliers are present in the model.

2.5. 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 $(\boldsymbol{X})$ [23]. The hypothesis for this mode is presentin the following form [24]:

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

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 $k \times 1$, $\boldsymbol{Z} = [\boldsymbol{1_n} \quad \boldsymbol{X} \quad \boldsymbol{WX}]$, $\boldsymbol{\delta}$=$[\alpha \quad \boldsymbol{\beta} \quad \boldsymbol{\theta}]^T$.

2.6. Spatial Outlier

Outliers are observations that show significant differences with other observations in a data set collection [25]. One graphical method fr identifying 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 [26]. Data plots in the upper left and lower right quadrants can be said to be spatial outliers. Spatial outlier can also be detected mathematically using the formula:

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

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.

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

When the error term of SDM contains outliers, Robust Spatial Durbin Model (RSDM) is applied [27]. 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 [28]:

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

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

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

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

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

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

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

where s is robust estimate scale obtained from:

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

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

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

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:

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

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.

2.8. Selection of The Best Model

The selection of the best model is based on the value of the coefficient of determination (R^2) and Mean Squared Error (MSE) [19]. 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.

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

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^2_{adj} value and a smaller MSE value.

3. MAIN RESULTS

The following is a graph showing the number of DHF cases on each province in Indonesia.

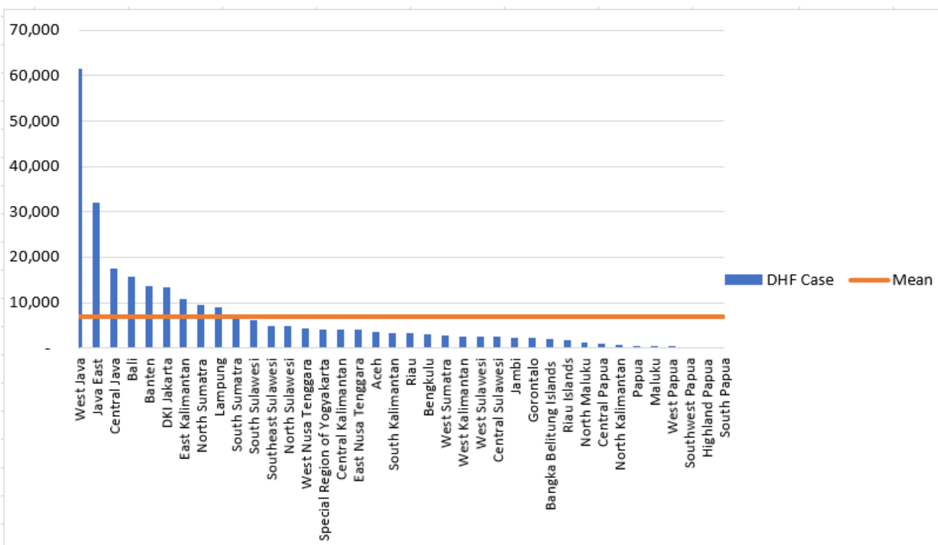

Figure 1. The number of Dengue Haemorrhagic Fever Cases in the provinces of Indonesia

Based on the graph above, it appears that the highest number of DHF cases is in West Java province and the lowest is in Highland Papua province. There were no cases of DHF found in South Papua. The following is a thematic map showing the pattern of DHF cases in Indonesia.

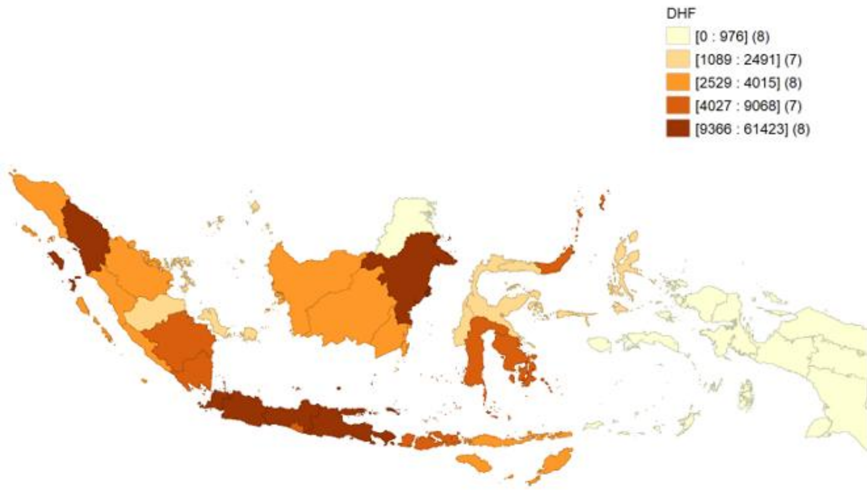

Figure 2. Map of DHF casescase distribution in Province of Indonesia

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 cases of DHF in Indonesia.

3.1. Spatial Dependence

In the initial least squares analysis, the resulting model violated the residual assumption as the data was not normally distributed. Consequently, the model produced by this method is invalid. Given the pattern of DHF distribution in Indonesia, spatial effects are suspected in the data. Therefore, tests were conducted to determine the suitability of spatial analysis for modeling this data. The following tests are necessary to confirm whether spatial analysis is an appropriate approach.

a. Moran Index Test

Spatial autocorrelation is examined via the Moran Index test, by using equation (4) above. As presented in Table 2, the test results indicate that variables Y, X_2, X_3 and X_5 have the p-value is lower than significant level ($\alpha = 0.05$), resulting in rejection of H_0. This indicates spatial autocorrelation in both the dependent and at least one independent variable, justifying the application of the Spatial Durbin Model (SDM) in this research.

Table 2. Moran Index Test

Variable	Moran Index Value	$Z_{statistic}$	*p-value*	Decision
Y	0.3615	2.4137	0.0079	Reject H_0
X_1	0.1153	0.9088	0.1817	Not Reject H_0
X_2	0.7384	4.7174	0.0000	Reject H_0
X_3	0.6960	4.4581	0.0000	Reject H_0
X_4	0.1051	0.8463	0.1987	Not Reject H_0
X_5	0.4381	2.8818	0.0019	Reject H_0

b. Lagrange Multiplier Test

The lagrange multiplier test, which includes LM lag and LM error test, was conducted. Table 3 presented the results, showing 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 3. Lagrange Multiplier Test

LM Test	LM Value	*p-value*	Decision
LM lag	3.8764	0.0088	Reject H_0
LM error	4.3423	0.0055	Reject H_0

3.2. Spatial Durbin Model (SDM)

The spatial dependence tests confirm that SDM model is appropriate for this data. The model was estimated using Estimation process in SDM is done by using Maximum Likelihood Estimation (MLE) method, and the results are shown in Table 4.

Table 4.

Parameter	Coefficient Mean	Standard Error	*p-value*
ρ	-0.3500*	0.1613	0.0300
α	-0.3736*	0.0756	0.0319
β_1	-0.1074	0.0679	0.1137
β_2	0.0573	0.0639	0.3694
β_3	0.8370*	0.0535	0.0000
β_4	0.3552*	0.0331	0.0095
β_5	0.0411	0.0692	05523
θ_1	0.6505*	0.1161	0.0000
θ_2	-0.0527	0.0528	0.3191
θ_3	-0.1579	0.1247	0.2055
θ_4	-0.0117	0.0613	0.8487
θ_5	0.0451	0.0379	0.2356

*significant at level $\alpha = 0.05$

As seen in the result of Table 4 that not all variables entered into the model have a significant effect on the response. Thus, the model was enhanced through the exclusion of non-significant variables and the inclusion of only those that were statistically significant. Table 5 presents the Stage 2 parameter estimation results. The proposed spatial regression model of the DHF casescase based on SDM method is formulated follows:

$$\hat{y} = -0.3165Wy - 0.1602 + 0.8459X_3 + 0.3973X_4 + 0.6847WX_1.$$

Table 5. Estimated Parameter Using SDM Stage 2

Parameter	Estimate Mean	Std. Error	p-value
ρ	-0.3165*	0.1665	0.0474
α	-0.1602*	0.0217	0.0460
β_3	0.8459*	0.0588	0.0000
β_4	0.3973*	0.0337	0.0238
θ_1	0.6847*	0.1109	0.0000

*significant at level $\alpha = 0.05$

3.2. Spatial Outlier Detection

Moran's Scatterplot was applied to the model's residuals to identify spatial outliers. As shown in Figure 3, the scatterplot revelas outliers in the 2nd, 3rd, 6th, 9th and 10th 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.

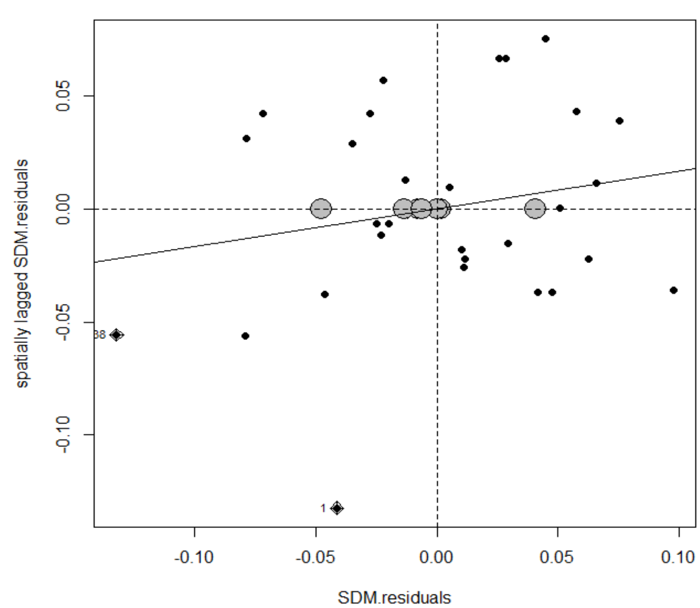

Figure 3. Outlier Detection on Proposed SDM Model.

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

The estimated RSDM parameter values are derived from the iterative process utilizing the IRLS method. The modeling outcomes achieved through this method are shown in Table 6.

Table 6. Iteration Process to Estimate Parameter Based on RSDM

Parameter	1th Iteration	2nd Iteration	3th Iteration	4th Iteration
α	0.004189	0.008448	0.0104736	0.0107447
β_3	0.687922	0.650838	0.6330299	0.6233739
β_4	0.010486	0.009991	0.0074774	0.0079907
θ_1	0.413688	0.361773	0.3701669	0.3755388

Parameter	5th Iteration	6th Iteration	7th Iteration	8th Iteration
α	0.0109049	0.0109847	0.0110275	0.0110596
β_3	0.6175131	0.6138533	0.6115236	0.6099618
β_4	0.0082369	0.0084374	0.0085861	0.0086767
θ_1	0.3788759	0.3809688	0.3823006	0.3831926

According to Table 6, it is known that the iteration stops at the 8th iteration where the difference between the values $\hat{\delta}^{(7)}$ and $\hat{\delta}^{(8)}$ is less than 0.001. The estimated values using the RSDM method based on the final results of the iteration process are presented in Table 7. The proposed model based on RSDM approach is as follows:

$$\hat{y} = -0.3165Wy + 0.0111 + 0.6099X_3 + 0.0087X_4 - 0.3832WX_1$$

Table 7. Estimated Parameter Using RSDM.

Parameter	Estimate Mean	Std. Error	p-value
ρ	-0.3165*	0.1125	0.0305
α	0.0111*	0.0114	0.0015
β_3	0.6099*	0.0606	0.0000
β_4	0.0087*	0.0247	0.0073
θ_1	-0.3832*	0.1072	0.0004

3.4. Best Model Selection

The goodness of fit for the models is evaluated using R^2_{adj} value and MSE value produced by both the SDM and RSDM model. A comparative analysis between the two models is presented below to identify the superior regression model. From Table 8, 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 8. Comparison of Model Goodness Measures Using SDM and RSDM

Model	R^2_{adj} Value	MSE
SDM	74.83%	0.0714
RSDM	82.88%	0.0013

4. CONCLUSIONS

This study seeks to develop a model of the DHF cases in Indonesia. Preliminary analysis revealed that the data was not normally distributed and exhibited significant spatial dependence on the response variable. Consequently, the Spatial Durbin Model (SDM) and Robust Spatial Durbin Model (RSDM) were employed for the analysis. This study confirms that the RSDM model is able to produce a model with a larger value and a smaller MSE value. The estimated model of DHF cases in Indonesia produced by the RSDM model with the M estimator is as follows:

$$\hat{y} = -0.3165Wy + 0.0111 + 0.6099X_3 + 0.0087X_4 - 0.3832WX_1$$

Factors that significantly influence the number of DHF cases notifications in Indonesia are the variables of the number of doctors in healthcare facilities (X_3), percentage of health insurance (X_4), population density and spatial lag of population density (WX_1).

Factors that significantly influence the number of DHF cases notifications in Indonesia are the variables of the number of doctors in healthcare facilities (X_3), percentage of the population covered by of health insurance (X_4), population density and spatial lag of population density (WX_1).

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] S.B. Halstead, Dengue Virus–Mosquito Interactions, Annu. Rev. Entomol. 53 (2008), 273-291. https://doi.org/10.1146/annurev.ento.53.103106.093326.

[2] E.P. Astuti, P.W. Dhewantara, H. Prasetyowati, M. Ipa, C. Herawati, et al., Paediatric Dengue Infection in Cirebon, Indonesia: A Temporal and Spatial Analysis of Notified Dengue Incidence to Inform Surveillance, Parasites Vectors 12 (2019), 186. https://doi.org/10.1186/s13071-019-3446-3.

[3] E.A. Mordecai, J.M. Cohen, M.V. Evans, P. Gudapati, L.R. Johnson, et al., Detecting the Impact of Temperature on Transmission of Zika, Dengue, and Chikungunya Using Mechanistic Models, PLOS Neglected Trop. Dis. 11 (2017), e0005568. https://doi.org/10.1371/journal.pntd.0005568.

[4] World Health Organization, Dengue and Severe Dengue, 2024. https://www.who.int/health-topics/dengue-and-severe-dengue.

[5] Kementerian Kesehatan Republik Indonesia, Profil Kesehatan Indonesia 2024. https://kemkes.go.id/id/profil-kesehatan-indonesia-2024.

[6] Z.Y. Yeo, H.C. Hapuarachchi, J. Ong, G. Yeo, C. Koo, et al., Inferring the Hidden and Long-Range Dengue Transmission Routes in Singapore, Mach. Learn.: Health 1 (2025), 015015. https://doi.org/10.1088/3049-477X/ae16ab.

[7] Y. Yang, O. Doğan, S. Taspinar, Model Selection and Model Averaging for Matrix Exponential Spatial Models, Econ. Rev. 41 (2022), 827-858. https://doi.org/10.1080/07474938.2022.2047507.

[8] S.A. Ania, N. Ilyas, E.T. Herdiani, Robust Spatial Temporal Analysis with Improved Geographically and Temporally Weighted Regression Model of Dengue Incidence Rate, Commun. Math. Biol. Neurosci. 2024 (2024), 22. https://doi.org/10.28919/cmbn/8376.

[9] Y. Notonegoro, Y. Andriyana, B.N. Ruchjana, Comparison of Distance-Based Spatial Weight Matrix in Modeling Internet Signal Strengths in Tasikmalaya Regency Using Logistic Spatial Autoregressive Model, Int. J. Data Netw. Sci. 8 (2024), 893-906. https://doi.org/10.5267/j.ijdns.2023.12.016.

[10] B. Haryanto, F. Dwirahmadi, T. Nurlambang, A. Asyary, Spatial Analysis on Dengue Fever Vulnerability in the Provinces of South Sulawesi and East Nusa Tenggara in Indonesia, Ann. Glob. Health 91 (2025), 86. https://doi.org/10.5334/aogh.4915.

[11] A. Fithriyyah, T. Purwaningsih, S. Konate, M.A.A. Abdalla, Mapping Dengue Vulnerability: Spatial Cluster Analysis Reveals Patterns in Central Java, Indonesia, Sci. Inf. Technol. Lett. 4 (2023), 109-122. https://doi.org/10.31763/sitech.v4i2.1203.

[12] F. Yanuar, D.P. Sari, Y. Asdi, F.K. Mutiya, Robust Spatial Durbin Approach in Modelling the Spreading of Tuberculosis in Indonesia, Commun. Math. Biol. Neurosci. 2024 (2024), 129.

https://doi.org/10.28919/cmbn/8920.

[13] F.F. Fauzi, Gabriella Hilary Wenur, Rochdi Wasono, Spatial Durbin Model of Unemployment Rate in Central Java, Parameter: J. Stat. 3 (2023), 7-18. https://doi.org/10.22487/27765660.2023.v3.i1.16423.

[14] L. Hermalia, Modeling the Open Unemployment Rate of Regency/City in West Java Province in 2021 Using Spatial Autoregresive Moving Average and Spatial Durbin Model, J. Stat. Data Sci. 2 (2024), 51-61. https://doi.org/10.33369/jsds.v2i2.30380.

[15] Y. Farida, M. Farmita, P.K. Intan, H. Khaulasari, A.T. Wibowo, Modeling Crime in East Java Using Spatial Durbin Model Regression, BAREKENG: J. Ilmu Mat. Dan Ter. 18 (2024), 1497-1508. https://doi.org/10.30598/barekengvol18iss3pp1497-1508.

[16] I. Yahya, B. Abapihi, M. Agusrawati, G.N.A. Wibawa, Spatial Error Model to Analyze Morbidity Rate in Indonesia, Int. J. Sci. Technol. Eng. Math. 3 (2023), 18-35. https://doi.org/10.53378/352998.

[17] W. Sanusi, H. Ihsan, N.H. Syam, Model Regresi Spasial dan Aplikasinya dalam Menganalisis Angka Putus Sekolah Usia Wajib Belajar di Provinsi Sulawesi Selatan, J. Math. Comput. Stat. 1 (2019), 183-192. https://doi.org/10.35580/jmathcos.v1i2.9241.

[18] X. Zhou, H. Lin, Spatial Weights Matrix, in: Encyclopedia of GIS, Springer, 2008: pp. 1113-1113. https://doi.org/10.1007/978-0-387-35973-1_1307.

[19] 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.

[20] M.M. Fischer, J.P. LeSage, Spatial Econometrics, Working Papers in Regional Science Vol. 02, Vienna University of Economics and Business, 2025. https://doi.org/10.57938/1019c1c5-41ba-4ffc-a61f-0c645889aa3b.

[21] S. Liu, P. Wu, The Impact of Urban Sprawl on Green Total Factor Productivity: A Spatial Econometric Analysis in China, Front. Environ. Sci. 11 (2023), 1095349. https://doi.org/10.3389/fenvs.2023.1095349.

[22] L. Anselin, Spatial Econometrics, in: A Companion to Theoretical Econometrics, Wiley, 2003: pp. 310-330. https://doi.org/10.1002/9780470996249.ch15.

[23] T. Rüttenauer, N. Kapelle, Panel Data Analysis, Preprint (2024). https://doi.org/10.31235/osf.io/3mfzq_v1.

[24] Y. Andriyana, A.N. Falah, B.N. Ruchjana, A. Sulaiman, E. Hermawan, et al., Spatial Durbin Model with Expansion Using Casetti's Approach: A Case Study for Rainfall Prediction in Java Island, Indonesia, Mathematics 12 (2024), 2304. https://doi.org/10.3390/math12152304.

[25] J. John, Outlier Detection and Spatial Analysis Algorithms, arXiv:2106.10669, 2021. https://doi.org/10.48550/arXiv.2106.10669.

[26] 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.

[27] H. Yasin, B. Warsito, A.R. Hakim, Regresi Spasial (Aplikasi dengan R), WADE Group, 2014. https://eprints2.undip.ac.id/id/eprint/3926/1/Buku%20Regresi%20Spasial%20Aplikasi%20dengan%20R.pdf.

[28] 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.