JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 9, No. 2, April 2025, pp. 686-698

Spatial Panel Regression Modelling of Rainfall in Indonesia

Fang You Dwi Ayu Shalu Saniyawati[1], Suci Astutik[1*], Henny Pramoedyo[1]
[1]Department of Statistics, Brawijaya University, Indonesia
suci_sp@ub.ac.id

Article History:
Received : 23-02-2025
Revised : 31-03-2025
Accepted : 07-04-2025
Online : 30-04-2025

Keywords:
Spatial;
Spatial Panel;
Regression;
Rainfall.

ABSTRACT

Rainfall is amount of water that falls to the earth's surface in the form of rain during a certain period of time, usually measured in millimeters. Rainfall data in Indonesia usually includes temporal and spatial dimensions, so the appropriate method for its analysis is spatial panel regression analysis. This study aims to identify factors that influence the amount of rainfall in Indonesia. This type of research is quantitative using secondary data from the central statistics agency website. The predictor variables used include air temperature, sunshine radiation, humidity, wind speed, and air pressure, while the response variable is amount of rainfall in 34 provinces in Indonesia. Spatial panel regression analysis is carried out using maximum likelihood estimation, which is used to estimate the regression coefficient and intercept that maximizes the likelihood of the existing data. Based on the lagrange multiplier test, spatial autocorrelation was found in the lag, so the appropriate model is SAR-FE. This model can overcome spatial autocorrelation by taking into account spatial interactions between locations, as well as controlling unobserved heterogeneity through fixed effects. The results show that sunshine radiation, humidity, and wind speed have significant effect on the amount of rainfall in Indonesia. The AIC value of SAR-FE model ($-4.352594 \times 10^{-13}$) is smaller than SEM-FE model ($-1.642001 \times 10^{-12}$), indicating that SAR-FE model is better at explaining the data.

Crossref
https://doi.org/10.31764/jtam.v9i2.29882

———————— ◆ ————————

A. INTRODUCTION

Rain is a natural phenomenon that shows the fall of water droplets from the atmosphere to the earth's surface (Ajr & Dwirani, 2019). Rainfall is one of the rain parameters that can be measured in millimeters. Rainfall is important in supporting various sectors, including agriculture and water availability. Rainfall in Indonesia often experiences unpredictable changes in patterns. Uncertainty in rainfall patterns results in an imbalance in water supply, which affects food production and increases vulnerability to natural disasters. To anticipate these natural disasters, a deep understanding of the factors that influence rainfall is needed. One method to determine the influencing factors is to use linear regression analysis. This analysis allows to identify the relationship between variables that influence rainfall statistically.

Linear regression analysis is one of the statistical methods commonly used in research, especially research involving observation. This analysis can provide valuable information if carried out with high accuracy (Roustaei, 2024). Regression analysis is usually used on data containing cross-sectional or temporal dimensions. The results of observations when applied not only contain various individuals/units at one time, but also a period of time (time series). Data that combines both dimensions is called panel data. If regression analysis is applied to panel data, it will cause bias in the estimation results. To overcome this problem, panel regression analysis can be used.

Panel regression analysis is a regression analysis used to determine the effect of one or more predictor variables on the response variable. This analysis has a data structure consisting of several periods. The advantage of panel regression analysis is that it can be used to identify and calculate influences that cannot be found in cross-sectional models. This analysis also takes into account the time dimension of the data. In addition, this analysis can overcome individual heterogeneity and produce more efficient and accurate estimates (Stefko et al., 2021). With panel regression analysis, variations between individuals and time can be analyzed simultaneously. This allows for a better understanding of the dynamics of the relationship between variables.

Panel regression analysis can be applied to various areas of life. A study applied panel regression to the case of entrepreneurship, financial sector development, and openness. The study used data from 15 upper-middle-income and high-income countries in 2001–2015. The results of this study indicate that trade openness, banking sector and capital market development, and FDI inflows have a significant impact on entrepreneurship (Bayar et al., 2018). In addition, panel regression can be applied to cases of child labor and rainfall deviations. The results show that above-average rainfall is closely related to children working in agriculture and those doing household chores (Trinh et al., 2020). However, the weakness of panel regression analysis is the difficulty in meeting the non-autocorrelation assumption.

In addition to having time and individual dimensions, data also has a location dimension. Data related to location is called spatial data. Regression analysis used when data contains location elements is called spatial regression analysis. The spatial regression model is a development of the simple regression model where this model is based on the influence of location on the data being analyzed (Tarigan, 2021). Tobler's first law states that everything is always related to something else, but something that is close will be more related than something that is far away (Yin et al., 2018). The advantages of using spatial regression analysis are that it can overcome the problem of spatial autocorrelation, consider the location dimension of the data, and can increase prediction accuracy. Tarigan's (2021) research on the spatial regression model on the development index of North Sumatra Province in 2020 shows that the spatial regression model is the right method to apply to this case because it takes location into account. The results of this study indicate that the chosen model is spatial sutoregressive with several assumptions that must be met, namely: the assumption of normality, multicollinearity, and heteroscedasticity. The weakness of this spatial regression analysis is in fulfilling the assumptions.

If the observation data contains location, time series, and cross-section information, then it can be analyzed using spatial panel regression analysis. Spatial panel regression analysis is able to accommodate spatial dependencies and provide more robust and efficient parameter estimates. This regression analysis uses Maximum Likelihood Estimation (MLE). MLE estimation is a parameter estimation method to estimate unknown population parameters by maximizing the likelihood of a probabilistic model. The advantage of this estimation is that the parameters will be efficient, consistent, and close to a normal distribution if in a large sample so that it can meet the assumption of normality. This spatial panel regression analysis can be applied to various fields. For example: in the economic field, modeling the crime rate in East Java with a spatial panel model results in a spatial dependency effect on the model and the model chosen is the spatial lag fixed effect with an R^2 of 71.54% (Yanti et al., 2023).

This study aims to determine the factors that influence the amount of rainfall in Indonesia through spatial panel regression analysis. By considering the significant variability of rainfall, which is influenced by meteorological phenomena such as El Nino and La Nina, this study identifies the factors that influence rainfall patterns in various regions. The results of this study are expected to provide a deeper understanding of the spatial and temporal dynamics of rainfall in Indonesia. This study also aims to help understand the impact of changes in rainfall patterns on important sectors such as agriculture, water resources management, and disaster mitigation. The main contribution of this study is to provide more accurate and data-based information that can be used to design more appropriate policies in dealing with floods and droughts. With a comprehensive analysis, government policies in disaster mitigation are expected to be more focused and effective. Therefore, this study has the potential to be an important reference in planning natural resource management in Indonesia.

B. METHODS

This study uses rainfall data in 34 provinces in Indonesia in 2019-2023 sourced from the website of the Central Statistics Agency. This study is explanatory because it uses statistical analysis to test hypotheses and explain the causes and effects of rainfall.

1. Chow Test

Chow test can be utilized to consider whether or not there is a diversity intercepts in the midst of cross-section units in the model (Yanti et al., 2023). The hypothesis used is:

$H_0 :\ \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$

H_1 : at least there is one $\alpha_i \neq 0$ where $i = 1,2,\dots,n$

Test statistics:

$$F_0 = \frac{(RRSS-URSS)/(N-1)}{(URSS)/(NT-N-K)} \tag{1}$$

Information:

RRSS : restricted residual sums of square from CEM
URSS : unrestricted residual sums of square from FEM
N : number of individuals
T : number of time periods
K : number of parameters in FEM

The decision criteria in drawing conclusions are H_0 is rejected if the p-value $< \alpha$, so the model used is FEM and continued to Hausman test.

2. Hausman Test

Hausman test can be utilized to determine whether the error is correlated with the dependent variable (Yanti et al., 2023). Hausman test is also based on diversity in variety of two estimators (Lakshmanasamy, 2015). The hypothesis used is:

$H_0 :\ E(w_{it}|X_{it}) = 0$

$H_1 : E(w_{it}|X_{it}) \neq 0$

Test statistics:

$$\chi^2_{hit} = \widehat{\boldsymbol{q}}'[Var(\widehat{\boldsymbol{q}})]^{-1}\widehat{\boldsymbol{q}} \tag{2}$$

with $\widehat{\boldsymbol{q}} = \widehat{\boldsymbol{\beta}}_{\boldsymbol{random}} - \widehat{\boldsymbol{\beta}}_{\boldsymbol{fixed}}.\widehat{\boldsymbol{\beta}}_{\boldsymbol{random}}$

Information:

i : number of individual units

t : number of time units

k : dimension of β vector

$\widehat{\boldsymbol{\beta}}_{\boldsymbol{random}}$: vector of independent variable coefficient of the model random effects

$\widehat{\boldsymbol{\beta}}_{\boldsymbol{fixed}}$: vector of independent variable coefficient of the model fixed effects

The decision criteria in drawing conclusions are H_0 is accepted if the p-value $> \alpha$, then the model used is REM. Meanwhile, if H_0 is rejected, then selected model is FEM. After that, do spatial modeling. Spatial panel regression analysis has three models, namely: spatial lag panel model, spatial error panel model, and spatial durbin panel model (Belotti et al., 2017). Previously, the formation of a spatial weighting matrix was carried out.

3. Spatial Weighting Matrix

Spatial weighting matrix or $\boldsymbol{W}$ that describes correlation of proximity between observation areas/locations measuring $n \times n$. The form of spatial weighting matrix can be formulated bellow (Cheung et al., 2020):

$$\boldsymbol{W} = \begin{bmatrix} 0 & \dots & W_{1n} \\ \vdots & \ddots & \vdots \\ W_{n1} & \dots & 0 \end{bmatrix}$$

4. Moran's I Test

Moran's I test aims to find out whether or not there is spatial dependency (Annur, 2019). Hypothesis of Moran's I test:

$H_0 : I = 0$

$H_1 : I \neq 0$

Test statistics (Wang & Yang, 2019):

$$Z_{hit} = \frac{I - E(I)}{\sqrt{Var(I)}} \sim N\ (0,1) \tag{3}$$

Decision criteria used are when $|Z(I)| > Z_{\alpha/2}$ then reject H_0. This means that there is spatial autocorrelation in the model or in other words, the characteristics of a region are related to other regions located nearby. After that, spatial effect test is carried out which is used to decide the spatial influence model in the data called the Lagrange Multiplier test (Arif et al., 2019). After the spatial influence in the data is known, a spatial panel regression model is formed.

5. Spatial Panel Regression

Spatial panel regression utilized to examine relationship among response variables and predictor variables in context of spatial panel data. Spatial panel data combines time dimensions and spatial dimensions, allowing analysis that considers the spatial structure in the data. In spatial panel regression analysis, there are several types of models that are commonly used to solve various problems (Elhorst, 2014):

a. Spatial Autoregressive Model (SAR)
   The general form of the SAR model in panel data can be expressed in equation (4) and equation (5).
   1) SAR *Fixed Effect Model*

$$y_{it} = \delta \sum_{j=1}^{N} w_{ij} y_{jt} + \boldsymbol{x_{it}\beta} + \mu_i + \varepsilon_{it} \tag{4}$$

   2) SAR *Random Effect Model*

$$y_{it} = \delta \sum_{j=1}^{N} w_{ij} y_{jt} + \boldsymbol{x_{it}\beta} + \theta + \varepsilon_{it} \tag{5}$$

b. Spatial Error Model (SEM)
   The general form of the SEM model in panel data can be expressed in equation (6) and equation (7).
   1) SEM *Fixed Effect Model*

$$y_{it} = \boldsymbol{x_{it}\beta} + \mu_i + \phi_{it} \tag{6}$$

   With $\phi_{it} = \lambda \sum_{j=1}^{N} w_{ij}\, \phi_{jt} + \varepsilon_{it}$
   2) SEM *Random Effect Model*

$$y_{it} = \boldsymbol{x_{it}\beta} + \theta + \phi_{it} \tag{7}$$

   With $\phi_{it} = \lambda \sum_{j=1}^{N} w_{ij}\, \phi_{jt} + \varepsilon_{it}$. Where:
   y_{it} : response variable at the i-th individual unit and t-th time unit
   y_{jt} : response variable at location unit j and time unit t
   δ : spatial autoregressive coefficient
   w_{ij} : elements of the spatial weighting matrix $\boldsymbol{W}$
   μ_i : spatial specific effects of individual unit i
   ε_{it} : *error of individual unit i and time unit t*
   θ : random effects parameters in spatial panel regression models
   λ : spatial autocorrelation coefficient
   ϕ_{it} : spatial autocorrelation error

   Next, best model is selected by utilize Akaike Information Criterion (AIC). Selected model is model that has smallest AIC value. After determining best model, normality and homogeneity assumptions are tested.

6. Assumptions Test

Normality assumption is one of important assumptions underlying the model in statistical analysis (Khatun, 2021). One of the normality tests is jarque bera test. This test is carried out by calculating the skewness and kurtosis values of the model residuals. Hypothesis:

$H_0 : \varepsilon_i \sim N(0, \sigma^2)$
$H_1 : \varepsilon_i \nsim N(0, \sigma^2)$
Test statistics (Gujarati, 2004):

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right] \quad (8)$$

Description:
n : sample size
S : skewness coefficient
K : kurtosis coefficient

Decision criteria used in this test is if the test statistic $JB \geq X^2_{\alpha,df}$ or $p\ value < \alpha$ then reject H_0. It means that residuals are normally distributed. After that, homogeneity assumption was carried out. Homogeneity assumption aims to decide whether absolute residual values of all observations of regression model are equal (Alita et al., 2021). Homogeneity test using breusch pagan. Breusch-Pagan tests whether residual variance is related or not to a set of explanatory variables (Klein et al., 2016). Hypothesis (Uyanto, 2022):

H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$
H_1: at least there is one $\alpha_p = 0$
Test statistics:

$$BP = \frac{1}{2}[f'Z(Z'Z)^{-1}Z'f] + \frac{1}{T}[\frac{e\prime We}{\sigma^2}]^2 \sim \chi^2_{(k+1)} \quad (9)$$

The decision criteria used in this test is if the test statistic $BP \geq X^2_{\alpha,df}$ or $p\ value < \alpha$ then reject H_0. This means that the residual variance is not homogeneous (heterogeneous). Finally, an interpretation of the model is carried out.

C. RESULT AND DISCUSSION

1. Descriptive Analysis

Table 1. Descriptive Statistics of Rainfall in Indonesia

Year	Minimum	Maximum	Average
2019	637.60	4072.70	1950.23
2020	954.30	4731.00	2748.26
2021	1000.80	5332.30	2932.21
2022	879.40	4950.50	2866.73
2023	587.50	5313.70	2135.57

Table 1 shows that the lowest rainfall in Indonesia was obtained by Central Sulawesi in 2023 at 587.50 mm with an average rainfall in that year of 2135.57 mm. Meanwhile, the highest rainfall in Indonesia was 5332.30 mm obtained by West Sumatra in 2021. This figure is above the toddler average for rainfall in 2021, which was 2932.21 mm.

2. Multicollinearity Examination

Multicollinearity examination aims to determine whether or not there is a linear correlation among two or more independent variables, so that the independent variables become mutually dependent. This detection is done by using the Variance Inflation Factor (VIF) value. If VIF value ≥ 10, then there is multicollinearity between independent variables. This study produced the following VIF values.

Table 2. Results of Multicollinearity Examination

Variable	VIF
X1 (Temperature)	4.979697
X2 (Sunshine Radiation)	2.012380
X3 (Humadity)	3.384352
X4 (Wind Speed)	1.959272
X5 (Air Pressure)	3.282626

Table 2 show that all VIF are less than 10. It means that all predictor variables used in this research have met assumption of non-multicollinearity.

3. Panel Regression Analysis

Panel regression analysis is an analysis utilized to model corellation among predictor variables and response variable with presence of temporal and cross section information. Panel data regression modeling is divided into three, namely CEM, FEM, and REM. Below is a panel regression model that is formed.

a. Common Effect Model (CEM)

$$\hat{Y}_{it} = 13998.53 + 429.63X_{1_{it}} - 6.58X_{2_{it}} + 149.81X_{3_{it}} + 99.64X_{4_{it}} - 35.19X_{5_{it}} \quad (10)$$

b. Fixed Effect Model (FEM)

$$\hat{Y}_{it} = 228.75X_{1_{it}} - 21.86X_{2_{it}} + 108.74X_{3_{it}} + 351.24X_{4_{it}} - 150.73X_{5_{it}} \quad (11)$$

c. Random Effect Model (REM)

$$\widehat{Y}_{it} = 16539.51 + 289.79X_{1_{it}} - 22.07X_{2_{it}} + 120.24X_{3_{it}} + 344.93X_{4_{it}} - 31,17X_{5_{it}} \quad (12)$$

After the panel regression model is formed, chow test aims to decide whether or not there is a difference in intercepts between cross-section units in the model. This test also determine selected model among CEM or FEM. Therefore, next step is to continue with the chow test.

4. Chow Test

Hypothesis:

$H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$

H_1 : at least there is one $\alpha_i \neq 0$ where $i = 1,2,\dots,n$

Decision Criteria:

a. Reject H_0 if $p - value < \alpha\ (0.05)$

b. Accept H_0 if $p - value > \alpha\ (0.05)$

Results of chow test showed that p-value of this test is $< 2.2 \times 10^{-16}$. Because the p-value is smaller than alpha (0.05), then reject H_0 and selected model is FEM and can be continued to Hausman test.

5. Hausman Test

Hypothesis:

$H_0: E(w_{it}|X_{it}) = 0$

$H_1: E(w_{it}|X_{it}) \neq 0$

Decision Criteria:

Reject H_0 if $p - value < \alpha\ (0.05)$

Accept H_0 if $p - value > \alpha\ (0.05)$

Results of hausman test showed that p-value of this test is 1.194×10^{-8}. Because p-value is smaller than alpha (0.05), then reject H_0 and selected model is FEM. Furthermore, we can compile a spatial weighting matrix.

6. Spatial Weighting Matrix

Spatial regression analysis requires a spatial weighting matrix. This matrix is used to describe a relationship of proximity between regions/observation locations. In this research, queen contiguity weighting matrix was used, which considers the intersection of sides and corners at the observation location. For example: North Sumatra based on the map of the region intersects sides and corners with Aceh, West Sumatra, and Riau so that it is given a weight of 1, while for other provinces it is worth 0. In 34 provinces, each is given a weight and then the matrix is standardized. The following is the form of the standardization matrix of the queen contiguity weighting.

$$\boldsymbol{W} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0.33333 & 0 & 0.33333 & 0.33333 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0.25 & 0 & 0.25 & 0.25 & 0 & 0.25 & 0 & \dots & 0 \\ 0 & 0.25 & 0.25 & 0 & 0.25 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0.2 & 0.2 & 0 & 0.2 & 0.2 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0 & 0.25 & 0.25 & \dots & 0 \\ 0 & 0 & 0.25 & 0 & 0.25 & 0.25 & 0 & 0.25 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

7. Moran's I Test

Table 3. Results of Moran's I Test

Variable	P-Value
Y	4.844×10^{-7}

Table 3 showed that the p-value obtained is less than alpha (0.05) on response variable, so H_0 is rejected. Therefore, in the model it can be said that there is spatial autocorrelation between regions.

8. Lagrange Multiplier Test

After spatial autocorrelation test is implement, it is continued with the Lagrange multiplier test. This test is carried out to decide spatial influence model in a data.

a. Lagrange Multiplier Test in the Lag
Hypothesis:
$H_0 : \delta = 0$
$H_1 : \delta \neq 0$

Table 4. Results of the Lagrange Multiplier Test in the Lag

LM_δ	P-Value
15.177	9.791×10^{-5}

Table 4 showed that results of lagrange multiplier test on the lag produced p-value (9.791×10^{-5}) less than alpha (0.05), so H_0 is rejected. It means that there is spatial dependency of lag in the fixed effect model.

b. Lagrange Multiplier Test in the Error
Hypothesis:
$H_0 : \lambda = 0$
$H_1 : \lambda \neq 0$

Table 5. Results of the Lagrange Multiplier Test in the Error

LM_λ	P-Value
4.2556	0.03912

Table 5 showed that results of lagrange multiplier test on the error produced p-value (0.03912) less than alpha (0.05), so H_0 is rejected. It means that there is spatial dependency of error in the fixed effect model.

9. Spatial Panel Regression Model

After the lagrange multiplier test, it can be continued by forming a spatial panel regression model. Because there is spatial dependency of lag and error, selected model is SAR-FE and SEM-FE. Significance of the parameters in each model with queen contiguity weighting can be shown in Table 6.

Table 6. Results of Parameter Estimation

Variable	SAR-FE		SEM-FE	
	Coefficient	P-Value	Coefficient	P-Value
X_1	79.9978	0.3731302	72.8260	0.4662788
X_2	-18.6959	0.0001184	-20.1944	0.0005126
X_3	74.2949	3.658×10^{-5}	70.0298	0.0002168
X_4	277.8011	3.988×10^{-6}	322.3330	3.904×10^{-6}
X_5	-57.2987	0.2728471	-55.8290	0.3343207
$\hat{\delta}$	0.337172	2.644×10^{-8}	-	-
$\hat{\lambda}$	-	-	0.304811	5.111×10^{-6}

Table 6 showed that if alpha (0.05), then the parameter $\hat{\delta}$ is significant in the model. The results of the parameter estimation above can be used to form a model.

a. SAR-FE

$$\hat{Y}_{it} = 0.337172 \sum_{j=1}^{N} w_{ij}\hat{Y}_{jt} + 79.9978\, X_{1_{it}} - 18.6959\, X_{2_{it}} + 74.2949\, X_{3_{it}} + 277.8011\, X_{4_{it}} - 57.2987\, X_{5_{it}} + \mu_i \quad (13)$$

b. SEM-FE

$$\hat{Y}_{it} = 72.8260\, X_{1_{it}} - 20.1944\, X_{2_{it}} + 70.0298\, X_{3_{it}} + 322.3330\, X_{4_{it}} - 55.8290\, X_{5_{it}} + \mu_i + \hat{\phi}_{it} \quad (14)$$

$$\hat{\phi}_{it} = 0.304811 \sum_{j=1}^{N} w_{ij}\hat{\phi}_{jt} \quad (15)$$

Table 7. AIC Value in Spatial Panel Regression Model

Model	AIC Value
SAR-FE	$-4.352594 \times 10^{-13}$
SEM-FE	$-1.642001 \times 10^{-12}$

Table 7 showed that best model is SAR-FE with queen contiguity weighting because it has the smallest AIC value. SAR-FE is used to model the amount of rainfall in Indonesia based on provinces in 2019-2023.

10. Assumption Testing

After the formation of the spatial panel regression model, normality assumption test was undertake on the error by lilliefors test on SAR-FE. The results of this test showed that p-value (0.3359) $>$ alpha (0.05) was obtained, then H_0 was accepted. Thus, it can be concluded that error is normally distributed. Furthermore, homogeneity assumption test was also carried out on the error using the Breusch-Pagan test statistic. The result of this test showed that p-value (0.5875) $>$ alpha (0.05) was obtained, then H_0 was accepted. Thus, it can be concluded that the error has a homogeneous variety.

11. Interpretation

Based on equation (13), it can be seen that the amount of rainfall in the i-th province is influenced by the amount of rainfall in the neighboring j-th province. When the amount of rainfall in the neighboring j-th province increases by 1 mm, then the amount of rainfall in the observed i-th province will increase by 0.337172 mm assuming other factors are considered constant. In addition, an example of an equation model can be taken for one of the regions, namely North Sumatra. North Sumatra has similar characteristics to Aceh, West Sumatra, and Riau.

$$(\hat{Y}_{it})_{North\ Sumatra} = 0.337172(\hat{Y}_{it})_{Aceh} + 0.337172\ (\hat{Y}_{it})_{West\ Sumatra} + 0.337172\ (\hat{Y}_{it})_{Riau} + 79.9978X_{1_{it}} - 18.6959\ X_{2_{it}} + 74.2949\ X_{3_{it}} + 277.8011\ X_{4_{it}} - 57.2987\ X_{5_{it}} + \mu_i \quad (16)$$

Interpretation of the model that has been formed shows that the increasing amount of rainfall in North Sumatra is influenced by Aceh, West Sumatra, and Riau. If the amount of rainfall in the province increases by 1 mm, then the amount of rainfall in North Sumatra which has similar characteristics increases by 0.337172 mm. The similarity of these characteristics shows that there is a relationship in a region that has close proximity to its neighbors in the region. This relationship can be shown through the diversity and similar potential between adjacent regions. In addition, if air temperature increases by 1 °C, rainfall will increase by 79.9978 mm, assuming other variables are constant. If duration of sunshine increases by 1%, rainfall will decrease by 18.6959 mm, assuming other variables are constant. If humidity increases by 1%, rainfall will increase by 74.2949 mm, assuming other variables are constant. If wind speed increases by 1 m/s, rainfall will increase by 277.8011 mm, assuming other variables are constant. If air pressure increases by 1 mbar, rainfall will decrease by 57.2987 mm, assuming other variables are constant.

D. CONCLUSION AND SUGGESTIONS

Based on this research, it can be concluded that best spatial panel regression model in the case of rainfall in Indonesia is Spatial Autoregressive - Fixed Effect (SAR-FE) with a queen contiguity weighting matrix. The model can relate spatial dependencies between adjacent locations (especially in rainfall data), thus providing more accurate estimates compared to conventional regression models. Based on results of estimation and significance parameters in the model, there are three out of five variables that have a significant effect on the case of rainfall, namely sunshine radiation (X_2), humadity (X_3), and wind speed (X_4). However, this model is not completely accurate due to the possibility of unaccounted for non-observed variables. Further research is needed to address this issue and test the sensitivity of the model to changes in parameters and assumptions. Beside that, further research can use a distance weighting matrix or other estimation methods such as Bayesian, so that spatial panel regression model is obtained with high level of model goodness.

ACKNOWLEDGEMENT
Praise be to Allah SWT for the abundance of blessings, grace, and guidance so that the author can complete the writing of this journal. Author want to thank the lecturers Department of Statistics in Brawijaya University, parents, all family, and friends who have provided advice, support, and motivation until the writing of this journal was completed.

REFERENCES
Ajr, E. Q., & Dwirani, F. (2019). Menentukan Stasiun Hujan dan Curah Hujan dengan Metode Polygon Thiessen Daerah Kabupaten Lebak. *JURNALIS (Jurnal Lingkungan Dan Sumberdaya Alam)*, *2*(2), 139–146. https://ejournal.lppm-unbaja.ac.id/index.php/jls/article/view/674/387

Alita, D., Putra, A. D., & Darwis, D. (2021). Analysis of Classic Assumption Test and Multiple Linear Regression Coefficient Test For Employee Structural Office Recommendation. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *15*(3), 295–306. https://doi.org/10.22146/ijccs.65586

Annur, M. F. (2019). Analisis Faktor Indeks Pembangunan Manusia di Kabupaten Landak Menggunakan Model Regresi Spasial. *RIEMANN (Research of Mathematics and Mathematics Education)*, *1*(1), 44–54. https://doi.org/10.38114/riemann.v1i1.14

Arif, A., Tiro, M. A., & Nusrang, M. (2019). Metode Boostrap dan Jackknife dalam Mengestimasi Parameter Regresi Linear Ganda (Kasus: Data Kemiskinan Kota Makassar Tahun 2017). *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, *1*(2), 66–76. https://doi.org/10.35580/variansiunm12895

Bayar, Y., Gavriletea, M. D., & Ucar, Z. (2018). Financial Sector Development, Openness, and Entrepreneurship: Panel Regression Analysis. *Sustainability (Switzerland)*, *10*(10), 1–11. https://doi.org/10.3390/su10103493

Belotti, F., Hughes, G., & Mortari, A. P. (2017). Spatial Panel-Data Models using Stata. *Stata Journal*, *17*(1), 139–180. https://doi.org/10.1177/1536867x1701700109

Cheung, T. K. Y., Wong, W. hung, Zhang, A., & Wu, Y. (2020). Spatial Panel Model for Examining Airport Relationships within Multi-Airport Regions. *Transportation Research Part A: Policy and Practice*, *133*(January), 148–163. https://doi.org/10.1016/j.tra.2019.12.011

Elhorst, J. P. (2014). Spatial Econometrics. In *SpringerBriefs in Regional Science* (pp. 1–119). https://doi.org/10.1007/978-3-642-40340-8_1

Gujarati, D. N. (2004). Basic Econometrics. In *The McGraw-Hill Companies* (pp. 1–1002). https://doi.org/10.1596/1813-9450-8096

Khatun, N. (2021). Applications of Normality Test in Statistical Analysis. *Open Journal of Statistics*, *11*(1), 113–122. https://doi.org/10.4236/ojs.2021.111006

Klein, A. G., Gerhard, C., Büchner, R. D., Diestel, S., & Schermelleh-Engel, K. (2016). The Detection of Heteroscedasticity in Regression Models for Psychological Data. *Psychological Test and Assessment Modeling*, *58*(4), 542–568. https://www.researchgate.net/publication/311518028

Lakshmanasamy. (2015). The Effect of Monsoon Rainfall Shortage on Farmer Suicides in India - A Panel Data Econometric Analysis. *The Indian Geographical Journal*, *90*(1), 14–25. www.igschennai.org

Roustaei, N. (2024). Application and Interpretation of Linear-Regression Analysis. *Medical Hypothesis, Discovery, and Innovation in Ophthalmology*, *13*(3), 151–159. https://doi.org/10.51329/mehdiophthal1506

Stefko, R., Gavurova, B., Kelemen, M., Rigelsky, M., & Ivankova, V. (2021). Relationships between Renewable Energy and the Prevalence of Morbidity in the Countries of the European Union: A panel Regression Approach. *International Journal of Environmental Research and Public Health*, *18*(12), 1–18. https://doi.org/10.3390/ijerph18126548

Tarigan, W. S. (2021). Analisis Regresi Spasial pada Indeks Pembangunan Manusia di Provinsi Sumatera Utara Tahun 2020. *Seminar Nasional Official Statistics*, *2021*(1), 403–408. https://doi.org/10.34123/semnasoffstat.v2021i1.896

Trinh, T. A., Posso, A., & Feeny, S. (2020). Child Labor and Rainfall Deviation: Panel Data Evidence from Rural Vietnam. *Developing Economies*, *58*(1), 63–76. https://doi.org/10.1111/deve.12215

Uyanto, S. S. (2022). Monte Carlo Power Comparison of Seven Most Commonly used Heteroscedasticity

Tests. *Communications in Statistics: Simulation and Computation*, *51*(4), 2065–2082. https://doi.org/10.1080/03610918.2019.1692031

Wang, H., & Yang, J. (2019). Total-Factor Industrial Eco-Efficiency and Its Influencing Factors in China: A Spatial Panel Data Approach. *Journal of Cleaner Production*, *227*(1), 263–271. https://doi.org/10.1016/j.jclepro.2019.04.119

Yanti, N. K. J. F., Susilawati, M., & Suciptawati, N. L. P. (2023). Pemodelan Tingkat Kriminalitas di Provinsi Jawa Timur dengan Regresi Data Panel Spasial. *Journal on Education*, *6*(1), 1508–1517. http://jonedu.org/index.php/joe

Yin, C., Yuan, M., Lu, Y., Huang, Y., & Liu, Y. (2018). Effects of Urban Form on the Urban Heat Island Effect based on Spatial Regression Model. *Science of the Total Environment*, *634*(1), 696–704. https://doi.org/10.1016/j.scitotenv.2018.03.350