CAUCHY –Jurnal Matematika Murni dan Aplikasi
Volume 7(4) (2023), Pages 535-547
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Spatial Analysis of Dengue Disease in Jakarta Province

Muhamad Sobari[1,4]*, I Gede Nyoman Mindra Jaya[2], Budi Nurani Ruchjana[3]

[1]Post-graduate program in Applied Statistics, Universitas Padjadjaran, Bandung
[2]Department of Statistics, Universitas Padjadjaran, Bandung
[3]Department of Mathematics, Universitas Padjadjaran, Bandung
[4]Central Bureau of Statistics of Tasikmalaya Regency, West Java

Email: muhamad21059@mail.unpad.ac.id

ABSTRACT

Dengue disease is a virus-borne illness spread by the bite of the female Aedes aegypti mosquito. Jakarta is at risk for dengue disease because it has a lot of people who live in urban slums. The objective of this study is to identify the risk factors that affect the number of dengue disease cases in Jakarta by considering spatial dependencies. This study used a spatial autoregressive (SAR) model with the queen contiguity spatial weight matrix to account for the spatial dependence. The number of dengue disease cases in Jakarta is strongly affected by the number of flood-prone areas, the number of slum neighborhood associations, the population density, the number of hospitals and the number of public health centers per 1,000 people also the spatial lag. When dengue disease cases increase in one sub-district, the number of dengue disease cases in the sub-districts around it will increase as well because of the positive and significant spatial lag coefficient. Based on the direct impact, each additional one percent of flood-prone points in one sub-district will increase the number of dengue disease cases in that sub-district by four cases. This research contributes that flood control is important in Jakarta to control dengue disease.

Keywords: Dengue disease; Spatial Autoregressive; Queen Contiguity

INTRODUCTION

The female Aedes aegypti mosquito bite is the primary method of transmission for the dengue virus, which causes dengue disease. These mosquitoes can breed in slum places such as there are pools of water that are not taken care of, dark, and damp [1]. Dengue disease is still one of the main health problems that threaten the Indonesian people because Indonesia is a tropical country that is vulnerable to vector-borne diseases, that is diseases that increase the likelihood and risk of occurrence due to changes in the weather cycle [2].

Dengue virus is highly sensitive to changes in average temperature, humidity, and increased rainfall which can affect the life cycle and reproduction of the Aedes aegypti mosquito that carries the virus [2], [3]. And Indonesia is a tropical country with changes in two seasonal cycles, that is the dry season and the rainy season, so that dengue disease will increase during the rainy season because the weather conditions become humid and waterlogging often occurs due to water channels that do not flow or post-floods that are

Submitted: August 30, 2022 Reviewed: February 25, 2023 Accepted: March 04, 2023
DOI: http://dx.doi.org/10.18860/ca.v7i4.17423

not cleaned immediately.

Based on research [1], slum areas have more potential to become a breeding ground for Aedes aegypti mosquitoes, so it is necessary to have a program to eradicate mosquito nests in slum areas to eradicate mosquitoes that carry the dengue virus to reduce dengue disease cases. Meanwhile, according to research [4], areas that often flood due to high rainfall have a vulnerability in the health sector, that is dengue disease because of climate change. So, it can be concluded that flooding has an indirect effect on dengue disease through climate change.

Several studies show that dengue disease is related to mobility and population density and people's living behavior. As done by [5], [6] shows that there will be a rise in dengue disease cases because of increased population density. This shows that dengue disease spreads more easily in areas with a high population density. Mosquitoes live in the tropics with warm temperatures, in areas below 1,000 meters sea level in Indonesia [7]. Jakarta, in particular, is located in a region with ideal conditions for mosquito breeding.

In 2019, the province with the highest population density in Indonesia which reached 15,328 people per km^2 was Jakarta [8]. Jakarta Province is also the highest percentage of urban slum households (the lowest 40 percent of the population) which reached 42.73 percent. Urban slum households are defined as households that: do not have access to safe drinking water sources; do not have access to proper sanitation; do not have access to floor area >= 7.2 m^2 per capita; and do not have access to proper roof, floor, and wall conditions [9].

Jakarta Province is vulnerable to the transmission of dengue disease due to high population density, percentage of urban slum households, some characteristics that may flood if the rainfall is high, and an optimum temperature for the breeding of the Aedes aegypti mosquito. This is evident in 2019 the dengue disease morbidity rate per 100,000 population Jakarta Province is the top 10 provinces in Indonesia which reached 82.45 [10].

Research on the spread of dengue disease with a spatial approach has been carried out by [11] using the Moran’s I method and the Local Indicators of Spatial Association (LISA) which shows that dengue disease cases, population, population density, temperature, rainfall, and wind speed have positive spatial autocorrelation between villages in Padang municipality on the six variables. Furthermore, research [12] using the Moran and Geary's C Index method shows that dengue disease transmission in Semarang municipality exhibits spatial autocorrelation. Both studies have the same conclusion, that is there is a positive spatial autocorrelation in the number of dengue disease cases.

Another study was conducted by [5] using the Spatial Autoregressive (SAR) model, and the study's findings show that factors that significantly influence the number of dengue disease cases in Central Java Province are number of protected spring facilities, population percentage access to sustainable drinking water, population density, number of village polyclinics per 1,000 population, number of public health centers per 1,000 population, and percentage of clean water quality free of bacteria, fungi, and chemicals. In addition, research by [13] comparing Spatial Durbin Model (SDM) and SAR model revealed that SAR performed better than SDM in predicting the factors that influence the transmission of dengue disease in Central Java Province. In general, the number of residents and the average length of schooling are factors that affect the spread of dengue disease in Central Java Province. The two studies have something in common, that is the unit observation and analysis is the regency and municipality in Central Java Province. This study looks at how various environmental and social issues in Jakarta influence the spread of dengue disease by accounting for spatial dependencies.

Relying on this background, the objective of this study is to identify risk factors that significantly affect the number of dengue disease cases in Jakarta Province, using sub-districts as observation and analysis units.

METHODS

Data and Variables

This study relied on secondary data from the Central Bureau of Statistics of Jakarta Province (BPS Jakarta Province) and the Provincial Government of Jakarta's Office of Communication, Information, and Statistics (Diskominfotik Jakarta Province). The data were obtained from publications published by BPS Jakarta [14] and from websites managed by Diskominfotik Jakarta Province [15]. The unit observation and analysis in this study covers 42 out of 44 sub-districts Jakarta Province in 2017. The data used are not all sub-districts in Jakarta Province and 2017 dataset is due to the limited data availability. The two sub-districts not included in this study are all sub-districts in Kepulauan Seribu Regency. The following variables were used in this study:

Table 1. Data Source and Variable Name

Notation	Variable Name	Data Source
Y	Number of Dengue Disease Cases	Diskominfotik Jakarta Province
X_1	Number of Flood-Prone Points	Diskominfotik Jakarta Province
X_2	Number of Slum Neighborhood Associations (RT)	BPS Jakarta Province
X_3	Population Density (People per km^2)	Diskominfotik Jakarta Province
X_4	Number of Hospitals per 1,000 Population	Diskominfotik Jakarta Province
X_5	Number of Public Health Centers per 1,000 Population	Diskominfotik Jakarta Province

Data Analysis Steps

Data processing was carried out using R software version 4.1.2 and thematic map creation using Q.GIS Desktop 3.16.15. Steps of data analysis were carried out as follows:

1. Exploring data for all variables using thematic maps so that the pattern of distribution of data between sub-districts can be known;
2. Before modeling with spatial regression, the classical assumptions of multiple linear regression models must be tested. [16];
3. Before performing Moran's I test, it is necessary to create a spatial weight matrix. The most common way to represent spatial data relationships is through the concept of Contiguity. That is, areas will be considered related if their boundaries have the same points. In the concept of Queen Contiguity every region that touches the boundary of another region, either a side or a corner, is considered a neighbor [17]. Queen Contiguity is the spatial weight matrix used in this study;
4. Checking whether there is an autocorrelation between sub-districts by conducting the Moran's I test (Moran index) [18]. The hypothesis of Moran's I test is as follows:

   H_0 : No spatial autocorrelation under given $\mathbf{W}$ (spatial weight matrix)

   H_1 : There is a spatial autocorrelation under the given $\mathbf{W}$

   Moran's I is defined:

$$I = \frac{N}{S_0}\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{1}$$

where N denotes the number of observations, $\bar{y}$ denotes the average value y_i from N Locations, y_i denotes the value at locations I, y_j denotes the value at locations j, w_{ij} denotes the spatial weight matrix element and Moran's I test statistic is defined:

$$Z_I = \frac{I - E(I)}{\sqrt{Var(I)}} \tag{2}$$

where Z_I denotes the Moran's I test statistic value, $E(I)$ denotes the expected value of Moran's I and $Var(I)$ denotes Variance of Moran's I which defined:

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

$$E(I) = -\frac{1}{n-1} \tag{4}$$

$$S_2 = \sum_{i=1}^{N}\left(\sum_{j=1}^{N} w_{ij} + \sum_{j=1}^{N} w_{ji}\right)^2 \tag{5}$$

$$S_1 = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(w_{ij} + w_{ji})^2 \tag{6}$$

$$S_0 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} \tag{7}$$

The decision criteria in making conclusions are to reject H_0 if $Z_I > Z_{\frac{\alpha}{2}}$;

5. Checking whether there is a spatial dependence on lag or error by using the Lagrange Multiplier (LM) test. The Lagrange Multiplier (LM) test is used to determine the type of spatial analysis that is appropriate to use [19], [20]. The general form of the SAR model is defined as follows [20]–[22]:

$$\boldsymbol{y} = \rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon};\ \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma_{\boldsymbol{\varepsilon}}^2\boldsymbol{I}) \tag{8}$$

where $\boldsymbol{y}$ denotes the response variable, $\boldsymbol{X}$ denotes the predictor variable, ρ denotes the spatial autocorrelation coefficient on the response variable, $\boldsymbol{W}$ denotes the spatial weight matrix, $\boldsymbol{\beta}$ denotes the intercept and regression coefficient and $\boldsymbol{\varepsilon}$ denotes the error. The hypotheses for the spatial dependence on lag are as follows:

$H_0 : \rho = 0$ (No spatial dependence on lag)

$H_1 : \rho \neq 0$ (Lag has a spatial dependence)

The test statistic for the spatial dependence on lag are as follows:

$$LM_{LAG} = \frac{[(e^T W_A y)/(e^T e/N)]^2}{\left[(W_A X\hat{\beta})^2 M(W_A X\hat{\beta})/(e^T e/N)\right] + \left[tr\left({W_A}^T W_A + {W_A}^2\right)\right]} \sim \chi^2_{(1-\alpha);df=1} \tag{9}$$

The decision criteria in making conclusions are to reject H_0 if $LM_{LAG} > \chi^2_{(1-\alpha);df=1}$. If H_0 is rejected, the Spatial Autoregressive (SAR) model is used. And here are the test statistics for spatial dependence on error:

$$LM_{ERR} = \frac{[(e^T W_A e)/(e^T e/N)]^2}{\left[tr\left({W_A}^T W_A + {W_A}^2\right)\right]} \sim \chi^2_{(1-\alpha);df=1} \tag{10}$$

The decision criteria in making conclusions are to reject H_0 if $LM_{ERR} > \chi^2_{(1-\alpha);df=1}$. If H_0 is rejected, the Spatial Error Model (SEM) is used. If both LM_{LAG} and LM_{ERR} are significant, comparing the Akaike Information Criterion (AIC) values allows one of the best models to be chosen. The best model is the one with the lowest AIC value [16]. The AIC calculated using the Maximum Likelihood Estimation (MLE) method is as follows [23]:

$$AIC = -2L_m + 2m \tag{11}$$

where L_m denotes the Maximum log-likelihood and m denotes the number of model parameters;

6. Estimate the parameters of the SAR model. The SAR model is a model whose dependent variables are spatially correlated. Parameter estimation using the Maximum Likelihood method is defined as follows [24], [25]:

$$\widehat{\boldsymbol{\beta}}_{ML} = (\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y} - \rho(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T W_p y} = \widehat{\boldsymbol{\beta}}_{OLS} - \rho\widehat{\boldsymbol{\beta}}_L \tag{12}$$

with $\widehat{\boldsymbol{\beta}}_L$ is a regression parameter estimator based on weight matrix ($\mathbf{W}$) and spatial autocorrelation ρ. Equation (12), however, cannot be directly solved because the value ρ is unavailable. As a result, the log-likelihood concentrated function (L_c) is employed, as defined below [18]:

$$\ln L_c(\rho) = C - \frac{n}{2} ln\left[\frac{1}{n}(\boldsymbol{e_0} - \rho\boldsymbol{e_L})^T(\boldsymbol{e_0} - \rho\boldsymbol{e_L})\right] + ln\left|\boldsymbol{I} - \rho\boldsymbol{W}_\rho\right| \tag{13}$$

with C is a constant. Equation (13) is a non-linear function in one parameter and is maximized using a numerical technique with direct search;

7. Interpretation of the obtained SAR model, including the direct impact of covariates;
8. Diagnostic testing of the SAR model.

RESULTS AND DISCUSSION

Descriptive Analysis

Figure 1 depicts the distribution of dengue disease cases number in the Jakarta province. Considering Figure 1 The sub-districts with a high dengue disease cases number are shown in dark red, with the majority located in the municipalities of East Jakarta and North Jakarta and a tiny portion in the municipalities of West Jakarta and South Jakarta. The sub-districts with a moderate number of dengue disease cases are denoted in light red, with the majority found in the municipalities of West Jakarta and South Jakarta and a tiny portion in the municipalities of East Jakarta and North Jakarta. And the sub-districts with a low number of dengue disease cases are marked in white, with the majority of these sub-districts being in the municipality of Central Jakarta and a tiny portion in the municipalities of South Jakarta and North Jakarta. This indicates that sub-districts with high dengue disease cases likely to be located near sub-districts with moderate dengue disease cases.

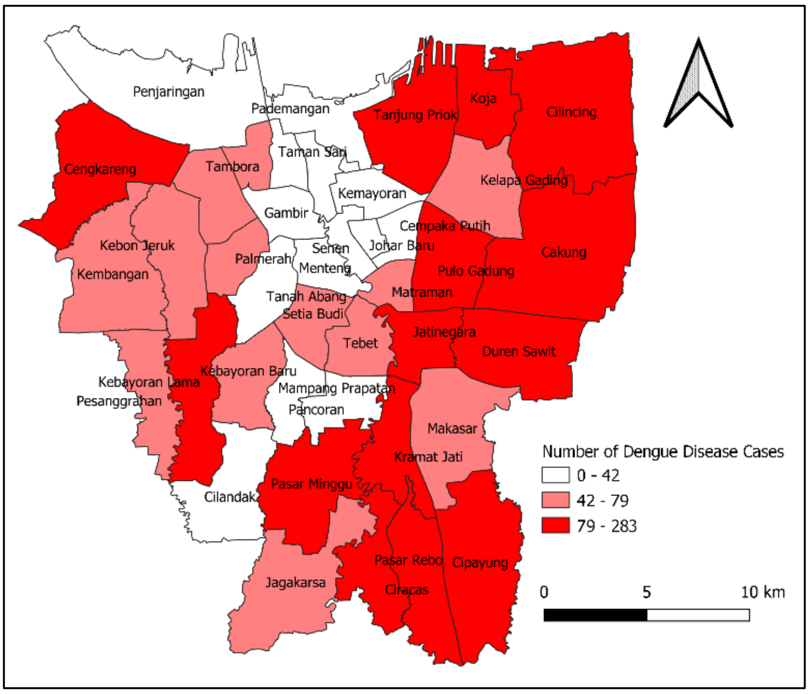


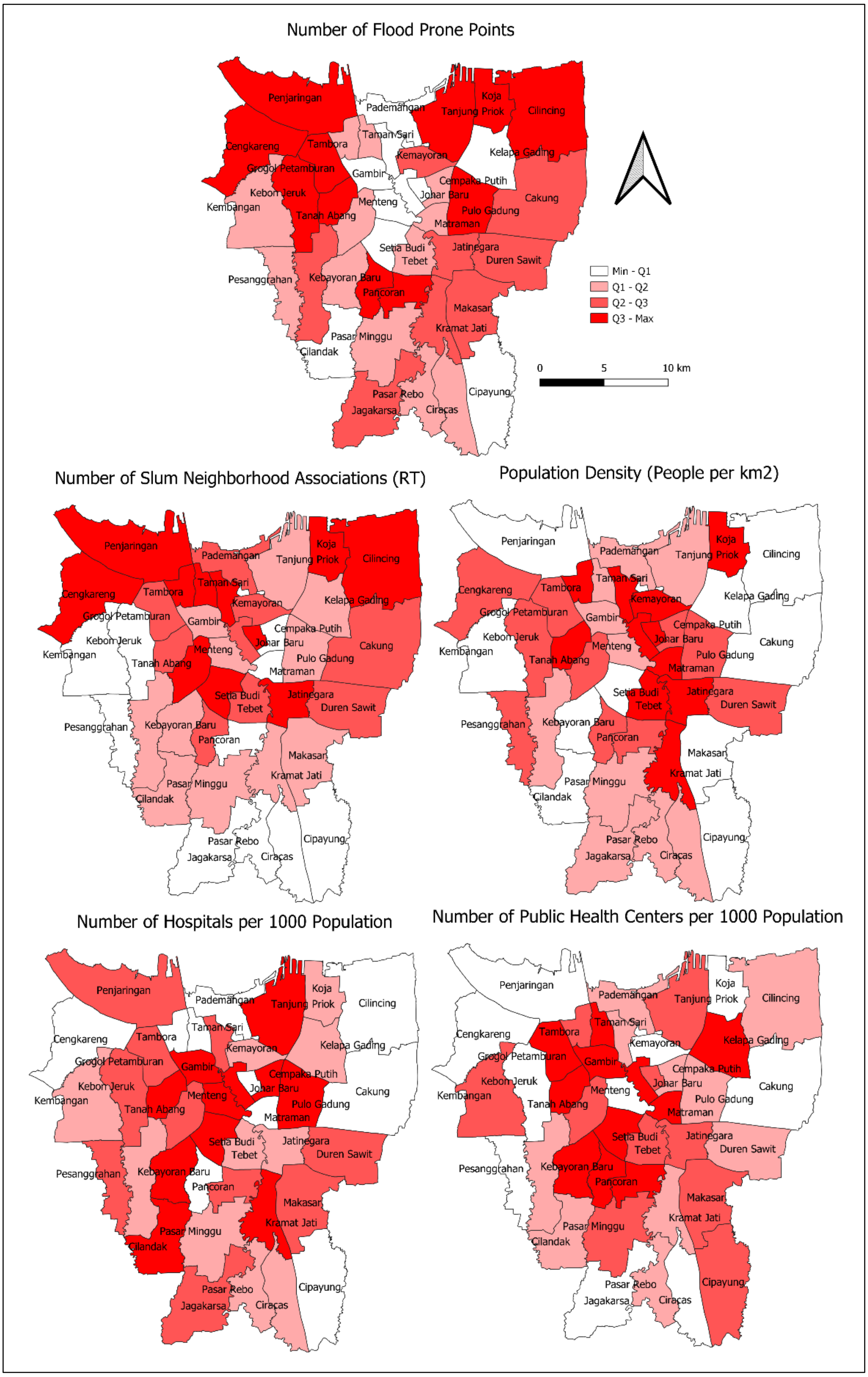
Figure 1. Map of the Distribution of Dengue Disease Cases in Jakarta Province

Figure 2 shows that Cilincing, Koja, Tanjung Priok, Cengkareng, and Pulo Gadung are the sub-districts with the highest number of flood-prone points. In comparison to dengue disease cases number in these sub-districts, the number of dengue disease cases is also high. In addition, it can be noticed that Pademangan, Taman Sari, Gambir, Senen, Menteng, Johar Baru, and Cilandak are the sub-districts with the lowest flood-prone points. Moreover, as compared to dengue disease cases number in these sub-districts, the number of dengue disease cases is comparatively low. This indicates that there is a positive correlation between the number of flood-prone points and the dengue disease cases number in Jakarta Province.

Figure 2 also shows that Cilincing, Koja, Cengkareng, and Jatinegara are the sub-districts with the highest number of slum neighborhood associations. In comparison to dengue disease cases number in these sub-districts, the number of dengue disease cases is also high. Cempaka Putih and Pancoran can be noted to have the fewest slum neighborhood associations. Comparatively to dengue disease cases number in these sub-districts, the number of dengue disease cases is comparatively low. This indicates a substantial correlation between the number of slum neighborhood associations and dengue disease cases number in Jakarta Province.

We can see in Figure 2 that Koja, Kramat Jati, and Jatinegara are the sub-districts with a high population density, and these sub-districts also have a significant prevalence of dengue disease cases. Penjaringan and Cilandak are the sub-districts with the lowest population density, and both sub-districts also have the lowest number of dengue disease. This indicates a positive correlation between population density and the number of cases of dengue disease in Jakarta Province.

We can also see in Figure 2 that Cilincing, Cakung, Cengkareng, and Cipayung are sub-districts with a low number of hospitals per 1,000 populations. However, when compared with dengue disease cases number, these sub-districts are classified as sub-districts with a high dengue disease cases number, indicating a negative relationship between the number of hospitals per 1,000 populations and the number of cases of dengue disease in Jakarta Province.



Number of Flood Prone Points

Number of Slum Neighborhood Associations (RT)

Population Density (People per km2)

Number of Hospitals per 1000 Population

Number of Public Health Centers per 1000 Population

Figure 2. Map of Distribution of Predictor Variable

And we can also notice from Figure 2 the sub-districts with the lowest public health centers per 1,000 people number are Cakung, Koja, Cengkareng, and Ciracas. However, when compared to the number of dengue disease cases, these sub-districts are classified as sub-districts with a high number of dengue disease cases. This indicates a negative correlation between the number of public health centers per 1,000 people and dengue disease case number in Jakarta Province.

Relying on Figures 1 and 2, We can conclude that descriptively there is a relationship between the number of cases of dengue disease and all predictor variables used in this study. However, in order to be more convincing, a spatial regression analysis must be performed.

Spatial Analysis

Table 2 shows the results of the linear regression model's classical assumption test. The p-values of the normality assumption test and the homoscedasticity assumption test are greater than 0.05, implying that the normality and homoscedasticity assumptions are fulfilled. Table 2 also shows that the Durbin Watson (d) value is between $4-dU<d<4-dL$ which can be concluded that the non-autocorrelation assumption is fulfilled, and the VIF value lower than 5 for all variables can be concluded that the non-multicollinearity assumption is fulfilled.

Table 2. The results of the classical assumption of linear regression model

Statistic test	Result
Normality (Shapiro-Wilk)	P-value = 0.3081
Non-autocorrelation (Durbin-Watson)	2.0914 (dL=1.2546 dU=1.7814)
Non- multicollinearity (VIF)	(X_1) 1.4457, (X_2) 1.4091, (X_3) 1.2851, (X_4) 1.1829, (X_5) 1.0831
Homoscedasticity (Breusch-pagan)	P-value = 0.3219

Even though the linear regression model has all of the assumptions fulfilled, it is still necessary to investigate whether there is an autocorrelation between sub-districts by performing the Moran's I test. Before carrying out the Moran's I test, a spatial weight matrix is needed. And the spatial weight matrix used in this study is the neighbor (Contiguity) spatial weight matrix with the neighboring type is (Queen). The following table shows the results of the spatial autocorrelation test using Moran's I test.

Table 3. Spatial Autocorrelation Test Results with Moran's I

Variable	Statistic Moran's I	P-Value (α=5%)
Number of Dengue Disease Cases	0.4140	1.893×10^{-6}*
Number of Flood-Prone Points	0.2926	0.0004*
Number of Slum Neighborhood Associations (RT)	0.1591	0.05191
Population Density (People per km^2)	0.1596	0.0363*
Number of Hospitals per 1,000 Population	0.1217	0.0877
Number of Public Health Centers per 1,000 Population	0.0919	0.2044

*) Significant

According to Table 3, there is a spatial autocorrelation in dengue disease cases number because the p-value is less than 0.05. Two of the five predictor variables also showed that there was a spatial autocorrelation. Moran's I statistical values are all between 0 and 1, indicating that the closer an area is, the more similar the variable values are.

Table 4. LR dan LM Test Results and AIC value

Test	Value	P-Value (α=5%)	AIC
LR	9.9310	0.0016	451.7152
LM_{LAG}	7.6763	0.0056	443.7842
LM_{ERR}	5.1326	0.0235	446.3620

The Likelihood Ratio (LR) test is used to determine which model performs better, spatial regression or linear regression. And the Lagrange Multiplier (LM) test is employed to find whether the spatial dependence is on the dependent variables (lag), on unresearched variables (error), or both (error and lag). Table 4 shows the output of the LR and LM tests that were performed.

According to the results in Table 4, the LR test is significant because the p-value is less than 0.05, indicating that there is a significant difference between the spatial regression model and the linear regression model. It can also be seen that the AIC of the linear regression model is the largest when compared to the two spatial regression models. And it is clear that both the spatial dependence in lag and the spatial dependence in error are significant, as indicated by the p-value less than 0.05. However, this study use the SAR model because its AIC value is lower than the AIC value in the SEM model. Table 5 below shows the estimation results of the SAR model parameters.

Table 5. Estimation of SAR model parameters

	Estimate	Std Error	Z-Value	Pr (>|z|)
(Intercept)	20.839	29.102	0.7161	0.4739
Number of Flood-Prone Points	3.5296	0.9152	3.8568	0.0001*
Number of Slum Neighborhood Associations (RT)	0.1752	0.1336	1.3116	0.1896
Population Density (People per km^2)	-0,0003	0.0006	-0.4603	0.6453
Number of Hospitals per 1,000 Population	2.8001	315.00	0.0089	0.9929
Number of Public Health Centers per 1,000 Population	-792.52	539.49	-1.4690	0.1418
Spasial lag (Rho)	0.56618	0.12682	4.4644	0.000008*
Statistic Wald: 19.931; p-value: 0.000008				

*) Significant

According to Table 5, the Spatial lag variable (Rho) has a statistically significant and positive coefficient, indicating that as dengue disease cases number in one sub-district increases, so will dengue disease cases number in neighboring sub-districts. And the Wald test p-value is less than 0.05, indicating a significant relationship between the number of dengue disease cases in Jakarta Province and all predictor factors.

Dengue disease cases number in Jakarta Province is significantly affected by the number of flood-prone points, the number of slum neighborhood associations, the population density, the number of hospitals and the number of public health centers per 1,000 populations, and the spatial lag. At a significance level of 5%, only the number of flood-prone points and spatial lag have a significant impact on the number of cases of dengue disease in Jakarta Province. This is consistent with Lilis Wijaya's 2018 research, which found that floods led to post-flood ailments such as diarrhea, dengue disease, Leptospirosis, Acute Respiratory Infection (ARI), intestinal worms, skin problems, and many more [26]. Wang et al. (2016) reported that the Aedes Aegypti mosquito will live longer if the humidity level is high, such as during the rainy season, particularly in areas prone to floods, where the huge volume of standing water will make the disease more likely to spread [2].

The number of slum neighborhood associations and the number of public health centers per 1,000 people has a same coefficient sign with previous studies [1], [5]. The sign of the coefficient indicates the direction of the relationship between the predictor variable and the number of cases of dengue disease. Although in this study the effect is almost significant with the p-value still below 0.2, which means the error rate is still below 20%. Meanwhile, the population density variable and the number of hospitals per 1,000 population have different coefficient signs from previous studies [5], [6]. However, the p-value shows that the effect is highly insignificant because the p-value exceeds 0.5, which means the error rate is above 50%. Based on the parameter estimation results in Table 5, the form of the SAR model in this study is as follows:

$$\hat{y}_i = 0.56618 \sum_{j=1, i\neq j}^{n} w_{ij} y_j + 20.839 + 3.5296\, X_{1i} + 0.1752\, X_{2i} - 0{,}0003 X_{3i} + 2.8001 X_{4i} - 792.52 X_{5i} \tag{14}$$

The impact of covariates in the SAR model could be classified into three categories: total impact; indirect impact; and direct impact; [16]. Total impact refers to the changes that occur in one sub-district as a consequence of changes in that sub-district and its surroundings. Indirect impact refers to the effect that takes place when the predictor factors in the bordering sub-district change. And impacts that occur locally in an area, which in this study is a sub-district, as a consequence of changes in predictor factors in that sub-district are referred to as direct impacts. The magnitude of direct and indirect impact of Table 6 shows the SAR model used in this study.

Table 6. Measures of direct, indirect and total impact of The SAR model

Variable	Direct Impact	P-Value Direct Impact	Indirect Impact	P-Value Indirect Impact	Total Impact	P-Value Total Impact
Number of Flood-Prone Points	3.8601	0.0002*	4.2760	0.0859	8.1361	0.0110*
Number of Slum Neighborhood Associations (RT)	0.1916	0.2004	0.2123	0.3637	0.4039	0.2751
Population Density (People per km^2)	-0.0003	0.6310	-0.0004	0.7318	-0.0007	0.6816
Number of Hospitals per 1,000 Population	3.0623	0.9833	3.3923	0.9584	6.4546	0.9671
Number of Public Health Centers per 1,000 Population	-866.74	0.1388	-960.11	0.2705	-1826.8	0.1908

*) Significant

Table 6 shows the number of flood-prone points in Jakarta Province has a substantial direct effect on the number of dengue disease cases. No variable has a significant indirect effect on dengue disease in Jakarta Province, based on an examination of indirect effects. The number of flood-prone points in Jakarta Province has a significant impact on the total number of dengue disease cases. The growth in the number of flood-prone points will have a direct impact on dengue disease cases number in Jakarta Province. Each one percent increase in flood-prone points in a sub-district will result in a rise of 3.86 cases of dengue disease in that sub-district.

Table 7. SAR Model Diagnostic Test

Statistic test	P-Value	Conclusion
Normality (Shapiro-Wilk)	0.8740	Residuals are normally distributed
LM test for residual autocorrelation	0.7339	There is no autocorrelation between residuals
Homogeneity (Breusch-pagan)	0.6246	The homogeneity assumption is met

It is important to conduct a diagnostic test that involves the assumptions of homogeneity, normality and non-autocorrelation to determine the quality of the SAR model. The SAR model satisfies all the assumptions, as shown in Table 7.

CONCLUSIONS

Moran's I test showed the dengue disease cases number, the number of flood-prone points, and the population density have spatial correlation whereas the number of slum neighborhood associations, the number of hospitals per 1,000 populations, and the number of Public Health Centers per 1,000 population have no geographical correlation.

Relying on the Lagrange Multiplier test, the best spatial model is the Spatial Autoregressive (SAR) model. The growth in the flood-prone points number in Jakarta Province will have a direct effect on the number of instances of dengue disease. Each additional one percent of flood-prone points in a sub-district will result in 3.86 additional cases of dengue disease in that sub-district. The derived SAR model is valid since it satisfies the assumptions of normality, absence of autocorrelation, and homogeneity.

Recommendations are made for the Jakarta Provincial Government to enhance flood management regulations in order to lower the incidence of dengue disease cases. Additional variables with a substantial association to the frequency of dengue disease cases can be included in future investigations. In addition, the Spatial Durbin Model (SDM) technique can be utilized for more research on the number of dengue disease cases if all predictor variables exhibit a spatial lag. Or Geographically Weighted Poisson Regression (GWPR) can also be utilized, which overcomes the presence of spatial heterogeneity in the response data in the form of count data (amount). Or the Conditional Autoregressive-Bessag York Mollie (CAR-BYM) can also be utilized, which can accommodate geographical and non-spatial features induced by the heterogeneity of cases between regions.

ACKNOWLEDGMENTS

The authors would like to thank the Rector of Universitas Padjadjaran, who offered financial help to disseminate studies reports. This research is part of PDKN contract DIKTI: 094/E5/PG.02.00.PT/2022 and DRPM: 1318/UN6.3.1/PT.00/2022 and also it is part of Academic Leadership Grant (ALG) contract 2023/UN6.3.1/PT.00/2022. We are also thankful for the discussion on social media analytics through the RISE_SMA project funded through the European Union year 2019-2024 and also studies center of modeling and computation Universitas Padjadjaran.

REFERENCES

[1] H. N. Avida *et al.*, “Pemberantasan Sarang Nyamuk (PSN) Sebagai Upaya Pencegahan Kasus Demam Berdarah Dengue (DBD) di Yogyakarta,” *Jurnal Pengabdian Serulingmas*, vol. 2, no. 1, pp. 23–25, Apr. 2022.

[2] R. K. Sari, I. Djamaluddin, Q. Djam’an, and T. Sembodo, “Pemberdayaan Masyarakat dalam Upaya Pencegahan Demam Berdarah Dengue DBD di Puskesmas Karangdoro,” *Jurnal ABDIMAS-KU: Jurnal Pengabdian Masyarakat Kedokteran*, vol. 1, no. 1, pp. 25–33, Jan. 2022, doi: 10.30659/abdimasku.1.1.25-33.

[3] L. Faridah *et al.*, “Temporal Correlation Between Urban Microclimate, Vector Mosquito Abundance, and Dengue Cases,” *J Med Entomol*, vol. 59, no. 3, pp. 1008–1018, May 2022, doi: 10.1093/jme/tjac005.

[4] D. Tarmana, “Potensi Peluang Demam Berdarah Dengue (DBD) Berdasarkan Proyeksi Perubahan Iklim (Study Kasus : DKI Jakarta),” *The Indonesian Journal of Infectious Disease*, vol. 1, no. 2, pp. 14–22, 2013.

[5] I. F. Fatati, H. Wijayanto, and A. M. Sholeh, “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, Dec. 2017, doi: 10.14710/medstat.10.2.95-105.

[6] Jajang, B. Pratikno, and Mashuri, “Disease Mapping Kasus DBD Menggunakan Model CAR-BYM,” in *Pengembangan Sumber Daya Perdesaan dan Kearifan Lokal Berkelanjutan XI*, Oct. 2021, pp. 1–10.

[7] I. G. N. M. Jaya, F. Kristiani, Y. Andriyana, and B. N. Ruchjana, “Modeling dengue disease transmission for juvenile in bandung, indonesia,” *Communications in Mathematical Biology and Neuroscience*, vol. 2021, 2021, doi: 10.28919/cmbn/5435.

[8] Badan Pusat Statistik, “Kependudukan Menurut Badan Pusat Statistik,” *Kepadatan Penduduk Menurut Provinsi*, 2019. https://bps.go.id/subject/12/kependudukan.html#subjekViewTab3 (accessed Jun. 12, 2022).

[9] Badan Pusat Statistik, “Kemiskinan Menurut Badan Pusat Statistik,” *Persentase Rumah Tangga Kumuh Perkotaan*, 2019. https://bps.go.id/subject/23/kemiskinan-dan-ketimpangan.html#subjekViewTab3 (accessed Jun. 12, 2022).

[10] Kementerian Kesehatan Republik Indonesia, “Kasus Demam Berdarah Dengue (DBD) Menurut Jenis Kelamin dan Provinsi di Indonesia,” in *Profil Kesehatan Indonesia 2019*, Jakarta: Kementerian Kesehatan Republik Indonesia, 2020, p. 458.

[11] R. Yuliana, M. Rahmaniati, I. Apriantini, and R. Triarjunet, “Analisis Autokorelasi Spasial Kasus Demam Berdarah Dengue di Kota Padang,” *JIK (Jurnal Ilmu Kesehatan)*, vol. 6, no. 1, pp. 34–42, Apr. 2022, doi: 10.33757/jik.v6i1.484.g210.

[12] N. Faiz, R. Rahmawati, and D. Safitri, “Analisis Spasial Penyebaran Penyakit Demam Berdarah Dengue Dengan Indeks Moran Dan Geary’s C (Studi Kasus Di Kota Semarang Tahun 2011),” *Jurnal Gaussian*, vol. 2, no. 1, pp. 69–78, 2013, [Online]. Available: http://ejournal-s1.undip.ac.id/index.php/gaussian

[13] A. P. N. Taryono, D. Ispriyanti, and A. Prahutama, “Analisis Faktor-Faktor yang Mempengaruhi Penyebaran Penyakit Demam Berdarah Dengue (DBD) di Provinsi Jawa Tengah dengan Metode Spatial Autoregressive Model dan Spatial Durbin Model,” *Indonesian Journal of Applied Statistics*, vol. 1, no. 1, pp. 1–13, May 2018.

[14] Badan Pusat Statistik Provinsi DKI Jakarta, "Jumlah RT Kumuh Menurut Kecamatan," in *Pendataan RW Kumuh DKI Jakarta 2017*, 1st ed., Jakarta: Badan Pusat Statistik Provinsi DKI Jakarta, 2017, pp. 190–194.

[15] Dinas Komunikasi Informatika dan Statistik Pemprov DKI Jakarta, "Data Terbuka Pemerintah Provinsi DKI Jakarta," *Jumlah Kasus DBD, Jumlah Titik Rawan Banjir, Kepadatan Penduduk, Jumlah RS per 1000 penduduk, Jumlah Puskesmas per 1000 penduduk*, 2017. https://data.jakarta.go.id/ (accessed Jun. 12, 2022).

[16] H. I. Zebua and I. G. N. M. Jaya, "Spatial Autoregressive Model of Tuberculosis Cases in Central Java Province 2019," *CAUCHY –Jurnal Matematika Murni dan Aplikasi*, vol. 7, no. 2, pp. 240–248, Dec. 2022, doi: 10.18860/ca.v7i2.13451.

[17] R. Dubin, "Spatial Weight," in *The SAGE Handbook of Spatial Analysis*, A. S. Fotheringham and P. A. Rogerson, Eds. London: SAGE Publications Ltd, 2009, pp. 125–130.

[18] L. Anselin, "Spatial Econometrics," in *A Companion to Theoretical Econometrics* , B. H. Baltagi, Ed. Blackwell Publishing Ltd, 2001, pp. 310–330.

[19] L. Anselin, "Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity," *Geogr Anal*, vol. 20, no. 1, pp. 1–17, Jan. 1988.

[20] C. Song *et al.*, "Disease relative risk downscaling model to localize spatial epidemiologic indicators for mapping hand, foot, and mouth disease over China," *Stochastic Environmental Research and Risk Assessment*, vol. 33, no. 10, pp. 1815–1833, Oct. 2019, doi: 10.1007/s00477-019-01728-5.

[21] H. Yasin, A. R. Hakim, and B. Warsito, "Spatial Autoregressive, Spatial Error Model," in *REGRESI SPASIAL (Aplikasi dengan R)*, 1st ed., Pekalongan: WADE Group, 2020, pp. 21–48.

[22] W. Wang *et al.*, "Reclaiming independence in spatial-clustering datasets: A series of data-driven spatial weights matrices," *Stat Med*, vol. 41, no. 15, pp. 2939–2956, Jul. 2022, doi: 10.1002/sim.9395.

[23] A. A. Grasa, "Criteria BIC and AIC," in *Econometric Model Selection: A New Approach*, vol. 16, Dordrecht: Springer Netherlands, 1989, p. 132. doi: 10.1007/978-94-017-1358-0.

[24] J. P. Lesage, "Lecture 1," in *Maximum likelihood estimation of spatial regression models*, Toledo: University of Toledo, 2004, pp. 1–16.

[25] S. S. Handajani, H. Pratiwi, and Y. Susanti, "Hybrid Ensemble Spatial Regression Model for Case Number of Dengue Hemorrhagic Fever in Central Java Province," in *International Symposium on Advanced Intelligent Informatics (SAIN) : "Revolutionize Intelligent Informatics Spectrum for Humanity,"* Aug. 2018, pp. 148–152.

[26] M. G. Saputra and F. Ummah, "Kesiapan Masyarakat dalam Menghadapi Penyakit Pasca Banjir di Dusun Lohgawe Desa Gawerejo Kecamatan Karangbinangun Kabupaten Lamongan," *Jurnal ASRI (Administrasi Rumah Sakit Indonesia)*, vol. 7, no. 2, pp. 54–62, 2021.