Estimation of the Probability of Earthquakes Return Period in Zimbabwe Using Gumbel's Extreme Value Theory Method

Abstract

Using Gumbel's extreme value theory method, this study looked into the probability of the largest earthquakes for different return periods in Zimbabwe. The Advanced National Seismic System (ANSS), the Northern California Earthquake Data Centre website, and UC Berkeley in the United States provided the data used for this study. The natural earthquakes with Mb ≥ 4.0 that occurred in the study area between January 1, 1901and December 31, 2001 (a period of 100 years) with a focal depth ranging from 0 to 700 km made up the selected data. The study area is between 15^0S - 22^0S and 25^0E -34^0E in coordinates. A total of 81 events were used in the investigation. According to the results, there is a 100% chance that earthquakes with a magnitude between 4.5 and 5.5 will occur during the time interval ofapproximately t 1 to 14 years and a 9.5%chance that within the next 100 years, there will be an earthquake with a magnitude of M_b = 6.0 or higher. For magnitudes Mb = 6.0 and above, the return period is relatively long - roughly 50–701 years. This suggests that there is little chance that Zimbabwe will experience earthquakes larger than magnitude 6.0. Despite being low, it is impossible to forecast with certainty because earthquake forecasting and prediction is still a complex topic.

Keywords:largest earthquakes, Zimbabwe, return period, Gumbel's extreme value theory.

1.Introduction

One technique used to assess the seismicity of a particular area is the seismic risk analysis utilizing past earthquake data. According to the elastic rebound theory, all earthquakes, regardless of fault or fault segment, are related to previous earthquakes. The seismic data that was collected during the earthquakes can be used to determine the frequency and timing of potential future earthquakes.

Based on statistical analysis of uncommon events, extreme value theory comprises the statistical laws of extreme values of a random variable. This method requires only a few straightforward computations and can accurately characterize the data's tail features. An invaluable resource for the study of natural disasters is extreme value theory.

This method requires only a few straightforward computations and can accurately characterize the data's tail features. An invaluable resource for the study of natural disasters is extreme value theory. Numerous disciplines, including hydrology, meteorology, and earth science, use this important model, which matches the tail distribution of catastrophe risk data [1]; [2]; [3]. “The extreme value theorem was developed by Fréchet, Fisher, and Tippett, who were the first to employ a statistical model to explain the behavior of a random variable's maximum and lowest values. This theorem demonstrated that the maximum or minimum value fits a three-parameter”[4].

Furthermore, [4] provided evidence for “the extreme value theorem, which was frequently used in the application domain. The probability distribution of extreme occurrences above a threshold is determined using the peaks-over-threshold (POT)method”. Within the seismic domain, [5] and [6] discovered that the POT approach has been utilized for tail parameter estimate and extreme value modeling.

The extreme values generated from model fitting can be used to anticipate future earthquake catastrophe events, and the related return period or high quantile can be used to estimate the risk of unusual events. Despite the fact that extreme value theory has evolved much since it first gained attention, certain debates persist. One of these claims is that the right tail goes to infinity when the extreme value theory model's form parameter is $\xi \geq 0$, which leads to irrationality in a number of application cases. For instance, the magnitude of an earthquake cannot tend to infinity.

"When compared to other statistical models, researchers have discovered that extreme value theory offers useful properties in explaining the characteristics of the right tail of earthquake magnitude data" [7]and [8]. In recent years, there has been a lot of focus on modeling extreme events. It is important to estimate the quantity of the extreme events or the corresponding return period in order to estimate the risk of rare occasions [9] and [10].

[11] computed earthquake probabilities and return periods for all the zone regions in western Anatolia by comparing four sophisticated statistical distributions namely,exponential (EXP), Log Pearson Type 3 (LP3), Generalized Pareto (GP), and Extreme Value Distribution Type 1 (Gumbel) (GUM). The study's conclusions demonstrated that, for the peaks over threshold earthquake series and yearly maximum earthquake data in western Anatolia,the GP and GUM distributions were best suited for explanation.

To ascertain the magnitude and frequency relationship for the Denizli region, [12] examined the GR model in comparison to the Exponential, Gumbel, and Poisson models. Their research showed that compared to the other models, the Poisson model produced return periods with larger values. [13] used the Gumbel III asymptotic distribution based on the GR model to estimate a few basic seismic parameters for various Turkish regions. [14] used the Gumbel's I asymptotic distribution to calculate the probabilities for various time periods at given magnitudes, the most likely maximum magnitudes, and the average return periods (in years) in order to estimate the seismicity of Turkey’s seismic regions.

[15] employedGumbel's extreme value distribution method to calculate the b-value of an earthquake in Bangladesh's Sylhet seismic region. The study's conclusions showed that there is an 85% chance of an earthquake of magnitude Mw≥6.5 occurring again in 100 years, with a 53-year return period, and a 48.5% chance of an earthquake of magnitude Mw≥7.0 occurring again in 151 years.

[16] used the Kumaraswamy value method to assess the magnitude of earthquakes that occurred in Iraq and the surrounding regions between 1944 and 2015. The study's conclusions demonstrated that, particularly for larger earthquakes, the Kumaraswamy-reflected Weibull model is quite flexible when it comes to data analysis and fits the observations better than either the Gumbel or the Reflected Weibull distribution.

[17]conducted“a study on probabilistic recurrence of earthquakes in Northeast India. Wei-bull, Gamm, and Lognormal models were employed along with the earthquake catalog spanning the period from 1846 to 1995. The resulted cumulative probability for a large earthquake of magnitude 7.0 could be reached to 0.8 after about 15–16yearsand0.9afterabout18–20yearsfromthetimeofthelastearthquake(1995)inthestudyarea”.

[18]estimated“recurrence periods of earthquakes in Zimbabwe using exponential distribution model. The return periods for magnitude 4.2 and 6.2 were estimated to be4.00 and 47.48 years respectively”.

The southernmost point of the East African Rift System is where Zimbabwe is located [19]. Three seismic zones make up Zimbabwe's seismic activity: the eastern region, the central region, and the Zambezi basin region. The Deka fault zone in the northwest, the Lake Kariba region in the Mid-Zambezi basin, and the border with Mozambique are the usual locations for earthquake activity.

This paper attempts to use Gumbel's Annual Extreme Values Method to estimate the probability of return period of earthquakes in Zimbabwe due to the increase in industrial development, urbanization, number of high rise buildings, construction of dams, and mining activities.

1.1 Seismicity of the study area

Zimbabwe's seismic activity is generally considered to be moderate, with a few notable occurrences around the country's northern region, which includes the Zambezi area, Nyamandlovu area, and

Zimbabwe-Mozambique border situated at the southernmost point of the East African Rift System [20], [21], [22], [19]. Three seismic zones make up Zimbabwe's territory: the eastern region, the central region, and the Zambezi basin. Its borders with Mozambique to the east, the Deka fault zone in the Hwange area, to the northwest, and over the Mid-Zambezi basin in the Lake Kariba area are the main locations for seismic activity [23].

The Mid-Zambezi basin has experienced seismic events with magnitudes greater than 5.0; these events exhibit normal faulting [24], [25]. Research on early rifting in the Mid-Zambezi Basin confirms that tectonic activity there is identical to that northward along the East African rift system [20], [21]. The Western flank of the rift extension from Lake Malawi is formed by Zimbabwe's border area in the southeast [26].

2.Data source

The Advanced National Seismic System (ANSS), a website maintained by the Northern California Earthquake Data Center in the United States, provided the data set used for this study. The natural earthquakes with Mb≥4.0 that occurred in the study region between January 1, 1901 and December 31, 2001 (a period of 100 years) with focal depths ranging from 0 to 700 km made up the selected data. The data set included the earthquake's date of occurrence, its origin time, its epicentre coordinates, its magnitude, its event identification, its focal depth, and its event type E. The study area is located between 15^0S and 22^0S and 25^0E and 34^0E in terms of coordinates (Fig.1).

For the study, a total of 81 events were used.

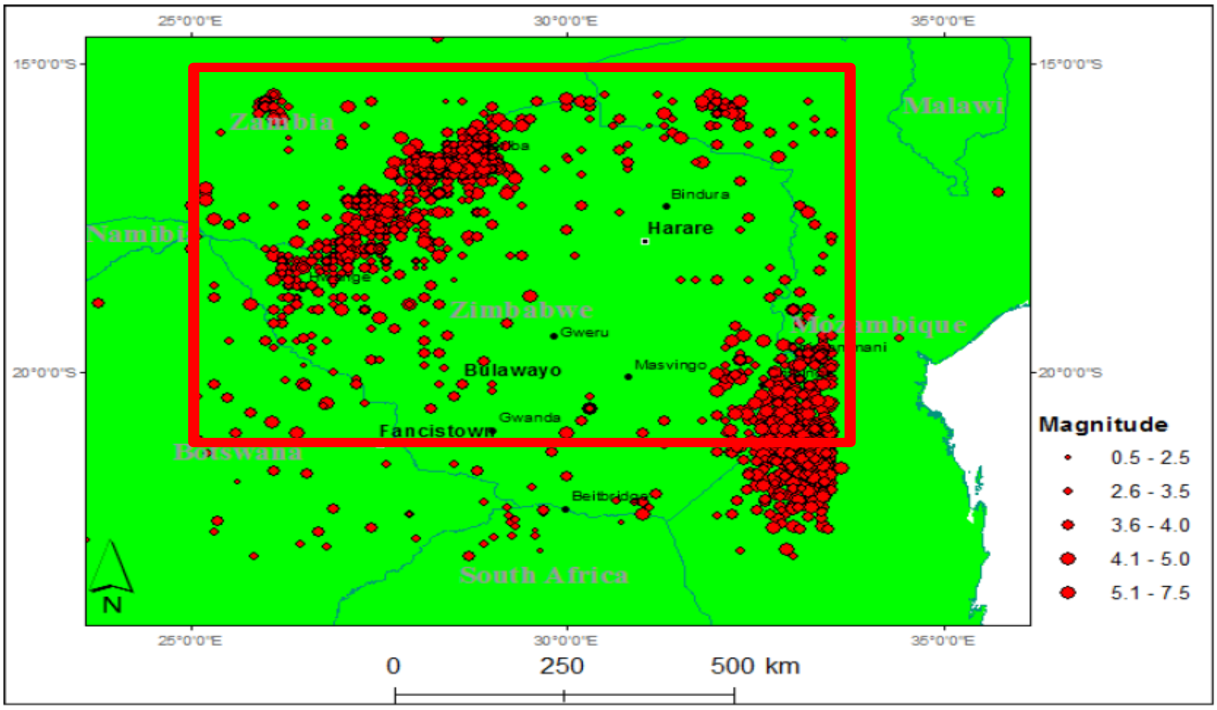

Fig.1: Map of Zimbabwe with rectangle showing the study area(Modified from [23]

3. Method of data analysis

3.1 Gumbel's Model

This model was created by Gumbel in 1935 and has a number of uses to provide answers to both practical and scientific issues. The foundation of this model is the presumption that current common conditions will persist into the future. The idea that observed values are independent of one another is also the second assumption.

If we consider an earthquake's magnitude (x) to be a random variable, the cumulative distribution function (CDF) is provided by [27]:

$$F(x) = 1 - e^{-x} \qquad x \geq 0 \qquad (1)$$

Using the theory of probabilities and Gumbel, the CDF of G(Mj) of maximum yearly magnitude gives rise to:

$$G(Mj) = \exp(- e^{Mj}) \qquad M_j \geq 0 \qquad (2)$$

[28]also added to the first assumption of Gumbel which states that the number of earthquakes in a year is a variable of Poisson distribution with an average of α the average number of earthquakes per year above magnitude 0.0then eqn(1) changes to:

$$F(x) = 1 - e^{-\beta x} \qquad x \geq 0 \qquad (3)$$

Where β isthe inverse of the average magnitude of earthquakes

The CDF of G(Mj) of the maximum yearly magnitude then becomes

$$G(Mj) = \exp(-\alpha e^{-\beta Mj}) \qquad M_j \geq 0 \qquad (4)$$

Where G(Mj) is the probability that an earthquake magnitude within a year to have a value $M_j$or lower. Eqn(4)now becomes

$$Ln[-Ln\ G(Mj)] = Ln\alpha - \beta M \qquad (5)$$

Eqn(5) corresponds to the Gutenberg – Richter relation and the parameters a and b of the Gutenberg – Richter have relationship with α and β.

$$a = \frac{Ln\ \alpha}{Ln 10}, \qquad b = \frac{\beta}{Ln 10}, \qquad (6)$$

substitutingeqn(6) into eqn(5) yields

$$Log[-Ln\ G(Mj)] = a - bM_j \qquad (7)$$

Eqn(7) is the mathematical expression of first – type Gumbel distribution.

Generally, the Gumbel's model Extreme Value Method is given by:

$$G(M) = e^{-\alpha e - \beta M} \qquad (8)$$

Where M is the magnitude of earthquake, α, β are the regression coefficients. G(M) is the probability of not exceeding the earthquakes having magnitudes more than M in one year, α is the average number of earthquakes per year above magnitude 0.0 and β is the inverse of the average magnitude of earthquakes in the region under investigation.

3.2 Frequency- Magnitude relationship

Gutenberg-Richter proposed the following relationship which relates the earthquake magnitude to the total number of earthquakes in one year [29].

LogN=a–bM (9)

For a given region and time interval, eqn (9) gives the cumulative number of earthquakes (N) with magnitude (M), where **a**and**b** are positive, real constants. The parameter **a** represents the seismic activity. It is determined by the event rate and for a given region depends on the volume and time window used. The **b** parameter is a tectonic parameter that describes the properties of the seismic medium.

The correlations between Gumbel and Richter formulations are as given by [29].

$$\alpha = 10^{a} \quad , \quad a = Log\alpha \tag{10}$$

$$\beta = \frac{b}{Loge} = \frac{b}{0.4343}, \; b = \beta * 0.4343 \tag{11}$$

$$N = \alpha e^{-\beta M} = -LnG \tag{12}$$

Relative frequency $f = \frac{j}{(n+1)}$ (13)

Where j is the number of earthquake and $n = \sum j$.

3.3 Calculation of annual risk (Probability)for different return periods

The earthquake number greater than a certain magnitude M for one year N(M), return period (Tr), risk for one year (R_1) and D^{th} year have been calculated by employing the following formulae[29].

$$N(M) = \alpha \exp^{-\beta M} \tag{14}$$

$$Tr = \frac{1}{N(M)} \tag{15}$$

$$R_1(M) = 1 - e^{-N(M)} \tag{16}$$

$$R_D(M) = 1 - e^{-DN(M)} \tag{17}$$

4.0 Results and discussion

The results of the study are as shown (Fig.3 – Fig.4 and Table 1 – Table 2)

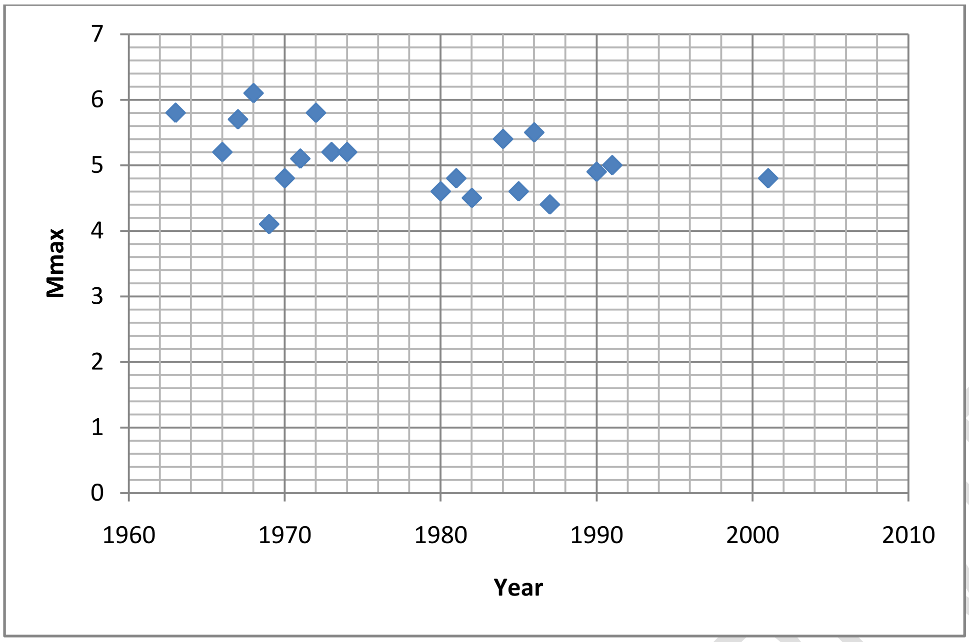

Fig.2: Variation of earthquakes of maximum magnitude during time interval (1901-2001)

The variation of maximum magnitude during time interval (1901-2001) shows that earthquake with magnitude 6.1 occurred in the study area (Fig.2).

Table 1 Calculations of Gumbel's annual maximum distribution

Body wave magnitude Mb	Number of earthquakes J	Relative frequency f	Probability of earthquake G(M)	Cumulative number of earthquakes N = -LnG	LogN
4.0	1	0.0122	0.0122	4.4067	0.6441
4.1	3	0.0366	0.0488	3.0204	0.4801
4.2	6	0.0732	0.1220	2.1041	0.3231
4.3	4	0.0488	0.1707	1.7677	0.2474
4.4	11	0.1341	0.3049	1.1878	0.0748
4.5	7	0.0854	0.3902	0.9410	-0.0264
4.6	12	0.1463	0.5366	0.6225	-0.2058
4.7	7	0.0854	0.6220	0.4749	-0.3234
4.8	5	0.0610	0.6829	0.3814	-0.4187
4.9	3	0.0366	0.7195	0.3292	-0.4826
5.0	5	0.0610	0.7805	0.2478	-0.6058
5.1	4	0.0488	0.8293	0.1872	-0.7277
5.2	3	0.0366	0.8659	0.1440	-0.8415
5.4	1	0.0122	0.8780	0.1301	-0.8859
5.5	4	0.0488	0.9268	0.0760	-1.1193
5.7	1	0.0122	0.9390	0.0629	-1.2013
5.8	3	0.0366	0.9756	0.0247	-1.6074

6.1	1	0.0122	0.9878	0.0123	-1.9112
Total number of earthquakes	81				

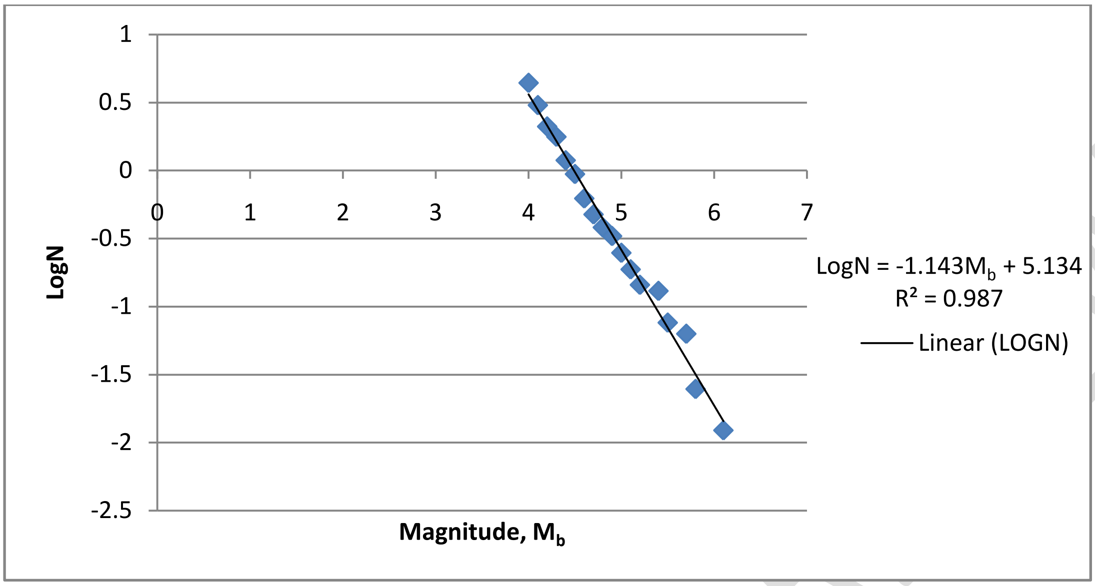

Fig.3:Relationship between Magnitude andLogN

Table 2: Probabilities of earthquake occurrence

Body wave magnitude (M_b)	**Cumulative number of earthquakes N(M)**	**Risk for one year (R_1)**	**Cumulative number of earthquakes for fifty years(50N)**	**Risk for fifty years (R50)**	**Cumulative number of earthquakes for one hundred years (100N)**	**Risk for one hundred years (R100)**	**Return period of earthquakes (Tr)**
4.5	1.00946	0.635580	50.473	1.000	100.946	1.000	0.99
5.0	0.27172	0.237917	13.585	1.000	27.17	1.000	3.68
5.5	0.07313	0.070520	3.6565	0.974	7.313	0.999	13.67
6.0	0.01968	0.019488	0.9840	0.632	1.968	0.865	50.81
6.5	0.005298	0.005284	0.2649	0.221	0.5298	0.393	188.75
7.0	0.001426	0.001420	0.0713	0.049	0.1426	0.095	701.26

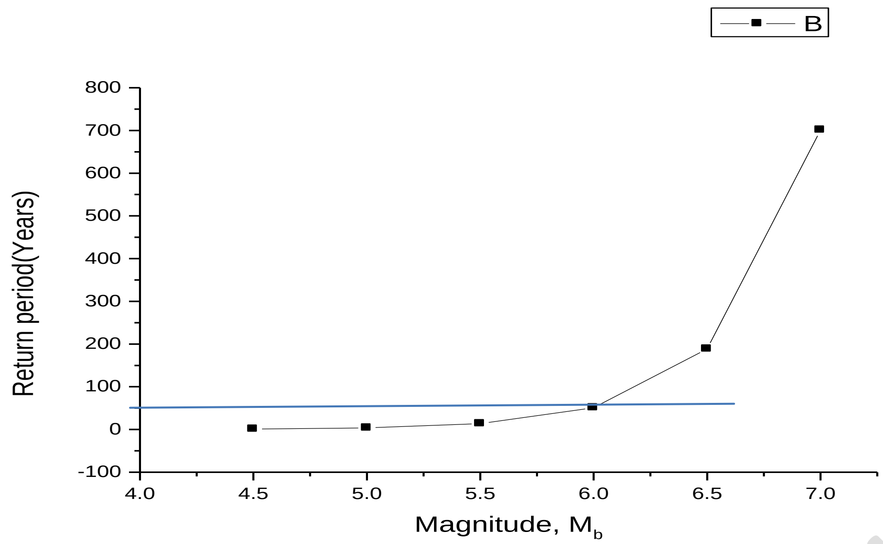

Fig.4: Magnitude distribution with respect to return period.

It was assumed that the earthquakes with magnitude 4.0 and above were used in the analysis because Gumbel'sannual extreme value method works well with only larger past earthquakes. The regression constants **a** and**b** were calculated to be **a**=5.134 and **b**= 1.14 using Least Square (LS) approximation method.Gumbel's regression coefficients α, β were calculated using the Gutenberg-Richter's relationship.The values were found to be α = 136and β=2.62eqn (3) and eqn (4).N, f and LogN values were calculated using equations (5) and (6) and the results are shown (Table 1).

The probabilities of earthquake occurrence for the seismic source were calculated for periods of T=1, 50 and 100 years and magnitude M_b = 4.5, 5.0, 5.5, 6.0, 6.5 and 7.0 using eqns (7),(8),(9) and (10) and the results are as shown (Table 2).As a result, there is a very high probability (100%) of occurrence of earthquakes of magnitude in the range 4.5 to 5.5 in the time interval of 0.99 -13.67years and 9.5% probability that an earthquake 6.0 and above will occur in the next 100years with return period of about 50.81-701years (Table 2).The results obtained are in consonance with similar studies [18] [17] and [15].This implies that Zimbabwe may not likely experience any serious earthquake of magnitude 6.0 and above till the years between 2052 - 2702 sincenow earthquake of magnitude 6.0 and above last occurred in 2001 with the probable return period of approximately 51 years and the estimated return period for 7.0 magnitude is approximately 701 years which is far greater than the seismic history taken into consideration. Fig.4 shows the trend of return period of different earthquake magnitude. It is observed that as the magnitude increases towards higher return period.

Conclusion

This study has indicated that there is a very high likelihood that earthquakes with magnitudes between 4.5 and 5.5 will occur, and the likelihood of earthquakes with magnitudes of 6.0 and higher will decrease. This suggests that Zimbabwe is unlikely to have another significant earthquake with a magnitude of 6.0 or higher until the years 2052–2702. Nevertheless, the concept of earthquake forecasting and prediction remains a complex one due to saturation of earthquake magnitudes and differences in seismic data collection by various seismic stations and networks, making it impossible to predict the occurrence of earthquakes with absolute certainty.

These results are important because they provide seismologists with knowledge about regions that have both short and long recurrence periods. They can use this to aid in their future planning and preparation.

References

[1] Tabari H. Extreme value analysis dilemma for climate change impact assessment on global flood and extreme precipitation. J. Hydrol. 2020; 593: 125932.

[2] Pendergrass D, Shen L, Jacob D, Mickley, L. Predicting the impact of climate change on severe winter time particulate pollution events in Beijing using extreme value theory. Geophys.Res. Lett., (2019); 46: 1824–1830.

[3] Caissie, D., Ashkar, F, El-Jabi, N. Analysis of air/river maximum daily temperature characteristics using the peaks over threshold approach. Ecohydrology, (2020); 13: e2176.

[4] Ma N, BaiY, Meng S. Return period evaluation of the largest possible earthquake magnitudes in Mainland China based on extreme value theory. Sensors, 2021; 21: 3519.

[5] Coles S, Bawa J, Trenner L, DorazioP. An introduction to statistical modelling of extreme values. Springer: Berlin/Heidelberg, Germany, 2001; 208.

[6] Embrechts P, Klüppelberg C, Mikosch. Modelling extremal events: for insurance andFinance; Springer Science & Business Media: Berlin, Germany, 2013;33.

[7] De Haan L, Ferreira A. Extreme value theory: An introduction. Springer Science & Business Media: Berlin, Germany, 2007.

[8] Pavlenko V. Estimation of the upper bound of seismic hazard curve by using the generalised extreme value distribution. Nat. Hazards, (2017; 89:19–33.

[9] Tuyls, DM, ThorndahlS, Rasmussen MR. Return period assessment of urban pluviall floods through modelling of rainfall–flood response. J. Hydroinform,2018; 20, 829–845.

[10] García-Bustos S, LandínJ, Moreno R, Chong A, Mulas M, Mite M, Cárdenas N. Statistical analysis of the largest possible earthquake magnitudes on the Ecuadorian coast for selected return periods. Georisk Assess. Manag. Risk Eng. Syst. Geohazards, 2020;14, 56–68.

[11] Firuzan E.Statistical earthquake frequency analysis for Western Anatolia.Turkish Journal of Earth Sciences, 2008; 17, 741-62.

[12] Cobanoğlu I., Bozdağ Ş., Dinçer I., Erol H.Statistical approaches to estimating the recurrence of earthquakes in the Eastern Mediterranean Region. Istanbul University Earth Sciences Review, 2006; 19: 91-100.

[13] BayrakY, Yılmaztürk A, Öztürk S. Relationships between fundamental seismic hazard parameters for the different source regions in Turkey", Natural Hazards, 2005;36, 445–62.

[14] Öztürk S., BayrakY., CınarH., Koravos GC, TsapanosTM. A quantitative appraisal of earthquake hazard parameters computed from Gumbel I method for different regions in and around Turkey. Natural hazards,2008; 47(3): 471-95.

[15] Ray SAlam, JBM, HaqueM, DasJK, Tanmoy BB, HasanMN. A study on b-value and investigation of seismic hazard in Sylhet seismic region, Bangladesh using Gumbel's extreme distribution method. SN Applied Sciences,2019; 1:435.

[16] Al AbbasiJN, RisanHK, Resan IA.Application of Kumaraswamy extreme value distributions to earthquake magnitudes in Iiraq and Conterminous regions. International Journal of Applied Engineering Research, 2018;13(11), 8971 – 8980.

[17]Yadav, R.B.S.; Tripathi, J.N.; Rastogi, B.K.; Das, M.C.; Chopra, S. Probabilistic Assessment of Earthquake Recurrence in Northeast India and Adjoining Regions. Pure Appl. Geophys. PAGEOPH 2010, 167, 1331–1342.

[18]Abong AA; Atsu JU; Anari H.A; Ushie, JA. The Use of Exponential Distribution Model to Estimate Recurrence Periods of Earthquakes in Zimbabwe. Journal of Geography, Environment and Earth Science International,12(4): 1-8, 2017[19] HlatywayoDJ. Seismic estimates in central Southern Africa.Geophysics Journal International,(1997; 130 (3):737-745.

[20] Vail JR. The southern extension of the East Africa rift system and related igneous activity. Geological Rands,1967; 57(1), 601-604.

[21] FairheadDJ, Girdler RW. Evolution of rifting in Africa Nature, 1969;1178-1182.

[22] FairheadDJ, HendersonNB.The seismicity of South Africa and incipient rifting.1977; 41:T19-T26.

[23] Mapuranga VP. Probabilistic Seismic Hazard Analysis for Zimbabwe.2014; M.Sc.Thesis University of Pretoria, South Africa.

[24] ShudofskyGN. Source Mechanisms and focal depths of East African earthquakes using Rayleigh-wave inversion and body-wave modeling. Geophysics Research Journal,1985; 83(1), 563-614.

[25] Hlatywayo D.J (1995). Fault-plane solutions of the Deka fault zone and mid-zambezi basin. Geophysics Journal International, 129(3), 567-576.

[26] Hlatywayo DJ..Seismicity of Zimbabwe during the period 1959-1990.1996; Institute of Geophysics, Seismological Department, University of Uppsala, Sweden.

[27] BayrakY, Ozturk S, KoravosGC, LeventakisGA, Tsapanos TM Seismicity assessment for different regions in and around Turkey on instrumental data:Gumbel first asymptotic distribution and Gutenberg – Richter cumulative frequency law. Journal of Geophysical Research,2008; 8: 109 -122.

[28] Gutenberg B, Richter CF EarthquakeMagnitude, intensity and Acceleration. Bulletin Seismological Society of America, 1956; 32(3).

[29] Tezcan SS. Probability analysis of earthquake Magnitude, Turkish Earthquake Foundation,1996; 26 P.