Estimation of Gutenberg-Richter seismicity parameters for the Bundaberg region using piecewise extended Gumbel analysis

Mike Turnbull

Central Queensland University

Abstract

The Gumbel statistics of extreme events is used to determine the Gutenberg-Richter seismicity parameters for the Central Queensland region within the ±3° geographical square centred on Bundaberg, based on the annual maximum magnitude events in available historic records from the 1870's through to the present time.

Introduction

It is common to characterize temporal and quantitative earthquake seismicity of a region by respectively specifying values for the a and b parameters of the Guttenberg-Richter seismicity model (the G-R model). Estimations of these parameters can be derived from a number of statistical processes. In situations where a comprehensively complete catalogue of earthquake events is not available, methods provided by the statistics of extreme events (the so-called extreme value theory (EVT)) have been applied, using reduced variate probability plotting.

The generalized EVT cumulative distribution function (cdf) reduces to one of three specific Fisher Tippett distributions (Fisher & Tippett, 1928), depending on the value chosen for its three parameters, ξ, $\theta(> 0)$, and $k(>0)$. These three distributions are summarized below (Johnson et al, 1995).

Fisher Tippett Type 1:

$$\Pr[X \le x] = \exp\{-\exp\{-1/\theta(x - \xi)\}\} \qquad \ldots \qquad \text{Eq. 1}$$

Fisher Tippett Type 2:

$$\Pr[X \le x] = 0, \quad \text{where } x < \xi \qquad \ldots \qquad \text{Eq. 2}$$
$$= \exp\{-\exp\{-(1/\theta(x - \xi))k\}\}, \text{where } x \ge \xi$$

Fisher Tippett Type 3:

$$\Pr[X \le x] = \{-\exp\{-(1/\theta(\xi - x))k\}\}, \quad \text{where } x \le \xi \qquad \ldots \qquad \text{Eq. 3}$$
$$= 1, \quad \text{where } x > \xi$$

The Type 2 distribution is often referred to as the Fréchet distribution. The Type 3 distribution is often referred to as the Weibull distribution. The Type 1 distribution is mostly referred to as the Gumbel distribution, but is sometimes referred to as the log-Weibull distribution. In this paper it will be referred to as the Gumbel distribution.

This paper uses Gumbel plotting of historical annual extreme earthquake magnitudes to estimate the G-R model parameters for the Central Queensland region in the ±3° geographical square centred on Bundaberg.

Using the Gumbel distribution to model extreme earthquakes

Cinna Lomnitz (1974) showed that if an homogeneous earthquake process with cumulative magnitude distribution

$$F(m; \beta) = 1 - e^{-\beta m}; \quad m \ge 0 \qquad \ldots \qquad \text{(Eq. 4)}$$

is assumed (compare with Eq. 24) , where β is the inverse of the average magnitude of earthquakes in the region under consideration; and α is the average number of earthquakes per year above magnitude 0.0; then y, the maximum annual earthquake magnitude, will be distributed according to the following Gumbel cdf.

$$G(y; \alpha, \beta) = \exp(-\alpha \exp(-\beta\, y)); \quad y \geq 0 \qquad \ldots \qquad \text{(Eq. 5)}$$

Using the probability integral transformation theorem (Bury, 1999, p 268) simulated maximum yearly earthquakes can be generated using the following inversion formula, where u_i is a random value in the (0, 1) closed interval.

$$y_i = -(1/\beta) \ln((1/\alpha) \ln(1/u_i)) \qquad \ldots \qquad \text{(Eq. 6)}$$

Manipulation of Eq. 6 produces the following linear relation.

$$-\ln(-\ln(p_i)) = \beta\, y_i - \ln(\alpha) \qquad \ldots \qquad \text{(Eq. 7)}$$

where p represents the probability plotting position, and the left hand expression is the reduced variate that can be used to plot data that is postulated as being drawn from a Gumbel distribution. Eq. 8 has been demonstrated (Turnbull & Weatherly, 2006) to be a suitable formula for determining plotting position values for analysis of extreme magnitude earthquakes, and will be used in this paper.

$$p_m = (m - 0.3) / (n + 0.4) \qquad \ldots \qquad \text{(Eq. 8)}$$

This formulation approximates the median of the distribution free estimate of the sample variate to about 0.1% and, even for small values of n, produces parameter estimations comparable to the results obtained by maximum likelihood estimations (Bury, 1999, p 43).

It has been demonstrated (Turnbull & Weatherly, 2006) that reduced variate probability plotting using the Gumbel distribution (Gumbel plotting) can be used to estimate α and β parameters from single data set calendars with 95% confidence that the estimated value will be within 15% and 5% respectively of the true value.

The average recurrence period T_y of an earthquake of magnitude y, is given by

$$T_y = 1/N_y \qquad \ldots \qquad \text{(Eq. 9)}$$

where N_y, the number of expected earthquakes per year exceeding magnitude y is given by

$$N_y = \alpha e^{-\beta y} \qquad \ldots \qquad \text{(Eq. 10)}$$

giving

$$T_y = (\alpha e^{-\beta y})^{-1} \qquad \ldots \qquad \text{(Eq. 11)}$$

G-R Parameter estimation using the Gumbel distribution

The Gutenberg-Richter (G-R) seismicity relation of earthquake frequency versus magnitude may be expressed as:

$$N(m \geq M) = 10^{(a - b\, m)} \qquad \ldots \qquad \text{(Eq. 12)}$$

where $N(m \geq M)$ is the number of earthquakes observed having magnitudes greater than or equal to M; and a and b are parameters specific to the observed data set. As a pragmatic mathematical and practical choice, the lower limit of M, M_0 is usually assigned the value zero. In that formulation the parameter a represents the logarithm to the base 10 of the number of independent earthquakes in the observation period with magnitude greater than or equal to zero.

$$a = \log_{10} N(m \geq M_0) \Rightarrow N(m \geq M_0) = 10^a \qquad \ldots \qquad \text{(Eq. 13)}$$

If it is assumed that all earthquakes included in the data set are independent, and that each event has equal probability of occurring, then Eq. 12 can be normalised to produce a frequency relation as follows,

$$\Pr(m \geq M) = N(m \geq M)/N(m \geq M_0) = 10^{(a - b\, m)}\, 10^{-a} = 10^{-bm} \qquad \text{(Eq. 14)}$$

It can be seen from Eq. 14 that the value of the parameter b determines the propensity for lower or higher magnitude earthquakes. Smaller values of b model a system that has a greater propensity for larger magnitude earthquakes. It also demonstrates that magnitude of the earthquakes is not dependent on the a parameter. The cdf formulation is as follows (compare with Eq. 4).

$$\Pr(m \le M) = 1 - 10^{-bm} \quad \ldots \quad \text{(Eq. 15)}$$

From Eqs. 4, 5, 13, 15 the relationships between the Gumbel parameters α and β and the G-R parameter a and b are seen to be

$$e^{-\beta} = 10^{-b} \quad => \quad b = \beta \log_{10} e \quad \ldots \quad \text{(Eq. 16)}$$
$$\alpha = 10^{a} \quad => \quad a = \log_{10}\alpha \quad \ldots \quad \text{(Eq. 17)}$$

Methodology

The Gumbel analysis methodology used for this paper is described below.

Basic method

The basic method assumes a complete catalogue. In the earthquake catalogue being analysed, the n annual extreme magnitudes are identified and isolated. The n annual extreme events are then sorted into ascending order, and ranked from 1 to n. Each m ordered event is assigned a plotting position using Eq. 8. From the plotting position value, the reduced variate value for each event is determined (see left hand of Eq. 7). The reduced variates are then plotted against the event magnitudes, using linear scales.

The values for α and β can be estimated from the slope and intercept using Eq. 7.

The values for the G-R a and b parameters can be estimated using Eqs. 16 and 17. Error bounds in a and b can be calculated by applying ±15% to the α estimation, and ±5% to the β estimation, and recalculating (Turnbull& Weatherly, 2006). In a similar manner, the recurrence periods for earthquakes of various magnitudes can be calculated using Eq. 11.

Censoring and imputation of data

Gumbel analysis should only be applied to catalogues, or to subsets of catalogues, that can be considered complete or very nearly so for earthquake magnitudes above a given lower value. For instance, both the Geoscience Australia (GA, 2006) and the Earth Systems Science Computational Centre (ESSCC, 2006) catalogues of earthquakes in the ±3° geographical square centred on Bundaberg have more than 50% of years with no entry. For entries prior to 1975 the lower limit of completeness varies depending on the instrumental coverage and availability of felt reports from time to time. For entries from 1975 to the present the author considers that both catalogues are complete within the area of interest for extreme magnitude 2.0 and above, with only minor absence of data. The author considers the full catalogues to be complete for extreme magnitude 5.5 and above.

Consequently, for Gumbel analysis purposes, two sets of useful data can be obtained: one set covering the years 1975 to 2005, for magnitude 2.0 and above; and a second set covering the years 1872 to 2005 for the ESSCC catalogue, and 1878 to 2005, for the GA catalogue, for magnitude 5.5 and above.

Missing data, or values below the completion magnitude, are imputed to some value less than the completion magnitude (in this case to magnitude zero). The imputed values are included in the ranking and calculation of plotting positions, but censored in the plotting.

Piecewise extension

The formulation used to determining plotting positions in the current study allows for extension of the analysis to include compatible data sets above and below the subset maximum and minimum magnitudes (See Turnbull & Weatherly, 2006, Section 3.2). This allows the two data sets identified in the previous section to be combined in a piecewise fashion, to extend the range of extreme magnitudes considered in the analysis. The lower range data set produces reduced variates from 31 data (including censored and imputed data) from 1975 to 2005. The upper range data set produces reduced variates from 128 and 134 data (including censored and imputed data) from 1878 to 2005 in the GA catalogue and 1872 to 2005 in the ESSCC catalogue respectively. The uncensored reduced variates and magnitudes from both data sets are combined to perform the piecewise extended Gumbel analysis.

Analysis of the Geoscience Australia earthquake catalogue

Figure 1 shows the piecewise extended Gumbel plot of extreme magnitude data extracted from the GA catalogue (GA, 2006).

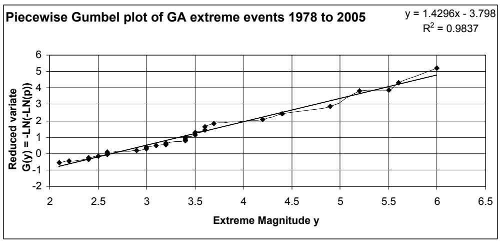

Figure 1: Piecewise extended Gumbel plot of GA extreme events 1878 to 2005

Intercept	-3.80	Error Bounds	
Slope	1.43	Lower	Upper
α	44.61	37.92	51.30
β	1.43	1.36	1.50
a	1.65	1.58	1.71
b	0.62	0.59	0.65

Table 1: Parameter estimation allowing for ±15% and ± 5% error bounds in α and β

Table 1 shows the statistical estimation of parameters, allowing for ±15% and ±5% relative error in the estimation of the Eq. 5 Gumbel α and β parameters respectively (see Turnbull & Weatherly, 2006). This results in an estimation of G-R parameter a within ±4.2%, and b within ±4.8% assuming that the source data is accurate.

Table 2 details the calculated recurrence periods for earthquakes in the magnitude range from 0.0 to 7.0. Extrapolation beyond magnitude 7.0 is not considered prudent until an analysis of the maximum expected magnitude based on the available data has been conducted. In Table 2 year, month and day values have been rounded to the nearest whole number.

	Nominal Parameter Value Expected return period			Lower Parameter Bounds Expected return period			Upper Parameter Bounds Expected return period		
Mag	Years	Months	Days	Years	Months	Days	Years	Months	Days
0.0			3			4			3
0.5			6			7			6
1.0		1	13		4	14		1	12
1.5		2	26		7	27		2	25
2.0		5	53		14	54		5	53
2.5		10	108		27	106		10	112
3.0	2	20		2	54		2	21	
3.5	3	40		3	106		4	45	
4.0	7	82		6	72		8	95	
4.5	14			12			17		
5.0	29			23			35		
5.5	58			46			75		
6.0	119			91			159		
6.5	243			180			337		
7.0	497			355			713		

Table 2: Expected Earthquake Recurrence Periods derived from GA catalogue 1878 to 2005

Figure 2 (below) shows a graph of the recurrence period relationship expressed by Eq.11.

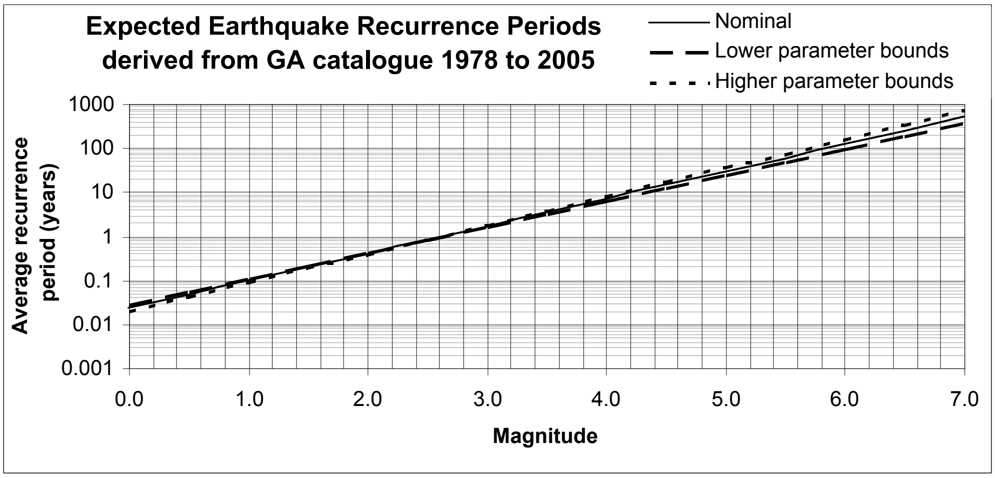

Figure 2: Expected Earthquake Recurrence Periods derived from GA catalogue 1878 to 2005

Analysis of the ESSCC earthquake catalogue

Figure 3 shows the piecewise extended Gumbel plot of extreme magnitude data extracted from the ESSCC catalogue (ESSCC, 2006).

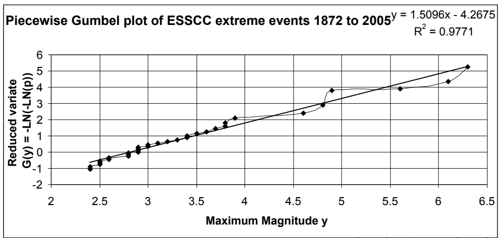

Figure 1 : Piecewise extended Gumbel

Figure 3: Piecewise extended Gumbel plot of ESSCC events 1872 to 2005

Intercept	-4.27	Error Bounds	
Slope	1.51	Lower	Upper
α	71.34	60.64	82.04
β	1.51	1.43	1.59
a	1.85	1.78	1.91
b	0.66	0.62	0.69

Tab

Table 3: Parameter estimations allowing for ±15% and ±5% error bounds in the estimation of α and β respectively

Table 3 shows the statistical estimation of parameters, allowing for ±15% and ±5% relative error in the estimation of the Eq. 5 Gumbel α and β parameters respectively (see Turnbull & Weatherly, 2006). This results in an estimation of G-R parameter a within ±3.8%, and b within ±6.0% assuming that the source data is accurate.

Table 4 details the calculated recurrence periods for earthquakes in the magnitude range 0 to 7.0. Extrapolation beyond magnitude 7.0 is not considered prudent until an analysis of the maximum expected magnitude based on the available data has been conducted. In Table 4 year, month and day values have been rounded to the nearest whole number.

	Nominal Parameter Value Expected return period			Lower Parameter Bounds Expected return period			Upper Parameter Bounds Expected return period		
Mag	Years	Months	Days	Years	Months	Days	Years	Months	Days
0.0			2			2			2
0.5			4			5			4
1.0		1	9		1	9		1	8
1.5		2	18		2	19		2	18
2.0		3	39		3	39		3	39
2.5		7	83		7	80		8	87
3.0	1	16		1	15		1	17	
3.5	3	33		2	30		3	38	
4.0	6	71		5	61		7	83	
4.5	12			10			15		
5.0	27			21			34		
5.5	57			44			74		
6.0	120			90			165		
6.5	256			184			363		
7.0	544			377			803		

Table 4: Expected earthquake recurrence periods from ESSCC catalogue 1872 to 2005

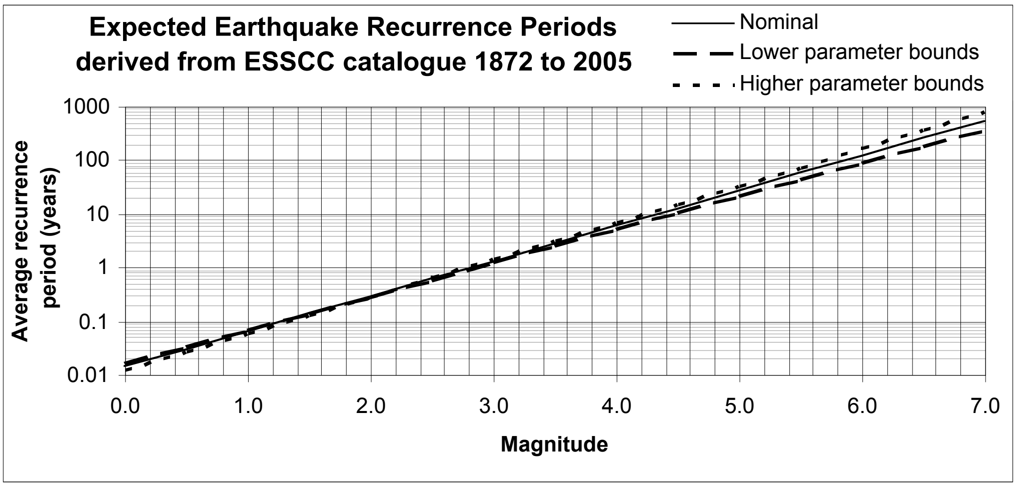

Figure 4 shows a graph of the recurrence period relationship expressed by Eq. 11.

Summary

Piecewise extended Gumbel analysis of the Geoscience Australia and Earth Systems Science Computational Centre catalogues of Australian earthquakes within the ±3° geographical square centred on Bundaberg indicates a Gutenberg-Richter b parameter value of 0.62 to 0.66. The analysis suggests that the region studied can expect on average to exhibit at least one earthquake of magnitude 6.0 or greater in any given 120 year period. This analysis supersedes similar analysis previously carried out by the author (Turnbull, 2001) which calculated a b value of 0.59, and recurrence period of 85 years.

References

Bury K. 1999, Statistical Distributions in Engineering, Cambridge University Press, ISBN 0 521 63506 3.

ESSCC, 2006, Queensland Earthquake Catalogue (1866-2005), the Earth Systems Science Computational Centre, University of Queensland, 2006. Accessed online 7 August 2006 at http://www.esscc.uq.edu.au/Seismology/html/Catalog.html.

Fisher R.A. & Tippett L.H.C. 1928, Limiting forms of the frequency distribution of the largest or smallest member of a sample, Proc. Cambridge Phil. Soc., 24:180.

GA, 2006, Earthquake Online Search Form, Geoscience Australia, 2006. Accessed online 2 August 2006 at http://www.ga.gov.au/oracle/quake/quake_online_input.jsp.

Gumbel E.J. 1958, Statistics of Extremes, Dover Publications Edition 2004, ISBN 0 486 43604 7, originally published Colombia University Press, 1958.

Johnson N.L., Kotz, S. & Balakrishnan N. 1995, Continuour Univariate Distributions, Volume 2, 2nd Ed., John Wiley & Sons. ISBN 0 471 58494 0.

Lomnitz C. 1974, Global Tectonics and Earthquake Risk, in Developments in Geotectonics 5, Elsevier Scientific Publishing Company.

Turnbull M.L. 2001, A Seismic Hazard Assessment and Microzonation of Bundaberg., Central Queensland University, Faculty of Engineering and Physical Systems, Master’s Degree Thesis, April 2001.

Turnbull M.L. & Weatherly D. 2006, Validation of using Gumbel Probability Plotting to estimate Gutenberg-Richter seismicity parameters, Proceedings of the AEES 2006 Conference, Canberra, Australia, November 2006 (in print).