BAREKENG: Journal of Mathematics and Its Applications
September 2023 Volume 17 Issue 3 Page 1221–1234
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol17iss3pp1221-1234

MODELLING EARTHQUAKE DISASTER DAMAGE DUE TO DEPTH OF EPICENTER AND MAGNITUDE USING SPATIAL REGRESSION

Dhea Laksmita Arsya Primananda [1], Muhammad Muhajir [2*]

[1,2*] *Department of Statistics, Faculty of Mathematics and Science, Indonesian Islamic University*
Kaliurang St KM 14.5, Sleman Regency, 55584, Indonesia

*Corresponding author's e-mail: ** ***mmuhajir@uii.ac.id***

Article History:

Received: 21st November 2022
Revised: 24th June 2023
Accepted: 3rd July 2023

Keywords:

Depth of epicenter;
Earthquake;
Earthquake magnitude;
House damage;
Spatial degression;
Spatial Durbin Model (SDM)

ABSTRACT

East Java Province is geographically close to the Eurasian and Indo-Australian Plate subduction zones, resulting in frequent earthquakes. East Java Province has a high population density, so it is very risky if a disaster occurs. One preventive solution to reduce this impact is estimating damage when an earthquake occurs. The purpose of this study was to determine the best modeling of damages due to earthquakes in East Java Province, using the amount of house damage as a response variable while the depth of the epicenter and the strength of the earthquake as predictor variables. It is suspected that there is a spatial dependency effect in this case. Hence, the solution is to use regression with an area approach, namely the Spatial Durbin Model (SDM). The amount of house damage collected from BNPB, the epicenter, and the magnitude of the earthquake collected from BMKG in 2021. The result shows that SDM is good at explaining the dependency relationship between response and predictor variables. The significant predictor variables are the depth of the epicenter and the strength of the earthquake. It means that the magnitude and the depth of the epicenter of the earthquake in an area have an impact on other adjacent area. There is a relationship between the amount of house damage in one area and other adjacent areas. The Regency will have a high number of damaged houses if it is adjacent to a Regency that has a high number of damaged houses.

How to cite this article:

D. L. A. Primananda and M. Muhajir., "MODELLING EARTHQUAKE DISASTER DAMAGE DUE DEPTH OF EPICENTER AND MAGNITUDE USING SPATIAL REGRESSION," *BAREKENG: J. Math. & App.*, vol. 17, iss. 3, pp. 1221-1234, September, 2023.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article • Open Access

1. INTRODUCTION

Earthquakes are one of the disasters that often occur in Indonesia due to the location of several regions of Indonesia that are above the Indo-Australian Plate, Eurasian Plate, and Pacific Plate, so Indonesia often experiences shallow earthquakes that cause huge losses [1], [2]. The precise location, timing, size, and depth of earthquakes cannot be accurately anticipated despite the fact that a majority of earthquakes, particularly those occurring at shallow depths (0 - 70 km), have a significant impact [2], [3]. Earthquakes resulting in significant damages have been documented to occur at depths ranging from 10 to 30 kilometers, with magnitudes ranging from 5.9 to 6.8 on the Richter scale. Notable instances include the earthquakes that struck Yogyakarta and Pangandaran in 2006, as well as the earthquake in Ambon in 2017 [2]. Given the potential for significant losses resulting from seismic events, it is imperative for households to secure insurance guarantees to mitigate these risks. Additionally, effective disaster management funding is crucial for the government, which can be achieved by engaging insurance firms in the process of post-disaster rehabilitation [4], [5].

The role of insurance is important in financing damage to buildings due to earthquakes because these disasters cannot be predicted, but the amount of damage can be predicted as the history of earthquake cases that have occurred where the combination of shallow earthquake depth and earthquake strength above 5 on the Richter scale caused major damage. In statistics, a problem can be modeled and its value predicted. Because earthquakes have an epicenter and spatial tend, the use of statistical methods needs to pay attention to the proximity factor. Spatial regression methods such as Spatial Autoregressive (SAR), Spatial Error Model (SEM), and Spatial Durbin Model (SDM) can be applied for modeling problems related to the risk of damage for insurance companies because each region can have different premium values based on differences in the risk of damage [6].

One of the best models for modeling disaster losses is the SDM. The SDM with Queen Contiguity weighted the best model compared to the linear regression model in identifying the factors that most influence the production of inland common water capture fisheries in Central Java [7]. The application of the SDM is better than linear regression in explaining the factors that affect the open unemployment rate in Central Java Province based on the smallest AIC value, with Queen Contiguity weighted [8]. The SDM is also capable of modeling many filings of insurance claims as a response variable, and each predictor parameter has significant influences [6]. Therefore, the SDM with Queen Contiguity weighted is a good model for modeling house damage due to earthquakes.

In this study, East Java Province was chosen as the research area because it is located on the Eurasian plate. Based on Badan Nasional Penanggulangan Bencana (BNPB) InaRISK data, it is estimated that physical losses due to the earthquake reached 56,912,821 billion with a risk area of 2,124,852 hectares covering 38 Districts/Cities. In 2021, East Java Province is recorded to have a high earthquake intensity compared to all Provinces on Java Island. Modeling of earthquake losses due to the depth of the epicenter and the strength of the earthquake in East Java Province in 2021 uses SDM with Queen Contiguity as spatial weights. The model is also used to describe the dependence of three variables: the depth of the epicenter, the strength of the earthquake, and the earthquake disaster losses. Based on [6], [9], and [10] research, SDM is capable of modeling disaster variables both as response and predictor variables with each significant parameter and low evaluation value. SDM is considered to be able to accommodate locality phenomena between regions. The resulting model is a real fit with R square close to 1 [9]. In relation to macroeconomics, the disaster variable as a predictor variable and the SDM model with a queen contiguity weighting matrix are able to explain the current Gross Regional Domestic Product (GRDP) phenomenon [10].

One of the benefits of employing SDM is its ability to incorporate spatial effects into the model. As a result, the equation is subject to variation in each location, contingent upon the characteristics of the neighboring area. The previous study combined the disaster variable with the social economics variable, but in this study, the disaster effect is predicted according to two predictor variables. The losses due to the earthquake meant damage to houses. The results of this modeling prediction are useful for local communities involved in participating in insurance policies. The government can provide a budget-subsidized assistance program to pay insurance premiums for areas with high predictive results, and the insurance company has the benefit of setting premiums based on areas with high risk of damage.

2. RESEARCH METHODS

2.1 Spatial Regression

The basis for the development of the spatial regression method is the classical linear regression method (multiple linear regression). The development is based on the presence of place or spatial influence on the data analyzed [11]. Spatial regression with an area approach that is commonly used is the Spatial Autoregressive Model (SAR), Spatial Error Model (SEM), and Spatial Autoregressive Moving Average (SARMA). In addition to these three models, the Spatial Durbin Model also includes spatial regression, which is a form of development of the SAR model that has a spatial lag in the response variable (Y). SDM has the characteristic of having a spatial lag in the predictor variable (X) [11]. The general equation of the Spatial Durbin Model is as follows:

$$\boldsymbol{y} = \boldsymbol{\rho W y} + \boldsymbol{\alpha 1_n} + \boldsymbol{X\beta_1} + \boldsymbol{WX\beta_2} + \boldsymbol{\varepsilon} \quad (1)$$

with,

$$\boldsymbol{\varepsilon} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\sigma^2 I_n})$$

where:

$\boldsymbol{y}$	:	n × 1 vector of response variable
$\boldsymbol{\rho}$	:	spatial lag coefficient of response variable
$\boldsymbol{Wy}$	:	weighting matrix of y
$\boldsymbol{\alpha}$	:	constant parameter
$\boldsymbol{1_n}$	:	**n** × **1** vector containing 1
$\boldsymbol{X}$	:	**n** × **p** matrix of predictor variable
$\boldsymbol{\beta}$	:	p × 1 vector of regression parameter
ε	:	n × 1 error vector, which is normally distributed with zero mean and variance of $\sigma^2 I_n$
n	:	number of observations $(i = 1,2, \dots, n)$
p	:	number of predictor variable $(i = 1,2, \dots, p)$

2.2 Spatial Weights Matrix

The definition of the spatial weights matrix $\boldsymbol{W}$, where the spatial topology of the spatial units is specified, is very important since estimation results may critically depend on the choice of this matrix. The contiguity matrix represents a $n \times n$ symmetric matrix $\boldsymbol{W}$ with elements $w_{ij} = 1$ when i and j are neighbors and 0 when they are not, while queen contiguity is considered as a matrix where regions that share a common border or a vertex are considered neighbors and for these $\boldsymbol{w_{ij}} = 1$ [12]. Queen Contiguity is considered suitable to be a spatial weighting matrix because Regency/City administrative boundaries have an irregular shape. The combination of different regression models with Queen Contiguity matrices gives the best results than the Rook Contiguity and Bishop Contiguity matrices [13] – [15]. The shape of the intersection of the Queen Contiguity can be seen in Figure 1.

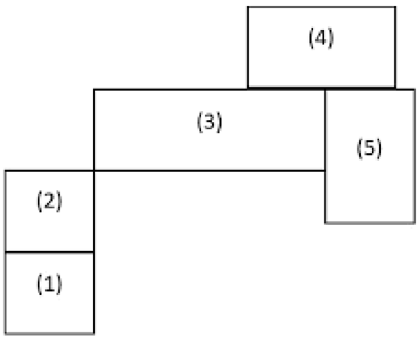


Figure 1. Intersection Method on Queen Contiguity

With the Queen Contiguity method for the 5 regions in **Figure 1**, the matrix formed is a symmetric matrix with the following description:

$$\boldsymbol{W}_{queen} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \quad (2)$$

The rows and columns represent the regions on the map. The spatial weighting matrix is a symmetric matrix with the rule that the main diagonal is always zero. The matrix is standardized to get the number of rows that are units. The number of rows is equal to one, so that the matrix becomes as follows:

$$\boldsymbol{W}_{queen} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0{,}5 & 0 & 0{,}5 & 0 & 0 \\ 0 & 0{,}3 & 0 & 0{,}3 & 0{,}3 \\ 0 & 0 & 0{,}5 & 0 & 0{,}5 \\ 0 & 0 & 0{,}5 & 0{,}5 & 0 \end{bmatrix} \quad (3)$$

2.3 Spatial Effect Test

TEhe presence of spatial effects can be tested in two ways: spatial dependency and heterogeneity. The dependent test uses the Moran's I method and the Lagrange Multiplier (LM), while the heterogeneity test uses the heteroscedasticity test with the Breusch-Pagan Test statistic **[8], [16]**. The formula for Moran's I is as follows **[17]**:

$$\boldsymbol{I_{MS}} = \frac{\boldsymbol{n}}{\boldsymbol{S_0}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \boldsymbol{W_{ij}}(\boldsymbol{x_i}-\bar{\boldsymbol{x}})(\boldsymbol{x_j}-\bar{\boldsymbol{y}})}{\sum_{i=1}^{n}(\boldsymbol{x_i}-\bar{\boldsymbol{x}})^2} \quad (4)$$

$$E(I_{MS}) = I_0 = -\frac{1}{n-1}$$

$$var(I_{MS}) = \frac{n[(n^2-3n+3)S_0 - nS_2 + 2{S_0}^2]}{(n-1)(n-2)(n-3){S_0}^2}$$

where,

$\boldsymbol{S_1} = \frac{1}{2}\sum_{i\neq 1}^{n}(\boldsymbol{w_{ij}} + \boldsymbol{w_{ji}})^2$ $\quad$ $\boldsymbol{S_2} = \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{w_{i0}} + \boldsymbol{w_{0i}})^2$

$\boldsymbol{S_0} = \sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{w_{ij}}$ $\quad$ $\boldsymbol{w_{0i}} = \sum_{i=1}^{n}\boldsymbol{w_{ij}}$

with,

x_i : observation data sequence-i $(i = 1,2,\dots n)$

x_j : observation data sequence-j $(j = 1,2,\dots n)$

$\bar{x}$: mean of observation data

$var(I_{MS})$: Moran's I variance

$E(I_{MS})$: expected value Moran's I

The Lagrange Multiplier (LM) test is one of the more specific spatial impact tests in testing whether there are dependencies in autoregression or errors **[18]**:

Hypothesis for SAR:

H_0 : $\rho = 0$ (no spatial lag dependencies in the model)

H_1 : $\rho \neq 0$ (there are spatial lag dependencies in the model)

The test statistic for the LM test is provided in **Equation (5)**:

$$\boldsymbol{LM_{SAR}} = \left(\frac{\boldsymbol{e\prime Wy}}{\boldsymbol{e'en^{-1}}}\right)^{\mathbf{2}} \frac{\mathbf{1}}{\boldsymbol{H}} \tag{5}$$

with,

$$\boldsymbol{H} = \left\{\left(\boldsymbol{WX\widehat{\beta}}\right)'\left[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X'X})^{-\mathbf{1}}\boldsymbol{X'}\right](\boldsymbol{W}\widehat{\boldsymbol{X\beta}})\widehat{\boldsymbol{\sigma}}^{\mathbf{2}}\right\} + \boldsymbol{\sigma}(\boldsymbol{W'W} + \boldsymbol{W}^{\mathbf{2}})$$

$$e = y_i - \frac{1}{n}\sum_{i=1}^{n} y_i$$

where,

n : number of observation
I : I_n is identity matrix

2.4 Data Transformation

Data that is not normally distributed can be transformed so that the data becomes normal [19]. The inverse transformation is appropriate for data that has a large range. If there is data with a value of 0, then there will be undefined results [20]. If the value 0 is in the sample data, then you can choose any small value (e.g., 0.0001) and replace all 0 values with that number. However, a statistician prefers to use the equation, i+1, in which, Xi is the i-th data value for small observational data [21]. Based on the results of the multiple linear regression assumption test using the original data, the residuals are not normal, so a transformation is needed. The appropriate transformation for the data is the inverse square root of all variables because the house damage data has a wide range of differences with intervals 1 – 3,446. The form of the transformation for each variable is as follows:

a. The formula for transforming the amount of house damage variable:

$$\boldsymbol{y'} = \frac{\mathbf{1}}{\sqrt{\boldsymbol{y}+\mathbf{1}}} \tag{6}$$

b. The formula for transforming the depth of epicenter variable:

$$\boldsymbol{X_1}' = \frac{\mathbf{1}}{\sqrt{\boldsymbol{X_1}+\mathbf{1}}} \tag{7}$$

c. The formula for transforming the magnitude of earthquake variable:

$$X_2' = \frac{1}{\sqrt{X_2+1}} \tag{8}$$

2.5 Analysis Method

The analytical method used to model earthquake losses is the spatial regression method. The data used are obtained from BNPB, BMKG, and BPBD of East Java Province. The response variable data (Y) for house damage was obtained from BNPB, and disaster report data from BPBD East Java Province with the amount of 24 Regencies that have an impact. The independent variable (X_1) data for the depth of the epicenter and the independent variable (X_2) for the magnitude of the earthquake obtained from the BMKG repo that is overlaid with the administrative boundaries of Districts/Cities in East Java Province. Software for processing research data uses Quantum GIS, R Studio, and Microsoft Excel. The flowchart of this research is seen in **Figure 2**.

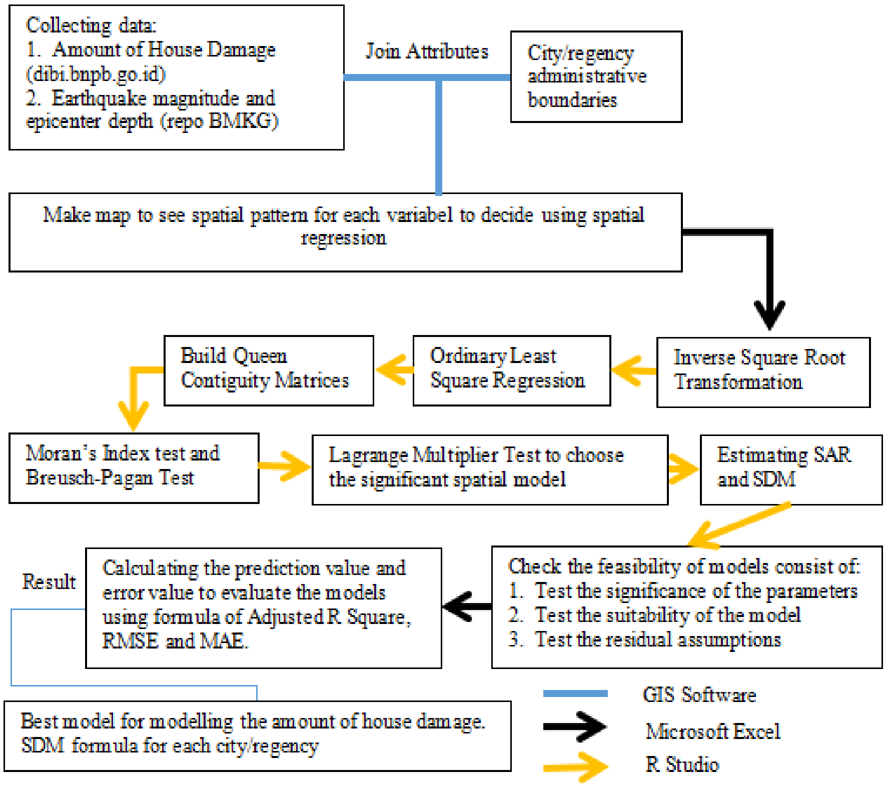


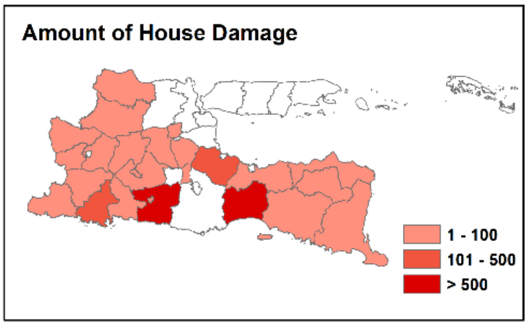
Figure 2. Research Method Flowchart

3. RESULTS AND DISCUSSION

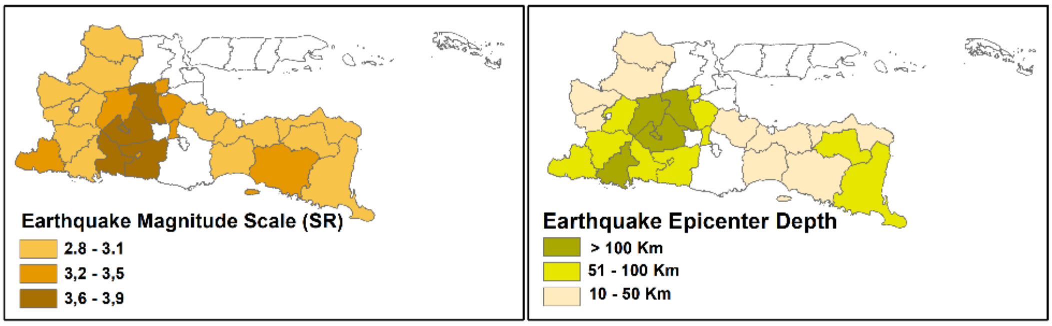
The earthquake disaster in East Java Province in 2021 is one of the dominant disasters that has a large impact on society. The number of earthquake events in East Java Province during 2021, according to the National Disaster Management Agency (BNPB), was 26 incidents, but there were 2 data that were outliers, there are Malang Regency and Malang City, so this study only used 24 incident data.

3.1 Description of the East Java Province Disaster Events in 2021

Visually, the pattern of distribution of damage to houses, the depth of the epicenter, and the magnitude of the earthquake in East Java Province show that. There is a grouping of data. Areas with damage under 100 in light colors are grouped, while high-damage houses marked with dark colors appear close together. The depth of the epicenter and the magnitude of the earthquake also appear to be grouped between light and dark colors.



(a)



(b) (c)

Figure 3. Regional Spatial Map of East Java Province Year 2021 for (a) Amount of House Damage, (b) Earthquake Magnitude, and (c) Epicenter Depth

As seen in Figure 3, there are patterns of distribution of the amount of house damage, the depth of the epicenter, and the magnitude of the earthquake, which tend to be in groups; it is suspected that there is a spatial dependency on the response and predictor variable. To strengthen the hypothesis that there is a spatial dependency, it is necessary to carry out a spatial effect test to accurately build a model for the problem of this house damage. The next step is to construct multiple linear regression equations and test the residual assumptions.

3.2 Data Transformation

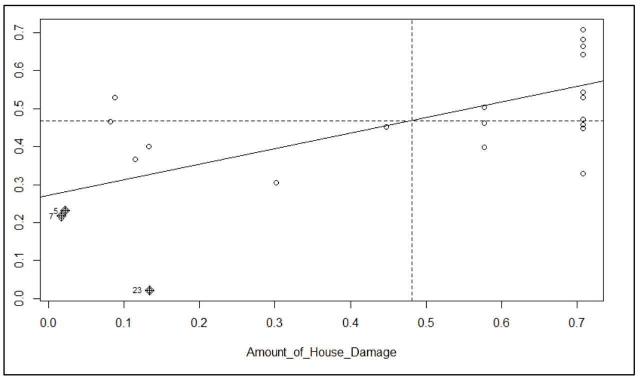
The multiple linear regression equations provide unsatisfactory outcomes, as the parameters either partially or simultaneously fail to influence the model, and the residual assumption test indicates non-normality. According to the Kolmogorov-Smirnov test, the p-value of $1.293e^{-11}$ suggests that the residual data does not exhibit a normal distribution. Then, the data is transformed so the residual normality tests can be fulfilled. The transformation in this study uses the inverse root square **Equation (6)**, **Equation (7)**, and **Equation (8)** with calculations using Microsoft Excel software. The data resulting from the transformation can be seen in **Table 1**. The range of the variable response to house damage changed from initially 1 – 3.446 to 0.01703 – 0.70711. The difference between the data is small, as well as the predictor variables, so that the assumption of normality in the residuals can be fulfilled.

Table 1. Data After Transformed

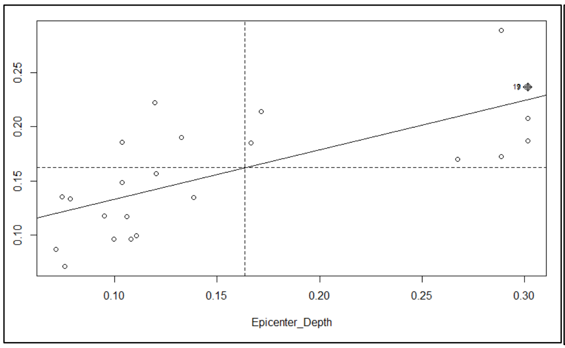
Regency/City	y'	X_1'	X_2'	Regency/City	y'	X_1'	X_2'
Pacitan	0.70711	0.10600	0.48224	Pasuruan	0.08839	0.26726	0.50637
Ponorogo	0.57735	0.13868	0.51299	Mojokerto	0.70711	0.10370	0.48224
Trenggalek	0.08165	0.09492	0.50637	Jombang	0.70711	0.07412	0.46127
Tulungagung	0.11471	0.10783	0.45175	Nganjuk	0.70711	0.07833	0.48795
Blitar	0.02224	0.09950	0.46625	Madiun	0.70711	0.11952	0.50000
Kediri	0.44721	0.07125	0.46625	Magetan	0.70711	0.30151	0.49386
Lumajang	0.01703	0.30151	0.51299	Ngawi	0.70711	0.30151	0.49386
Jember	0.13245	0.17150	0.47140	Bojonegoro	0.70711	0.28868	0.50000
Banyuwangi	0.70711	0.12039	0.50637	Tuban	0.70711	0.28868	0.50000
Bondowoso	0.57735	0.13245	0.50637	Kediri City	0.70711	0.07559	0.46625
Situbondo	0.70711	0.16667	0.50637	Blitar City	0.13363	0.11043	0.47673
Probolinggo	0.30151	0.30151	0.51299	Batu City	0.57735	0.10370	0.48224

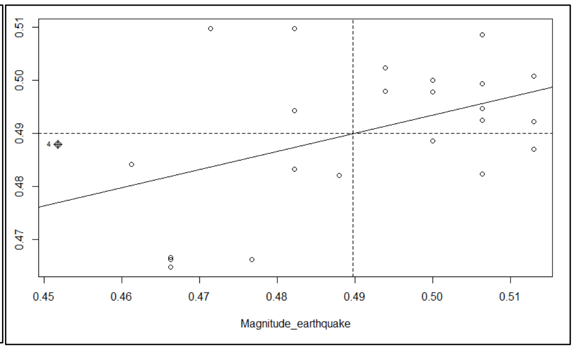
3.3 Spatial Effect Test

From Figure 1, it can be seen that the data is grouped by region. So, it is necessary to test the spatial effect of each variable. The first spatial effect test is the Moran index test to determine whether or not spatial dependencies exist in the data. Spatial dependencies can be visualized using Moran's I scatter plot to see the pattern of data grouping and hypothesis testing to determine the significance of the results. The results scatter plot can be seen in **Figure 4**.



(a)





(b) **(c)**

Figure 4. Moran's I Scatter Plot for (a) Amount of House Damage, (b) Epicenter Depth, and (c) Earthquake Magnitude

Figure 4 shows that the pattern for the amount of house damage variable data is in quadrants I and III. Quadrant I High-High means that a high category of house damage is also surrounded by a high category area. Meanwhile, quadrant III, Low-Low, is the opposite of quadrant I, where areas with low scores appear to be clustered. The epicenter depth variable is almost the same as the house damage variable. Earthquakes are more dominantly centered on the southern coast of East Java Province. So, the data with high observation values are grouped together as in Quadrant I, and low observation values are grouped in the central and northern parts of East Java Province, which are included in Quadrant III. The distribution pattern of values for the earthquake strength variable is in quadrant III Low-Low because the majority of earthquake magnitudes are low and in groups. Earthquakes with high magnitude more commonly occur in the Indian Ocean, so they are not recorded in the administration boundaries of East Java Province.

The results of the Moran's I value for the house damage variable, the epicenter depth variable, and the magnitude of earthquake variable with the spatial Queen Contiguity weighted matrix are as follows:

Table 2. Result of Spatial Dependencies Test using Moran's I

Variable	Moran's I	P_{value}	Z_{value}
y'	0.40792	0.00182	2.90855
X_1'	0.45487	0.00066	3.21106
X_2'	0.34129	0.00658	2.47918

By using a confidence level of 95%, the variable number of damaged houses (y'), depth of the epicenter (X_1'), and the magnitude of the earthquake (X_2') shows that $I_{MS} > I_0$, which means there is positive spatial autocorrelation and shows a grouped data pattern, then the conclusion is obtained that there is relationship between regions for each variable, both response and predictor.

The second spatial effect test is heterogeneity test, using the Breusch-Pagan test. The results are BP_{value} = 1.3675; ${\chi^2}_{(\alpha,2)}$= 5.991455; *p-value* = 0.5047. The test criteria used are if $BP_{value} > {\chi^2}_{(\alpha;p)}$ or *p-value* < α, then H_0 is rejected. The result failed to reject H_0, which means that there was no tendency of heterogeneity in the spatial data, there is no variation between Districts, and there is no need to incorporate the influence of location into the model, so the GWR method is not necessary to be used. Still, SAR, SEM, SCR, SDM, or SARMA methods are selectable.

The results of the LM test for the house damage variable, the epicenter depth variable, and the magnitude of earthquake variable with the spatial Queen Contiguity weighting matrix are as follows:

Table 3. Result of Lagrange Multiplier Test

Spatial Dependencies Test	LM_{value}	P_{value}	Conclusion
LM *error*	5.80761	0.01596	Reject H_0
LM *lag*	6.23394	0.01253	Reject H_0
Robust LM *error*	0.32767	0.56703	Failed to Reject H_0
Robust LM lag	0.75400	0.38521	Failed to Reject H_0
SARMA	6.56160	0.03760	Failed to Reject H_0

By using a confidence level of 95%, the LM error, LM lag, and SARMA result in the rejection of H_0, meaning that there is a spatial dependency for house damage as the impact of earthquakes. Thus, it is obtained that $\rho \neq$0 and $\lambda \neq$0, meaning that the Spatial Autoregressive Model (SAR), Spatial Error Model (SEM), and Spatial Autoregressive Moving Average (SARMA) can be selected. However, because the result of the smallest p-value is in the LM lag, the SAR and SDM models were chosen to model the problem of house damage due to earthquakes.

3.4 Spatial Regression Model

The parameter estimation results for the SAR model can be seen in **Table 4**, and the parameter estimates for the SDM model can be seen in **Table 5**.

Table 4. Parameter Estimation for SAR Model

Parameter	Variable	Coefficient	Z_{value}	P_{value}
$\boldsymbol{\rho}$	Wy	0.64907	4.5652	4.9904e-06
$\boldsymbol{\beta_0}$	Intercept	-0.82375	-0.6181	0.5365
$\boldsymbol{\beta_1}$	X_1’	-0.78618	-1.3247	0.1853
$\boldsymbol{\beta_2}$	X_2’	2.30712	0.8181	0.4133

Equation for SAR model from transformed data:

$$\hat{y}_i{}' = 0.64907 \sum_{j=1}^{n} W_{ij} y_j{}' - 0.82375 - 0.78618 X_{1i}{}' + 2.30712 X_{2i}{}' \qquad (9)$$

The ρ parameter as a parameter of the spatial lag coefficient of the dependent variable has a significant effect. This shows that cases of house damage in one regency/city have an influence on cases of house damage in other regencies/cities. Even though the independent variable is not significant, it can still be included in the calculation because it has a spatial effect.

The SDM model gives results that the depth of the epicenter (X_1), the magnitude of the earthquake (X_2), the lag of the depth of the epicenter, and the lag of the magnitude do not significantly affect the number of cases of house damage in East Java Province. However, the ρ parameter as a parameter of the spatial lag coefficient of the dependent variable has a significant effect. Parameter estimation value β_{11} and β_{12} are non-spatial regression coefficient and parameter estimation value β_{21} and β_{22} are the regression coefficient of the spatial lag parameter on the independent variable. The estimated value of the parameter ρ shows the effect of spatial lag on the dependent variable.

Table 5. Parameter Estimation for SDM Model

Parameter	Variable	Coefficient	Z_{value}	P_{value}
ρ	Wy	0.62183	4.1720	3.0191e-05
β_0	Intercept	-2.33529	-0.9981	0.3182
β_{11}	X_1’	-1.13150	-1.5709	0.1162
β_{12}	X_2’	1.13621	0.3896	0.6968
β_{21}	WX_1	0.39709	0.2836	0.7767
β_{22}	WX_2	4.26532	0.9498	0.3422

Equation of Spatial Durbin Model (SDM) which obtained:

$$\hat{y}_i{}' = 0.62183\sum_{j=1}^{n} W_{ij}y_j{}' - 2.33529 - 1.13150X_{1i}{}' + 1.13621X_{2i}{}' + 0.39709\sum_{j=1}^{n} W_{ij}\,X_{1j}{}' + 4.26532\sum_{j=1}^{n} W_{ij}\,X_{2j}{}' \tag{10}$$

To determine the feasibility of the SDM model, it is necessary to test the significance of the parameters, the suitability of the model, and the residual assumptions. The parameter significance test uses the Wald test, and the model suitability test uses the LR (Likelihood Ratio) test. The results of the parameter significance test and model suitability are shown in **Table 6**.

Table 6. Result of Parameter Significance Test and Suitability Test for SAR and SDM

Model	LR_{value}	$Wald_{value}$	$\chi^2{}_{0.05;1}$
SAR	8.2476	20.841	3.481
SDM	7.4356	17.406	3.481

The results of the analysis show that the value of the LR test and the Wald test is greater than $\chi^2{}_{0.05;1}$, so it can be concluded that the two variables have a significant effect on the model and also the model is suitable for estimating house damage due to earthquakes.

The residual assumption test uses the residual normality test and heteroscedasticity test. The residual normality test uses the Kolmogorov-Smirnov test, and the heteroscedasticity test uses the Breusch Pagan test. The results of the residual assumption test can be seen in **Table 7**.

Table 7. Result of Residual Assumption Test for SAR and SDM

Model	D_{value}	D_{table}	BP_{value}	$\chi^2{}_{0.05;1}$
SAR	0.19342	0.1619	0.1109	3.481
SDM	0.20196	0.1619	1.9259	3.481

The normality test results for SAR and SDM modeling show $D_{value} < D_{table}$, the existing data fails to reject H_0, meaning that the residuals of all these models are normally distributed. The results of the heteroscedasticity test showed $BP_{value}<\chi^2{}_{0.05;1}$, it failed to reject H_0, meaning that there is no heteroscedasticity problem in these models.

The results of the parameter significance test and the suitability of the model are satisfactory, as well as the assumption of normal residuals and no heteroscedasticity occurs. The SAR and SDM models that were constructed using transformation data are feasible to use to provide an estimate of the house damage due to the depth of the epicenter and the magnitude of the earthquake.

After obtaining two feasible models from multiple linear regression with maximum likelihood estimation, the best models must be chosen using four criteria. In these two models, the researchers used the best model to estimate the amount of house damage in East Java Province in 2021 by looking at the smallest R square, Akaike Information Criterion (AIC), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) values. Error data input comes from actual data and predictive data that has been returned to the form before the transformation.

Table 8. Selection of the Best Model using Actual Data

Model	Adjusted R^2	AIC	RMSE	MAE
SAR	90.89%	396.90	813.38	245.36
SDM	90.92%	393.43	811.99	244.01

A good model is a model that has a greater coefficient of determination, in which the value indicates the strength of the variable number of house damages (y') explained by the independent variables (X'). Based on the model evaluation values in Table 8, the AIC value shows a measure of the goodness of fit of the model. The smaller the AIC value, the better the model is for estimating. The coefficient of determination of the SDM Based on the table above is greater than the SAR model, and the AIC value of the SDM is smaller than the other model.

The smaller result of the sum square error indicates the better model. The smaller the RMSE, MAE, and Mean Absolute Percentage Error (MAPE) values indicate good term [17]. The RMSE value is calculated by squaring the error divided by the average number of data and then taking the roots. The RMSE value has no units. The smaller the RMSE value, the closer the predicted and observed values are. The MAE value is the average absolute difference between the actual (actual) value and the predicted (forecasting) value. In Table 7, the Adjusted R square value shows that the largest value in the SDM model and the smallest AIC, RMSE, and MAE values in the SDM model. It can be concluded that the SDM model is the best model for modeling house damage due to earthquakes in East Java Province.

3.5 Spatial Durbin Interpretation

Based on Equation (9), the SDM model can be interpreted that if other factors are considered constant, then when the earthquake depth increases by 99 units, then the damage to houses decreases by 77.11 units. If the strength of the earthquake increases by 8 units, it will increase the damage to the house by 5.97 units. The estimated parameter ρ is 0.62183, and the parameter coefficient is positive, which indicates that the regency/city will have a large number of damaged houses if it is adjacent to a regency/city which has a large number of damaged houses too. The estimated value of the β_{11} parameter is -1.13150, whereas the projected value of the β_{21} parameter is 0.39709. The coefficient associated with the depth parameter of the epicenter exhibits a negative sign, indicating an inverse relationship between earthquake depth and the extent of house damage. The positive coefficient of the lag parameter pertaining to the depth of the epicenter suggests that the Regency/City in question exhibits a substantial epicenter depth and is geographically proximate to another Regency/City with a similarly significant epicenter depth. The estimated value for the β_{12} parameter is 1.13621, whereas the estimated value for the β_{22} parameter is 4.26532. The positive coefficient of the parameter lag of the magnitude suggests that there is a correlation between high earthquake magnitudes in regency/cities and neighboring districts/cities, and the subsequent occurrence of a significant number of damaged houses. The earthquake strength parameter coefficient is assigned a positive weight. This observation demonstrates a positive correlation between the magnitude of an earthquake and the extent of damage inflicted upon residential structures.

As an example, we will look for the predicted value of the number of damaged houses (y) for Batu City, using the SDM model, because Batu City (24) is a neighbor of Pasuruan Regency (13) and Mojokerto Regency (14), by transforming the house damage data, the depth of the epicenter, and the magnitude of the earthquake, then substitutes in each variable, then the equation is as follows:

SDM equation for Batu City:

$$\hat{y}_i{}' = 0.62183\sum_{j=1}^{n} W_{ij}y_j{}' - 2.33529 - 1.13150X_{1i}{}' + 1.13621X_{2i}{}' + 0.39709\sum_{j=1}^{n} W_{ij}\,X_{1j}{}' + 4.26532\sum_{j=1}^{n} W_{ij}\,X_{2j}{}'$$

$$\Leftrightarrow \hat{y}_{24}{}' = 0.62183\big((0.5)(y_{13}{}' + y_{14}{}')\big) - 2.33529 - 1.13150X_{1(24)}{}' + 1.13621X_{2(24)}{}' + 0.39709\left((0.5)\big(X_{1(13)}{}' + X_{1(14)}{}'\big)\right) + 4.26532\left((0.5)\big(X_{2(13)}{}' + X_{2(14)}{}'\big)\right)$$

$$\Leftrightarrow \hat{y}_{24}{}' = 0{,}62183\big((0{,}5)(0{,}08839 + 0{,}70711)\big) - 2{,}33529 - 1{,}13150(0.10369) + 1.13621(0.48224) + 0.39709\big((0.5)(0.26726 + 0.10369)\big) + 4.26532\big((0.5)(0.50637 + 0.48224)\big)$$

$$\Leftrightarrow \hat{y}_{24}{}' = 0.24733 - 2.33529 - 0.11733 + 0.54793 + 0.07365 + 2.10837$$

$\Leftrightarrow \hat{y}'_{24} = 0.52466$

The predicted value of Batu City is still in the form of a transformation, so it needs to be returned to the normal prediction form with the following formula:

$$\hat{y}_i{}' = \frac{1}{\sqrt{\hat{y}+1}} \Leftrightarrow \hat{y} = \frac{1}{\hat{y}'^2} - 1 \tag{11}$$

The normal prediction value for Batu City after being calculated using formula (11) produces a value $y_{predicted}$= 2.63275, whereas in the actual incident, Batu City has 2 house damages, so the evaluation value for the SDM model in Batu City can be calculated as follows: $|y_{actual} - y_{predicted}| = |2 - 2.63275| =$ 0.63275.

4. CONCLUSIONS

Several conclusions can be written according to the problem formulated and based on the results of the analysis, namely: a dependency between District/City locations in East Java Province for the variable house damage, the depth of the epicenter, and the magnitude of the earthquake. The amount of damage to houses in Regencies or Cities in East Java Province uses SDM with Queen Contiguity weighted is as follows:

$$\hat{y}_i{}' = 0.62183 \sum_{j=1}^{n} W_{ij} y_j{}' - 2.33529 - 1.13150 X_{1i}{}' + 1.13621 X_{2i}{}' + 0.39709 \sum_{j=1}^{n} W_{ij} X_{1j} + 4.26532 \sum_{j=1}^{n} W_{ij} X_{2j}{}'$$

Based on the criteria for selecting the best model, it was found that the SDM is the best model that describes the condition of the loss or the amount of house damage due to earthquakes in 24 Regencies/Cities in East Java Province in 2021. Modeling house damage due to earthquakes is very useful for the government to determine the proportion of the disaster insurance premium subsidy assistance budget per region. This is done to prevent the premium payment from being counted as a loss.

REFERENCES

[1] N. Cholis Madjid, "Analisis Metode Penghitungan dan Alokasi Anggaran Bencana Alam". *Simposium Nasional Keuangan Negara*, pp. 1046-1065, 2018.
[2] Tirto, "Mengapa Gempa Dangkal Lebih merusak daripada Gempa Dalam," 2018. [Online] Diakses: https://tirto.id/mengapa-gempa-dangkal-lebih-merusak-daripada-gempa-dalam-c6mg
[3] Widodo Pawirodikromo, *Seismologi Teknik dan Rekayasa Kegempaan*. Yogyakarta: Pustaka Pelajar, 2012.
[4] R. P. Plate, Firdaniza, and D. Susanti, "Penentuan Kerugian Tahunan dengan Pemodelan Kebencanaan Alam dan Premi Asuransi Pada Kerusakan Rumah Akibat Banjir," *Jurnal Matematika Integratif*, vol. 15, no. 2, pp. 121–128, 2019.
[5] E. A. Ash-Shidiqqi, "Alternatif Strategi Pembiayaan Asuransi Bencana Alam di Indonesia," *Journal of Governance and Policy Innovation* (JGPI) vol. 1, no. 2, pp. 111–119, 2021.
[6] M. V. Rivas-Lopez, R. Minguez-Salido, M. M. Garcia, and A. E. Rey, "Contributions from spatial models to non-life insurance pricing: An empirical application to water damage risk," *Mathematics*, vol. 9, no. 19, Oct. 2021.
[7] P. Retnowati, R. Rahmawati, and A. Rusgiyono, "Analisis Faktor-Faktor Perikanan Tangkap Perairan Umum Daratan di Jawa Tengah menggunakan Regresi Berganda dan Model Durbin Spasial," *Jurnal Gaussian*, vol. 6, no. 1, pp. 141–150, 2017.
[8] S. E. Triliani and R. Dwi Bekti, "Spatial Durbin Model untuk Mengidentifikasi Faktor-Faktor yang Mempengaruhi Pengangguran di Provinsi Jawa Tengah," *Jurnal Statistika Industri dan Komputasi*, vol. 2, no. 2, pp. 93–103, 2017.
[9] H. R. S. Saefulhakim, "Pemodelan spasial penentuan instrumen strategis penataan ruang untuk pengendalian risiko banjir di wilayah Jabodetabek". *Jurnal Ilmu Pertanian Indonesia*, Vol 16, no.1, pp 49-58, 2011.
[10] H. Kusumaningrum, D. Endah Kusrini dan, D. Susilaningrum, "Pemodelan spasial ekonometrik kerugian makroekonomi akibat bencana alam". Tugas Akhir S1 yang tidak dipublikasikan, 2014.
[11] H. Yasin Budi Warsito Arief Rachman Hakim, *Regresi Spasial (Aplikasi dengan R)*. Pekalongan: WADE Publish, 2020.
[12] P. Simoes, I. Natario, M. L. Carvalho, S. Aleixo, and S. Gomes, "Health Line Saude24: An Economic Spatial Analysis of Its Use," in *Geospatial Technology for Human Well-Being and Health*, F. S. Faruque, Ed., Swiss: Springer, 2022, pp. 91–118.
[13] A. Arif, M. A. Tiro, and M. Nusrang. "Perbandingan matriks pembobot spasial optimum dalam Spatial Error Model (SEM) kasus: indeks pembangunan manusia kabupaten/kota di Provinsi Sulawesi Selatan Tahun 2015" *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, vol. 1, no. 2, p. 32, Jul. 2019.
[14] M. Sholikhah, *Penggunaan Matriks Pembobot Spasial Tipe Queen Contiguity dan Rook Contiguity pada Regresi Tobit Spasial*, Undergraduate [Thesis]. Malang: Universitas Brawijaya, 2014. [Online]. Diakses: UB Repository.
[15] Y. Sari, A. Rizki, and A. H. Wigena, *Penggunaan Matriks Queen Contiguity dan K-Nearest Neighbor pada Data Panel Spasial Kemiskinan Sumatera Utara*. Undergraduate [Thesis]. Bogor: IPB, 2020. [Online]. Diakses: IPB Repository.
[16] K. Yeni Widiyanti and H. Yasin, "Pemodelan Proporsi Penduduk Miskin Kabupaten dan Kota di Provinsi Jawa Tengah Menggunakan Geographically and Temporally Weighted Regression," vol. 3, no. 4, pp. 691–700, 2014.
[17] S. Riyanto and A. R. Putera, *Metode Riset Penelitian Kesehatan dan Sains*. Yogyakarta: Deepublish, 2022.
[18] E. Caroline, *Aplikasi Ekonometrika Spasial dengan Software STATA*. Surabaya: Scopindo Media Pustaka, 2020.

[19] I. Ghozali, *Ekonometrika Teori, Konsep dan Aplikasi dengan SPSS 17*. Semarang: Badan Penerbit Universitas Diponegoro, 2009.
[20] J. Osborne, *Best Practices in Data Cleaning: A Complete Guide to Everything You Need to Do Before and After Collecting Your Data*. SAGE Publications, Inc., 2014.
[21] J. E. De Muth, *Basic Statistics and Pharmaceutical Statistical Applications 2nd Edition*, 2nd ed. United Kingdom: Taylor & Francis, 2006.