Pakistan Journal of Statistics and Operation Research

Implementation of Bayesian Simulation for Earthquake Disaster Risk Analysis in Indonesia based on Gutenberg Richter Model and Copula Method

Pratnya Paramitha Oktaviana [1*], Kartika Fithriasari [2]

* Corresponding Author

1. Institut Teknologi Sepuluh Nopember, Indonesia, paramita@statistika.its.ac.id
2. Institut Teknologi Sepuluh Nopember, Indonesia, kartika_f@statistika.its.ac.id

Abstract

Indonesia is a country prone to earthquakes because it is located in the Pasific ring of fire area. The earthquakes caused a lot of damages and casualties. In this paper, we use Bayesian Simulation on Gutenberg Richter model and Copula method to estimate the risk parameters of earthquake, specifically the probability and the recurrence (return) period of an earthquake occurrence in Indonesia. Those risk parameters are estimated from dependence structure of frequency and magnitude of earthquakes. The dependence structure can be determined by using Gutenberg Richter model and Copula method. The Gutenberg Richter model is a model based on linear regression used to determine dependence structure, while the Copula method is a statistical method used to determine dependence structure that ignores linearity and normality assumptions of data. Bayesian Simulation is a method used to estimate parameters based on simulation. The data used is an annual data of frequency and magnitude (magnitude ≥ 4 Richter Scale) of earthquakes occur in Indonesia for 4 years from Meteorological, Climatological, and Geophysical Agency of Indonesia. There are several steps of analysis to be performed: firstly, we perform regression analysis of frequency and magnitude of the earthquakes to determine Gutenberg Richter Model; secondly, we perform Copula analysis; thirdly, we estimate probability and the recurrence (return) period of an earthquake occurrence using Bayesian Simulation based on the result of step one and two. The result indicates Bayesian Simulation can estimate risk parameters very well.

Key Words: Earthquake, Bayesian, Gutenberg Richter, Copula, Risk Parameter.

1. Introduction

Indonesia is located in disaster prone area and the intensity of disasters is increasing and becoming more complex (Wibowo *et al*, 2013). Law of the Republic of Indonesia Number 24 of 2007 concerning Disaster Management defines natural disasters as disasters caused by events or a series of events caused by nature including earthquakes, tsunamis, volcanic eruptions, floods, droughts, hurricanes, and landslides. Earthquake is one of the natural disasters that often occur in Indonesia. Indonesia experiences an average of 20 earthquakes per day (most are too weak to be felt). There are also about 500 volcanoes, of which 128 are active and have been recorded in history to have erupted; while 21 are considered to be the most active (Susilastuti, 2016). The earthquakes caused a lot of damages and casualties. Modeling or analysis of earthquake risk needs to be done so that useful information can be found about the earthquake disaster.

Disaster risk analysis is an important aspect used for estimating the disaster risk parameters. In this study, the researchers focus the analysis on estimating the risk parameter of earthquake in Indonesia. The method used in this analysis is Bayesian Simulation based on Gutenberg Richter Model and Copula Method. The data used in this study are frequency and magnitude (magnitude ≥ 4 Richter Scale) of the earthquake in Indonesia in period of January 2014 to December 2017 (4 years). The data is collected from the website of Meteorology, Climatology and Geophysics Agency in Indonesia from April 5 until May 8, 2019.

Copula is one of the statistical methods used to determine the relationship of two or more variables, in which case the distribution is normal or not (Sklar, 1959). To describe the relationship between two variables, correlation or regression can be used, but for extreme events, Copula is superior because it can clearly depict dependencies on extreme points. Gutenberg Richter Model is a method used to obtain the pattern of frequency-magnitude relationships in earthquake risk analysis (Kara, 2017). Bayesian Network Simulation is a network simulation technique used to estimate parameters so that the results obtained are better. There are two conditions to continue Bayesian analysis; those are the posterior distribution of parameters should be stationary and the parameters should be convergence (Oktaviana and Fithriasari, 2017).

The first step of the analysis is performing regression of frequency (as dependent variable) and magnitude (as independent variable), and its result is determined as Gutenberg Richter Model. After that, Copula analysis is performed. Then, Bayesian simulation is used to estimate probability and the recurrence (return) period of an earthquake occurrence based on the results of Gutenberg Richter Model and Copula method. Its result risk parameters of Eartquake occurance very well.

2. The Gutenberg Richter Model

Gutenberg Richter Model is a method used to obtain the pattern of frequency-magnitude relationships in earthquake risk analysis (Kara, 2017). The Gutenberg Richter Model equation is defined as follows:

$$y = \log N = a - bM \tag{1}$$

where M is the earthquake magnitude and N is the cumulative frequency of the earthquake. The cumulative frequency of earthquakes shows the number of earthquakes that have magnitudes equal to or greater than M. Parameters a and b in the model is described as regression parameters calculated using Least Square as follows:

$$a = \bar{y} + b\widehat{M} \tag{2}$$

$$b = \frac{\sum_{i=1}^{k} M_i y_i - k\bar{M}\bar{y}}{\sum_{i=1}^{k} M_i^2 - k\bar{M}^2} \tag{3}$$

In equation (3), k is the number of groups or classes, $\bar{y}$ and $\bar{M}$ is the mean of frequency and magnitude, respectively. To obtain the frequency value, the cumulative frequency value (N) is divided by the time period t (year) and then the annual logarithmic frequency value $y = \log\left(\frac{N}{t}\right)$ is obtained from the logarithmic value.

Gençoğlu (1972) and Ünal et al. (2014) in Kara (2017) provides earthquake risk parameters using the Poisson method as follows:

$$n(M) = 10^{a-bM} \tag{4}$$

$$R(M) = 1 - e^{-n(M)T} \tag{5}$$

$$Q = \frac{1}{n(M)} \tag{6}$$

where $n(M)$ is the annual average earthquake; $R(M)$ is the risk of an earthquake with magnitude M during T years, in certain regions during the t-year observation interval; Q is the recurrence period (return).

3. Copula Analysis

According to Schölzel and Friederichs (2008), if there is a random vector $\boldsymbol{X}$ with m-dimensions which has marginal distribution function $F_{X_1}, F_{X_2}, \ldots, F_{X_m}$ in non-decreasing domain, $\boldsymbol{R}$, that are $F_{X_1}(-\infty) = 0$ and $F_{X_1}(\infty) = 1$, then according to Sklar's Theorem (1959), the joint distribution is:

$$F_X(x) = C_X(F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_m}(x_m)) \tag{7}$$

where $C_X : [0,1] \times \ldots \times [0,1] \rightarrow [0,1]$ is joint distribution of transformation random variable $U_j = F_{X_j}(x_j)$ for $j = 1, 2, \ldots, m$, with U_j has marginal ditribution of Uniform.

If marginal distribution $F_{X_j}(x_j)$ is continuous, then C_X is unique (Nelsen, 2005), and C_X could be written as:

$$C_X(u_1, \ldots, u_m) = \int_0^{u_1} \ldots \int_0^{u_m} c_X(u_1, \ldots, u_m) du_1 \ldots du_m \quad (8)$$

where C_X is Copula and c_X is density equation of Copula. Copula is joint distribution function or multivariate distribution function for a uniform marginal distribution (Nelsen, 2005).

Normal Copulas

The copula function of Normal Copulas is:

$$C_X(u_1, \ldots, u_m) = F_{N(0,\Sigma)}(F^{-1}_{N(0,1)}(u_1), \ldots, F^{-1}_{N(0,1)}(u_m)) \quad (9)$$

The density of Normal Copula is:

$$c_X(u_1, \ldots, u_m) = \frac{\partial}{\partial u_1} \ldots \frac{\partial}{\partial u_m} C_X(u_1, \ldots, u_m)$$

$$= \frac{f_{N(0,\Sigma)}((F^{-1}_{N(0,1)}(u_1), \ldots, F^{-1}_{N(0,1)}(u_m))}{\prod_{j=1}^{m}(f_{N(0,1)}(F^{-1}_{N(0,1)}(u_j)))} \quad (10)$$

According to Schölzel and Friederichs (2008), if Normal Copulas is used in a multivariate normal distribution, it is assumed that the variables have a linear relationship.

Bivariate Return and Risk Period (Probability Occurred) based on Copula

The bivariate risk value (R^{OR}) is associated with the period of joint return (Q^{OR}) as defined by Yen (1970) and Fan et al. (2015) in Kara (2017) as follows:

$$R^{OR} = 1 - \left(1 - \frac{1}{Q^{OR}}\right)^t \quad (11)$$

where t is the number of years of observation. The equation can also be written as follows:

$$R = 1 - (1 - p)^t \quad (12)$$

where p is a probability exceeded. Bivariate risk values can be written by P^{OR} probability Copula as follows:

$$R = 1 - C_{U,V}(u, v) \quad (13)$$

4. Bayesian Simulation

Bayesian probability theory in relation to extreme models is used to develop the Bayesian distribution of extreme earthquake events by assuming that an earthquake is a Poisson process with magnitude having an exponential distribution (Campbell, 1982). The Bayesian distribution represents the probability that M_{max}, the highest estimated earthquake magnitude occurring in the t year period, will exceed several magnitudes m, and is calculated based on the pattern of relationships:

$$\tilde{P}(M_{max} > m|t) = 1 - \left(\frac{t''}{t'' + t[1 - \tilde{F}(m)]}\right)^{n''} \quad (14)$$

where n'' and t'' represent the posterior Bayesian estimate updated from the number of earthquakes and time periods, is the Bayesian distribution of magnitude, which is updated from the prior estimate using the earthquake history that occurs. Through the application of Bayesian probability theory, the distribution contains two features that make it more powerful and reliable than the extreme values of conventional distributions.

5. Earthquake Analysis in Indonesia using Bayesian Simulation

The data used in this study are the frequency and magnitude of earthquakes that occurred in Indonesia. Data is selected by selecting earthquakes that have magnitude ≥ 4 on the Richter Scale in the period January 2014 to December 2017 (4 years). The data was obtained through the Climatology and Geophysics Meteorological Agency website. The data is collected from April 5 until May 8, 2019.

Table 1 : Research Variable

Variable	Definition	Scale
Y	Frequency of Earthquake	Ratio
M	Magnitude of Earthquake	Ratio

Gutenberg Richter Model

Bayesian simulation for this method is performed by using WinBUGS software. The Directed Acyclic Graph (DAG) is created as Figure 1. The model of DAG is shown in Figure 2. The model used is Gutenberg Richter Model, i.e.:

$$y = \log N = a - bM \tag{15}$$

where y is frequency of earthquake and M is magnitude of earthquake that is ≥ 4 Richter Scale. While a and b are regression parameters we estimate by using bayesian simulation. We used three markov chains in simulation process.

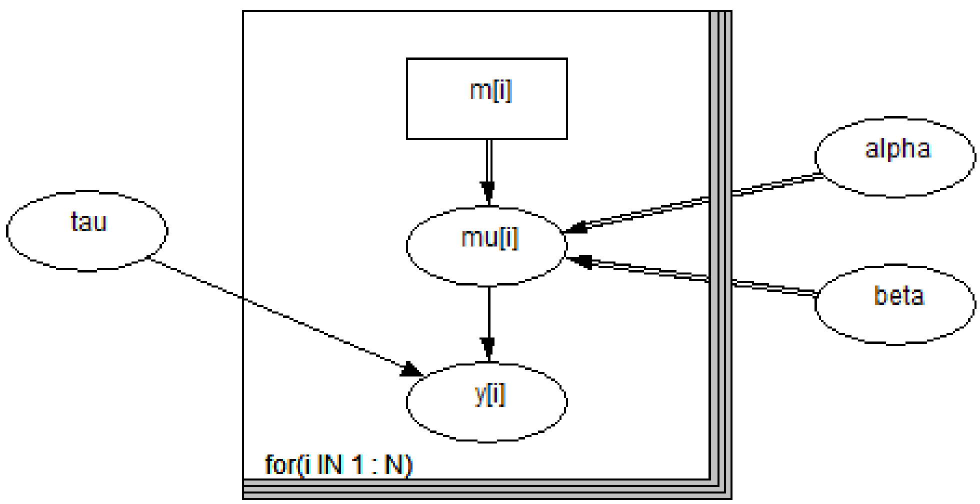

Figure 1 : The DAG of Bayesian Simulation of Gutenberg Richter Model

```
model;
{
  for( i in 1 : N ) {
    y[i] ~ dnorm(mu[i],tau)
    mu[i] <- alpha + beta * m[i]
  }
  alpha ~ dnorm( 0.0,1.0E-4)
  beta ~ dnorm( 0.0,1.0E-4)
  tau ~ dgamma(1.0E-4,1.0E-4)
}
```

Figure 2 : The Model of DAG

Figure 1 and Figure 2 denote that $y[i]$, where i is 1, 2, ..., N observations, is normally distributed with mean $mu[i]$ and variance *tau*. As regression model before, $mu[i]$ is estimated by using formula *alpha* + *beta***m*[i], where *alpha* is a, *beta* is b, and *m*[i] is magnitude of earthquake. As a common regression, we use normal distribution with mean 0

 DOI: http://dx.doi.org/10.18187/pjsor.v19i2.3089

and variance 0.0001 as the prior of *alpha* and *beta*. While we use gamma distribution with α 0.0001 and β 0.0001 as the prior of *tau*.

Figure 3 shows that the time series plot of history of three chains markov produce the stationary posterior parameters of earthquake. Figure 4 shows that the Gelman Rubin statistics of parameters produce the convergence parameters.

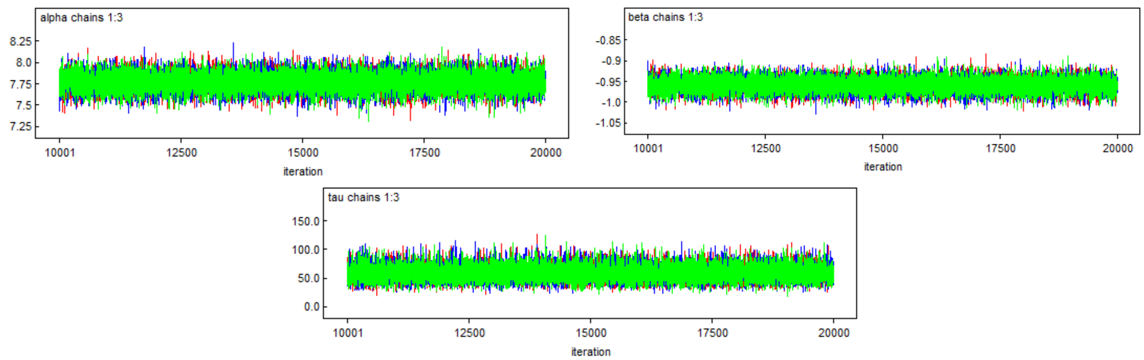

Figure 3 : Time Series Plot of History Chain

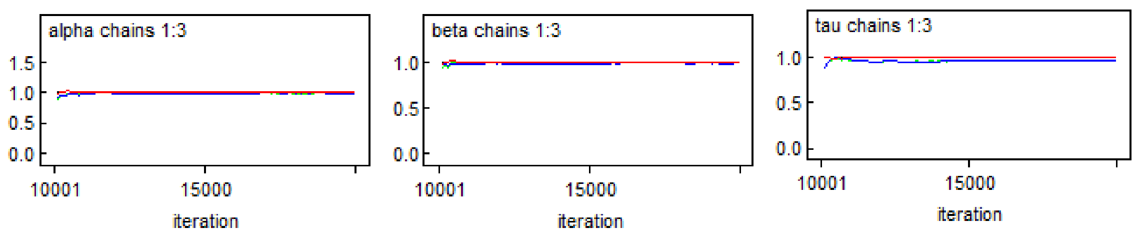

Figure 4 : Gelman Rubin Statistics

The result of bayesian simulation is the posterior summaries of *alpha* and *beta*. We obtain that the posterior mean for each *alpha* and *beta* is 7.769 and -0.959. Then we have the model of Gutenberg Richter as: $\hat{Y} = 7.769 - 0.959M$. We used this result to calculate *n*(*M*), the annual average earthquake; *R*(*M*), the risk of an earthquake; and *Q*, the recurrence period (return). The result of those risk parameters is indicated in Table 2:

Table 2 : Risk Parameters of Earthquake Based on Gutenberg Richter Model

M	*n(M)*	*R(M)=1 year*	*R(M)=4 years*	*R(M)=10 years*	*R(M)=15 years*	*Q*
4	8570.3785	1.0000	1.0000	1.0000	1.0000	0.0001
4.5	2841.1882	1.0000	1.0000	1.0000	1.0000	0.0004
5	941.8896	1.0000	1.0000	1.0000	1.0000	0.0011
5.5	312.2482	1.0000	1.0000	1.0000	1.0000	0.0032
6	103.5142	1.0000	1.0000	1.0000	1.0000	0.0097
6.5	34.3163	1.0000	1.0000	1.0000	1.0000	0.0291
7	11.3763	1.0000	1.0000	1.0000	1.0000	0.0879
7.5	3.7714	0.9770	1.0000	1.0000	1.0000	0.2652
8	1.2503	0.7136	0.9933	1.0000	1.0000	0.7998
8.1	1.0025	0.6331	0.9819	1.0000	1.0000	0.9975
8.2	0.8039	0.5524	0.9599	0.9997	1.0000	1.2439

Table 2 indicates that the recurrence period (Q) of earthquake with magnitude 4 is 0.0001 year. It means that earthquake with magnitude 4 is often occur in Indonesia. While the recurrence period (Q) for each earthquake with magnitude 8.1 and 8.2 is 0.9975 and 1.2439 years. It means that earthquake with magnitude 8.1 and 8.2 will be occur in 1-2 years in Indonesia. The probability of earthquake with magnitude ≥ 8.2 will be occur in 1 year is 0.5524, while it is 1.000 in 15 years.

Copula

Bayesian simulation for this method is performed by using package 'sbgcop' in R software because the model is too complicated if we use Winbugs. The same with Gutenberg Richter before, in Copula we use 30000 sampling. The Copula used is Normal Copula, based on the previous analysis. The output is shown in Figure 5 :

```
 #### MCMC details         ####

 number of saved samples: 1000

 average effective sample size: 50.95151

 effective sample sizes
  M*Y
50.95

 #### Parameter estimation ####
 Posterior quantiles of correlation coefficients:
    2.5% quantile 50% quantile 97.5% quantile
M*Y            -1           -1          -0.99

 Posterior quantiles of regression coefficients:
    2.5% quantile 50% quantile 97.5% quantile
M~Y            -1           -1          -0.99
Y~M            -1           -1          -0.99
```

Figure 5 : Bayesian Simulation for Copula

Table 3 : The Copula function $C(u,v)$; Survival Copula function $\hat{C}(u,v)$; and The Cumulative Survival Copula function $\bar{C}(u,v)$

Variable						Normal Copula ($\rho = -0.99$)		
M	Y	u	v	1-u	1-v	$C(u,v)$	$\hat{C}(u,v)$	$\bar{C}(u,v)$
4	3.6192	0.0208	0.8958	0.9792	0.1042	0.00000	0.00000	0.08340
4.5	3.4720	0.1250	0.7917	0.8750	0.2083	0.00009	0.00009	0.08339
5	3.0954	0.2292	0.6875	0.7708	0.3125	0.00069	0.00069	0.08399
5.5	2.5109	0.3333	0.5833	0.6667	0.4167	0.00137	0.00137	0.08477
6	2.0212	0.4375	0.4792	0.5625	0.5208	0.00173	0.00173	0.08502
6.5	1.5855	0.5417	0.3750	0.4583	0.6250	0.00157	0.00157	0.08487
7	1.0607	0.6458	0.2708	0.3542	0.7292	0.00098	0.00098	0.08438
7.5	0.6767	0.7500	0.1667	0.2500	0.8333	0.00028	0.00028	0.08358
8	-0.3010	0.8542	0.0417	0.1458	0.9583	0.00000	0.00000	0.10410
8.1	-0.3010	0.8750	0.0417	0.1250	0.9583	0.00000	0.00000	0.08330
8.2	-0.3010	0.8958	0.0417	0.1042	0.9583	0.00000	0.00000	0.06250

Figure 5 indicates that the parameter estimate of Normal Copula is -0.99. The Copula function $C(u,v)$ and Survival Copula function $\hat{C}(u,v)$ are shown in Table 3. Those results are obtained by Cumulative Distribution Function (CDF) calculation of Normal Copula by using parameter -0.99 and imputing the value of u and v. The Cumulative Survival Copula function $\bar{C}(u,v)$ is obtained by using the same way but we input the value of 1-u and 1-v. By using the results in Table 3, then we can calculate the risk parameters as shown in Table 4.

Table 4 : Risk Parameters of Earthquake Based on Copula (t = 4 years)

M	Y	N	$\mu = t/N$	Q	R
4	3.6192	16645	0.0002	0.0002	1.0000
4.5	3.4720	11858	0.0003	0.0003	0.9999
5	3.0954	4983	0.0008	0.0008	0.9993
5.5	2.5109	1297	0.0031	0.0031	0.9986
6	2.0212	420	0.0095	0.0095	0.9983
6.5	1.5855	154	0.0260	0.0260	0.9984
7	1.0607	46	0.0870	0.0870	0.9990
7.5	0.6767	19	0.2105	0.2106	0.9997
8	-0.3010	2	2.0000	2.0000	1.0000
8.1	-0.3010	2	2.0000	2.0000	1.0000
8.2	-0.3010	2	2.0000	2.0000	1.0000

The risk parameters result in Table 4 is quite same as the result in Table 2. Based on Copula result, the probability of earthquake with magnitude 4 will be occur in 4 years is 1.0000, while the recurrence period is 0.0002. It means that earthquake with magnitude 4 will be occur in almost everyday in Indonesia. The probability of earthquake with magnitude 8, 8.1, and 8.2 will be occur in 4 years is 1.0000, while the recurrence period is 2.0000. It means that earthquake with magnitude 8, 8.1, and 8.2 will be occured in 2 years later in Indonesia.

6. Conclusions

The aim of this study is to determine the risk parameter of earthquake. Based on the analysis that has been done, it can be concluded that Gutenberg-Richter method and Copula method are the useful methods for earthquake case studies analysis. Based on Gutenberg-Richter method, earthquake with magnitude 8.1 and 8.2 will be occur in 1-2 years in Indonesia. The probability of earthquake with magnitude ≥ 8.2 will be occur in 1 year is 0.5524, while it is 1.000 in 15 years. Based on Copula method, the probability of earthquake with magnitude 8, 8.1, and 8.2 will be occur in 4 years is 1.0000, while the recurrence period is 2.0000. It means that earthquake with magnitude 8, 8.1, and 8.2 will be occured in 2 years later in Indonesia. The Bayesian simulation is a good method to use to estimate the risk parameters well both for Gutenberg Richter model and Copula. For the future research, it is highly recommended for using the last update data of earthquake occurrence and the other modelling analysis like general linear model.

Acknowledgment

We are grateful to Directorate of Research and Community Service, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia, for supporting this research.

References

1. Campbell, K.W. (1982). Bayesian Analysis of Extreme Earthquake Occurrences. Part I. Probabilistic Hazard Model. Bulletin of the Seismological Society of America, Vol. 72, No. 5, pp. 1689-1705.
2. Kara, E.K. (2017). The Earthquake Risk Analysis Based on Copula Models for Turkey. Sigma J Eng & Nat Sci 35 (2), 2017, 187-200.
3. Oktaviana, P.P. and Fithriasari, K. (2017). Bayesian Network Inference in Binary Logistic Regression: A Case Study of Salmonella sp Bacterial Contamination on Vannamei Shrimp. Journal of Mathematics and Statistics 2017, 13 (4): 306.311.
4. Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris 8: 229–231.
5. Susilastusi. (2016). Indonesia Country Profile. National Disaster Management Authority.
6. Wibowo, A., Surbakti, I., dan Yunus, R. (2013). Presentation file: Indonesia Disaster Database, Data Informasi Bencana Indonesia (DIBI). (http://dibi.bnpb.go.id/).