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

Adil Rehman*, Huai Zhang
Key Laboratory of Computational Geodynamics, University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: adil.geologist5@gmail.com

Received: 26 August 2023 / Accepted: 20 November 2023 / Published: 02 December 2023

Abstract: The long-term pattern of severe incidents is one of the most crucial and fascinating topics of seismic events. This work aims to analyze the maximum annual earthquake magnitude in the Makran subduction zone using extreme value theory by implementing the block maxima method. The seismic data utilized for the current study was collected from the International Seismological Center (ISC) ranging from 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 parameter (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. The estimated return levels for different 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. The significance of this research is to inform decision-makers to implement suitable risk-mitigation methods.

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

INTRODUCTION

According to Mokhtari et al. (2019), the Makran subduction zone has a complex tectonic background that lies close to the triple junction where the Eurasian, Arabian, and Indian plates are connecting. The seismicity of Makran subduction compared to other subduction zones in the world is low (Mokhtari et al., 2008). According to Rajendran et al. (2013), the eastern portion of Makran is more seismically active than the western portion. The Makran subduction has a low angle of dip. Numerous earth scientists have studied the seismic behavior and tectonic features of the Makran area and suggest that this region is active (Kopp et al., 2000; Byrne et al., 1992; Wiedicke et al., 2001; Vernant et al., 2004). A recent tsunamogenic earthquake hit the eastern Makran region on November 28, 1945, with a moment magnitude of 8.1(Mw) and caused high fatalities. While the western portion of Makran has not yet experienced any large earthquake, the reason may be that this region is locked.

The motivation behind this study is to find accurate earthquake magnitude for risk assessment to warn the people in the study region early. Extreme value theory focuses on the study of rare incidents. The use of extreme value theory is necessary to study catastrophic events. Fisher & Tippett (1928) and Gnedenko (1943) established the basis for Extreme value theory. Based on experimental and measured records and some principal molds, an extreme value approach is used to assess the probability of the occurrence of extreme values. The extreme value analysis (EVA) approach is applied to handle extreme deviations from the probability distribution’s median. 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. In the block maxima approach, the values are disseminated as GEV distribution. The block maxima technique focuses on the maximum observation during each observation period by splitting the observation time into equal-sized, non-overlapping intervals. Two key distributions are created in EVA, i.e., Generalized Extreme Value (Jenkinson, 1955) and Generalized Pareto distribution (Wals & Kelleher, 1971; Pickands, 1975).
Some of the advantages of using the block maxima approach over others are the only available information may be block maxima (e.g., yearly maxima). In cases when the observations are not perfectly independent and uniformly distributed, the block maxima technique may be preferred, the block maximum strategy could be easier to employ given that many situations naturally occur block periods. Maruyama (2020) employed the extreme value method to analyze Japan’s annual maximum magnitude of earthquakes. Simulating only the block maxima would be inefficient if more extreme values were available. Kjikko & Sellevoll (1981) and Epstein & Lomnitz (1966) utilized extreme value theory to model significant earthquake events. Garcia-Bustos et al. (2018) used an extreme method to calculate the return period for large-magnitude 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. Zimbidis et al. (2007), Pisarenko et al. (2010), and Al Abbasi et al. (2018) utilized the Extreme value method for earthquake prediction in different regions. Lavenda & Cipollone (2000) utilized extreme values for the thermodynamics of the aftershock sequence. 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. 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), and Haan & 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) and Embrechts et al. (1997) utilize the extreme value theory technique for flood, drought, and financial incident forecasting. Mothupi et al. (2012) operated a generalized Pareto distribution for monthly maximum temperature modeling in Shakawe in Botswana. However, we used the GEV model for maximum earthquake magnitude analysis for the current work. Jenkinson (1955) utilized generalized extreme distribution for temperature extreme demonstrating. Various researchers utilized an extreme approach for temperature extremes (Chikobvu & Singake, 2013; Hasan et al., 2012; Parey et al., 2013; Siliverstovs et al., 2008; Wen et al., 2015). Arreydip & 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 & 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 & Ma, 2014). Several authors used the maximum likelihood approach to study shape, location, and scale parameters (Jenkinson 1969; Prescott & 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 the 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.

The main objective of this work is to utilize extreme value theory for forecasting the annual maximum earthquake magnitude and the return periods of 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.

METHOD

The data utilized in the current work for the maximum earthquake magnitude analysis in the Makran subduction zone range between 1934 and 2022 (67 years) has a magnitude ≥5.0. The data used in this study was collected from the International Seismological Center (ISC), and a homogenized seismic catalog was prepared (Scordilis, 2005). 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 disagreement among researchers over the best approach. The block maxima (BM) approach is more
effective than the peak-over-threshold method under typical circumstances (Ferreira & de Haan, 2015). 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.](image)

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 strategy focuses on the statistical patterns of

\[ M_n = \max(X_1, X_2, \ldots, X_n) \]  

where \( X_1, X_2, \ldots, X_n \) are random independent variables with uniform distribution \( F \). While the \( M_n \) indicate the maximum of the process \( n \) time unit over 67 years of observation. In theory, the distribution of \( M_n \) can be derived exactly for all values of \( n \).

\[ Pr(M_n \leq z) = Pr(X_1 \leq z, \ldots, X_n \leq z) = Pr(X_1 \leq z) \cdot K \cdot Pr[X_n \leq z] = (F(z))^n \]  

The GEV distribution's cumulative distribution function is given as

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

where \( Z \) is the block extreme value, \( \mu \) is location, \( \xi \) is the shape, and \( \sigma \) is scale parameters. The shape parameter determines the tail decay rate. The probability distribution \( G(z) \) are assessed through block maxima approach. Based on the shape parameter, \( (\xi) \) 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 with

\[ Z_p = \begin{cases} 
\mu - \frac{\sigma}{\xi} \left[ 1 - (\log(1 - p))^{-\xi} \right], & \text{For } \xi \neq 0 \\
\mu - \sigma \log(-\log(1 - p)), & \text{For } \xi = 0 
\end{cases} \]
where $Z_p$ is the return level linked with return period $1/p$. While the probability that an earthquake of magnitude $Z_p$ would exceed once each year has been given by $p$.

**RESULTS & DISCUSSION**

Statistical results from the analysis of the GEV model are 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.**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>5.38</td>
<td>0.04</td>
<td>(5.30, 5.47)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.30</td>
<td>0.03</td>
<td>(0.23, 0.37)</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.29</td>
<td>0.11</td>
<td>(0.05, 0.52)</td>
</tr>
</tbody>
</table>

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

The parameters of GEV distribution have been 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 ($\xi$) calculated is 0.29 with a 95% confidence interval (0.05, 0.52), shown in Figure 3.
Return periods with return levels, uncertainty, and 95% CI predicted for the maximum magnitude of the earthquake in the Makran subduction zone are shown in Table 2. The return level calculated for the return period 10 is 6.35 with a standard error of 0.1862 and a 95% confidence interval (CI) of 5.98 and 6.71. The return level also increases with the increase in the return period (Figure 4).

Table 2. Return Period calculated of GEV Model.

<table>
<thead>
<tr>
<th>Return period</th>
<th>10</th>
<th>20</th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return level</td>
<td>6.35</td>
<td>6.81</td>
<td>7.58</td>
<td>8.31</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.1862</td>
<td>0.326</td>
<td>0.637</td>
<td>1.00</td>
</tr>
<tr>
<td>95% CI</td>
<td>(5.98, 6.71)</td>
<td>(6.17, 7.45)</td>
<td>(6.33, 8.83)</td>
<td>(6.34, 10.28)</td>
</tr>
</tbody>
</table>

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

Figure 4. Profile Log likelihood plot for a (a) 10, (b) 20, (c) 50, and (d) 100-year return period.
The return level estimated for return period 20 is 6.81 with uncertainty of 0.326 and 95% CI of 6.17 and 7.45. The return levels for return periods 50 and 100 are estimated at 7.58 and 8.31 with standard error and CI of (0.637, 6.33, 8.33) and (1.00; 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, it 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 recorded earthquakes, with a moment magnitude of 8.1 (Mw). It happens once after one century. The 2013 Awaran earthquake, with a magnitude of 7.7 (Mw), occurred once after 50 years.

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 implementing early warning systems and improving infrastructure in the Makran subduction zone to mitigate the impact of extreme seismic events. We recommend adopting the required safety measures to guard against natural catastrophes caused by catastrophic events to safeguard people and property. As discussed earlier, the eastern portion is more active than the western portion of Makran. The reason may be that the western portion is locked, and stress accumulates in this zone and may have the possibility of tsunamigenic earthquake occurrences. Therefore, we suggest using several other approaches, such as seismicity rate variation, Bayesian extreme and extreme quantile methods, etc., to predict devastating tsunamogenic events for the Makran subduction zone.

ACKNOWLEDGMENT

This research is supported by the National Science Foundation of China (U2239205, 41725017) and the National Key R&D Program of the Ministry of Science and Technology of China (2020YFA0713401). This research also is supported by the National Key Scientific and Technological Infrastructure project "Earth System Science Numerical Simulator Facility" (EarthLab).

Declaration of Competing Interest

The author declares no potential competing interests.

REFERENCES


Mothupi, T., Thupeng, W. M., Mashabe, B., & Mokoto, B. (2016). Estimating extreme quantiles of the maximum surface air temperatures for the Sir Seoretse Khama International Airport using the Generalized Extreme Value
Estimation of the upper bound of seismic hazard curve by using the generalised extreme value distribution. *Natural Hazards, 89*(1), 19-33. https://doi.org/10.1007/s11069-017-2950-z


Scordilis, E. M. (2005). Globally valid relations converting Ms, mb and MJMA to Mw. In *Meeting on earthquake monitoring and seismic hazard mitigation in Balkan Countries, NATO ARW, Borovetz, Bulgaria* (pp. 11-17).


