Journal of Science and Science Education
Vol. 2 No. 2 November 2018 63–67
ISSN: 2598-3830 (Online)

Comparison of Parameter Estimation Methods to Determine the Frequency Data Magnitude of Aftershock in Nabire, Papua

Alvian Sroyer*, Tiku Tandiangnga, Felix Reba, Ishak Semuel Beno

Department of Mathematics, Cenderawasih University, Papua, Indonesia

ABSTRACT

Many researchers use Weibull distribution to analyze wind speed data parameters and so forth, but in this research Weibull distribution is used to analyze the frequency data magnitude of aftershock. We use different methods for estimating Weibull distribution parameters. Some methods are compared according to the mean square error (MSE) criteria to select the best method.The parameter estimation results from the data are then used to determine the mean and magnitude of the earthquake. Further data is depicted in a curve for analysis. The case study in this study uses an aftershock frequency data in Nabire district, Papua. After the test of conformity with Kolmogorov-Smirnov test, it is obtained that the data follows the Weibull distribution pattern. Further research results show that the best method among the three methods is the maximum likelihood method (MLE).

Keywords: Weibull, Maximum Likelihood Method (MLE), Least Squares Method (LSM), Graphical Method, Kolmogorov Smirnov, Frequency data magnitude.

Received: September 4, 2018; **Accepted:** October 8, 2018; **Published online:** December 3, 2018
*Corresponding author: sroyeralvian@gmail.com

Citation: Sroyer, A., Tandiangnga, T., Reba, F., & Beno, I. S. (2018). Comparison of parameter estimation methods to determine the frequency data magnitude of aftershock in Nabire, Papua. *Journal of Science and Science Education*, *2*(2), 63–67

1. Introduction

Indonesia is one of the countries prone to earthquake since Indonesia is situated a meeting area of three tectonic plates. This condition makes the region of Indonesia an active tectonic area with high seismicity (Gaillard *et al.*, 2008). One of them is in the Papua and West Papua provinces. This setting causes the area of Papua and West Papua are both vulnerable to the occurrence of earthquake (Anton & Gibson, 2008). Many previous studies used Poisson, Exponential, Gamma and Weibull distributions to analyze the probability of earthquakes. Greenhough & Main (2008) used poisson distribution in the investigation of the hazard of earthquake frequency. Meanwhile Hagiwara (1974) used Weibull distribution to determine the occurrence of earthquake. In determining the probability of a distribution, it should be noted how much error is generated. Thus, we need to try some methods in estimating the parameters of a distribution. Werapun *et al.* (2015) and Saleh *et al.* (2012) compared several methods in estimating the Weibull distribution parameters. The best method is the method that generates the minimum error using the mean square error (MSE) criteria. However, not many have studied the determination of earthquake magnitude which has Weibull distribution with some estimation method. It is therefore in this study after determining the best method in estimating the parameters, then the data parameter is utilised to determine the magnitudes of the earthquake. The general purpose of this research is to mitigate future earthquake disaster, especially in Nabire District-Papua, where the occurrence of the 2004 earthquake led to the paralysis of the entire city, the electricity and telephone connections were completely disconnected and even the Nabire airport also experienced massive damage.

2. Estimation Method of Weibull Distribution Parameters

In estimating Weibull distribution with two parameters, the Least Square Method (LSM), Maximum Likelihood Method (MLE) and Graphical Method are employed in the study.

2.1 Distribution of two parameter Weibull

A continuous *T* random variable which follows a Weibull distribution with the shape parameter $c > 0$ and scale parameter $b > 0$ has the following probability density function (Rinne (2008)):

$$f(t|b,c) = \begin{cases} \frac{c}{b}\left(\frac{t}{b}\right)^{c-1} \exp\left\{-\left(\frac{t}{b}\right)^{c}\right\} & ,t > 0, c > 0, b > 0. \\ 0 & ,\text{for others,} \end{cases}$$

The cumulative distribution function is as follows:

$$F(t) = 1 - S(t) = 1 - \exp\left\{-\left(\frac{t}{b}\right)^{c}\right\}, t > 0. \tag{1}$$

In survival analysis, the survival function $S(t)$ can be obtained from Eq. (1), that is, the result of the division of the density function by the hazard function:

$$S(t) = \frac{f(t)}{h(t)} = \frac{\frac{c}{b}\left(\frac{t}{b}\right)^{c-1} \exp\left\{-\left(\frac{t}{b}\right)^{c}\right\}}{\frac{c}{b}\left(\frac{t}{b}\right)^{c-1}} = \exp\left\{-\left(\frac{t}{b}\right)^{c}\right\}, t > 0$$

After the shape and scale parameters are identified, the earthquake mean and magnitude can be determined. The expected value *T* is symbolized by $E(T)$ and defined as follows:

$$E(T) = \int_{-\infty}^{\infty} t\, f(t)dt = b\Gamma\left(\frac{1}{c} + 1\right).$$

Furthermore, the magnitude frequency relationship is given by, Bormann Peter (2002),

$$\log N(M) = p - qM,$$

where M is the Richter scale, $N(M)$ is the number of an earthquake with its Richter scale M, and (p,q) is a constants vector which is referred to as seismicity parameters.

2.2 Maximum Likelihood Estimation (MLE) method

Parameter estimation requires numerical iteration. The shape parameter $c > 0$ and scale parameter $b > 0$ is estimated by the following two equations, Johnson (1996),

$$c = \left[\frac{\sum_{i=1}^{n} t_i{}^{c}\ln(t_i)}{\sum_{i=1}^{n} t_i{}^{c}} - \frac{\sum_{i=1}^{n}\ln(t_i)}{n}\right]^{-1},$$

$$b = \left[\frac{1}{n}\sum_{i=1}^{n} t_i{}^{c}\right]^{\frac{1}{c}},$$

where t_i is the frequency of the aftershocks at i and n represents the amount of earthquake frequency data. Iteration is required to determine parameter c.

2.3 Least Squares Method (LSM)

The second estimation technique to be discussed is the least squares method. This method is very commonly applied in physics, engineering and mathematics problems. The use of the LSM for Weibull distribution is as follows (Osarumwense & Rose, 2014):

$$\hat{c} = \frac{\left\{n.\sum_{i=1}^{n}(\ln t_i).(\ln\{\ln[\frac{1}{1-\frac{i}{n+1}}]\}\right\} - \left\{\sum_{i=1}^{n}\ln(\ln[\frac{1}{1-\frac{i}{n+1}}]).\sum_{i=1}^{n}\ln t_i\right\}}{\{n.\sum_{i=1}^{n}(\ln t_i)^2\} - \{\sum_{i=1}^{n}(\ln t_i)\}^2},$$

$$\hat{b} = e^{\left(\frac{\bar{X}-\bar{Y}}{c}\right)}.$$

2.4 Graphical Method (GM)

Usually, there are physical methods that are used because of their simplicity and speed. However, sometimes some methods produce a greater probability of errors. Next, we will discuss two main things in graphical method. If both sides of the cumulative distribution function in Eq. (1) are transformed by $\ln\left(\frac{1}{1-t}\right)$, then we obtain:

$$\ln\left(\frac{1}{1-F(t_i)}\right) = \left(\frac{t_i}{b}\right)^c$$

$$\ln\left[\ln\left(\frac{1}{1-F(t_i)}\right)\right] = c\ln t_i - c\ln b.$$

3. Assessment and Selection Method

We use the Mean Square Error (MSE) to determine the best method of guessing the Weibull distribution parameters with two parameters. Meanwhile, the Average Square Error (ASE) is used to measure the accuracy of estimator.

3.1 Goodness-of-fit test

Kolmogorov Smirnov test is generally based on the cumulative distribution function. Suppose a random variable $T_1, T_2, \dots, T_n$comes from an unknown distribution $F(t)$ and $t_{(1)} < t_{(2)} < \cdots < t_{(n)}$ are in statistical order. We will test the hypothesis that $F(t)$ is equal to a particular distribution $F_0(t)$. Kolmogorov Smirnov statistical test, D_n, is defined by

$$D_n = max(D_+, D_-),$$
$$D_+ = max[F_n(t) - F_0(t_i)],$$
$$D_- = max[F_0(t_i) - F_{n-1}(t)].$$

for $i = 1, 2, \dots, n$, and $F_n(t)$ is an empirical distribution function. The empirical distribution function is useful as a predictor of an unknown distribution function $F(t)$.

3.2 Comparison of Estimation Methods

Criteria of Mean Squared Error (MSE) is as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}[\tilde{F}(t_i) - F(t_i)]^2,$$

where $\tilde{F}(t_i)$ is obtained by substituting the parameter values *c* and *b* (for each method) in Eq. (1), whereas $\hat{F}(x_i) = 1 - e^{\left(\frac{-t_i}{b}\right)^c}$ and $F(x_i) = \frac{i-0.3}{n-0.4}$ are empirical distribution functions. A method which has minimum square error MSE_{min} becomes the best method to estimate the Weibull distribution parameters.

4. Results and Discussion

Based on the Table 1, the MSE indicates that the best method in estimating the parameters of subsequent earthquake frequency data in Nabire, Papua is the MLE, with the magnitude is generated at 5.9584 at 15:00 to 21:00 and 6,0253 at 03:00 to 09:00.

Table1. Parameter scale and shape, mean, Richter scale.

Method	Hour	Parameter		Richter	Mean	MSE
		Scale	Shape	Scale		
MLE	15.00 – 21.00	54.8117	1.3889	5.9584	50.0199	6.876443
	03.00 – 09.00	47.4489	1.3941	6.0253	43.2749	7.938585
LSM	15.00 – 21.00	49.489	1.6307	5.9633	49.489	7.643076
	03.00 – 09.00	43.716	1.5075	6.0206	43.716	8.354454
GM	15.00 – 21.00	47.2601	1.7389	7.5158	47.2601	8.696648
	03.00 – 09.00	40.9555	1.6714	7.5349	40.9555	8.88584

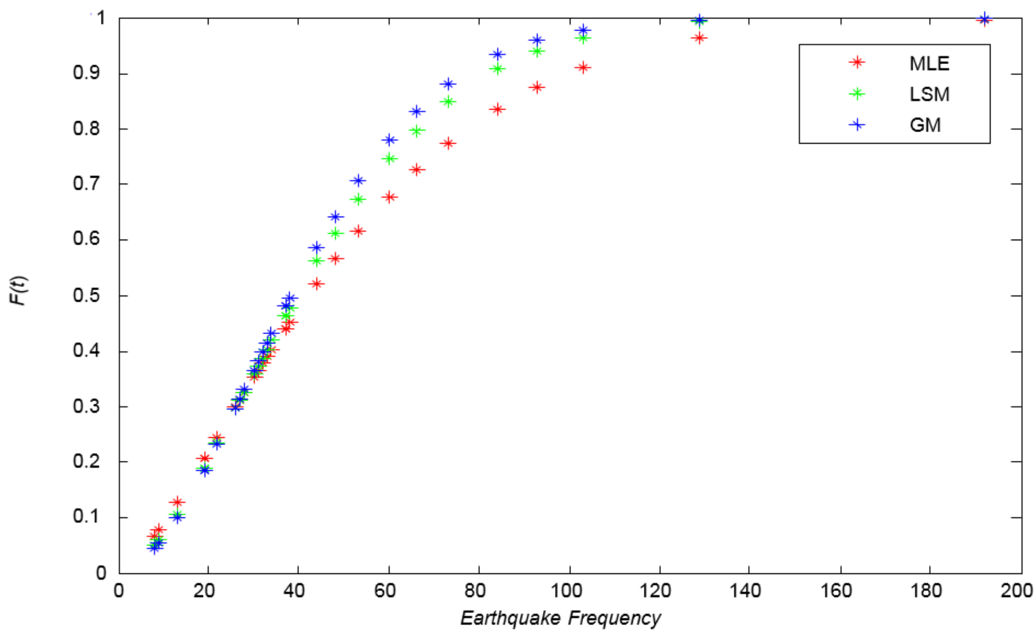

Figure 1. Earthquake frequency vs distribution $F(t)$.

Figure 1 clearly show that the hours of 03.00–09.00 appears higher earthquake frequency compared to other hours, where the parameters of shape and scale are produced in a row as (c = 1.3941 and b = 47.4489). Thus the aftershocks that occurred in Nabire, Papua in 2004 at 03:00 to

09:00 instigated everyone to get out of their rooms, even people who were driving could also felt the shock, a building that has poor construction was collapsed.

5. Conclusion and Suggestion

5.1 Conclusion

In this research we compared the least squares, maximum likelihood estimation, and graphical methods to estimate Weibull distribution parameters for frequency data magnitude of aftershock in Nabire district, Papua. We showed that the best method in estimating parameters of aftershock data is the MLE.

5.2 Suggestion

Further research can be done with the different estimation method and distribution. The earthquake parameters also need to be estimated independently on the basis of the data.

References

Anton, L., & Gibson, G. (2008). Analysing earthquake hazard in Papua New Guinea. *Earthquake Engineering in Australia, AEES2008, Ballarat Victoria, Australia*, 21–23.

Bormann, P. (2002). New manual of seismological observatory practice (NMSOP-1). *GeoForschungZentrum, Potsdam, 1252p*.

Gaillard, J. C., Clavé, E., Vibert, O., Denain, J. C., Efendi, Y., Grancher, D., & Setiawan, R. (2008). Ethnic groups' response to the 26 December 2004 earthquake and tsunami in Aceh, Indonesia. *Natural Hazards*, *47*(1), 17–38.

Greenhough, J., & Main, I. G. (2008). A Poisson model for earthquake frequency uncertainties in seismic hazard analysis. *Geophysical Research Letters*, *35*.

Hagiwara, Y. (1974). Probability of earthquake occurrence as obtained from a Weibull distribution analysis of crustal strain. *Tectonophysics*, *23*(3), 313–318.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1996). Continuous univariate distributions. *Journal of the Royal Statistical Society-Series A Statistics in Society*, *159*(2), 349.

Osarumwense, O. I., & Rose, N. C. (2014). Parameters estimation methods of the Weibull distribution: A comparative study. *Elixir Statistics*, 69, 23177–23184.

Rinne, H. (2008). *The Weibull distribution: a handbook*. Chapman and Hall/CRC.

Saleh, H., Aly, A. A. E. A., & Abdel-Hady, S. (2012). Assessment of different methods used to estimate Weibull distribution parameters for wind speed in Zafarana wind farm, Suez Gulf, Egypt. *Energy*, *44*(1), 710–719.

Werapun, W., Tirawanichakul, Y., & Waewsak, J. (2015). Comparative study of five methods to estimate Weibull parameters for wind speed on Phangan Island, Thailand. *Energy Procedia*, *79*, 976–981.