Generalized Extreme Value Distribution for Modeling Earthquake Risk in Makran Subduction Zone Using Extreme Value Theory

Adil Rehman[1*], Huai Zhang[1]

[1] Key Laboratory of Computational Geodynamics, University of Chinese Academy of Sciences, Beijing 100049, China

*Corresponding author, E-mail: adil.geologist5@gmail.com

Abstract

The long-term pattern of severe incidents is one of the most crucial and fascinating topics of seismic events. The annual maximum earthquake magnitude in the Makran subduction zone is estimated using the extreme value theory method by implementing the block maxima method. The seismic data utilized for the current study range between 1934 to 2022. The extreme parameters have fitted utilizing the generalized extreme value distribution. Numerous plots of the generalized extreme value distribution approach gave the accuracy of the used model when fitted to seismic data of the Makran subduction zone. Using the profile likelihood approach, the shape parameters (ξ) calculated is 0.29. According to the model fit, the Fréchet distribution is the best model for predicting the annual maximum earthquake magnitude in the makran subduction zone. We also calculate the maximum return period, i.e., 10, 20, 50, and 100 years for seismic data of Makran subduction with 95% confidence intervals. We also computed the profile likelihood to achieve a precise confidence interval. The values calculated for the return periods 10, 20, 50, and 100 are 6.35, 6.81, 7.58, and 8.31, respectively, indicating that an earthquake's maximum magnitude is increasing across the future 100 years. Thus the 1945 earthquake of the Makran region with magnitude 8.1(Mw) has one of the most significant events in this area and occurred once every 100 years.

Keywords: Makran subduction zone, Earthquake, Generalized extreme distribution, Extreme value theory, Return Period

Introduction

Extreme rainfall has a severe negative impact on human activity and infrastructure and can cause fatalities. Extreme events have a high risk of catastrophe, are difficult to predict, and have profound economic repercussions. Extreme value theory is a significant statistical subject in the scientific field. The earthquake's magnitude affected several places in the Makran region, i.e., the 1945 earthquake with magnitude 8.1 occurred and caused high fatalities. According to Coles (2001), extreme value methods have become broadly used in various disciplines, i.e., as telecommunication for traffic forecasting, engineering, finance, hydrology, and insurance. According to Beirlant et al.(2006); Coles (2001); De Haan and Ferreira (2006), there are several predicting techniques, extreme statistical approaches have attracted much interest, and the statistical analysis strategy is the same. Statistical methods that concentrate on predicting extreme events in finance and earth sciences have produced encouraging outcomes. Katz et al. (2002); Embrechts et al. (1997) utilize the extreme value theory technique for flood, drought, and financial incident forecasting. Maruyama (2020) employed the extreme value method to analyze Japan's annual maximum magnitude of earthquakes. The extreme data values are routinely simulated in practice using two different approaches. The first method is block maxima, and the second is the peak-over-threshold approach. The GEV distribution can model the block maxima. Simulating only the block maxima would be inefficient if more extreme values were available. Kijko and Sellevoll (1981), Epstein and Lomnitz (1966) utilized extreme value theory to model significant earthquake events. García-Bustos et al. (2018) used an extreme method to calculate the return period for large-magnitude of earthquakes on the Ecuadorian coast. In contrast, Pavlenko (2017) utilized a generalized extreme value approach for calculating Japan's seismic hazard curves. Bustos (2021) used extreme value and Poisson regression methods to characterize the seismicity of five zones of Ecuador through cluster analysis. Mothupi et al. (2012) operated generalized Pareto distribution for monthly maximum temperature modeling in Shakawe in Botswana. Although for the current work, we used the GEV model for maximum earthquake magnitude analysis. Jenkins (1955) utilized generalized extreme distribution for temperature extreme demonstrating. Various researchers utilized an extreme approach for temperature extremes (Chikobvu and Sigauke, 2013; Hasan et al., 2012; Parey et al., 2013; Siliverstovs et al., 2008; Wen et al., 2015). Zimbidis et al. (2007), Pisarenko et al. (2010), and Abbasi et al. (2018) utilized the Extreme value method for earthquake prediction in different regions. Lavenda and Cipollone (2000) utilized extreme values for the thermodynamics of the aftershock sequence. Arreyndip and Joseph (2015) used the GEV model to predict extreme temperatures in Mbonge, Cameroon, and in comparison to other models, the GEV model best fit the data. Buishand and de Haan (2008) utilized extreme value theory to calculate Holland's daily average rainfall. While several other researchers also used the extreme method for rainfall extreme in different regions (Li et al., 2005; Varathan et al., 2010; Shahid, 2011; Chu et al., 2013; Carreau et al., 2013; Ender and Ma, 2014). Several authors used the maximum likelihood approach to study shape, location, and scale parameters (Jenkinson 1969; Prescott and Walden 1980, 1983; Macleod 1989). Although the use of the GEV model to identify extreme seismic magnitude in the makran subduction zone has received little to no attention in literature. In the Makran subduction zone, this work appears to be the first to use the GEV model for modeling extreme seismic events. The return level of the specified period calculated must be stationary. A stationary model must be used to predict the return levels.

This study uses extreme value approaches to examine the annual maximum earthquake magnitude of extreme seismic data for the Makran subduction zone. The main objective of this work is to utilize extreme value theory for forecasting extreme seismic events in the Makran subduction zone. We used the block maxima approach to predict the annual maximum earthquake magnitude in the Makran subduction zone. The GEV model has utilized the maximum likelihood approach to find the best-fit model (Coles, 2001). We also estimated the return level for the return period of 10, 20, 50, and 100 years. This work's outcome will help to provide early warning due to earthquakes to save lives.

Methodology

The data utilized in the current work for the maximum earthquake magnitude analysis in the Makran subduction zone range between 1934 to 2022 has a magnitude ≥5.0. Although a magnitude five earthquake does not cause massive destruction, it can damage buildings and infrastructure. Figure 1 shows the yearly maximum earthquake magnitude of Makran subduction zone seismic data utilizing the block maxima approach. There is no apparent trend in the data. We model the data as discrete observations from the GEV distribution, supposing from Figure 1 that the pattern of variation has remained constant across the observation period. The large magnitude occurred in the year 1945 with a magnitude of 8.1. It is essential to find the distribution that best suits the data to understand the pattern of significant earthquakes in a particular region.

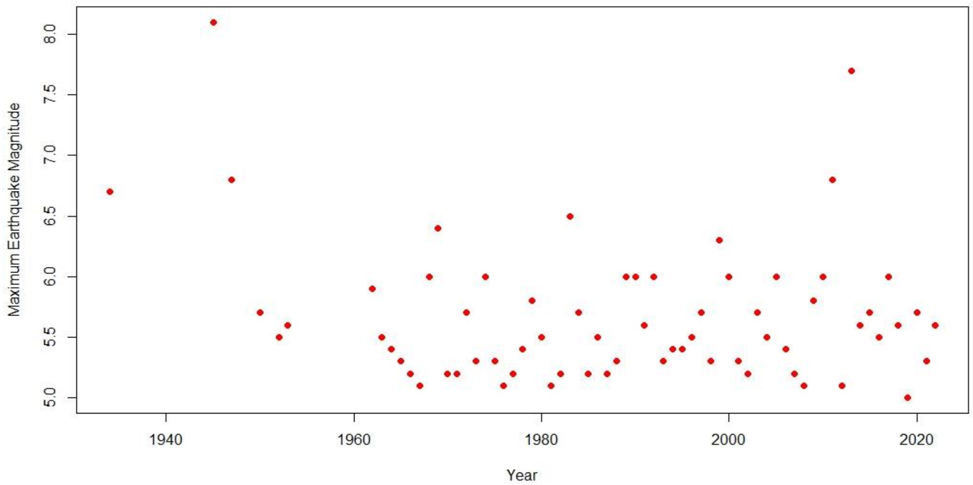

Figure 1. Scatter Plot of Maximum earthquake and time.

In the current work, we utilized GEV distribution for Makran subduction zone seismic data using the extreme package in R (Gilleland et al., 2013). The GEV distribution's cumulative distribution function has given in Equation 1.

$$G(z)=\begin{cases}\exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, For\ \xi\neq 0\\ \exp\left\{-\exp\left[-\left(\frac{z-\mu}{\sigma}\right)\right]\right\}, For\ \xi=0\end{cases} \quad (1)$$

Where Z is the block extreme value, μ is location, ξ is the shape, and σ is scale parameters. The shape parameter determines the tail decay rate. Based on the shape parameter, (ξ) there are three families of GEV distribution. For the shape parameter, $\xi=0$ we get the Gumbel distribution, and $\xi>0$ we get Frechet, while $\xi<0$ then we get the Weibull distribution. Once the data has fitted the GEV model, the focus is on determining the seismic data's return level. The GEV distribution return period and return level can be estimated using equation 1.2.

$$Z_p = \begin{cases} \mu - \frac{\sigma}{\xi}\left[1 - \{-\log(1-p)\}^{-\xi}\right], For\ \xi \neq 0 \\ \mu - \sigma \log\{-\log(1-p)\}, For\ \xi = 0 \end{cases}$$

(1.2)

Where Z_p is the return level linked with return period 1/p. While the probability that an earthquake of magnitude Zp would exceed once each year has given by p.

Result and Discussion

Statistical results from the analysis of the GEV model as described in this section. Figure 2 represents several GEV distribution plots fitted to the maximum annual magnitude seismic data of the Makran subduction zone. The probability and the quantile plot show that the fitted model is accurate since both displayed points are nearly linear. The return curve plot is not linear because the value estimated is more significant than zero. In comparison, the density plot appears compatible with the data. Table 1 represents the maximum annual earthquake magnitude parameters with confidence intervals of GEV distribution calculated using the block approach.

Table 1. Parameters of GEV distribution estimated.

	μ	σ	ξ
Parameters Estimate	5.38	0.30	0.29
Standerd Error	0.04	0.03	0.11
95% CI	(5.30, 5.47)	(0.23, 0.37)	(0.05, 0.52)

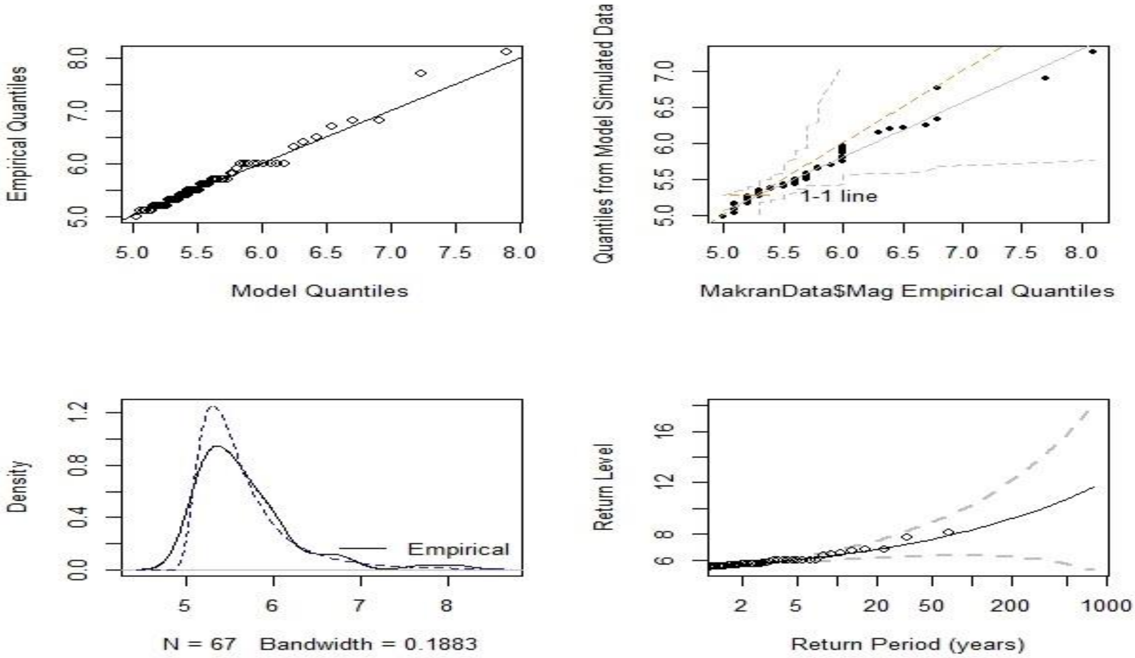

Figure 2. A plot of the GEV distribution Model fitted to earthquake data of the Makran subduction zone.

The parameters of GEV distribution has calculated by utilizing the maximum likelihood approach. Using profile log-likelihood, Confidence intervals estimated can often be more accurate. The profile log-likelihood of the shape parameter (ξ) calculated is 0.29 with a 95% confidence interval (0.05, 0.52), shown in Figure 3.

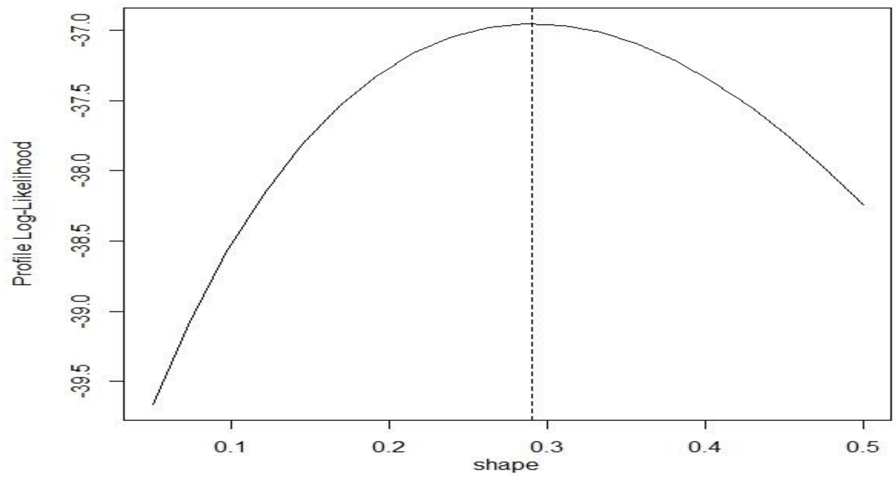

Figure 3. Annual Maximum seismic magnitude profile log-likelihood plot of shape Parameter (ξ) .

In Table 1, the shape parameter value $(\xi = 0.29)$ represents that the Frechet distribution might be the more accurate model in the GEV distribution. The negative log-likelihood value estimated for the GEV model is 36.95. Return periods with return levels and 95% predicted for the maximum magnitude of the earthquake in the Makran subduction zone has shown in Table 2. The return level calculated for the return period 10 is 6.35 with a 95% confidence interval (CI) of 5.98 and 6.71. The return level also increases with the increase in the return period.

Table 2. Return Period calculated of GEV Model.

Return period	10	20	50	100
Return level	6.35	6.81	7.58	8.31
95% CI	(5.98, 6.71)	(6.17, 7.45)	(6.33, 8.83)	(6.34, 10.28)

The return level estimated for return period 20 is 6.81 with a 95% CI of 6.17 and 7.45. While the return level for return periods 50 and 100 are estimated 7.58 and 8.31 with CI of (6.33, 8.33) and

(6.34, 10.28) respectively. Another way to explain is that if the magnitude of an earthquake exceeds 6.35, then there is approximately a 10% chance of earthquake occurrences each year. The profile log-likelihood estimates the return periods 10, 20, 50, and 100 years to get a more precise CI. The profile log-likelihood value calculated for return period 10 is 6.35, for 20 years is 6.8, and for 50 years is 7.58. while for the return period 100, it is 8.31. The result of the profile log-likelihood 95% confidence interval is the same as we previously calculated. The confidence interval for the return periods 10, 20, 50, and 100 are (5.98, 6.71; 6.17, 7.45; 6.33, 8.83; and 6.34, 10.28) respectively. The 1945 earthquake of the makran region is one of the most significant magnitudes recorded earthquakes having a moment magnitude of 8.1 (Mw). It happens once after one century. The 2013 Awaran earthquake having a magnitude of 7.7 (Mw), occurs once after 50 years.

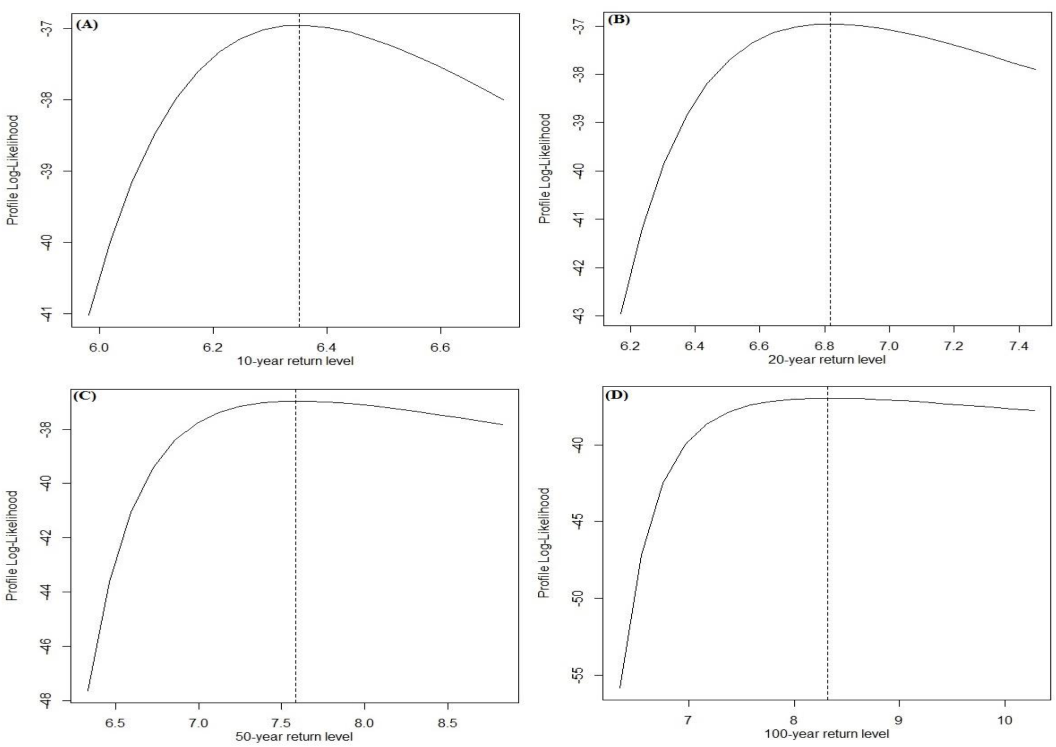

Figure 4. Profile Log likelihood plot for a (a) 10, (b) 20, (c) 50, and (d) 100-year return period.

Conclusion

In the current, we utilized the extreme value approach for estimating the maximum annual earthquake magnitude in the Makran subduction zone using the seismic data from 1934 to 2022. The generalized extreme value distribution is fitted to the seismic data of Makran using the block maxima approach. Numerous plots give the accuracy of the GEV model fitted to seismic data. Results show that the return level of earthquake magnitude increases with the return period. We also analyzed the return level of the different return periods using the maximum likelihood approach. This study shows how extreme value theory might apply as a modeling tool for extreme events. This study will help understand the Makran subduction zone's significant earthquake event. This work aims to inform decision-makers in this region by using statistical analysis about extreme earthquake events so that they may adopt appropriate methods for risk reduction. We recommend adopting the required safety measures to guard against natural catastrophes caused by catastrophic events to safeguard people and property. We also recommend further use of other extreme value approaches, such as Bayesian extreme and extreme quantile, to investigate devastating events in this zone and other regions.

References

Al Abbasi, J. N., Risan, H. K., & Resen, I. A.: Application of Kumaraswamy Extreme Values Distributions to Earthquake Magnitudes in Iraq and Conterminous Regions. *International Journal of Applied Engineering Research*, *13*(11), 8971-8980 (2018).

Arreyndip, N. A., & Joseph, E.: Extreme temperature forecast in mbonge, cameroon through return level analysis of the Generalized Extreme Value (GEV) distribution. *International Journal of Physical and Mathematical Sciences*, *9*(6), 347-352 (2015).

Beirlant, J., Goegebeur, Y., Segers, J., & Teugels, J. L.: Statistics of extremes: theory and applications. John Wiley & Sons (2006).

Buishand, T. A., de Haan, L., & Zhou, C.: On spatial extremes: with application to a rainfall problem (2008).

Carreau, J., Neppel, L., Arnaud, P., & Cantet, P.: Extreme rainfall analysis at ungauged sites in the South of France: comparison of three approaches. *Journal de la Société Française de Statistique*, *154*(2), 119-138 (2013).

Chikobvu, D., & Sigauke, C.: Modelling influence of temperature on daily peak electricity demand in South Africa. *Journal of Energy in Southern Africa*, *24*(4), 63-70 (2013).

Chu, L. F., McAleer, M., & Wang, S. H.: Statistical Modeling of Recent Changes in Extreme Rainfall in Taiwan (2012).

Coles, S.: An Introduction to Statistical Modeling of Extreme Values. Springer Verlag, Berlin (2001).

Embrechts, P., Klüppelberg, C., & Mikosch, T.: *Modelling extremal events: for insurance and finance* (Vol. 33). Springer Science & Business Media (2013).

Ender, M., & Ma, T.: Extreme value modeling of precipitation in case studies for China. *International Journal of Scientific and Innovative Mathematical Research (IJSIMR)*, *2*(1), 23-36 (2014).

Gilleland, E., Ribatet, M., & Stephenson, A. G.: A software review for extreme value analysis. *Extremes*, *16*, 103-119 (2013).

Haan, L., & Ferreira, A.: Extreme value theory: an introduction (Vol. 3). New York: Springer (2006).

Hasan, H., Radi, N. A., & Kassim, S.: Modeling of extreme temperature using generalized extreme value (GEV) distribution: A case study of Penang. In *Proceedings of the world congress on engineering* (Vol. 1, pp. 181-186) (2012, July).

Jenkinson, A. F.: The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, *81*(348), 158-171 (1955).

Katz, R. W., Parlange, M. B., & Naveau, P.: Statistics of extremes in hydrology. *Advances in water resources*, *25*(8-12), 1287-1304 (2002).

Lavenda, B. H., & Cipollone, E.: Extreme value statistics and thermodynamics of earthquakes: aftershock sequences (2000).

Li, Y., Cai, W., & Campbell, E. P.: Statistical modeling of extreme rainfall in southwest Western Australia. *Journal of climate*, *18*(6), 852-863 (2005).

Li-Ge, C., Jun, Z., Bu-Da, S., Jian-Qing, Z., & Gemmer, M.: Probability distribution and projected trends of daily precipitation in China. *Advances in climate change research*, *4*(3), 153-159 (2013).

Maposa, D., Cochran, J. J., Lesaoana, M., & Sigauke, C.: Estimating high quantiles of extreme flood heights in the lower Limpopo River basin of Mozambique using model based Bayesian approach. *Natural Hazards and Earth System Sciences Discussions*, *2*(8), 5401-5425 (2014).

Mothupi, T., Thupeng, W. M., Mashabe, B., & Mokoto, B.: Estimating extreme quantiles of the maximum surface air temperatures for the Sir Seretse Khama International Airport using the Generalized Extreme Value Distribution. American Journal of Theoretical and Applied Statistics, 5(6), 365-375 (2016).

Parey, S., Hoang, T. T. H., & Dacunha-Castelle, D.: The importance of mean and variance in predicting changes in temperature extremes. *Journal of Geophysical Research: Atmospheres*, *118*(15), 8285-8296 (2013).

Pisarenko, V. F., Sornette, D., & Rodkin, M. V.: Distribution of maximum earthquake magnitudes in future time intervals: application to the seismicity of Japan (1923–2007). *Earth, planets and space*, *62*, 567-578 (2010).

Shahid, S.: Trends in extreme rainfall events of Bangladesh. *Theoretical and applied climatology*, *104*, 489-499 (2011).

Siliverstovs, B., Ötsch, R., Kemfert, C., Jaeger, C. C., Haas, A., & Kremers, H.: Climate change and modelling of extreme temperatures in Switzerland. *Stochastic environmental research and risk assessment*, *24*, 311-326 (2010).

Varathan, N., Perera, K., & Wikramanayake, N.: Statistical modeling of daily extreme rainfall in Colombo (2010).

Wen, X., Fang, G., Qi, H., Zhou, L., & Gao, Y.: Changes of temperature and precipitation extremes in China: past and future. *Theoretical and Applied Climatology*, *126*, 369-383 (2016).

Zimbidis, A. A., Frangos, N. E., & Pantelous, A. A.: Modeling earthquake risk via extreme value theory and pricing the respective catastrophe bonds. *ASTIN Bulletin: The Journal of the IAA*, *37*(1), 163-183 (2007).

Jenkinson, A.F.: Statistic of Extremes. *Technical Note* **98**, World Meteorological Organization. Chapter 5, 183-227 (1969).

Prescott, P., & Walden, A. T.: Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika*, *67*(3), 723-724 (1980).

Prescott, P., & Walden, A. T.: Maximum likelihood estimation of the parameters of the three-parameter generalized extreme-value distribution from censored samples. *Journal of Statistical Computation and Simulation*, *16*(3-4), 241-250 (1983).

Macleod, A. J.: Comment on maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. *Applied Statistics*, *38*, 198-199 (1989).

Maruyama, F.: Analyzing the Annual Maximum Magnitude of Earthquakes in Japan by Extreme Value Theory. Open Journal of Applied Sciences, 10(12), 817 (2020).

Epstein, B., and Lomnitz, C.: A model for occurrence of large earthquakes. Nature, 211, 954–956 (1966).

Kijko, A., & Sellevoll, M. A.: Triple exponential distribution, a modified model for the occurrence of large earthquakes. *Bulletin of the Seismological Society of America*, *71*(6), 2097-2101 (1981).

Pavlenko, V. A.: Estimation of the upper bound of seismic hazard curve by using the generalised extreme value distribution. *Natural Hazards*, *89*(1), 19-33 (2017).

García-Bustos, S., Landín, J., Moreno, R., Chong, A. S. E., Mulas, M., Mite, M., & Cárdenas, N.: Statistical analysis of the largest possible earthquake magnitudes on the Ecuadorian coast for selected return periods. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 14*(1), 56-68 (2020).

Bustos, S. L. G., Navarrete, S., Chancay, A., Mendoza, M., Pincay, M., & Teran, M.: Zoning of Ecuador According to Maximum Magnitudes of Earthquakes and their Frequency of Occurrence using Statistical Models Estimated by Maximum Likelihood. Gazi University Journal of Science, 34(3), 916-935 (2021).