

SPATIAL AUTOREGRESSIVE (SAR) POISSON MODELING IN DENGUE FEVER CASES ON LOMBOK ISLAND IN 2021

Ririn Robiatul Husnaeni¹, Siti Hauliati², Imroatun Sholihah³, Bq Tia Ayu Lasmiani⁴,
Siti Hariati Hastuti^{5*}, Muhammad Gazali⁶

^{1,2,3,4,5}Statistic Study Program, Universitas Hamzanwadi

Jl. Cut Nyak Dien No.85, Pancor, Selong District, East Lombok Regency, 83611

West Nusa Tenggara, Indonesia

Corresponding Author's E-mail: siti.hariatih@hamzanwadi.ac.id

Abstract: Indonesia, the fourth most populous country in the world with 275.5 million people, faces increasing human activity that can lead to negative impacts such as the spread of infectious diseases. One of these diseases is Dengue Hemorrhagic Fever (DHF), which is particularly susceptible in residential areas with poor environmental hygiene. The rising number of DHF cases on Lombok Island is a significant concern. This study employs a spatial analysis modeling approach, specifically the Spatial Autoregressive Poisson (SAR Poisson) model, which considers the spatial dependence of dengue cases assumed to follow a Poisson distribution. The objective is to model and map the potential distribution of DHF cases on Lombok Island in 2021. The analysis reveals spatial autocorrelation in the data based on Moran's I. Significant variables affecting DHF cases include the number of permanent sanitation facilities (X2) and the number of drinking water facilities (X3). Mapping results based on the SAR Poisson model indicate that the distribution of DHF cases is relatively uniform across most sub-districts, with the highest incidence suspected in Tanjung Sub-district.

Keywords: DHF, Lombok Island, SAR Poisson, Spatial

1. INTRODUCTION

Indonesia, with a population of 275.5 million, is the fourth most populous country globally [1]. As the population grows, human activities and interactions increase correspondingly. These heightened activities can result in negative consequences that threaten human life, including spreading infectious diseases.

[2] The issue of infectious diseases remains a significant health challenge in Indonesia. These diseases are caused by microorganisms such as bacteria, viruses, and fungi, which can be transmitted from infected individuals to healthy individuals, thereby causing illness. One example of such an infectious disease is Dengue Hemorrhagic Fever (DHF).

DHF is caused by the dengue virus. The dengue virus is transmitted through the bites of *Aedes Aegypti* and *Aedes Albopictus* mosquitoes, which have previously bitten individuals infected with the dengue virus. These two types of *Aedes* mosquitoes are widespread throughout Indonesia and generally prefer locations associated with clean water to lay their eggs individually [3].

Dengue Hemorrhagic Fever is highly susceptible to occurring in residential areas with poor environmental hygiene. In Indonesia, DHF has high morbidity and mortality rates. Data from the Indonesian Ministry of Health in 2021 show there were 73,518 cases of DHF, with 705 deaths. Additionally, approximately 43.25% of the total DHF cases in Indonesia were dominated by the 5-14-year age group [4]. This places West Nusa Tenggara Province in the fifth position for the highest DHF morbidity rate, at 50.9% per 1,000 population [5].

DHF remains a significant public health issue in the West Nusa Tenggara (NTB) Province. This is due to its rapid spread and the potential for mortality, and all regencies/cities in the province have experienced dengue outbreaks [6]. According to [4], in 2018, there were 535 reported cases of dengue fever in NTB, which increased to 2,971 cases in 2019. The number of cases surged dramatically in 2020, reaching 4,731 cases. However, in 2021, dengue fever's incidence decreased slightly, with 2,719 cases reported.

One of the strategies for managing the spread of DHF involves a modeling approach using spatial analysis. Spatial analysis is a method used to explore relationships between observational variables based on their geographic positions. A commonly used technique in this analysis is Spatial Autoregressive (SAR) modeling. SAR models are designed to analyze data using a spatial regression model where the autoregressive process occurs only on the lag of the dependent variable. The SAR method has been developed to accommodate the distribution of specific data types, one of which is the SAR Poisson model [7]. Poisson regression is a function with a response variable (Y) following a Poisson distribution [8].

Several studies have been conducted, including one titled: "Analysis of Factors Affecting the Spread of DHF in Central Java Province Using Spatial Autoregressive Model and Spatial Durbin Model" [9]. In this study, the weighting function employed was the Queen's contiguity weights, with the best model being the Spatial Autoregressive Model. The results indicated that population density and the average length of schooling generally affect the spread of DF in Central Java Province.

A subsequent study on Dengue Fever titled: "Spatial Autoregressive Method in Analyzing Dengue Fever Vulnerability in Gorontalo City" [10]. The results of this study indicate that the factors generally affecting the vulnerability to dengue outbreaks in Gorontalo City are population density, the number of impoverished individuals, healthcare facilities, and the overall elevation of the region above sea level, as represented in the SAR model.

Another similar study titled: "Spatial Autoregressive Poisson Model for Detecting Factors Affecting the Number of HIV/AIDS Cases in East Jakarta" [8]. The results of this study indicate that the number of HIV/AIDS cases is influenced by the proximity of regions and all significant independent variables.

The increasing incidence of Dengue Fever (DF) on Lombok Island warrants significant attention. Lombok Island comprises East Lombok Regency, West Lombok Regency, Central Lombok Regency, North Lombok Regency, and Mataram City. This study aims to examine the spatial distribution patterns of DF cases on Lombok Island using spatial analysis, as well as to explore the relationship between dengue fever and influencing variables such as population density, sanitation levels, and drinking water availability, employing the Spatial Autoregressive Poisson (SAR Poisson) method. Thus, the researcher uses the SAR Poisson method to conduct an in-depth spatial regression analysis concerning the distribution of DF cases on Lombok Island in 2021.

2. METHODOLOGY

2.1. Data and Research Variable

This study is quantitative as it utilizes numerical data. The data sources for this research were obtained from the official websites of the Health Offices of East Lombok, Central Lombok, North Lombok, West Lombok, and Mataram City. The data used in this study are secondary data at the sub-district level across Lombok Island for the year 2021. The variables employed in this study include the dependent variable, which is the number of DF cases (Y), and independent variables, which consist of population density (X_1), the number of permanent sanitation facilities (X_2), and drinking water facilities (X_3). The analysis for this study was conducted using R-Studio 4.3.2 software.

2.2. Data Analysis Technique

The following presents the stages of data analysis conducted in this study:

2.2.1 Data Exploration

Data exploration involves the use of thematic maps to understand the distribution of each variable from a geographical perspective.

2.2.2 Multicollinearity Test

The purpose of the multicollinearity test is to determine whether there is a correlation among the independent variables in the regression model. A good regression model should not exhibit correlations among independent variables. To assess multicollinearity, the Variance Inflation Factor (VIF) and Tolerance (TOL) are

evaluated. Commonly used cutoff values indicating multicollinearity are a tolerance value ≤ 0.10 or a VIF value ≥ 10 . The VIF is calculated using the following formula[11]:

$$VIF = \frac{1}{1 - R_j^2} \quad (1)$$

Where R_j^2 represents the coefficient of determination for the j -th variable.

2.2.3 Poisson Regression Modeling

Poisson regression is a regression method used to model the relationship between a dependent variable (Y) that follows a Poisson distribution and one or more independent variables (X), which can be either discrete or continuous [12]. Generally, the Poisson regression model is expressed as follows [12]:

$$y_i = \mu_i + \varepsilon_i, i = 1, 2, \dots, n \quad (2)$$

Where y_i represents the observed value of the i -th independent variable, each of which follows a Poisson distribution. μ_i denotes the mean of the dependent variable (Y) influenced by the independent variables at the i -th observation, while ε_i represents the error term for the i -th observation.

2.2.4 Calculating Euclidean Distance

Euclidean Distance is a method used to calculate the distance between two coordinate points[13]. The Euclidean distance is given by the following formula [14]:

$$d = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} \quad (3)$$

Where d represents the distance. X_1 denotes the latitude coordinate of the first point, while X_2 denotes the latitude coordinate of the second point. Y_1 represents the longitude coordinate of the first point, and Y_2 represents the longitude coordinate of the second point.

2.2.5 Calculating Spatial Weight Matrix Using Inverse Distance Weight (IDW)

One method for determining the spatial weight matrix based on distance is the Inverse Distance Weight (IDW) [10]. The principle of this calculation is based on the actual distance between locations. The sum of each row in the IDW matrix equals one. Mathematically, the IDW matrix is calculated as follows:

$$W = [w_{ij}] = \begin{cases} 0, & i = j \\ \frac{1}{d_{ij}}, & \neq j \\ \frac{1}{\sum_{j=1}^n d_{ij}}, & \end{cases} \quad (4)$$

Where w_{ij} represents the weight element between locations i and j . The conditions $\sum_{j=1}^n w_{ij} = 1$, $\sum_{j=1}^n w_{ij} = n$, and n represents the number of observation locations, with d_{ij} being the distance between locations i and j .

2.2.6 Spatial Effect Test

The Moran's Index is a statistical test used to assess spatial autocorrelation, which helps identify spatial clustering within a location [15]. Moran's Index values range from -1 to 1 (-1 indicating perfect negative autocorrelation and 1 indicating perfect positive autocorrelation). The Moran's Index can be calculated using the following formula:

$$\log it[\pi(x)] = \ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i \quad (5)$$

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})^2 (x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (6)$$

with n representing the number of observations, \bar{x} is the mean value of x_i from n locations. x_i denotes the observation value at the location i , x_j represents the observation value at the location j , and w_{ij} is the element of the spatial weight matrix at row i and column j .

The hypotheses used are as follows:

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

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

Moran's Index Test Statistic:

$$Z(I) = \frac{I - E(I)}{\sqrt{Var(I)}} \sim N(0,1) \tag{7}$$

The test statistic value where $Z(I)$ follows a normal distribution, meaning that H_0 will be rejected if $|Z(I)| > Z_{\alpha/2}$ or $P - Value < \alpha$ [16].

2.2.7 SAR Poisson Modeling

The estimation of the parameter coefficients in the SAR Poisson model is performed using the Maximum Likelihood Estimation (MLE) method [8]. The SAR Poisson model is nonlinear and does not have a closed-form solution; thus, the estimation of the regression coefficients involves iterative methods using the Newton-Raphson technique. The core of this method is to find the statistics by maximizing the likelihood function, which is then formulated through ε in the SAR Model. The process begins with formulating the general equation for spatial regression, which is expressed as follows:

$$y = \rho W y + X\beta + \varepsilon \tag{8}$$

The equation written above can be transformed into the following form:

$$\varepsilon = y - \rho W y - X\beta = (1 - \rho W) y - X\beta \tag{9}$$

The value of the likelihood function for the error term ε is:

$$L(\sigma^2; \varepsilon) = \left(\frac{1}{2\pi\sigma^2}\right) \exp\left(\frac{1}{2\pi\sigma^2} (\varepsilon^T \varepsilon)\right) \tag{10}$$

Next, differentiating the equation (8) with respect to y yields the Jacobian function:

$$J = \left|\frac{\partial \varepsilon}{\partial y}\right| = |1 - \rho W| \tag{11}$$

Subsequently, substituting equations (9) and (11) into equation (10) results in equation (12). Thus, the likelihood function for the variable y is given by:

$$\begin{aligned} L(\rho, \beta, \sigma^2 | y) &= \left(\frac{1}{2\pi\sigma^2}\right) (J) \exp\left(\frac{1}{2\pi\sigma^2} (\varepsilon^T \varepsilon)\right) \\ &= \left(\frac{1}{2\pi\sigma^2}\right) |1 - \rho W| \exp\left(-\frac{1}{2\sigma^2}\right) (y - \rho W y - X\beta) \end{aligned} \tag{12}$$

Thus, the natural logarithm (ln) of the likelihood function obtained is:

$$\begin{aligned} \ln(L) &= \frac{n}{2} \ln\left(\frac{1}{2\pi\sigma^2}\right) + \ln|1 - \rho W| - \frac{1}{2\sigma^2} \left(\left(\frac{1}{2\sigma^2}\right)^T (y - \rho W y - X\beta)\right) \\ \ln(L) &= \frac{n}{2} (2\pi) - \frac{n}{2} (\sigma^2) + \ln|1 - \rho W| \frac{1}{2\sigma^2} ((y - \rho W y - X\beta) (y - \rho W y - X\beta)) \end{aligned} \tag{13}$$

The parameter β can be estimated by maximizing the ln Likelihood function through differentiation. By differentiating equation (12) with respect to β , the parameter estimates are given as follows:

$$\begin{aligned} \frac{\partial \ln(L)}{\partial \beta} &= 0 \\ \frac{\partial \ln(L)}{\partial \beta} &= \frac{\partial\left(\frac{1}{2\pi\sigma^2} ((y - \rho W y - X\beta)^T (y - \rho W y - X\beta))\right)}{\partial \beta} \end{aligned} \tag{14}$$

Assuming the following:

$$R(\beta) = (y - \rho W y - X\beta)^T (y - \rho W y - X\beta) \tag{15}$$

Since $\beta^T X^T (y - \rho W y)$ is a 1×1 matrix and $\beta^T X^T (y - \rho W y)^T = (y - \rho W y)^T X \beta$ yields the same scalar result, it follows that $R(\beta) = (y - \rho W y)^T (y - \rho W y) - 2\beta^T X^T (y - \rho W y) + \beta^T X^T X \beta$. Thus, we obtain:

$$0 = \frac{\partial(\frac{1}{2\pi\sigma^2} (y - \rho W y)^T - 2\beta^T X^T (y - \rho W y))\beta^T X^T X \beta}{\partial \beta} \tag{16}$$

$$= \frac{-2X^T (y - \rho W y) + 2X^T X \hat{\beta}}{2\sigma^2}$$

$$= X^T X \hat{\beta} X^T (y - \rho W y)$$

$$X^T X \hat{\beta} = X^T (y - \rho W y) \tag{17}$$

$$\hat{\beta} = (X^T X)^{-1} X^T (I - \rho W) y$$

The probability mass function of the Poisson distribution is [8]:

$$f(y_i | X, W^*; y_1, \dots, y_n) = \prod_{i=1}^n \left\{ \frac{\mu_i^{SAR} \exp(-\mu_i^{SAR})}{y_i!} \right\} \tag{18}$$

with $\mu_i^{SAR} = \exp(a_i, X\beta)$, The log-likelihood function is:

$$\ln L(\beta, \rho | X, W^*; y_1, \dots, y_n) = y' A^{-1} X \beta - \sum_{i=1}^n \exp([a_i, X\beta]) \sum_{i=1}^n \ln (y_i!) \tag{19}$$

2.2.8 Parameter Significance Test

The significance of parameters is tested using the Wald test. The hypothesis test for ρ is:

- H_0 : $\rho = 0$ (no spatial correlation)
- H_1 : $\rho \neq 0$ (there is a spatial correlation)

$$G_\rho = \left\{ \frac{\hat{\rho}_0}{\widehat{se}(\hat{\rho}_0)} \right\}^2 \tag{20}$$

where G_ρ follows a chi-squared distribution with 1 degree of freedom. The decision criterion is to reject H_0 if $G_\rho > X^2$.

The hypothesis test for the coefficient parameter β_k is

- H_0 : $\beta_k = 0$
- H_1 : $\beta_k \neq 0$

$$G_\beta = \left\{ \frac{\hat{\beta}_0}{\widehat{se}(\hat{\beta}_0)} \right\}^2 \tag{21}$$

where G_β follows a chi-squared distribution with 1 degree of freedom. The decision criterion is to reject H_0 if $G_\rho > X^2$. The standard error is obtained using the Fisher Information matrix $I(\theta)$, with the formula as follows:

$$I(\theta) = - \begin{bmatrix} \frac{\partial^2 \ln L(\beta^*)}{\rho^2} & \frac{\partial^2 \ln L(\beta^*)}{\partial \beta_0 \partial \rho} & \dots & \frac{\partial^2 \ln L(\beta^*)}{\partial \beta_k \partial \rho} \\ & \frac{\partial^2 \ln L(\beta^*)}{\partial \beta_0^2} & \dots & \frac{\partial^2 \ln L(\beta^*)}{\partial \beta_0 \beta_k} \\ & & \ddots & \vdots \\ & & & \frac{\partial^2 \ln L(\beta^*)}{\partial \beta_k^2} \end{bmatrix} \tag{22}$$

The variance of $\hat{\theta} = [I(\theta)]^{-1}$, thus the standard error is $\sqrt{[I(\theta)]^{-1}}$

2.2.9 Model Goodness of Fit

After estimating the parameters and testing the significance of each parameter estimator, a measure of the coefficient of determination is needed to describe the strength of the relationship between the response variable and the explanatory variables. The coefficient of determination, or R^2_{DEV} , is a measure of the proportion of the variability in the response variable that the explanatory variables can explain. Based on the residual deviance and the development by Cameron and Windmeijer, it can be calculated as follows [8]:

$$R^2_{DEV} = 1 - \frac{\ln L(y) - \ln L(\hat{\mu})}{\ln L(y) - \ln L(\bar{y})} \tag{23}$$

where $\ln L(y)$ represents the natural logarithm of the maximum likelihood function when all parameters β_j and y_i are included, $\ln L(\hat{\mu})$ is the natural logarithm of the maximum likelihood function when all parameters β_j dan $\hat{\mu}_j$ are included, and $\ln L(\bar{y})$ is the natural logarithm of the maximum likelihood function when only β_0 and \bar{y} are included.

2.2.10 Mapping and Interpretation

The final step involves mapping, which provides a visual representation based on the results from the previous steps. This is followed by interpretation to explain the visualized results obtained.

3. RESULTS AND DISCUSSION

3.1. Data Exploration

The Data Exploration section is used to understand the characteristics or general overview of the distribution of DHF in Lombok Island for the year 2021. The following provides an explanation of the percentage distribution of DHF based on sub-districts in Lombok Island for the year 2021:

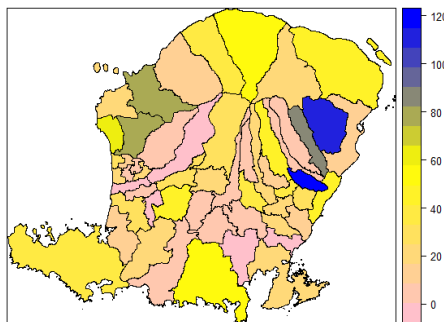


Figure 1. Map of DHF Distribution in Lombok Island for the Year 2021

Based on Figure 1 above, the characteristics of DHF case distribution with the highest incidence are indicated by the blue color, representing a range of 100 to over 120 DHF cases in those sub-districts. Conversely, the areas with the lowest DHF case distribution are marked in pink, corresponding to 0 to 20 DHF cases.

East Lombok's DHF case distribution varies from the lowest to the highest. Central Lombok and West Lombok's distribution ranges from low to moderate. However, the DHF case distribution ranges from low to nearly high in Mataram City and North Lombok.

In East Lombok, the DHF case distribution varies from the lowest to the highest. In Central Lombok and West Lombok, the distribution ranges from low to moderate. However, in Mataram City and North Lombok, the DHF case distribution ranges from low to nearly high.

3.2. Multicollinearity Test

The results of the Variance Inflation Factor (VIF) test for multicollinearity of each explanatory variable are as follows:

Table 1. Results of the Multicollinearity Test

Independent Variables	VIF
Population density (X_1)	1.003
The number of permanent sanitation facilities (X_2)	1.021
Drinking water facilities (X_3)	1.021

Based on Table 1 above, the VIF values are all less than 10. Therefore, it can be concluded, according to equation (1), that there is no multicollinearity among the explanatory variables.

3.3. Poisson Regression

Poisson regression analysis is used to examine the relationship between the number of DHF cases and the influencing factors. The DHF cases in the sub-districts of Lombok Island are assumed to follow a Poisson distribution. Although DHF cases are considered rare, they are a serious issue due to their potential impact on mortality. To address the high number of DHF cases, it is essential to understand the factors affecting the number of cases. One of the regression models that can be used is the Poisson regression model.

The Poisson regression model is constructed using three explanatory variables simultaneously. The estimated parameter values for this model are presented in Table 2.

Table 2. Estimated Parameters of the Poisson Regression Model

Parameter	Estimated Value	Standard Error	G Value	χ^2
β_0 (intercept)	3.044	3.309×10^{-2}	8462.443***	
β_1 (Population density)	2.207×10^{-6}	3.309×10^{-6}	0.445	3.841
β_2 (The number of permanent sanitation facilities)	1.269×10^{-5}	7.467×10^{-7}	2888.822***	
β_3 (Drinking water facilities)	7.752×10^{-6}	9.657×10^{-7}	64.438***	

Note: $\alpha = 0.05$

Based on Table 2 above, the explanatory variables that are significant with respect to the response variable (Y) are the number of permanent sanitation facilities and the number of drinking water facilities. The model indicates that an increase in the population density (X_1), the number of permanent sanitation facilities (X_2), and the number of drinking water facilities (X_3) will lead to an increase in the number of DHF cases. The estimated parameters from the Poisson regression will be used as initial values to obtain parameter estimates in the SAR Poisson model.

3.4. Inverse Distance Weighted (IDW) Weighting Matrix

The following presents the results of the weighting matrix using Inverse Distance Weighted (IDW), which involves assigning weights based on the distance.

$$W = \begin{bmatrix} 0 & 0.0075 & \dots & \dots & 0.0187 & 0.0841 \\ 0.0069 & 0 & \dots & \dots & 0.0088 & 0.0063 \\ \vdots & \vdots & \ddots & \dots & \dots & \dots \\ \vdots & \vdots & \dots & \ddots & \dots & \dots \\ \vdots & \vdots & \dots & \dots & \ddots & \dots \\ 0.0084 & 0.0229 & \dots & \dots & 0.0133 & 0.0077 \end{bmatrix}$$

Based on the matrix above, it can be observed that as the location gets closer to the area with the highest number of DHF cases, the weight assigned increases. Conversely, locations further from the area with the highest number of cases receive lower weights.

3.5. Spatial Effect Test

The spatial effect test is used to determine whether there is spatial dependence between observations. One way to assess this is by using the spatial dependence test, specifically by examining Moran's I Index, as outlined in equations (6), (7), and (8), to determine whether there is spatial autocorrelation among locations. The results of the Moran's I test are as follows:

$$H_0: I = 0 \text{ (no spatial autocorrelation among locations)}$$

$$H_1: I \neq 0 \text{ (there is spatial autocorrelation among locations)}$$

Table 3. Results of Moran's I Index Test

	Value
Moran I Statistic Standard Deviate	2.246
p-value	0.012
Moran I Statistic	0.026
Expectation	- 0.018
Variance	0.004

Based on the Moran's I Index test table, $p - value < 0.05(\alpha)$, leading to the rejection of H_0 . This indicates that there is spatial autocorrelation among locations. This suggests that the Poisson regression model alone is insufficient to model DHF cases, as there is a relationship among sub-districts. One suitable model for this scenario is the SAR Poisson model.

3.6. SAR Poisson

Parameter estimation for the SAR Poisson model is performed using the maximum likelihood estimation method. To obtain the estimated regression coefficients, the Newton-Raphson iterative method is used. By the third or fifth iteration, the coefficient estimates typically converge.

Table 4. Estimated Parameters of the SAR Poisson Regression Model

Parameter	Estimated Value	Standard Error	G Value	χ^2
ρ (Spatial)	0.1	2.31668×10^{-10}	18632361.48***	
β_0 (intercept)	3.044	6.07421×10^{-11}	25.11***	3.841
β_1 (Population density)	-4.7599×10^{-6}	2.519944×10^{-6}	3.568	
β_2 (The number of permanent sanitation facilities)	9.6818×10^{-6}	7.427479×10^{-7}	169.914***	
β_3 (Drinking water facilities)	2.7887×10^{-6}	1.010306×10^{-7}	761.992***	

Note: $\alpha = 0,05$

The analysis of the SAR Poisson model for Pulau Lombok, considering the sub-districts, indicates that the number of DHF cases is influenced by proximity between areas and several significant explanatory variables. The table above shows the significance of each parameter estimate using the Wald test. The spatial correlation obtained is $\rho = 0.1$ with a test statistic $G_p = 18632361.48$. The G_p value exceeds the critical value $\chi^2 = 3.841$, leading to the rejection of H_0 . This signifies a significant spatial correlation in the model at $\alpha = 0.05$. Thus, it can be concluded that the number of DHF cases in a sub-district is influenced by the number of cases in neighboring locations.

The parameter test for each variable X shows that all variables included are significant predictors of Y except for variable X_1 . This is indicated by the test statistic G_β being greater than the critical value $\chi^2 = 3.841$. Therefore, based on the signs of the coefficients for each variable, it can be interpreted that the population density (X_1) decreases in a sub-district, the number of DHF cases in that area increases. Conversely, an increase in the number of permanent sanitation facilities (X_2) and drinking water facilities (X_3) leads to an increase in the number of DHF cases in the sub-district.

The goodness-of-fit of the model can be assessed by the R^2 devians. Based on the coefficient of determination, it was found that the variability in the number of DHF cases can be explained by the explanatory variables with R^2 devians (R^2_{DEV}) = 10.29%. The SAR Poisson model can be expressed as follows:

$$\mu_i^{SAR} = \exp[a_i X \beta]$$

with $\rho = 0.1$ and $\beta = \begin{bmatrix} 3.044 \\ -4.7599 \times 10^{-6} \\ 9.6818 \times 10^{-6} \\ 2.7887 \times 10^{-6} \end{bmatrix}$

Based on the obtained model, a decrease of one person per km² in the sub-district (X_1) results in an increase in DHF cases by $\exp(-4.7599 \times 10^{-6}) = 0.99$ times assuming that other factors in the model remain constant. This means that every decrease of 100 people per km² in the sub-district will increase the expected number of DHF cases by 99 people, assuming other factors in the model remain constant. The population density in a sub-district does not necessarily guarantee the cleanliness of sanitation and drinking water conditions. Therefore, it is crucial to implement innovations that ensure the provision of permanent sanitation and adequate drinking water for the local community.

An increase of one unit in the number of permanent sanitation facilities in a sub-district or village (X_2) results in an increase in DHF cases by $\exp(9.6818 \times 10^{-6}) = 1$ times, assuming that other factors in the model remain constant. This means that every increase of 100 units in permanent sanitation facilities in a sub-district or village will lead to an expected increase of 100 DHF cases, assuming other factors in the model remain constant. Permanent sanitation facilities alone do not guarantee that the sanitation conditions will be clean and suitable for community use. Furthermore, permanent sanitation helps the community avoid improper disposal of human waste and contributes to better environmental management.

An increase of one unit in the number of drinking water facilities in a sub-district (X_3) results in an increase in DHF cases by $\exp(2.7887 \times 10^{-6}) = 1$ times, assuming that other factors in the model remain constant. This means that every increase of 100 units in the number of drinking water facilities in a sub-district or village will lead to an expected increase of 100 DHF cases, assuming other factors in the model remain constant. Drinking water facilities include PDAM, Non-PDAM, dug wells, and bore wells. Water in these facilities can become a breeding ground for mosquitoes, regardless of quality. When water becomes stagnant, it creates an ideal environment for mosquitoes to breed.

3.7. Mapping of Estimated Values Based on the SAR Poisson Model

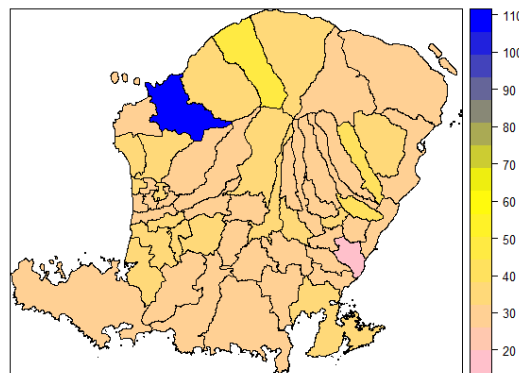


Figure 2. Mapping of Estimated Values of Variable Y Based on the SAR Poisson Model

Figure 2 illustrates the mapping of estimated values for variable Y based on the SAR Poisson model. As depicted by this model, the distribution of DHF cases across Lombok Island shows a relatively uniform pattern among adjacent districts. This suggests that neighboring districts exhibit similar patterns in the spread of DHF cases. The areas with the highest incidence of DHF are represented in blue, corresponding to 100 to over 110 cases. Conversely, the areas with the lowest incidence are shown in pink, corresponding to 0 to 20 cases. The map highlights that the district with the highest number of DHF cases, as indicated by the blue color, is Tanjung in North Lombok.

Figures 1 and 2 show a noticeable difference in the highest incidence of DHF cases. According to the actual data presented in Figure 1, the highest DHF cases are observed in the districts of Suela and Suralaga in East

Lombok. In contrast, the estimated DHF cases based on the SAR Poisson model, as shown in Figure 2, indicate that the highest incidence is found in the Tanjung district of North Lombok. This discrepancy may be attributed to potential issues with data meeting the assumption of equidispersion and the limited number of predictor variables used in the model. Therefore, future research should address these issues by verifying the equidispersion assumption and incorporating a broader range of predictor variables, along with a larger dataset, to enhance the reliability of the SAR Poisson model.

4. CONCLUSION

Based on the conducted analysis, it was found that there is significant spatial autocorrelation among the districts on Lombok Island regarding the distribution of DHF cases. Consequently, the SAR Poisson model was applied. The results from the SAR Poisson modeling indicate that all model parameters are significant except for the parameter X_1 . Additionally, the variability in the number of DHF cases can be explained by the predictor variables with an R^2 devians (R_{DEV}^2) = 10.29%.

REFERENCES

- [1] N. Izzati, H. K. Wardana, and J. W. Leksono, “Sosialisasi Earth Hour Sebagai Upaya Gerakan Penghematan Listrik Melalui Siaran Radio Suara Tebuireng Jombang,” *Sewagati*, vol. 8, no. 1, pp. 1061–1070, 2023, doi: 10.12962/j26139960.v8i1.695.
- [2] A. A. Sidharta, F. Diniarti, and D. Darmawansyah, “Analisis Spasial Faktor Risiko Kejadian Demam Berdarah Dengue Di Kota Bengkulu,” *J. Vokasi Kesehat.*, vol. 2, no. 2, pp. 43–56, 2023, doi: 10.58222/jvokes.v2i2.162.
- [3] R. U. Hurit, A. O. B. Puka, and M. Y. Maing, “Penerapan Metode Euler Dan Heun Pada Penyebaran Penyakit Demam Berdarah,” *Prox. J. Penelit. Mat. dan Pendidik. Mat.*, vol. 7, no. I, pp. 18–26, 2024.
- [4] D. NTB, Kesehatan, “Jumlah Kasus Demam Berdarah Dangu (DHF) di Provinsi,” 2020. [Online]. Available: <https://data.ntbprov.go.id/dataset/jumlah-kasus-demam-berdarah-dangu-DHF-di-provinsi-ntb>.
- [5] S. Sukardin, N. M. Sumartyawati, and N. Nurhidayah, “Edukasi tentang Pencegahan Penularan Demam Berdarah Dengue (DHF) pada Siswa di SDN 1 Kekeru Lombok Barat,” *J. Abdimas Kesehat.*, vol. 5, no. 1, p. 139, 2023, doi: 10.36565/jak.v5i1.479.
- [6] M. Maya Utari and E. Retno Setyowati, “Hubungan Sgpt Dan Sgot Dengan Jumlah Trombosit Pada Pasien Demam Berdarah Dengue (DHF) Di Rsudp Ntb,” *Nusant. Hasana J.*, vol. 2, no. 5, pp. 113–117, 2022.
- [7] N. Amni, “ANALISIS AREA GEMPA BUMI DI PROVINSI LAMPUNG DENGAN METODE SPATIAL AUTOREGRESSIVE (SAR),” p. 6, 2021.
- [8] S. R. Rohimah and A. M. Riyantobi, “Model Spasial Otoresif Poisson untuk Mendeteksi Faktor-Faktor yang Berpengaruh terhadap Jumlah Penderita HIV AIDS di Jakarta Timur,” *J. Stat. dan Apl.*, vol. 3, no. 2, pp. 35–44, 2019, doi: 10.21009/jsa.03205.
- [9] A. P. N. Taryono, D. Ispriyanti, and A. Prahutama, “Analisis Faktor-Faktor yang Mempengaruhi Penyebaran Penyakit Demam Berdarah Dengue (DHF) di Provinsi Jawa Tengah dengan Metode Spatial Autoregressive Model dan Spatial Durbin Model,” *Indones. J. Appl. Stat.*, vol. 1, no. 1, p. 1, 2018, doi: 10.13057/ijas.v1i1.24026.
- [10] T. S. Mahading, R. Resmawan, L. Yahya, and I. A. Rizkyani, “Metode Spatial Autoregressive Dalam Analisis,” *JMPM J. Mat. dan Pendidik. Mat.*, vol. 5, no. 2, pp. 9–19, 2020.
- [11] Y. D. Pratiwi, S. Mariani, and P. Hendikawati, “Pemodelan Regresi Spasial menggunakan Geographically Weighted Regression,” *UNNES J. Math.*, vol. 8, no. 2, pp. 32–41, 2019, [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/ujme>

- [12] N. P. Sukmawati, *METODE REGRESI POISSON DENGAN SPATIAL LAG PADA PEMODELAN HIV/AIDS DI KABUPATEN DAN KOTA DI PROVINSI JAWA TIMUR TAHUN 2017*. Doctoral dissertation, Universitas Brawijaya, 2019.
- [13] A. P. Kusuma and A. D. Oktavianto, “Analisis Metode Euclidean Distance dalam Menentukan Koordinat Peta pada Alamat Rumah,” *J. Teknol. dan Manaj. Inform.*, vol. 8, no. 2, pp. 108–115, 2022, doi: 10.26905/jtmi.v8i2.8871.
- [14] Y. Miftahuddin, S. Umaroh, and F. R. Karim, “PERBANDINGAN METODE PERHITUNGAN JARAK EUCLIDEAN , HAVERSINE , (STUDI KASUS : INSTITUT TEKNOLOGI NASIONAL BANDUNG),” vol. 14, no. 2, pp. 69–77, 2020.
- [15] W. A. K. Saputri and K. Suryowati, “Analisis Faktor-Faktor Yang Mempengaruhi Gini Ratio Di Provinsi Papua Dengan Model Spasial Data Panel,” *Jurnal Statistika Industri dan Komputasi*, vol. 03, no. 2. pp. 1–11, 2018.
- [16] R. Ambarwati and S. P. Hastono, “Pemodelan Faktor-faktor yang Mempengaruhi Balita Gizi Buruk di Provinsi Jawa Barat Tahun 2017 dengan Pendekatan Spatial Autoregressive Model (SAR),” *J. Ilmu Kesehat. Masy.*, vol. 9, no. 02, pp. 121–128, 2020, doi: 10.33221/jikm.v9i02.507.

