Seismic risk in the east of the Bayan Har block based on the POT model

YanFang Zhang\textsuperscript{a,b}, YiBin Zhao\textsuperscript{a} and QingQing Ren\textsuperscript{a}

\textsuperscript{a}Department of Mathematics, Institute of Disaster Prevention, SanHe, HeBei; \textsuperscript{b}China Seismological Bureau, Key Laboratory of Building Failure Mechanism and Defense, SanHe, Hebei

ABSTRACT
This article elaborates on a peaks over threshold (POT) model with earthquake samples based on the Pareto distribution. We analyze the earthquakes using the POT model in the eastern Bayan Har block. The result shows that the area experiences an earthquake above $M_s 7.5$ every 30 years and a large earthquake above $M_s 7.8$ every 100 years. The average magnitude of earthquakes above $M_s 5.2$ that occur in the area every 5 years is $M_s 7.1$ or greater. The upper limit of the theoretical magnitude predicted by the model is $M_s 8.7$. Therefore, the risk of geological hazards is high in this area.

ARTICLE HISTORY
Received 31 May 2022
Accepted 22 September 2022

KEYWORDS
Pareto distribution; POT model; return period; return level; earthquake prediction

1. Introduction
The extreme value theory is an effective method for early warning and risk assessment of extreme events with low frequency but great harm. The generalized extreme value distribution (GEV) (Jenkinson 1955) was first introduced into earthquake early warning by Nordquist (1945). Other scholars (Epstein and Lomnitz 1966; Chen and Lin 1973; Tuncel et al. 1974) systematically studied the maximum magnitude prediction methodology.

Coles et al. (1999) showed the dependence measures for extreme value analyses. Because only the maximum value of the group is used for the analysis, the data information is lost. This affects the effectiveness of the prediction. People pay more attention to large earthquakes that will cause heavy losses but do not pay much attention to small earthquakes with little destructive power. Therefore, the Pareto distribution is reasonable for early earthquake warning analysis.

The scholars (Balkema and Haan 1974; Pickands 1975) proved that the limit distribution of excess of a given threshold is the generalized Pareto distribution (GPD). It is also called the POT model (peaks over threshold) and is widely used in risk assessment (Castillo 1988; Coles and Walsh 1994; Coles and Tawn 1996; Embrechts and Mikosch 2003; Hussain 2021; Wen et al. 2021). The researchers improved the model...
in several aspects (Hosking and Wallis 1987; Davison and Smith 1990; Beirlant and Teugels 1996; Juárez & Schucany 2004). Scholars applied GEV and GPD to seismic risk analysis (Castillo and Hadi 1997; Sornette 2003; Pisarenko et al. 2014; Dutfoy 2019). Pisarenko et al. (2008) proposed a new approach to characterize the tail distribution of earthquake magnitudes.

Qian et al. (2013) studied the seismic risk index based on the Pareto distribution. Dutfoy (2021) built a seismic prediction model based on the generalized Pareto distribution with different observation periods. Barra and Vega-Jorquera (2021) discussed the properties of the q-Pareto distribution and its application in earthquake prediction.

Due to complex geological conditions and frequent earthquakes, positive earthquake warnings and risk assessments are necessary to ensure people’s lives and minimize economic losses. The east of the Bayan Har block is an area with dense fault zones and intense activity. Three earthquakes occurred after 2000 (Wenchuan Ms 8.0 in 2008; Ya’an Ms 7.0 in 2013; ABA Ms 7.0 in 2017) in this area. The intense earthquake attracted the attention of some scholars (Li et al. 2018). Chen et al. (2013) studied the division and deformation characteristics of the east of the Bayan Har block. Wen et al. (2011) discussed the movement and deformation of the north and east boundary fault systems. He proposed that the 2008 Wenchuan earthquake was the latest of the large earthquake sequences. Li et al. (2016) used the fault source model of the seismogenic structure and the time-dependent seismicity model to predict the seismic risk of the Bayan Har block. He identified earthquake-prone areas in the next 50 years.

Due to the effectiveness of the POT model in medium- and long-term earthquake prediction, the earthquake data in the east of Bayan Har block are studied based on the POT model in this article.

2. Extreme value distribution

The theory of extreme value distribution is discussed in detail (Coles 2001; Shi 2006). Extreme value analysis is based on the Fisher–Tippett theorem (Fisher and Tippett 1928) as follows.

Let \( X_1, X_2, \ldots, X_n \) be a series of independent and identically distributed random variables with a distribution function \( F(x) \). Let \( M_n = \max\{X_1, X_2, \ldots, X_n\} \), if there exist constants 

\( \{\alpha_n > 0\}, \{\beta_n\}, \) and non-degenerate function \( \Lambda(x) \), then

\[
\lim_{n \to \infty} P \left( \frac{M_n - \beta_n}{\alpha_n} \leq x \right) = \Lambda(x), x \in R.
\]

Then \( \Lambda(x) \) must be the distribution function of either Gumbel, Fréchet or Weibull. These three distributions can be merged into the GEV distribution. The GEV has a distribution function denoted by:
\[
\text{Gev}(x, \xi, \mu, \sigma) = \exp \left\{ - \left( \frac{x - \mu}{\sigma} \right)^{-\frac{1}{\xi}} \right\}, 1 + \frac{(x-\mu)}{\sigma} > 0, \tag{2}
\]

where \( \xi \) is the shape parameter, \( \mu \) is the location parameter and \( \sigma \) is the scale parameter. The Fréchet case is obtained when \( \xi > 0 \), the Weibull when \( \xi < 0 \), while the Gumbel case is defined when \( \xi = 0 \).

Because we are interested in the tail distribution of the magnitude, the distribution of the excess \( X-u \) is important. Scholars (Balkema and Haan 1974; Pickands 1975; Leadbetter 1984) proved that if the distribution function \( F \) belongs to the attraction domain of a GEV distribution, the limiting distribution of the excesses is the generalized Pareto distribution (GDP). That is, if \( X \) is a random variable that holds Equation (1), then the distribution function of the excess is:

\[
F_u(x) = P\{X-u \leq x|X>u\} = \frac{F(y+u)-F(u)}{1-F(u)} \rightarrow \text{Gp}(x, \xi, \sigma), \tag{3}
\]

where,

\[
\text{Gp}(x, \xi, \sigma) = 1 - \left(1 + \frac{x}{\sigma} \right)^{-\frac{1}{\xi}}, x \geq 0, 1 + \frac{x}{\sigma} \geq 0, \sigma = \sigma + \xi (u-\mu). \tag{4}
\]

It is the distribution function of the two-parameter generalized Pareto distribution. The shape parameter \( \xi \) is the same as that of the GEV distribution. On the basis of the above theory, we build a POT model in the next section.

### 3. The POT model based on Pareto distribution

In this section, estimations of the GPD parameters and seismic risk indexes are given.

#### 3.1. Determination of the threshold

The first set of conditions to construct a POT model is the determination of a reasonable threshold.

For a reasonable threshold \( u \), the mean excess function is

\[
e(u) = E(X-u|X>u) = \frac{\sigma}{1-\xi} + \frac{\xi}{1-\xi} u. \tag{5}
\]

and \( e(u) \) is the linear function of \( u \). Let \( N_u = \sum_{i=1}^{n} I_{[X_i>u]} \) be the number of samples over the threshold. The mean excess function of the samples is

\[
e_n(u) = \frac{1}{N_u} \sum_{i \in \Delta_n(u)} (X_i - u), \Delta_n(u) = \{ i : X_i > u \}.
\]

If the mean excess function of the samples \( e_n(u) \) swings near a straight line, we can select \( u \) as the threshold.
3.2. Estimation of the GPD parameters

When $\xi \neq 0$, the log-likelihood function is

$$\ln L(\xi, \sigma) = -n \ln \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{n} \ln \left(1 + \frac{\xi x_i}{\sigma}\right),$$

where $x_i \in D(\xi, \sigma) = \left\{\begin{array}{ll}
[0, +\infty) & \xi > 0 \\
[0, -\sigma/\xi) & \xi < 0
\end{array}\right.$

The maximum likelihood estimation of the parameters $(\hat{\xi}, \hat{\sigma})$ can be calculated using the numerical method. Estimates of parameters $(\hat{\xi}, \hat{\sigma})$ are $(\hat{\xi}, \hat{\sigma})$. The confidence interval for each parameter can be calculated by the Fisher information matrix:

$$\hat{\xi} \pm Z_{1-\alpha} \sqrt{M_{\text{cov}}(1,1)}, \quad \hat{\sigma} \pm Z_{1-\alpha} \sqrt{M_{\text{cov}}(2,2)},$$

where $Z_{1-\alpha}$ is the quantile of the standard normal distribution and $M_{\text{cov}}$ is the covariance matrix of $\xi$ and $\sigma$.

3.3. Estimation of the distribution function and the quantile of the POT model

We can get the estimation of the distribution function of the excess as follows. From Equation (3),

$$F_u(x) = \frac{F(x+u) - F(u)}{1-F(u)},$$

which leads to

$$\hat{F}_{u}(x) = 1 - F_u(x) = 1 - \frac{F(x+u) - F(u)}{1-F(u)} = \frac{1-F(u) - F(x+u) + F(u)}{1-F(u)} \quad \text{and} \quad \hat{F}(u+x)$$

$$= \hat{F}_u(x) \hat{F}(u).$$

From Equation (4),

$$\hat{F}_{u}(x) = (1 + \frac{\hat{\xi} x}{\hat{\sigma}})^{-\hat{\xi}}.$$

The frequency $N_u/n$ above the threshold $u$ is the estimation of $\hat{F}(u)$, where $N_u$ is the number of samples over the threshold. Then, $\hat{F}(u) = \frac{N_u}{n}$. We can get the estimation of $\hat{F}(u+x)$ as follows:

$$\hat{F}(u+x) = \frac{N_u}{n} \left(1 + \frac{\hat{\xi}}{\hat{\sigma}} \frac{X}{\hat{\sigma}}\right)^{-1/\hat{\xi}}.$$  \hspace{1cm} (8)

Then, we can get the estimation of $F(x)$:

$$\hat{F}(x) = 1 - \frac{N_u}{n} \left(1 + \frac{\hat{\xi}}{\hat{\sigma}} \frac{X-u}{\hat{\sigma}}\right)^{-1/\hat{\xi}}.$$  \hspace{1cm} (9)

$$= 1 - \left[1 + \hat{\xi} \frac{x-u-h}{\hat{\sigma}^2} \left(\frac{N_u}{n} \hat{\xi} - 1\right)\right]^{-1/\hat{\xi}}$$

$$= 1 - \left(1 + \hat{\xi} \frac{x-h}{\hat{\sigma}^2}\right)^{-1/\hat{\xi}}$$

where $\hat{\sigma}' = \hat{\sigma} \left(\frac{N_u}{n}\right) \hat{\xi}$, $h = u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\frac{N_u}{n} \hat{\xi} - 1\right)$.
From Equation (9), we get the estimation of the p-quantile $x_p$:

$$x_p = u + \frac{\hat{\sigma}}{\xi} \left\{ \left[ \frac{n}{N_u} (1 - p) \right]^{-\frac{1}{\xi}} - 1 \right\}, \quad (11)$$

where $F(x_p) = p, 0 < p < 1$.

The confidence level of $x_p$ is calculated using the Delta method (Pisarenko et al. 2008):

$$\hat{x}_p \pm Z_{1-\alpha/2} \sqrt{\text{Var}(\hat{x}_p)}. \quad (12)$$

If $\xi < 0$ and $p \to 1$, the estimation of the upper limit point $x^*$ of the support $F$ is:

$$\hat{x}^* = u - \frac{\hat{\sigma}}{\xi}. \quad (13)$$

### 3.4. Seismic risk indexes and their estimation

Based on the above conclusions, the seismic risk indexes are defined as follows.

#### 3.4.1. Return level and return period

Under the assumption of seismic elastic rebound theory (Wood 1911): large earthquakes occur periodically. Combined with the research results of Qian (2013), the return period and the return level were defined as follows.

The meaning of the return level with the return period $T$ is that the average number of earthquakes over the threshold is 1. For a given magnitude $u$, $E(\sum_{i \leq T} I_{[X_i \geq u]}) = 1$. That is, $u = F^{-1}(1 - 1/T)$ and $u$ is the quantile of $F$. Here, we can denote $u$ by $x_p$.

$$P\{X > x_p\} = 1 - F(x_p) = 1/T, \quad T = 1/F(x_p)$$

is the theoretical return period with the return level $x_p$ and $F(x_p) = p$ $(0 < p < 1)$. If the sample is daily-observed data, there are 365 days per year, and the return period $T$ (unit: year) with the return level of $x_p$ satisfied:

$$F(x_p) = \frac{1}{365T},$$

which leads to $T = \frac{1}{365(1 - p)}$,

where $F(x_p) = p$ $(0 < p < 1)$.

From Equation (8), the return period estimate can be obtained:

$$\hat{T} = \frac{n}{365N_u} \left( 1 + \frac{\hat{x}_p - u}{\hat{\sigma}} \right)^{1/\xi}. \quad (14)$$

If the given return period is $T$ years, bring $p = 1 - 1/(365T)$ into Equation (11), and we can estimate the return level:
3.4.2. Expected return magnitude

Let $x = x(T)$ be the return level, to describe the magnitude exceeding $x(T)$, Qian (2013) proposed the expected return magnitude.

$$\hat{x}_{1-1/365} = u + \frac{\hat{\sigma}}{\xi} \left( \left( \frac{n}{N_u} (1 - p) \right)^{-\tilde{\xi}} - 1 \right) = u + \frac{\hat{\sigma}}{\xi} \left( \left( \frac{365T \cdot N_u}{n} \right)^{\tilde{\xi}} - 1 \right).$$  \hspace{1cm} (15)
\[ \hat{x}(T) = E(X|X>x(T)) = x(T) + E(X-x(T)|X>x(T)). \]

From Equation (5), \( E(X-x(T)|X>x(T)) = \frac{\hat{\sigma}}{1-\xi} + \frac{\hat{\xi}}{1-\xi}[x(T)-u] \). The estimate of the expected return magnitude is as follows:

\[ \hat{x}(T) = \hat{x}(T) + \frac{\hat{\sigma}}{1-\xi} + \frac{\hat{\xi}}{1-\xi}[\hat{x}(T)-u] = \frac{\hat{x}(T)}{1-\xi} + \frac{\hat{\sigma}-\hat{\xi}u}{1-\xi}. \] (16)

4. Seismic risk analysis of the east of Bayan Har block based on the POT model

4.1. Earthquake case data

The data in this article are from the National Seismic Science Data Sharing Center (https://data.earthquake.cn/). Nineteen thousand two hundred twenty-one records before December 2019 are the research samples in the scope of the article (Pisarenko & Sornette 2003). The magnitude of the surface wave \( M_s \) is used to describe the magnitude of the earthquake. If the magnitude of the surface wave is missing, we converted the magnitude of the local earthquake \( M_l \) to \( M_s \) according to the equation \( M_s = 1.13M_l - 1.08 \) (Wang et al. 2010).

The late Pleistocene and Holocene main faults in the east of the Bayan Har block (data provided by the Seismic Active Fault Survey Data Center) and the spatial distribution of the earthquake are shown in Figure 1. This figure can help people...
understand the geological features (Singh et al. 2012). Earthquakes below $M_s$ 4.0 in the east of the Longmenshan fault zone and in the southwest of the Sichuan–Yunnan fault are frequent. Along the fault zones of Yushu, Xianshui River, and Anning River, the earthquake frequency over $M_s$ 7.0 is very high. In 1879 and 2008, earthquakes of $M_s$ 8.0 occurred, respectively, in the south of Wudu, Gansu Province, and Wenchuan. The fault zones in the east of the Bayan Har block are dense. The geological conditions are complex, and the whole fault is in a geological activity period.

Earthquakes above $M_s$ 5.0 in the east of the Bayan Har block have been recorded since 1920 (Huang et al. 1994). All earthquake cases have been recorded since the establishment of the China Seismic Network in 1970. It shows that the earthquakes above $M_s$ 4.0 and $M_s$ 3.0 in the area were completed in 1950 and 1965, respectively, as shown in Figure 2. Because the large earthquakes above the threshold are studied in this article, 19,065 earthquake data after 1960 are selected as samples.

### 4.2. Threshold of the POT model

The threshold $u$ is crucial. If $u$ is too large, there will be few samples, and the estimation variance is significant. If it is too small, the assessment of the Pareto distribution is partial.

If the excess obeys the Pareto distribution, the value of the shape parameter $\xi$ does not change as the threshold increases. For the qualified threshold $u$, the mean excess
function $e(u)$ is a linear function of $u$, and the mean excess function of the samples swings near a straight line with a fixed slope.

The relationships between shape parameter and threshold, as well as mean residual life and threshold, are shown, respectively, in Figures 3 and 4.

When the threshold $u \in [5, 6.4]$, the estimation of the shape parameter $\xi$ remains unchanged, as shown in Figure 3. Empirical residual life fluctuates slightly along a straight line with a fixed slope in Figure 4. It is reasonable to select $u = 5.2$ as the threshold. The number of samples at the point is 96, representing about 0.5% of the total data, which meets the need for tail modelling.

### 4.3. Parameter estimation and the POT model

According to the POT model in Section 3, we calculated the parameters using the maximum likelihood estimation method based on Matlab. Package ‘POT’ in R (Ribatet and Dutang 2022) is convenient for applying the POT methodology. The estimations of the shape parameter $\xi$ and the scale parameter $\sigma$ of the Pareto distribution are shown in Table 1.

The density function curve of the Pareto distribution shown in Figure 5 is consistent with the outline of the histogram. The theoretical quantile is consistent with the empirical quantile from the Q–Q chart shown in Figure 6.

The return level on the logarithmic coordinate axis is convex, as shown in Figure 7. The scatter of order statistics based on samples over the threshold swings near the
return level, which is in the middle of the confidence interval of the return level. So, it is reasonable to describe the statistical law of the excess with the POT model. For comparison, GEV and exponential distribution are used to analyze the same data. We built the GEV model with the annual maximum as the samples. The final results of the parameters obtained by the GEV method can be summarized as follows.

\[ \hat{\xi} = -0.1542, \hat{\mu} = 5.0594, \hat{\sigma} = 0.8805. \]

The upper limit point \( x^* \) of the support \( F \) is \( x^* = \hat{\mu} - \frac{\hat{\sigma}}{\xi} \) for the GEV model. And the theoretical maximum magnitude is about \( M_s = 10.8 \) from the above equation. This result is not in line with reality. Because the annual maxima are the research data in the GEV model, the information may be lost. If two or more magnitudes more significant than the threshold are observed in a given \( T \) interval, the GEV method keeps only the largest one, while the GDP method will use all of them.

Seismicity can be deduced from the research results in the article (Chen and Lin 1973) that the magnitude approximately obeys the exponential distribution with the distribution function \( F(x) = 1 - e^{-1.056x} \). From Equation (12), the approximate CDF of the POT model is:

| Parameter | Shape parameter | Scale parameter |
|-----------|-----------------|-----------------|
| Estimation | -0.261          | 0.914           |
| Confidence interval | [-0.4099, -0.1124] | [0.7165, 1.1666] |

**Figure 5.** Frequency histogram of the magnitude.
Figure 6. Q–Q chart (plot quantile of model versus empirical quantile).

Figure 7. Plot of confidence of the return level.
\[ F_p(x) = 1 - 0.005(1 - \frac{26}{91}(x - 5.2))^{-\frac{1}{\xi}}. \]

A comparison between the two distributions is shown in Figure 8, in which the statistical law of the sample tail in the POT model is more consistent with the sample information. The upper limit of the confidence interval of \( \xi \) in Table 1 is less than 0, indicating that the return level of the magnitude predicted by the model has a theoretical upper limit and the theoretical maximum magnitude is about \( M_s 8.7 \) according to Equation (13). The above analysis shows that the POT model based on Pareto distribution is suitable for seismic risk analysis of the east of Bayan Har block and the prediction conclusion is reasonable.

### 4.4. Seismic risk assessment in the east of Bayan Har block

We assess the risk of an earthquake by the return period and the return level. From Section 4.3, for the given return level \( x_p \), the estimation of the return period \( \hat{T} \) can be obtained by Equation (14). For a given return period, the corresponding return level can be calculated according to Equation (15). We summarize the estimates of the return level for the given return period in Table 2.

The 100-year return level is \( M_s 7.8 \) and the confidence interval is (7.4, 8.3), indicating that earthquakes with a magnitude of about \( M_s 8.0 \) will occur in the east of Bayan Har every 100 years. This result is consistent with the fact that there was an earthquake of Ms 8.0 in 2008. There are four samples that exceed the lower limit of
the confidence interval ($M_s$ 7.4). They are the $M_s$ 7.4 earthquakes in Dari in 1947, $M_s$ 7.5 in Kangding in 1955, $M_s$ 7.6 in Luhuo in 1973 and $M_s$ 8.0 in Wenchuan in 2008. In the prediction results, the return level of the 30-year return period is $M_s$ 7.5. In the last 100 years, earthquakes exceeding $M_s$ 7.5 occurred 3 times in 1947, 1973 and 2008, respectively. The time interval between the three of them is about 30 years. The above conclusions show that the earthquake predicted by the POT model is consistent with the sample information.

The expected return magnitudes of the eastern Bayan Har block within 10 years are higher than the upper limit of the corresponding confidence interval, which indicates that the seismic frequency exceeding the predicted result in this area is high. The average earthquake magnitude exceeding the threshold within 3 years is $M_s$ 7.0 and the average extent within 5 years is not less than $M_s$ 7.1. These two results, combined with the theoretical upper limit of $M_s$ 8.7 predicted above, further show that geological activities in the east of the Bayan Har block are not only frequent but also violent.

### 5. Conclusions

1. The construction of the POT model for seismic risk analysis is described in detail. Compared to the GEV and exponential distributions, the POT model based on the generalized Pareto distributions can objectively describe the seismic risk in the east of the Bayan Har block. Based on parameter estimation, the strong earthquake magnitude distribution of the Bayan Har block is:

$$F_p(x) = 1 - 0.005\left( \frac{26}{91} (x-5.2) \right)^{-\frac{1}{\gamma}}.$$ 

The model can effectively predict an occurrence that exceeds a certain magnitude threshold.

2. The tail distribution of the data can be explained by the POT model.

The predicted results of the strong earthquake distribution, the seismic risk in a given period, the return period, and the return level in the east of the Bayan Har block can provide a valuable reference for formulating emergency management plans. It gives an idea to determine the upper limit of the magnitude of the earthquake.

If the earthquake data are focused on one or more interrelated fault zones, the research will be more targeted, and the results will be more meaningful.
Funding

This work was supported by the Self-Financing Project of Scientific Research and Development Plan of Lang Fang Science and Technology Bureau under grant number 2022011019; the Fundamental Research Funds for Central Universities under grant number ZY20215140; and the Colleges and Universities in Hebei province science and technology research project under grant number Z2020224.

Data availability statement

Some or all data, models, or codes that support the findings of this study are available from the corresponding author YFZ, upon reasonable request.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Balkema AA, Haan LD. 1974. Residual life time at great age. Ann Probab. 2(5):792–804.
Barra E, Vega-Jorquera P. 2021. On q-Pareto distribution: some properties and application to earthquakes. Eur Phys J B. 94(1):1–16.
Beirlant J, Teugels VJL. 1996. Tail index estimation, Pareto quantile plots, and regression diagnostics. Publ Am Stat Assoc. 91(436):1659–1667.
Castillo E. 1988. Extreme value theory in engineering. San Diego: Academic Press.
Castillo E, Hadi AS. 1997. Fitting the generalized Pareto distribution to data. J Am Stat Assoc. 92(440):1609–1620.
Chen CY, Ren JW, Meng GJ, Li PX, Xiong RW, Hu CZ, Shu XN, Su JF. 2013. Division, deformation and tectonic implication of active blocks in the eastern segment of Bayan Har block. Chin J Geophys. 56(12):4125–4141. (in Chinese).
Chen PS, Lin BH. 1973. An application of statistical theory of extreme values to moderate and long interval earthquake prediction. Acta Geophys Sin. 16(1):6–24. (in Chinese).
Coles S. 2001. An introduction to statistical modelling of extreme values. Springer series in statistics. London: Springer-Verlag.
Coles S, Heffernan J, Tawn J. 1999. Dependence measures for extreme value analyses. Extremes. 2(4):339–365.
Coles SG, Tawn JA. 1996. Modelling extremes of the areal rainfall process. J R Stat Soc B Methodol. 58(2):329–347.
Coles SG, Walshaw D. 1994. Directional modeling of extreme wind speeds. J R Stat Soc. 43(1):139–157.
Davison AC, Smith RL. 1990. Models for exceedances over high thresholds. J R Stat Soc B Methodol. 52(3):393–442.
Dutfoy A. 2019. Estimation of tail distribution of the annual maximum earthquake magnitude using extreme value theory. Pure Appl Geophys. 176(2):527–540.
Dutfoy A. 2021. Earthquake recurrence model based on the generalized Pareto distribution for unequal observation periods and imprecise magnitudes. Pure Appl Geophys. 178(5):1549–1561.
Embrechts P, Mikosch T. 2003. Modelling extremal events for insurance and finance. Berlin (Heidelberg): Springer-Verlag.
Epstein B, Lomnitz C. 1966. A model for the occurrence of large earthquakes. Nature. 211(5052):954–956.
Fisher RA, Tippett LHC. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Math Proc Camb Phil Soc. 24(2):180–190.
Hosking JRM, Wallis JR. 1987. Parameter and quantile estimation for the generalized Pareto distribution. Technometrics. 29(3):339–349.

Huang WQ, Li WX, Cao XF. 1994. Research on completeness of earthquake data in the Chinese Mainland(II)—the regional distribution of the beginning years of basically complete earthquake data. Acta Seismol Sin. 7(4):529–538.

Hussain S, Bhatti SH, Ahmad T, Shehzad MA. 2021. Parameter estimation of the Pareto distribution using least squares approaches blended with different rank methods and its applications in modeling natural catastrophes. Nat Hazards. 107(2):1693–1708.

Jenkinson AF. 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Q J R Meteorol Soc. 81(348):158–171.

Juárez SF, Schucany WR. 2004. Robust and efficient estimation for the generalized Pareto distribution. Extremes. 7(3):237–251.

Leadbetter MR, Lindgren G, Rootzén H. 1984. Extremes and related properties of random sequences and processes. New York: Springer Science + Business Media, LLC Springer Verlag.

Li CL, WJ XU, Wu J, Gao MT. 2016. Probabilistic seismic hazard analysis of eastern Bayan Har Block with new models. Prog Geophys. 31(6):2370–2379. (in Chinese).

Li CT, Su XN, Meng GJ. 2018. Heterogeneous strain rate field in the northeast margin of Bayan Har Block from GPS observations and its relationship with the 2017 Jiuzhaigou Ms7.0 earthquake. Earthquake. 38(2):37–50. (in Chinese).

Nordquist JM. 1945. Theory of largest values applied to earthquake magnitudes. Trans AGU. 26(1):29–31.

Pickands J. 1975. Statistical inference using extreme order statistics. Ann Stat. 3(1):119–131.

Pisarenko VF, Sornette A, Sornette D, Rodkin MV. 2008. New approach to the characterization of mmax and of the tail of the distribution of earthquake magnitudes. Pure Appl Geophys. 165(5):847–888.

Pisarenko VF, Sornette A, Sornette D, Rodkin MV. 2014. Characterization of the tail of the distribution of earthquake magnitudes by combining the GEV and GPD descriptions of extreme value theory. Pure Appl Geophys. 171(8):1599–1624.

Qian XS, Cai XG, Ren QQ. 2013. Characteristics of the great earthquake magnitude distributions for active tectonic boundaries in Chinese mainland. J Earthq Eng Eng Vib. 33(1):212–220. (in Chinese).

Qian XS, Wang FC, Sheng SZ. 2013. Characterization of tail distribution of earthquake magnitudes via generalized Pareto distribution. Acta Seismol Sin. 35(3):340–350. (in Chinese).

Ribatet M, Dutang C. 2022. POT: generalized Pareto distribution and peaks over threshold. R Package Version 1.1–10. [accessed 2022 Apr 14]. https://CRAN.R-project.org/package=POT.

Shi DJ. 2006. Practical extremum statistical method. TianJin: Tianjin Science and Technology Press.

Singh AP, Mishra OP, Yadav R, BS, Kumar D. 2012. A new insight into crustal heterogeneity beneath the 2001 Bhuj earthquake region of northwest India and its implications for rupture initiations – sciedirect. J Asian Earth Sci. 48:31–42.

Tuncel M, Yegulalp J, Kuo T. 1974. Statistical prediction of the occurrence of maximum magnitude earthquakes. Bull Seismol Soc Am. 64(2):393–414.

Wang SY, Gao AJ, Feng YJ, He R. 2010. Comparison and standardization of the Chinese earthquake catalogs. Earthquake. 30(2):38–45.

Wen LM, LI JX, Wang ZW, Li W. 2021. The statistical analysis of risk measure in Pareto risk model. J Jiangxi Norm Univ. 45(2):211–216.

Wen XZ, Du F, Zhang PZ, Long F. 2011. Correlation of major earthquake sequences on the northern and eastern boundaries of the Bayan Har block, and its relation to the 2008 Wenchuan earthquake. Chin J Geophys. 54(3):706–716. (in Chinese).

Wood HO. 1912. The elastic-rebound theory of earthquakes by H. F. Reid, University of California Publications, Bulletin of the Department of Geology. Bull Seismol Soc Am. 2:98–100.