

Seismic Risk of Circum-Pacific Earthquakes: II. Extreme Values Using Gumbel's Third Distribution and the Relationship with Strain Energy Release

PAUL W. BURTON¹ and KOSTAS C. MAKROPOULOS²

Abstract—In a previous paper (MAKROPOULOS and BURTON, 1983) the seismic risk of the circum-Pacific belt was examined using a 'whole process' technique reduced to three representative parameters related to the physical release of strain energy, these are: M_1 , the annual modal magnitude determined using the Gutenberg–Richter relationship; M_2 , the magnitude equivalent to the total strain energy release rate per annum, and M_3 , the upper bound magnitude equivalent to the maximum strain energy release in a region.

The risk analysis is extended here using the 'part process' statistical model of Gumbel's IIIrd asymptotic distribution of extreme values. The circum-Pacific is chosen, being a complete earthquake data set, and the stability postulate on which asymptotic distributions of extremes are deduced to give similar results to those obtained from 'whole process' or exact distributions of extremes is successfully checked. Additionally, when Gumbel III asymptotic distribution curve fitting is compared with Gumbel I using reduced chi-squared it is seen to be preferable in all cases and it also allows extensions to an upper-bounded range of magnitude occurrences. Examining the regional seismicity generates several seismic risk results, for example, the annual mode for all regions is greater than $m(1) = 7.0$, with the maximum being in the Japan, Kurile, Kamchatka region at $m(1) = 7.6$. Overall, the most hazardous areas are situated in this northwestern region and also diagonally opposite in the southeastern circum-Pacific. Relationships are established between the Gumbel III parameters and quantities $m_1(1)$, X_2 and ω , quantities notionally similar to M_1 , M_2 and M_3 although ω is shown to be systematically larger than M ; thereby giving a physical link through strain energy release to seismic risk statistics. In *all* regions of the circum-Pacific similar results are obtained for M_1 , M_2 and M_3 and the notionally corresponding statistical quantities $m_1(1)$, X_2 and ω , demonstrating that the relationships obtained are valid over a wide range of seismotectonic environments.

Key words: Seismic risk; extreme values; strain energy; circum-Pacific.

1. Introduction

This paper examines the relationship between assessments of seismic risk for the circum-Pacific belt, principally obtained using the statistical 'part process' provided by Gumbel's third asymptotic distribution of extreme values (referred to as Gumbel III), and that obtained from the 'whole process' analysis of strain energy release described

¹ Natural Environment Research Council, Institute of Geological Sciences, Murchison House, West Mains Road, Edinburgh EH9 3LA, Scotland.

² Seismological Laboratory, University of Athens, Panepistimiopoli, Athens, Greece.

in our previous paper (MAKROPOULOS and BURTON, 1983, hereafter referred to as Paper I).

A major purpose of this work is to demonstrate the usefulness and veracity of extreme value theory when applied to the estimation of earthquake occurrence and compared with results obtained using other methods. It should be borne in mind that the extreme value method has certain clear and obvious advantages as far as the requisite data are concerned when compared with methods requiring the whole data set, which is rarely completely reported: thus it is appropriate to check the stability postulate, on which asymptotic distributions of extremes are derived from exact distributions of extremes, by analysing an area for which there is a complete earthquake data set. This will allow direct comparison of these two categories of methods, namely those which need the whole data (all earthquakes) and the category which uses extreme value statistics which need only part of the data (the largest earthquakes). The selection of an area representative of well documented and complete seismicity inevitably leads to an area of high seismicity, the circum-Pacific is such a seismic zone. Paper I examined seismic risk in the circum-Pacific using the whole data set and related the parametric results to strain energy release; here the same data set will be examined using the largest values only to estimate the seismic risk and demonstrate the reliability of the method, and the parametric results will also be related to strain energy release giving a direct link with the physical process expressed principally through the average annual strain energy release. The physically realistic imposition of an upper limit to the size of extreme earthquakes (indeed to all earthquake occurrence) implies the use of the third asymptote of extreme values generating results which may be compared directly with the physical process of strain energy release implicit in the Benioff-type diagrams of Paper I.

The occurrence of earthquakes in space and time falls under the general category of stochastic processes, that is, mathematical models of a given physical system that changes in accordance with the laws of probability (LOMNITZ, 1974). Various statistical models have been applied to the analysis of earthquake occurrence with differing degrees of success, and results are often unconvincing because of incompleteness in the data sets or because inherent uncertainties in the distribution parameters are simply ignored.

Earthquake occurrence models have usually incorporated the Poisson distribution, or extended to clustering of events using Markovian models of non-independent events. The usual expression linking earthquake magnitudes with their rates of occurrence is due to GUTENBERG and RICHTER (1944)

$$\log N_c(M) = a - bM. \quad (1)$$

This assumes knowledge of the whole process above a threshold magnitude, with N_c being the annual cumulative number of earthquakes with magnitude equal or greater than M . An obvious weakness in application can be that there is a lack of accuracy, homogeneity and completeness of data sets analysed, particularly in the lower

magnitude ranges. Such perturbations mainly arise from the sensitivity and temporal variation of deployed seismograph networks monitoring seismicity.

Seismic risk and related earthquake engineering purposes usually require estimation of return periods or probabilities of exceedance of specific levels of design load criteria or extremal safety conditions. Thus what is of primary importance in earthquake engineering is compatible with a need to consider extreme value distributions separately from the statistics of the whole process. Extreme value statistical theory seems to satisfy most of the above problems and since GUMBEL's (1935, 1967) developments the theory has been applied to hydrological computations, climatic evaluations (JENKINSON, 1955; GRINGORTEN, 1963a; KRUMBEIN and LIEBLEIN, 1956) as well as to the analysis of earthquake occurrence (starting with NORDQUIST, 1945, and many subsequent authors, for example: EPSTEIN and LOMNITZ, 1966; KARNIK and HUBNEROVA, 1968; SCHENKOVA and KARNIK, 1970, 1977, 1978; YEGULALP and KUO, 1974; RADU and APOPEL, 1977; BURTON, 1978a, 1979; BURTON, MCGONIGLE, MAKROPOULOS and UCER, 1984; ROCA, ARROYO and SURINACH, 1984 (the last an application for earthquake intensity rather than magnitude recurrence)). Practical advantages of extreme value methods are known. Extreme values of a geophysical variate are usually better known than the smaller events in a time series of data: detailed knowledge of the parent distribution is not needed because the distributions of extremes depend on common asymptotic properties of the rare events in the tail of possible distributions of the variate.

Thus the use of extreme values now has a lengthy history in several branches of science, including application to the problem of earthquake recurrence estimation. There are several justifications for adopting extreme value distributions. There is the practical consideration in any investigation of seismic risk that it is the extreme or maximum events which are of most interest. Secondly, and more fundamentally, there is considerable theoretical justification and knowledge of the behaviour of distributions of extremes. The adoption of a distribution of extremes for the earthquake process has additional attractions when it is realised that these are often the more reliable data available to the seismologist. GUMBEL's (1967) treatment is lengthy, however, the more pertinent points can readily be extracted. Exact distributions of extremes are easily obtained for samples of size n when the initial distribution is known. Distributions of extremes are not characterised adequately simply by medians, modes or means: an average called the asymptotic value or characteristic extreme is introduced, which ultimately leads to the asymptotic distributions required. First consider exact distributions of a variate; if n independent samples of the variate are taken then the probability that *all* are less than x is $F^n(x)$. Therefore, the probability $\Phi_n(x)$ that x is an extreme or maximum value is

$$\Phi_n(x) = F^n(x). \quad (2)$$

Put alternatively this means that $\Phi_n(x)$ is the probability the largest of the n samples is less than or equal to x . Clearly, $1 - \Phi_n(x)$ is the probability that x may be exceeded

and it is assumed that the variate is continuous. If $F(x)$ is known then $\Phi_n(x)$ may be calculated directly. Exact distributions of extremes have some simple properties: for example if x is multiplied by a constant then so is the extreme value; addition of a constant to x adds equally to the extreme value. Expanding $\ln F(x)$ as a Taylor series gives the approximation for exact distributions of extremes, that

$$\Phi_n(x) \approx e^{-n(1 - F(x))}. \tag{3}$$

The characteristic largest value u_n may now be introduced for samples of size n (usually assuming reasonably large n), given that $n(1 - F(x))$ samples are expected to be equal to or larger than x then u_n may be introduced as

$$F(u_n) = 1 - 1/n, \tag{4}$$

which leads to the approximation that

$$\Phi_n(u_n) \approx e^{-1}. \tag{5}$$

The implication for any set of extremes is that approximately 36.8% of them will be below this characteristic value, about which Φ_n is skew. GUMBEL (1967) emphasises for exponential type distributions that most probable extremes *converge* toward these characteristic extremes; for exponential type distributions the mode converges towards these characteristic largest values. The theoretical justification extends to both asymptotic distributions of extreme values and also to the inclusion of variates, x , which are limited to the right which is appropriate to the consideration of observational estimates of earthquake magnitude. FRECHET's (1927) stability postulate facilitates extension to the asymptotic distributions of extremes: a sample of size n will have a largest value; N samples of size n will have N largest values; the largest of the nN samples will also be the extreme of the N largest values and, more fundamentally, both the distribution of the largest value of the set of samples and of the individual samples will be asymptotic to the same distribution. This implies that a linear transformation of x does not change the form of the probability distribution, $F(x)$, that is for the extreme values

$$F^n(x) = F(a_n x + b_n), \tag{6}$$

where a_n and b_n are functions of n . It can be shown that

$$\ln(-\ln F(x)) - \frac{x \ln n}{b_n} \tag{7}$$

is constant, from which Gumbel's first asymptotic distribution of extremes, or asymptote, follows in the notation used for our present purpose in (8) below. Additionally, if the variate x is upper bounded to the right by $x \leq \omega$ then the condition is introduced that $F(\omega) = 1$. Gumbel's third asymptote may then be deduced analytically or obtained from (8) by transformation of both the variable and the parameters and inclusion of the new condition $F(\omega) = 1$; the third asymptote is

expressed in terms of earthquake magnitude as the variate in (9). A final point to note before proceeding with the application of the asymptotes is that although independence of observations is required for the exact distributions of extremes, this may not be required for the asymptotes (WATSON, 1954) where inter-dependence between large values of the variate in the initial distribution (excluding aftershocks) may be weak or may disappear.

Gumbel's first asymptotic distribution of extreme values (Gumbel I) arising from (7) is of the form

$$G^I(m) = \exp \{ -\exp[-a(m - u)] \}, a > 0, \quad (8)$$

having two parameters: a , and the characteristic or modal extreme u . G is the probability that a magnitude m is an annual extreme (of course other intervals other than the annual may be used as convenient). This equation can easily be fitted to data using standard linear least squares regression, and EPSTEIN and LOMNITZ (1966) have demonstrated its relationship to the whole process of magnitude recurrence specified by (1). However, this open ended linear form is not always seen to be borne out by experimental observation of earthquakes leading to, for example, CORNELL and VANMARCKE'S (1969) truncation of (1) at a limiting largest magnitude which RICHTER (1958) suggests is about 8.5 to 9.0 for the whole world, and also to the introduction of a quadratic term in magnitude by SACUIU and ZORILESCU (1970) and MERZ and CORNELL (1973). The existence of an upper bound to the earthquake magnitude that can be generated by a finite volume of strain energy storage is physically inescapable (ESTEVA, 1976), and Paper I derives an estimate of this upper bound based on an analysis of the whole process. The part process extreme value distribution which has an upper bound to magnitude occurrence is Gumbel's third asymptotic distribution, and is of the form

$$G^{III}(m) = \exp \left[-\left(\frac{\omega - m}{\omega - u} \right)^k \right], k > 0, \quad (9)$$

with three parameters: the upper bound magnitude ω , the characteristic extreme magnitude value u (not the modal value), and $k (= 1/\lambda)$ which relates to curvature of the distribution. Inclusion of an upper bound to the variate leads naturally to the form of (9) when asymptotic extremes are considered, this is considered here to be an advantage compared to alternatives which, for example, directly truncate the Gutenberg-Richter relation by imposing an arbitrary cut-off magnitude. A further expected advantage of (9) arises simply from the fact that it is a three parameter distribution; values of reduced chi-squared will bear this out.

This paper is a sequel to Paper I in which the whole process was used to define seismic risk in terms of three parameters related to the physical release of strain energy, in Paper I these are: M_1 , the largest earthquake expected in a year; M_2 , the magnitude equivalent to the total strain energy release rate per annum; and M_3 , the upper bound to magnitude in a region. The major purpose of this paper is to examine

seismic risk in the circum-Pacific belt using the Gumbel III distribution with careful assessment of all three of its parameters and to relate these to physical release of strain energy through M_1 , M_2 , and M_3 of the whole process; thereby giving a physical link through strain energy release to seismic risk statistics.

2. Strain energy release and Gumbel III

The method of fitting Gumbel III to data is outlined briefly below. The method of Paper I will then be followed to relate (ω, u, λ) sets to the parameters of strain energy release M_1 , M_2 and M_3 , thereby generating a link between the Gumbel III parameters and the physical release of strain energy.

2.1. Evaluation of Gumbel III and forecasting

Equation (9) may be rearranged as

$$m = \omega - (\omega - u)[\ln(P(m))]^2 \quad (10)$$

where $G^{III}(m)$ has been replaced by $P(m)$, denoting the probability that magnitude m is an annual extreme. This non-linear function has to be fitted to the observational extreme value data. The generalised technique of curve fitting used here relies on LEVENBERG (1944) as developed by MARQUARDT (1963) and expounded on by BEVINGTON (1969). The method used here for curve fitting and ensuing forecasting, with modifications to be compatible with earthquake data, largely follows BURTON (1979) and MAKROPOULOS (1978).

Annual extreme magnitudes m_i are extracted from a catalogue of n -years duration, ranked $m_1 \leq m_2 \dots \leq m_n$ where m_n is the largest earthquake magnitude in the catalogue, and GRINGORTEN's (1963b) 'plotting point' probability assigned at each m_i

$$P(m_i) = (i - 0.44)/(n + 0.12), \quad i = 1 \dots n. \quad (11)$$

There may be j years for which there is no entry in m_i and in practice (5) is then calculated over $i = j + 1, \dots, n$; following YEGULALP and KUO (1974). Therefore the plotting point probability for the lowest observed extreme value becomes, using (11), the value $(j + 1 - 0.44)/(n + 0.12)$. BURTON (1979) noted that the procedure is satisfactory as long as $j \leq n/4$, which may often be achieved by using extreme intervals other than the annual. stability in Gumbel III forecasting obtained using different extreme intervals was demonstrated by BURTON (1981) in an area of low seismicity, although continual reduction in the extreme intervals to short periods will inevitably fail as both j and the level of incompleteness increases.

The manner in which the method is actually applied has several useful and desirable properties. A weight δm_i may be assigned to each extreme datum m_i , thus taking into

account reliability. The method is iterative and goodness of fit obtained between the current (ω, u, λ) set in (10) and the data (m_i, P_i) may be inspected after each iteration through ρ , the reduced chi-squared value on v degrees of freedom; $\rho = \chi^2/v$. A major advantage is obtained after acceptable goodness of fit has been achieved by then calculating the complete symmetrical error or covariance matrix ε amongst the final parameter set (ω, u, λ) , where

$$\varepsilon = \begin{bmatrix} \sigma_\omega^2 & \sigma_{u\omega}^2 & \sigma_{\lambda\omega}^2 \\ \sigma_{\omega u}^2 & \sigma_u^2 & \sigma_{\lambda u}^2 \\ \sigma_{\omega\lambda}^2 & \sigma_{u\lambda}^2 & \sigma_\lambda^2 \end{bmatrix} \tag{12}$$

Knowledge of ε is used to assess uncertainties in (ω, u, λ) through its diagonal elements. The importance of knowledge of all ε_{ij} is emphasized when it is seen that $\sigma_{\omega\lambda}^2$ is usually large and negative (BURTON, 1978a), and all ε_{ij} should be taken into account when evaluating uncertainties on statistical predictions or forecasts. The condition for a modal extreme magnitude is $d^2P/dm^2 = 0$, and for the next T -years the modal extreme magnitude $m_1(T)$ is given by

$$m_1(T) = \omega - (\omega - u)[(1 - \lambda)/T]^\lambda. \tag{13}$$

The uncertainty σ_m on $m(T)$ is evaluated using all ε_{ij} through

$$\sigma_m^2 \approx \sigma_\omega^2 \left(\frac{\partial m}{\partial \omega}\right)^2 + \sigma_\lambda^2 \left(\frac{\partial m}{\partial \lambda}\right)^2 + \sigma_u^2 \left(\frac{\partial m}{\partial u}\right)^2 + 2\sigma_{\omega\lambda}^2 \left(\frac{\partial m}{\partial \omega}\right) \left(\frac{\partial m}{\partial \lambda}\right) + \dots \tag{14}$$

σ_m in (14) may be calculated using partial derivatives obtained from (13), or it could equally well be obtained using the median, mean or 'return period' estimate of $m(T)$ at sufficiently lengthy T -years (BURTON, 1979). For lengthy T -year predictions it follows that

$$\sigma_m^2 \Big|_{T \rightarrow \infty} \rightarrow \sigma_\omega^2. \tag{15}$$

2.2 Strain energy release

Recall that (Paper I): M_1 is the most probable annual maximum magnitude (mode) determined from the Gutenberg–Richter whole process equation (1), M_2 is the magnitude equivalent to the mean annual rate of energy release, and M_3 is the upper bound magnitude equivalent to the maximum strain energy release in a region. By following the method of Paper I it is now possible to relate the parameters (ω, u, λ) of Gumbel III to M_1 and to the physical quantities M_2 and M_3 , which represent the whole process of strain energy release in a region.

The mode. It is clear that the whole process annual mode M_1 obtained from (1) is

a/b , and this should be an equivalent or similar (\approx) quantity to $m_1(1)$ of Gumbel III obtained by setting $T = 1$ in (13), that is

$$M_1 = a/b \approx m_1(1) = \omega - (\omega - u)(1 - \lambda)^{\lambda}. \tag{16}$$

T -year modes could be compared similarly using

$$M_T = (a + \log T)/b \approx m_1(T) = \omega - (\omega - u)[(1 - \lambda)/T]^{\lambda}, \tag{17}$$

but these would be expected to diverge at large T as $m_1(T) \rightarrow \omega$, thus reflecting to little purpose the unbounded nature of (1) in relation to the upper bounded form of (9). *The mean annual energy release.* The general methodology of Paper I is followed to relate M_2 to (ω, μ, λ) , using the equation linking energy and magnitude of an earthquake

$$\ln E = A + Bm. \tag{18}$$

The annual number of earthquakes exceeding magnitude m is given by (JENKINSON, 1955)

$$N(x \geq m) = \left(\frac{\omega - m}{\omega - u} \right)^k, \tag{19}$$

which implies

$$\begin{aligned} \frac{dN}{dm} &= -k \frac{(\omega - m)^{k-1}}{(\omega - u)^k} \\ dN &= -C(\omega - m) dm, \end{aligned} \tag{20}$$

where

$$C = \frac{k}{(\omega - u)^k}. \tag{21}$$

Annual energy release dE attributed to earthquakes in the range dm is using (18)

$$dE = e^{A+Bm} dN. \tag{22}$$

Total annual energy release TE is obtained using (20) in (22) and integrating over all possible magnitudes to give

$$TE = Ce^A \int_{-\infty}^{\omega} e^{Bm} (\omega - m)^{k-1} dm. \tag{23}$$

Changing variables to $x = \omega - m$,

$$TE = Ce^{A+B\omega} \int_0^{\infty} e^{-Bx} x^{k-1} dx, \tag{24}$$

and recognising the integral as $\Gamma(k)/B^k$ where $\Gamma(k)$ is the normal symbol for the

Gamma function, then

$$TE = Ce^A e^{B\omega} \frac{\Gamma(k)}{B^k}. \quad (25)$$

Using (18) to express this annual total energy release as an equivalent magnitude X_2 finally gives

$$M_2 \approx X_2 = \omega + \frac{1}{B} \ln \left(\frac{CE\Gamma(k)}{B^k} \right). \quad (26)$$

This magnitude X_2 defined in terms of (ω, u, λ) should be an equivalent or similar quantity to M_2 of Paper I. Evaluations of X_2 later in this paper use BÅTH'S (1958) constants in (18), but note that natural logarithms have been used which means that $B = 1.44 \ln 10$ is appropriate in (26).

The upper bound earthquake ω is the upper bound to magnitude from the (ω, u, λ) set. It is clear that ω is notionally equivalent to M_3 of Paper I, no matter whether M_3 is determined analytically (equation (17) of Paper I), or graphically from the cumulative strain energy release diagrams of Paper I.

However, there is a conceptual difference between M_3 and ω . The latter corresponds to a theoretical infinite return period corresponding to the statistical upper limit to the variate whereas M_3 corresponds to finite, rather than infinite, waiting time. Although the uncertainties on ω , which are usually large (BURTON, 1979), may be found to encompass M_3 , the conceptual difference between ω and M_3 leads us to expect

$$\omega - M_3 \geq 0. \quad (27)$$

3. Application of Gumbel III to circum-Pacific earthquakes

3.1. The data and their analysis

Seismicity of the circum-Pacific belt is analysed here in two time periods from 1897 to 1964 as in Paper I, and secondly from 1897 to 1975 inclusive. The data come from DUDA'S (1965) catalogue, supplemented by GUTENBERG and RICHTER (1954) and the BGS Seismicity File (BURTON, 1978b) since 1956 in those years for which Duda has no entry. Duda's sources are principally GUTENBERG and RICHTER (1954) for 1904–1952, with the revised surface wave magnitudes of RICHTER (1958) converted from the unified magnitude. Comments on magnitude accuracy in these original sources lead us to estimate weights on the extracted annual extremes as indicated in Table 1.

Annual extreme magnitudes are extracted for each of the seven regions of the circum-Pacific belt, ranked, and 'plotting point' probabilities assigned as in the manner associated with (11). The parameters (ω, u, λ) of Gumbel III are then estimated using

Table 1

Weights estimated on annual extremes of magnitude dependent on the original source material used by DUDA (1965). Annual extremes of magnitude are mostly extracted from DUDA's (1965) catalogue

Sub-period years	Duda's source	Weight assigned i.e. δm_i
1897–1903	GUTENBERG (1956)	± 0.6
1904–1917	GUTENBERG and RICHTER (1954)	± 0.6
1904–1917	DUDA's (1965) addition of 146 events	± 0.4
1918–1953	GUTENBERG and RICHTER (1954)	
	Case <i>a</i> (when magnitude assigned to a tenth of a unit)	± 0.3
	Case <i>b</i> (when magnitude assigned to the nearest quarter)	± 0.4
	Case <i>c</i> (when as Case <i>a</i> but with the addition of \pm)	± 0.4
	Case <i>d</i> (when as Case <i>b</i> but with the addition of \pm)	± 0.5
1954–1975	DUDA (1965), BURTON (1978b) ¹	± 0.3

¹ Not used by Duda.

the methodology of section 2.1, and the parameters ($u, 1/a$) of Gumbel I are estimated using standard least squares regression for comparison purposes.

3.2. Discussion of parametric results

The parameters with uncertainties for Gumbel I and III applied to seven regions of the circum-Pacific are listed in Tables 2 and 3, for the two time periods, 1897 and 1964 and 1897–1975 respectively. Observed annual extreme magnitudes and the two corresponding extreme value distribution curves fitted are shown in Figure 1 for South America (that is Region 1) and summarized for all seven regions and the world as a whole in Figure 2. Tables 2 and 3 each contain two additional columns of information beyond the Gumbel I and III parameters: the number of 'missing years' for which no extreme value is available is entered and is consistently considerably less than $n/4$ (see (11)), and the difference between the goodness-of-fit to the data obtained using Gumbel I and III is expressed as the difference in respective reduced chi-squares as $\rho_1 - \rho_3$. Table 3 also lists the largest observed magnitude in each of the regions during 1897–1975.

$\rho_1 - \rho_3$ values show in all cases that Gumbel III produces lower reduced chi-squared than Gumbel I, and is preferable on these grounds alone. It would have been a surprising result if a three parameter distribution failed to show better fit than one of two parameters. The lowest $\rho_1 - \rho_3$ of 0.05 is observed in Region 6 (New Hebrides, Solomon, New Guinea) and not surprisingly corresponds to minimum curvature or λ . Maximum $\rho_1 - \rho_3$ is 0.348 for Region 4 which shows a well formed curved third asymptotic distribution. u of Gumbel III is well determined with low σ_u in all cases. When a region shows little curvature, accompanied by small λ and high ω , then the parameter uncertainties are larger. When the period examined extends to 1975 in Table 3 then these values tend to show increasing stability. Statistical stability is examined in detail in Table 4 because it is a basic assumption of this

Table 2
Estimated Parameters of Asymptotic distributions (1897-1964)

Region	Third type			First type				Missing years	Chi-square $\rho_1 - \rho_3$			
	ω	σ_ω	u	σ_u	$\lambda = \frac{1}{k}$	σ_λ	u			σ_u	$\frac{1}{k}$	σ_k^1
(1) South America	10.16	1.13	7.08	0.04	0.197	0.091	7.08	0.04	0.477	0.028	8	+0.094
(2) North America	9.14	0.51	7.14	0.04	0.320	0.110	7.11	0.03	0.501	0.031	3	+0.186
(3) Aleutians, Alaska	9.66	0.63	6.78	0.04	0.260	0.083	6.84	0.04	0.486	0.029	13	+0.216
(4) Japan Kurile Kamchatka	9.30	0.34	7.38	0.03	0.327	0.076	7.34	0.03	0.437	0.024	2	+0.348
(5) New Guinea, Banda Sea Celebes, Moluccas Philippines	10.00	1.11	7.42	0.03	0.194	0.098	7.40	0.03	0.425	0.027	2	+0.069
(6) New Hebrides Solomon New Guinea	9.44	1.11	7.23	0.03	0.220	0.125	7.22	0.04	0.397	0.031	6	+0.060
(7) New Zealand Tonga Kermadec	8.95	0.35	6.89	0.04	0.357	0.091	6.95	0.04	0.441	0.027	10	+0.316
(8) World	9.23	0.25	8.17	0.03	0.358	0.056	8.12	0.03	0.285	0.025	0	+0.299

Table 3

Estimated Parameters of Asymptotic distributions (1897-1975_{sep})

Region	Third type			First type				Chi-square $\rho_1 - \rho_3$	Missing years	Observed maximum			
	ω	σ_ω	u	σ_u	$\lambda = \frac{1}{k}$	σ_λ	u				σ_u	$\frac{1}{a}$	σ_a^2
(1) South America	9.92	0.94	7.09	0.03	0.208	0.086	7.10	0.03	0.457	0.025	+0.101	10	8.9
(2) North America	9.01	0.39	7.11	0.03	0.352	0.094	7.10	0.03	0.471	0.027	+0.257	6	8.6
(3) Aleutians, Alaska	9.77	0.69	6.79	0.04	0.235	0.077	6.85	0.04	0.469	0.027	+0.178	16	8.7
(4) Japan Kurile Kamchatka	9.33	0.34	7.38	0.03	0.307	0.069	7.34	0.03	0.429	0.022	+0.329	2	8.9
(5) New Guinea, Banda Sea Celebes, Moluccas Philippines	9.58	0.72	7.41	0.03	0.235	0.097	7.38	0.03	0.414	0.025	+0.091	4	8.7
(6) New Hebrides Solomon New Guinea	9.56	1.14	7.25	0.03	0.198	0.112	7.24	0.03	0.381	0.027	+0.050	6	8.6
(7) New Zealand Tonga Kermadec	8.91	0.33	6.87	0.04	0.359	0.089	6.95	0.04	0.429	0.025	+0.294	14	8.7
(8) World	9.13	0.15	8.12	0.03	0.395	0.054	8.07	0.03	0.299	0.022	+0.326	0	8.9

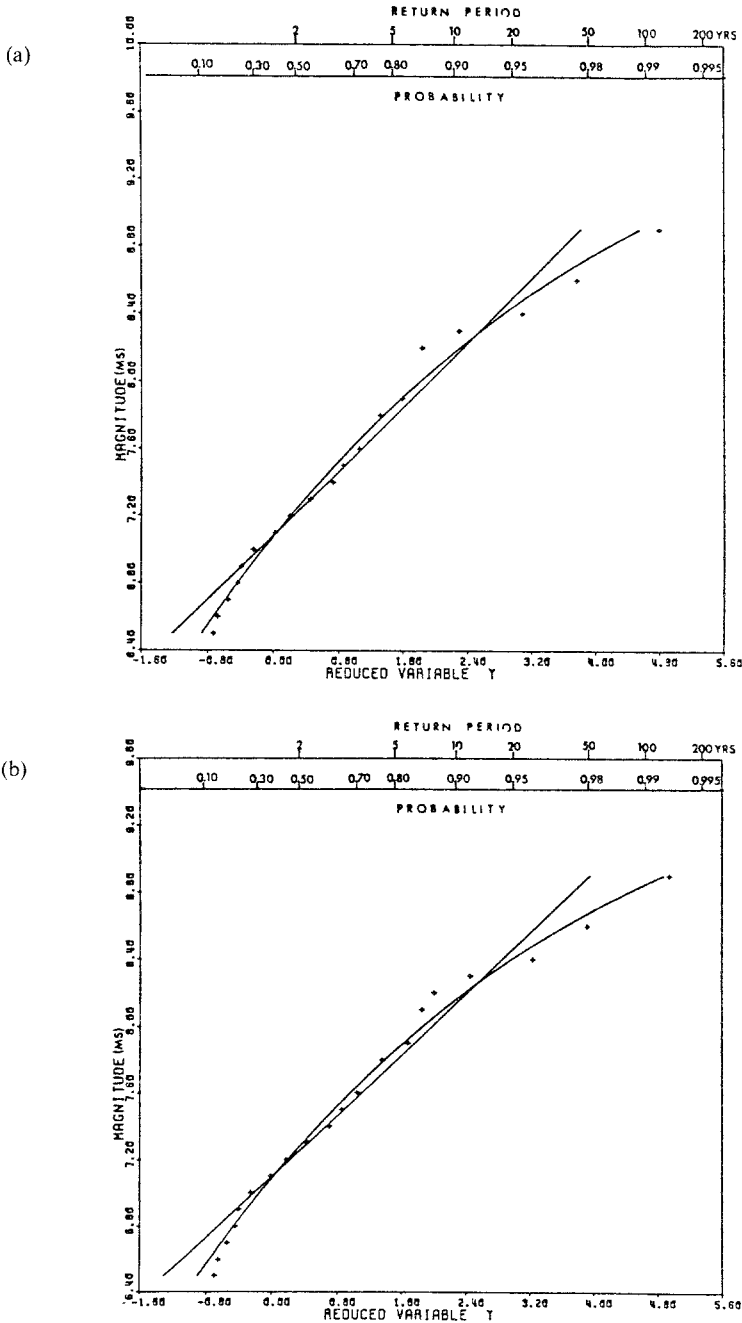


Figure 1

(a) Asymptotic distribution curves of extreme values of magnitude for South America (Region I) for the period 1897–1964. Straight line indicates the first type of extreme value distribution, curved line indicates the third type, + indicates observed annual maximum magnitude. Subsidiary x axis represents the probability of a magnitude being an annual extreme, and its return period in years, and the reduced variable Y in the x axis is $-\ln(-\ln P)$. (b) Asymptotic distribution curves of extreme values of magnitude for South America (Region I), for the period 1897–1975. (Explanation of symbols as in (a)).

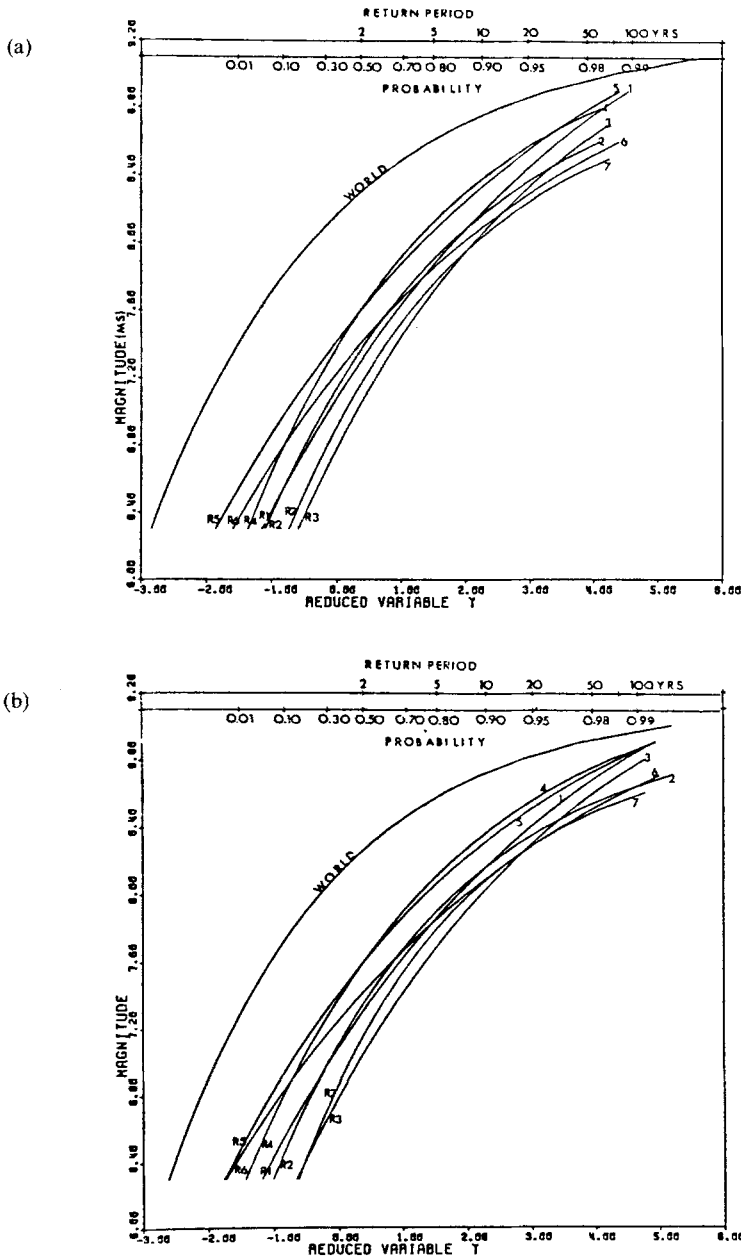


Figure 2

(a) Asymptotic distribution curves of extreme values of magnitude summarized for the seven regions (DUDA, 1965) of the circum-Pacific belt, labelled R1-R7, and the world as a whole, for the period 1897-1964. (b) Asymptotic distribution curves of extreme values of magnitude summarised for the seven regions (DUDA, 1965) of the circum-Pacific belt, labelled R1-R7, and the world as a whole, for the period 1897-1975. (Explanation of symbols as in Figure 1(a)).

Table 4
Test of statistical stability on $m_1(75)$

Region	Sample period (years)				
	35	45	55	65	75
	$m_1(75)$				
1	$9.2 \pm 1.3^*$	9.1 ± 1.4	9.0 ± 1.1	8.9 ± 1.0	8.9 ± 1.0
2	8.9 ± 1.0	8.8 ± 1.1	8.7 ± 0.8	8.7 ± 0.7	8.7 ± 0.8
3	8.7 ± 0.9	8.9 ± 1.2	8.8 ± 0.9	8.8 ± 0.9	8.8 ± 1.0
4	9.0 ± 0.8	8.9 ± 0.8	8.9 ± 0.9	8.9 ± 0.9	8.9 ± 0.7
5	9.0 ± 0.9	9.0 ± 0.8	8.9 ± 1.0	8.9 ± 0.9	8.9 ± 1.0
6	8.8 ± 0.9	8.9 ± 1.0	8.7 ± 1.0	8.7 ± 1.1	8.7 ± 1.2
7	8.8 ± 1.0	8.8 ± 0.9	8.7 ± 0.9	8.6 ± 0.9	8.6 ± 0.6

* The uncertainties are the ranges in which the mode with return period $T = 75$ years $m_1(75)$ will lie with probability 95%.

model that future seismicity will be similar to that in the past. Table 4 shows the 75-year modal earthquake $m_1(75)$ determined using increasing periods of time sampled from the catalogue, and it is clear that although statistical stability increases with the sample period, it is effectively stable over the whole range of sample periods used.

However, values of ω for almost all regions are higher in Tables 2 and 3 than that for the world as a whole, but when individual values of σ_ω are taken into account all are compatible with a World ω in the range 9.0–9.5 using surface wave magnitude, M_s , data. The circum-Pacific belt is the most active in the world with observed magnitudes as high as 8.9 and no region with an observed maximum less than 8.6. A magnitude range of 9.0–9.5 as an upper bound to future events seems realistic; the influence of magnitude saturation is discussed below. It should be noted that ω is obtained as one parameter of a distribution, fitted to data, designed for seismic risk estimating whereas M_3 is obtained as a direct measure of the limit to potential strain energy release. Figures 2(a) and (b) show that the regional Gumbel III curves are upper bounded by the world curve and lower bounded by Region 7 (New Zealand, Tonga, Kermadec), which shows the lowest seismicity and least number of shallow earthquakes of all the seven regions.

3.3. Regional variations in seismicity

Prediction or forecasts. Given the established statistical stability of the data, the Gumbel III distribution derived from the longest available period (1897–1975) may now be used to establish a contemporary view of the seismic risk equivalent to prediction of future occurrences of large earthquakes if time invariance is accepted. Note that it is prediction or forecasting using the entire Gumbel III distribution which is relevant, rather than individual parameters ω , u , or λ . Table 5 lists modal magnitudes expected to be exceeded once during the next 1, 10, 20, 50 and 100 years in each of the seven circum-Pacific regions and in the World as a whole.

Table 5

Predicted most probable largest earthquake magnitude (mode $m_1(T)$) for return periods T of 1, 10, 20, 50 and 100 years

Region		Return period (years)				
		1	10	20	50	100
(1) South America	a	7.2 ± .1	8.3 ± .1	8.5 ± .1	8.8 ± .1	9.0 ± .2
	b	7.2 ± .1	8.3 ± .1	8.5 ± .1	8.7 ± .1	8.9 ± .1
(2) North America	a	7.4 ± .1	8.3 ± .1	8.5 ± .1	8.6 ± .2	8.7 ± .2
	b	7.4 ± .1	8.3 ± .1	8.4 ± .1	8.6 ± .1	8.7 ± .2
(3) Aleutians, Alaska	a	7.0 ± .1	8.2 ± .1	8.4 ± .1	8.7 ± .1	8.9 ± .2
	b	7.0 ± .1	8.2 ± .1	8.4 ± .1	8.7 ± .1	8.8 ± .1
(4) Japan Kurile Kamchatka	a	7.6 ± .2	8.5 ± .1	8.7 ± .1	8.8 ± .1	8.9 ± .1
	b	7.6 ± .1	8.5 ± .1	8.6 ± .1	8.8 ± .1	8.9 ± .1
(5) N Guinea, Banda Sea Celebes, Moluccas Philippines	a	7.5 ± .1	8.5 ± .1	8.6 ± .1	8.8 ± .2	9.0 ± .2
	b	7.5 ± .1	8.4 ± .1	8.6 ± .1	8.8 ± .1	8.9 ± .1
(6) N. Hebrides Solomon N. Guinea	a	7.4 ± .1	8.2 ± .1	8.4 ± .1	8.6 ± .2	8.7 ± .3
	b	7.4 ± .1	8.2 ± .1	8.4 ± .1	8.5 ± .2	8.7 ± .2
(7) N. Zealand Tonga Kermadec	a	7.2 ± .1	8.2 ± .1	8.4 ± .1	8.5 ± .1	8.6 ± .2
	b	7.2 ± .1	8.2 ± .1	8.3 ± .1	8.5 ± .1	8.6 ± .1
World	a	8.3 ± .1	8.8 ± .1	8.9 ± .1	9.0 ± .1	9.0 ± .1
	b	8.3 ± .1	8.8 ± .1	8.9 ± .1	9.0 ± .1	9.0 ± .1

a: Using parameters estimated from sample period: 1897–1964
 b: Using parameters estimated from sample period: 1897–1975^{epx}

Magnitude saturation of the surface-wave scale has been pointed to as a possibility by KANAMORI (1978) for great earthquakes with fault lengths exceeding 60 km, and any impact of this on the above forecasts can be easily estimated. Kanamori found four giant earthquakes in the circum-Pacific belt which on his M_w scale exceed 9.0:

Chile (Region 1)	1960 May 22	9.6 M_w	(8.3 M_s)
Alaska (Region 3)	1964 March 28	9.2 M_w	(8.4 M_s)
Aleutian Islands (Region 3)	1957 March 9	9.1 M_w	(8.25 M_s)
Kamchatka (Region 4)	1952 November 4	9.0 M_w	(8.4 M_s)

The largest discrepancy in $M_w - M_s$ is given by the Chile earthquake and re-analysing Region 1 using 9.4 M_w as the largest extreme value generates an (ω, u, λ) set of $(11.05 \pm .34, 7.07 \pm .03, .158 \pm .018)$ which generates modal earthquakes: $m_1(1) = 7.2 \pm .1$, $m_1(10) = 8.4 \pm .1$, $m_1(50) = 8.9 \pm .2$, and $m_1(100) = 9.2 \pm .2$ for one, ten, 50

and 100 years respectively. Comparing these results with Table 5 shows they do not differ significantly from those obtained without $9.6 M_w$ for the 1960 earthquake. Considering that the adjustment $8.3 M_s$ to $9.6 M_w$ is the most dramatic adjustment to magnitude indicates that these few cases of saturation do not produce a significant bias in the prediction procedure using Gumbel III. Although the prediction capability of Gumbel III is not likely to show significant bias over none infinite forecasting durations using M_s it is clear for very large magnitude earthquakes that $M_w > M_s$. A real difficulty arises in that simple correlations between M_s and M_w will not provide reliable and consistent results for all available M_s : KANAMORI'S (1977) original data show at least as many decreases in M_w as there are increases when compared with the corresponding M_s value. Scaling laws related to earthquake spectra and physical dimensions of earthquake fault length had previously demonstrated (AKI, 1972) divergence between surface wave magnitudes, M_s , and body wave magnitudes, m_b , compared with the 'ω-square' model of earthquake spectra. Divergence increases with increasing M_s and ultimately the seismic moment, M_0 , takes over from M_s as a better representation of large earthquake size. Until complete seismic moment estimates are available for a lengthy period, thus facilitating complete M_w catalogues (although the analysis would then be performed preferably on a simple function of M_0), a more refined analysis will not be forthcoming. Estimates of upper bounds to earthquake magnitude, rather than seismic risk forecasts, might be modified for example by converting M_3 to M_w using any acceptable relation between M_s and M_w at large values of M_s only.

Regional seismicity. Several brief conclusions may be drawn from Table 5 on the regional seismicity:

- i) The annual mode for all the regions is greater than $m = 7.0$ with the maximum being in Region 4 (Japan, Kurile, Kamchatka) with $m(1) = 7.6$.
- ii) During the next 10 years (after 1975) a maximum magnitude earthquake exceeding $m(10) = 8.2$ is expected in almost every region in the circum-Pacific belt. This may be as high as $m(10) = 8.5$ for Region 4. Likewise, for the next 20 years a maximum magnitude earthquake is expected which may exceed 8.3 (Regions 6 and 7), 8.4 (Regions 2 and 3), 8.5 (Regions 1 and 5) and 8.6 (Region 4).
- iii) The regions in which events with predicted maximum magnitude expected to exceed 8.8–8.9 during the next 100 years are: Region 1, Region 4 and Region 5. These regions are situated in the northwestern (Region 4 and Region 5) and southeastern (Region 1) part of the circum-Pacific belt, diagonally opposite each other. This is compatible with the results of Paper I using the strain energy release method.

4. Conclusions: strain energy release and Gumbel III compared

For each of the seven regions both the annual mode, and the magnitude which

Table 6

Comparison between the parameters derived from the strain energy release and third type asymptotic methods of Gumbel III

Region	M_1	$m_1(1)$	M_2	X_2	M_3	$\omega \pm \sigma_\omega$	ρ_3
1	7.25	7.21	8.03	8.06	9.10	10.16 ± 1.2	0.07
2	7.34	7.37	7.94	7.95	9.00	9.14 ± 0.5	0.03
3	7.07	7.00	7.89	7.93	8.78	9.66 ± 0.6	0.09
4	7.52	7.61	8.20	8.16	8.96	9.30 ± 0.4	0.07
5	7.48	7.53	8.12	8.16	9.04	10.00 ± 1.1	0.02
6	7.35	7.35	7.90	7.91	8.83	9.44 ± 1.1	0.08
7	7.08	7.19	7.86	7.78	8.97	8.95 ± 0.4	0.05
World	8.09	8.30	8.68	8.49	9.52	9.23 ± 0.25	0.03

corresponds to the mean annual energy release, are calculated using (16) and (26) respectively. These are tabulated in Table 6 with the upper limit ω . This table also lists values of M_1 , M_2 and M_3 , calculated in Paper I for each region. A comparison can be made between M_1 and $m_1(1)$, M_2 and X_2 , whereas M_3 is expected to be within the range of ω , although less than ω .

Table 6 reveals some of the most significant features of this study. From the remarkably similar results for M_2 and X_2 as well as for M_1 and $m_1(1)$ we can draw several conclusions. The relations obtained between the parameters of the two different procedures used to describe the same phenomenon, are valid over a wide range of seismotectonic environments. Equation (16) simply records the annual mode determined from the Gutenberg–Richter frequency-magnitude and Gumbel III distributions. More importantly (26) provides a direct link between the physical process of strain energy release and the parameters of Gumbel III, expressed through the mean annual energy release in the region under investigation. In all regions M_3 is less than ω (Region 7 being an exception where $M_3 \approx \omega$), although this is not statistically significant in an individual region: this is compatible with our physical interpretation of ω and M_3 in that ω corresponds to a magnitude with a theoretically infinite return period (a weakness of Gumbel III) whereas M_3 is seen to have (Paper I) a finite ‘waiting time’ between expected occurrences. This physical interpretation is strengthened further by testing the means of regional $\omega - M_3$ values against an assumed mean of zero; this may be rejected beyond the 1% significance level demonstrating that ω systematically exceeds M_3 as indicated in (27) for the regions. A footnote to this is provided by considering all great earthquakes for the world on a non-regional basis when it is seen that calculation of the potential limit to strain energy release with a finite waiting time between such events now becomes similar to ω .

Region 3 has its own significance. It shows the highest value of reduced chi-squared ρ_3 , that is the worst fit, of all the regions examined. DUDA (1965) also found this to be the case using the Gutenberg–Richter frequency-magnitude formula and

commented 'this may be caused by the superposition of two natural populations of earthquakes'. Gumbel III for Region 3 (Aleutians, Alaska) might be similarly influenced by such a seismic feature producing this relatively poor fit, despite the fact that this distribution involves fitting using three rather than two parameters in the distribution.

The centres of highest seismic activity in the circum-Pacific belt are diagonally opposite each other, and this presumably relates to the tectonic movement of the Pacific plate (DUDA, 1965). The seismicity of this region is expected to generate great earthquakes which may exceed 8.2 in almost every region in the circum-Pacific belt during 10 years. Regions 1, 4, and 5 are the regions in which an earthquake with magnitude 8.8 to 8.9 is expected to be exceeded at least once during 100 years.

In general we conclude that the methodology of the Gumbel III asymptotic extreme value distribution described here, and the strain energy release method described in Paper I, provide a mutually compatible description of the seismic features of a region (which here is one of very high seismicity). In particular we conclude that X_2 of (26) gives the mean annual rate of seismic strain energy release in a region in terms of the Gumbel III parameter set (ω, u, λ) ; thus compatibly linking Gumbel III with parameters of the physical process. Finally, the use of complete earthquake data sets in both Paper I and here leads to compatible results between whole- and part-process statistics, this demonstrates clear consistency with the stability postulate from which asymptotic distributions of extremes are deduced and provides further confidence in this increasingly used method.

Acknowledgments

KCM is grateful to Professor J. Drakopoulos for leave of absence from the University of Athens. The work of PWB was supported by the Natural Environment Research Council and is published with the approval of the Director of the Institute of Geological Sciences (NERC).

REFERENCES

- AKI, K. (1972), *Scaling law of earthquake source time-function*. Geophys. J. R. astr. Soc., 31, 3–25.
- BÄTH, M. (1958), *The energies of seismic body waves and surface waves*. In: H. Benioff, M. Ewing, B. F. Howell Jr. and F. Press (Editors), *Contributions in Geophysics*, Pergamon, London, 1, 1–16.
- BEVINGTON, P. R. (1969), *Data reduction and error analysis for the physical sciences*. McGraw-Hill, New York, 336pp.
- BURTON, P. W. (1978a), *The application of extreme value statistics to seismic hazard assessment in the European area*. Proc. Symp. Anal. Seismicity and on Seismic Risk, Liblice, 17–22 October 1977. Academia, Prague 1978, 323–334.
- BURTON, P. W. (1978b), *The IGS file of seismic activity and its use for hazard assessment*. Inst. Geol. Sciences, Seism. Bul. No. 6, 13pp.

- BURTON, P. W. (1979), *Seismic risk in Southern Europe through to India examined using Gumbel's third distribution of extreme values*. Geophys. J. R. astr. Soc., 59, 249–280.
- BURTON, P. W. (1981), *Variation in seismic risk parameters in Britain*, In: Proc. 2nd Int. Symp. Anal. Seismicity and on Seismic Risk. Liblice, Czechoslovakia, 1981 May 18–23, Czechoslovak Academy of Sciences, 2, 495–531.
- BURTON, P. W., MCGONIGLE, R. W., MAKROPOULOS, K. C. and UCER, S. B. (1984), *Seismic risk in Turkey, the Aegean and the eastern Mediterranean: the occurrence of large magnitude earthquakes*. Geophys. J. R. astr. Soc., 78, 475–506.
- CORNELL, C. A. and VANMARCKE, E. H. (1969), *The major influences on seismic risk*. 4th World Conf. on Earth. Engineering, Chile.
- DUDA, S. J. (1965), *Secular seismic energy release in the Circum-Pacific Belt*. Tectonophysics, 2, 409–452.
- EPSTEIN, B. and LOMNITZ, C. (1966), *A model for the occurrence of large earthquakes*. Nature, 211, 954–956.
- ESTEVA, L. (1976), *Seismicity*. In: Seismic Risk and Engineering Decisions, Lomnitz, C. and Rosenblueth, E., editors, Elsevier Scient. Publ. Comp., Amsterdam, 425pp.
- FRECHET, M. (1927), *Sur la loi de probabilité de l'écart maximum*. Ann. de la Soc. polonaise de Math, Cracow, 6, 93–116.
- GRINGORTEN, I. I. (1963a), *Envelopes for ordered observations applied to meteorological extremes*. J. Geophys. Research, 68, 815–826.
- GRINGORTEN, I. I. (1963b), *A plotting rule for extreme probability paper*. J. Geophys. Research, 68, 813–814.
- GUMBEL, E. J. (1935), *Les valeurs extrêmes des distributions statistiques*. Ann. Inst. Henri Poincaré, 5, 115–158.
- GUMBEL, E. J. (1967), *Statistics of extremes*. Columbia Univ. Press, New York, 375pp.
- GUTENBERG, B. (1956), *Great earthquakes 1896–1903*. Trans. Am. Geophys. Union, 37, 608–614.
- GUTENBERG, B. and RICHTER, C. F. (1944), *Frequency of earthquakes in California*. Bull. Seism. Soc. Am., 34, 185–188.
- GUTENBERG, B. and RICHTER, C. F. (1954), *Seismicity of the earth and associated phenomena*. Princeton Univ. Press, Princeton, N.J., 310pp.
- JENKINSON, A. F. (1955), *The frequency distribution of the annual maximum (or minimum) values of meteorological elements*. Q. Jour. Roy. Meteor. Soc., 87, 158–171.
- KANAMORI, H. (1977), *The energy release in great earthquakes*. J. Geophys. Res., 82, 2981–2987.
- KANAMORI, H. (1978), *Quantification of earthquakes*. Nature, 271, 411–414.
- KARNIK, V. and HUBNEROVA, Z. (1968), *The probability of occurrence of largest earthquakes in the European area*. Pure appl. Geophys., 70, 61–73.
- KRUMBEIN, W. C. and LIEBLEIN, J. (1956), *Geological application of extreme value methods to interpretation of cobbles and boulders in gravel deposits*. Trans. Amer. Geophys. Union, 37, 313–319.
- LEVENBURG, K. (1944), *A method for the solution of certain non-linear problems in least squares*. Q. Appl. Math., 2, 164–168.
- LOMNITZ, C. (1974), *Global tectonics and earthquake risk*. Elsev. Scient. Publ. Comp. Amsterdam, 320pp.
- MAKROPOULOS, K. C. (1978), *The statistics of large earthquake magnitude and an evaluation of Greek seismicity*. Ph.D. thesis, University of Edinburgh.
- MAKROPOULOS, K. C. and BURTON, P. W. (1983), *Seismic risk of circum-Pacific earthquakes: I Strain energy release*. Pure appl. Geophys., 121, 247–267.
- MARQUARDT, D. W. (1963), *An algorithm for least-squares estimation of nonlinear parameters*. J. Soc. Indust. App. Maths., 11, 431–441.
- MERZ, H. and CORNELL, C. A. (1973), *Seismic risk analysis based on a quadratic magnitude-frequency law*. Bull. Seism. Soc. Am., 63, 1999–2006.
- NORDQUIST, J. M. (1945), *Theory of large values applied to earthquake magnitudes*. Trans. Amer. Geophys. Union, 26, 29–31.
- RADU, C. and APOPEI, G. (1977), *Application of the largest values theory to Vrancea earthquakes*. Inst. Geophys., Polish Acad. Sci., A5 (116), 229–243.
- RICHTER, C. F. (1958), *Elementary Seismology*. Freeman and Co., San Francisco, 766pp.
- ROCA, A., ARROYO, A. L. and SURINACH, E. (1984), *Application of the Gumbel III law to seismic data from southern Spain*. Eng. Geol., 20, 63–71.
- SACUIU, I. and ZORILESCU, D. (1970), *Statistical analysis of seismic data on earthquakes in the area of the Vrancea focus*. Bull. Seism. Soc. Am., 60, 1089–1099.

- SCHENKOVA, Z. and KARNIK, V. (1970), *The probability of occurrence of largest earthquakes in the European areas—Part II*. Pure appl. Geophys., 80, 152–161.
- SCHENKOVA, Z. and KARNIK, V. (1977), *Statistical prediction of the maximum magnitude earthquakes in the Balkan region*. Inst. Geophys., Polish Acad. Sci., A-5 (116), 245–250.
- SCHENKOVA, Z. and KARNIK, V. (1978), *The third asymptotic distribution of largest magnitudes in the Balkan earthquake provinces*. Pure appl. Geophys., 116, 1314–1325.
- WATSON, G. S. (1954), *Extreme values in samples from m -dependent stationary stochastic processes*. Ann. Math. Stats., 25, 798–800.
- YEGULALP, T. M. and KUO, J. T. (1974), *Statistical prediction of the occurrence of maximum magnitude earthquakes*. Bull. Seism. Soc. Am., 64, 393–414.

(Received 21st February 1984, revised/accepted 17th June 1986)
