Applied Mathematical Sciences, Vol. 12, 2018, no. 24, 1191 - 1199
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/ams.2018.87110

Best Fit Probability Model for Runup Height Tsunami in Aceh Using Some Mixture Distribution

Desta Wahyuni and Arisman Adnan

Department of Mathematics, Faculty of Mathematics and Science
University of Riau, Pekanbaru 28293, Indonesia

Rado Yendra[1,2], Ari Pani Desvina[1] and M. N. Muhaijir[1,*]

[1] Deparment of Mathematics, Faculty of Science and Technology
UIN Sultan Syarif Kasim Riau, Pekanbaru 28293, Indonesia

[2] Rado Research Center, Pekanbaru 28293, Indonesia
[*]Corresponding author

M. Arrafie Abduh

Faculty of Ushuluddin, UIN Sultan Syarif Kasim Riau
Pekanbaru 28293, Indonesia

Abstract

Choosing best fit probability distribution to represent the height wave tsunami has been a long topic of interest in hydrology. In this study, Lognormal distribution (L), Gamma distributions (G), Weibull (W), and mixture of two lognormal (ML), two gamma (MG) or two weibull (MW) distributions were applied to data tsunami runnup heights in Aceh. Parameter for each distribution are estimated by maximum likelihood techniques. For selecting the best fit model, graphical inspection (probability density function (pdf)) and numerical criteria (Akaike's information criterion (AIC), Bayesian information criterion (BIC)) were used. In most the cases, graphical inspection gave the same result but their AIC and BIC result differed. The best fit result was chosen as the distribution with the lowest values of BIC and AIC. Tsunami that happened in Aceh on 26 December 2004 at latitude 3.32 N and longitude 95.85 E, G and W distributions was not suitable to explain the tsunami heights wave data distribution. Result show that MG and MW are better alternatives to discribe height wave tsunami characteristics. We also show that MG best fit probability model in comparison to either G, M, or MW.

Keywords: Lognormal distribution, Gamma distributions, Weibull distribution, and mixture of two lognormal distribution, goodness of fit tests

1 Introduction

Tsunamis are long-period oceanic gravity waves generated by a large disruption of the entire water column, most notably via seafloor deformation resulting from a seismic event. Typically tsunamis are generated by large, shallow submarine earthquakes, but submarine landslides or volcanic activity and rarely, even an asteroid impact can generate a tsunami.

The massive earthquake on December 26, 2004 with magnitude 9.3, launched huge tsunami waves affected many coastal countries across Indian Ocean. It is important to first define the variable that defines the size of a tsunami. Although runup is the measurement most often associated with tsunamis, because it is defined as the wave height with respect to ambient sea level at the maximum inundation distance, runup will occur at different geographic locations for different tsunamis. Tide gauges, on the other hand, record wave amplitude at a fixed location.

For most probability problems, comparisons are made over broad geographic regions that may include both runup and wave amplitude measurements. Throughout this study, we will refer to runup as the tsunami size or hazard variable, although this may include other amplitude measurements of tsunamis as well. Tsunamis can be considered a stochastic process. As a result, tsunami probabilities can be defined by the frequency distribution of sizes.

Various statistical methods have been applied to analyze the observed tsunami runup height sheights or tsunami amplitudes to determine the maximum expected water level near the shore. Soloviev [21] investigated tsunami frequency-size distribution in terms of tsunami intensity and determined a classification of tsunami intensity based on the runup. Houston [7] determined frequency-of-occurrence curves based on the maximum expected wave elevation near the shore. Severnl authors have investigated the spatial distribution of tsunami heights along coastlines with extensive historical records, which tend to follow a log-normal distribution [1, 13, 25].

There are many variations in the existing methods for the probabilistic analysis of earthquake occurrence itself, as highlighted by Utsu [24]. Several examples of many variations of the famous and widely-used Gutenberg-Richter magnitude-frequency relation introduced in 1944 [4, 5, 6, 10, 11, 14, 15, 17, 22], the Weibull distribution [19], the Lognormal distribution [16], variations of the Gamma distribution [8, 9], variations of the Pareto distribution [18], the use of Bayesian statistics [2, 23] and many others. In addition, some authors believe that the factors such as the b-value are universal [9, 12], while others consider regional differences in the distribution to be significant. Given the many distributions applied with reasonable results, it suggests frequency of occurrence variations between seismic zones will be a significant factor in probabilistic analysis.

In recent past, mixture distribution were used to estimate runup heights tsunami that a quite acurate in discribing heights wave tsunami characteristics. Smit et al. [20] used mixture distribution to estimate the probabilities of exceedance and return periods for tsunamis in the tsunamigenic regions of Japan, Kuril–Kamchatka, and South America. Geist and Parsons [3] used mixture distribution to estimate the largest tsunami that can be expected in a given time period at a station.

In probability theory and statistics, the concept of mixture distribution is the combination of two or more probability distributions to create a new probability distribution. The objective of this study is to proposeGamma distributions (G), Weibull (W), and mixture of two gamma (MG) or two weibull (MW) distributions for runup height wave tsunami that happened in Aceh on 26 December 2004 forecasting. Comparison of the proposed mixture distributions with existing ditribution functions is done to demonstrate their suitability in describing runup heights wave tsunami characteristics.

2 Study Area and Data

107 data of Aceh heights wave tsunami had been collected that happened in Aceh archipelago, Banda Aceh, Seashore and Sigli. Some examples of wave position and height data are shown in Table 1.

Table 1. Aceh Tsunami Height Wave Data on Various Points

No.	Poitns of Sample	Latitude		Longitude		Heights Wave Tsunami
		Degree	Minute	Degree	Minute	
1	Sigli	5	23.052	95	58.082	4.07
6	Krueng Raya port	5	35.768	95	31.560	5.10
9	Sabang	5	49.574	95	20.837	3.02
12	Center of Banda Aceh	5	33.356	95	17.044	12.0
67	West Coast of Banda Aceh	5	28.638	95	14.696	12.42
107	West Coast of Banda Aceh	5	27.000	95	14.585	20.07

3 Methods

Runup height wave tsunami modelling requires analysis of height wave tsunami data, it is desirable to use statistical distribution function for describing the runup height wave tsunami variations. The primary tools to describe runup height wave tsunami characteristics are probability distribution functions. For selecting the best fit model, choice of the model definition, parameter estimation tools are important. The parameter estimation of the distribution function are calculated using maximum likelihood method. The procedure of goodness of fit tests for model selection, both numerically and graphically, is discussed. Some distributions used for data analysis are presented in Table 2.

Table 2 List of Distributions Used in This Study

Distribution	Formula
Lognormal (L)	$q(\log h;\ \mu,\sigma) = \frac{1}{\sqrt{2\pi \log h\, \sigma^2}} exp\left[-\frac{1}{2}\left(\frac{\log h - \mu}{\sigma}\right)^2\right]$
Weibull (W)	$f(h,k,c) = \frac{k}{c}\left(\frac{h}{c}\right)^{k-1} exp\left[-\left(\frac{h}{c}\right)^k\right]$
Gamma (G)	$g(h,\alpha,\beta) = \frac{h^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} exp\left(-\frac{h}{\beta}\right)$
Mixture Lognormal (ML)	$q(h_i; \mu_1, \sigma_1, \mu_2, \sigma_2, p) = pq(\log h_i\,; \mu_1, \sigma_1) + (1-p)q(\log h_i\,; \mu_2, \sigma_2, p)$
Mixture Weibull (MW)	$g(h_i; k_1, c_1, k_2, c_2, p) = pf(h_i; k_1, c_1) + (1-p)f(h_i; k_2, c_2)$
Mixture Gamma (MG)	$f(h_i; \alpha_1, \beta_1, \alpha_2, \beta_2, p) = pg(h_i; \alpha_1, \beta_1) + (1-p)g(h_i; \alpha_2, \beta_2)$

4 Result and Discussion

Runup height wave data were used in evaluating different probability density function, the data used for this paper are presented in table 1. In this resecearch, runup height wave tsunami data histogram on figure 1 gives a description about data that probably have more than one distribution. The visual technique of plotting data is one of the important methods for selecting a probability density function, this includes examining a histogram with the distribution overlaid and comparing the empirical model to the theoretical model.

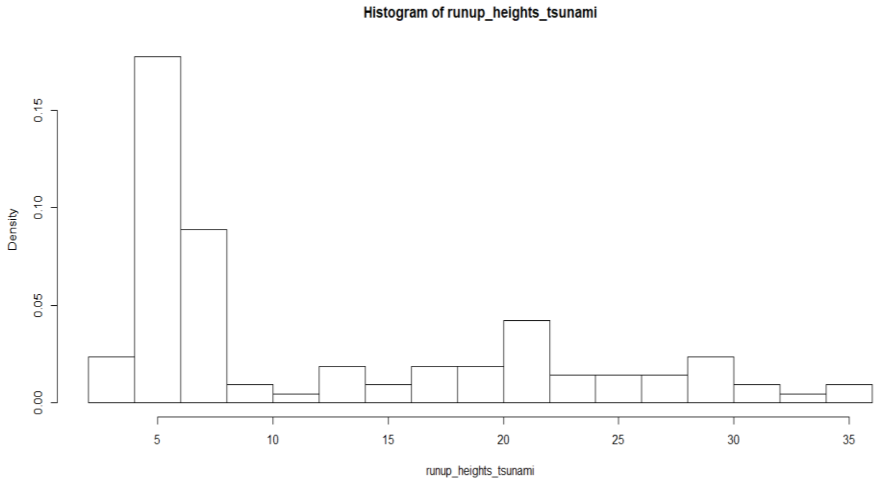

Figure 1. Observed Runup Height Wave Tsunanami Frequencies of Aceh

Computed parameter values of different Probability Density Function (PDF) are presented in Table 3. The statistical parameters for fitness evaluation of PDF currently analyzed are presented in Table 4.Statistical GOF namely AIC and BIC given in Table 4 from the various components of the distribution. Given a set of candidate models for a data set, the best fit model is taken as the minimum value for every case of AIC and BIC. In this research, indicate that proposed MG distribution provides best fit for the observed runup height wave tsunami frequency distribution, which is closely followed by ML and MW distribution. Convention PDF such as L, W, and G over predicted runup height wave tsunami aceh's data. As seen from Figure 2 and statistical parameters from table, MG provided the best fit for observed runup height wave data, closely followed by proposed ML and MW. Figure 2 show that mixture PDF fit much better than the conventional Log-Normal, Weibull and Gamma distributions.

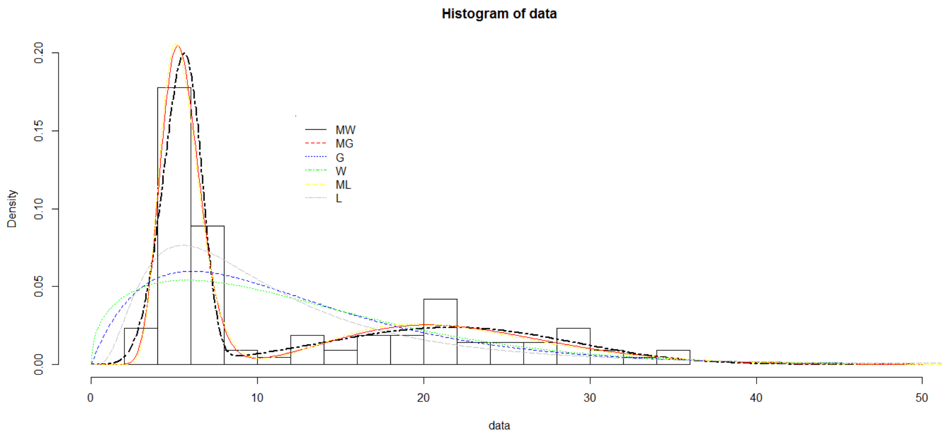

Figure 2. Predicted and Observed Runup Height Wave Tsunami Frequency of Aceh

Table 3. Computed Parameter Values of Different Probability Density Functions

	L	W	G	ML	MW	MG
p	-	-	-	0.5971988	0.5723286	0.5939644
μ	2.2341940	-	-	-	-	-
σ	0.7281442	-	-	-	-	-
ak	-	1.417413	-	-	-	-
c	-	13.605415	-	-	-	-
α	-	-	1.9829590	-	-	-
β	-	-	0.1616177	-	-	-
μ_1	-	-	-	1.6745425	-	-
σ_1	-	-	-	0.2223441	-	-
μ_2	-	-	-	3.0639412	-	-
σ_2	-	-	-	0.3002022	-	-
k_1	-	-	-	-	5.3650879	-
c_1	-	-	-	-	5.8103044	-
k_2	-	-	-	-	3.4472168	-
c_2	-	-	-	-	23.9537009	-
α_1	-	-	-	-	-	21.2558594
β_1	-	-	-	-	-	3.9000931
α_2	-	-	-	-	-	11.2609375
β_2	-	-	-	-	-	0.5062223

Table 4. Statistical and Best-fit Result of the Runup Height Wave Tsunami

Test	L	W	G	ML	MW	MG
Log Likelihood	-356.938	-365.800	-363.127	-314.133	-314.428	-313.294
AIC	717.877	735.601	730.254	638.266	638.857	636.589
BIC	723.223	740.946	735.600	651.631	652.222	649.954

5. Conlusions

In the present article, we conclude that the measured tsunami runup heights in Aceh as a whole are described by the finite mixture distributions, especially the Log Normal mixture distribution (ML), Weibull mixture distribution (MW) and Gamma mixture distribution (MG). A number of graphical (the frequency of probabilities density function) and numerical performance criteria (AIC and BIC) were used to select the best-fit model for each of finite mixture distribution. The two-component Gamma mixture distribution (MG) gives the best-fit result of runup tsunami heights wave on Aceh Island.

References

[1] B. H. Choi, E. Pelinovsky, A. Ryabov and S. I. Hong, Distribution Functions of Tsunami Wave Heights, *Natural Hazards*, **25** (2002), no. 1, 1–12. https://doi.org/10.1023/a:1013379705323

[2] O. C. Galanis, T. M. Tsapanos, G. A. Papadopoulos and A. A. Kiratzi, Bayesian extreme values distribution for seismicity parameters assessment in South America, *Journal of the Balkan Geophysical Society*, **5** (2002), no. 3, 77–86.

[3] E. L Geist and T. Parsons, Reconstruction of far-field tsunami amplitude distributions from earthquake sources, *Pure and Applied Geophysics*, **173** (2016), no. 12, 3703–3717. https://doi.org/10.1007/s00024-016-1288-x

[4] B. Gutenberg, The energy of earthquakes, *Quarterly Journal of the Geological Society*, **112** (1956), no. 1–4, 1–14. https://doi.org/10.1144/gsl.jgs.1956.112.01-04.02

[5] B. Gutenberg and C. F. Richter, *Seismicity of the Earth and Associated Phenomena*, Princeton University Press, Princeton, New Jersey, 1954.

[6] B. Gutenberg and C. F. Richter, Frequency of earthquakes in California, *Bulletin of the Seismological Society of America*, **34** (1994), no. 4, 185–188.

[7] J. R. Houston, R. D. Carver and D. G. Markle, *Tsunami-Wave Elevation Frequency of Occurrence for the Hawaiian Islands*, U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi, 1977.

[8] Y. Y. Kagan, Seismic moment distribution, *Geophysical Journal International*, **106** (1991), no. 1, 123-134. https://doi.org/10.1111/j.1365-246x.1991.tb04606.x

[9] Y. Y. Kagan, Seismic moment-frequency relation for shallow earthquakes: Regional comparison, *Journal of Geophysical Research: Solid Earth*, **102** (1997), 2835–2852. https://doi.org/10.1029/96jb03386

[10] Y. Y. Kagan and D. O. Jackson, Probabilistic forecasting of earthquakes, *Geophysical Journal International*, **143** (2000), no. 2, 438–453. https://doi.org/10.1046/j.1365-246x.2000.01267.x

[11] Y. Y. Kagan, Seismic moment distribution revisited: I. Statistical Results, *Geophysical Journal International*, **148** (2002), 520–541. https://doi.org/10.1046/j.1365-246x.2002.01594.x

[12] Y. Y. Kagan, Universality of the seismic moment-frequency relation, *Pure and Applied Geophysics,* **155** (1999), no. 2, 537–573. https://doi.org/10.1007/s000240050277

[13] K, Kajiura, Some statistics related to observed tsunami heights along the coast of Japan, in *Tsunamis-Their Science and Engineering*, Terra Scientific Publishing Company, Tokyo, Japan, 1983, 131–145.

[14] H. Kanamori and D. L. Anderson, Theoretical basis of some empirical relationsin seismology, *Bulletin of the Seismological Society of America,* **65** (1975), no. 5, 1073–1095.

[15] E. A. Okal, Seismic parameters controlling far-field tsunami amplitudes: A review, *Natural Hazards*, **1** (1988), 67–96. https://doi.org/10.1007/bf00168222

[16] K. Orfanogiannaki and G. A. Papadopoulos, Conditional probability approachof the assessment of tsunami potential: Application in three tsunamigenic regions of the Pacific Ocean, *Pure and Applied GeophySiCS,* **164** (2007), no. 2, 593–603. https://doi.org/10.1007/s00024-006-0170-7

[17] J. F. Pacheco and L. R. Sykes, Seismic moment catalog of large shallowearthquakes, 1900 to 1989, *Bulletin of the Seismological Society of America*, **82** (1992), no. 3, 1306-1349.

[18] V. F. Pisarenko and D. Somette, Characterization of the frequency of extremeearthquake events by the generalized Pareto distribution, *Pure and Applied Geophysics,* **160** (2003), 2343–2364. https://doi.org/10.1007/s00024-003-2397-x

[19] T. Rikitake and I. Aida, Tsunami hazard probability in Japan, *Bulletin of the Seismological Society of America*, **78** (1988), 1268–1278.

[20] A. Smit, A. J. Kijko dan A. Stein, Probabilistic tsunami Hazard assessment from incomplete and uncertain historical catalogues with application to tsunamigenic regions in the Pacific Ocean, *Pure and Applied Geophysics*, **174** (2017), 3065–3081. https://doi.org/10.1007/s00024-017-1564-4

[21] S. L. Soloviev, *Recurrence of tsunamis in the Pacific, Tsunamis in the Pacific Ocean*, East West Center Press, Honolulu, Hawaii, 1970.

[22] D. Sornette, L. Knopoff, Y. Y. Kagan and C. Vanneste, Rank-ordering statisticsof extreme events: Application to the distribution of large earthquakes, *Journal of Geophysical Research: Solid Earth*, **101** (1996), no. B6, 13883– 13894. https://doi.org/10.1029/96jb00177

[23] T. M. Tsapanos, A. A. Lyubushin and V. F. Pisarenko, Application of a Bayesian approach for estimation of seismic Hazard parameters in some regions of the circum Pacific belt, *Pure and Applied Geophysics*, **158** (2001), no. 5, 859–875. https://doi.org/10.1007/pl00001210

[24] T. Utsu, Representation and analysis of the earthquake size distribution: A Historical review and some new approaches, *Pure and Applied Geophysics*, **155** (1999), no. 2-4, 509–535. https://doi.org/10.1007/s000240050276

[25] W. G. Van Dorn, Tsunamis, in *Advances in Hydroscience*, Vol. 2, edited by V.T. Chow, Elsevier, New York, 1965, 1–48. https://doi.org/10.1016/b978-1-4831-9933-7.50007-9

Received: August 15, 2018; Published: September 27, 2018