ISET Journal of Earthquake Technology, Paper No. 388, Vol. 36, No. 1, March 1999, pp. 15-26

APPLICATION OF EXTREME-VALUE DISTRIBUTION FOR ESTIMATING EARTHQUAKE MAGNITUDE - FREQUENCY RELATIONSHIPS

R.D. Sharma, Sushil Gupta and Sanjeev Kumar
Seismology Group, Nuclear Power Corporation,
Vikram Sarabhai Bhavan, Anushaktinagar, Mumbai – 400094

ABSTRACT

The magnitude–frequency relationship of earthquakes has been investigated in view of its applications in seismic hazard assessment. For areas, where earthquake occurrences are frequent, the magnitude – frequency relationship can be established on the basis of regression analysis. For other areas, where earthquake occurrences are infrequent, quantification of seismicity has to be based on some assumptions and theoretical models. Application of the extreme event method helps in arriving at some estimates, which are verifiable in a few cases. This application is discussed in detail to illustrate method, which could be adopted for dealing with situations of inadequate data.

KEYWORDS: Magnitude-Frequency Relationship, Extreme Value Distribution, Seismic Hazard Assessment.

INTRODUCTION

Quantifying seismicity of a geographical region implies defining the distribution of the earthquake (magnitudes) population in space and time, mathematically. The Gutenberg's magnitude-frequency relation, namely:

$$\log_{10} N(M) = a - bM \quad (1)$$

has been commonly used for this purpose (Richter,1958). Here $N(M)$ is the number of earthquakes of magnitude equal to or greater than M occurring annually in area, and a and b are constants called seismicity parameters of the region. Values of the parameters a and b can be estimated from data on past earthquakes. In the intermediate range of magnitudes (say, between 5 and 7) global earthquake occurrences can be described, reasonably well, by such a relationship. Outside this range the number of observed earthquakes is not as large as predicted by the above relationship (Oliver et al., 1966; Shlien and Toksoz, 1970; Merz and Cornell, 1973). The effect of this deviation on the probabilities of ground motion is, relatively, small because of the negligible effects (in terms of hazardous ground motions) of smaller earthquakes on the one hand and the very large recurrence intervals of large earthquakes on the other.

Magnitude–frequency relationships for earthquakes provide a basis for applying statistical techniques to seismicity issues, under the basic assumption that such a relationship is also applicable on regional basis (Kaila and Narain, 1971). For some areas, where frequencies of past earthquake occurrences can be subjected to regression analysis, reliable estimates of a and b can be obtained. For other areas, where earthquake data are inadequate for carrying out such analysis, seismicity parameters will have to be based on some assumptions. Possibilities of quantifying the seismicity of such areas are examined in this paper.

GLOBAL EARTHQUAKE MAGNITUDE-FREQUENCY DATA

Frequencies of earthquakes grouped in different magnitude ranges for a period of 32 years (1964-95) are listed in Table 1. Values of $\log_{10} N(M)$ for these data are plotted against those of M in Figure 1. The $\log_{10} N(M)$ and M relationship of the type of Equation (1) fitted to these data gives the following magnitude-frequency relation for these data:

$$\log_{10} N(M) = 8.44 - 1.06M \quad (2)$$

Table 1: Frequencies of Earthquakes in Different Magnitude Ranges for Global Data (NOAA Earthquake Data File)

Year	Magnitude								Total
	4.5-4.9	5.0-5.4	5.5-5.9	6.0-6.4	6.5-6.9	7.0-7.4	7.5-7.0	≥ 8	
1964	1473	777	273	87	19	0	0	1	2630
1965	2040	1123	408	98	56	9	3	0	3737
1966	1699	911	216	56	25	4	2	0	2913
1967	1637	903	236	75	27	1	1	0	2880
1968	1774	974	270	63	15	9	3	0	3108
1969	1698	1026	264	74	26	14	3	1	3106
1970	1631	950	266	80	38	15	5	0	2985
1971	1643	1065	276	78	35	13	6	1	3117
1972	1528	1021	299	85	25	12	3	0	2973
1973	1584	1038	299	65	30	6	7	0	3029
1974	1669	1052	263	69	30	12	2	0	3097
1975	1809	1151	309	71	37	7	7	1	3392
1976	1858	1324	340	82	32	12	3	2	3653
1977	1825	1338	330	71	18	8	3	2	3595
1978	1729	1207	318	68	25	9	7	1	3364
1979	1877	1112	254	71	25	8	5	0	3352
1980	2095	1060	239	73	32	11	2	1	3513
1981	2027	950	228	48	17	8	4	0	3282
1982	2085	1164	250	56	15	4	2	0	3576
1983	2319	1376	331	70	18	5	3	0	4122
1984	2458	1305	284	47	11	5	1	0	4111
1985	2798	1408	284	47	21	9	6	1	4574
1986	2776	1369	284	47	11	3	2	1	4493
1987	2490	1131	297	55	22	7	4	1	4007
1988	2451	1184	290	48	16	6	2	0	3997
1989	2495	1170	254	51	13	4	1	1	3989
1990	2724	1303	308	60	16	8	3	0	4422
1991	2534	1224	240	63	19	9	1	0	4090
1992	2721	1201	321	57	29	8	4	0	4341
1993	2640	1165	249	65	20	8	2	1	4150
1994	2538	1216	289	83	35	7	3	2	4173
1995	2395	689	223	53	19	9	5	1	3394
Total	**67020**	**35887**	**8992**	**2116**	**777**	**250**	**105**	**18**	**115165**

Figure 1 represents an average relationship for the entire area of the earth's surface, which includes areas of varying seismicity. If the global data is subdivided into such categories, different magnitude-frequency relationships will emerge for different regions (Sharma, 1989). While it is possible to derive reasonably representative expressions for the highly seismic and moderately seismic areas, data from areas of infrequent earthquake occurrences cannot be subjected to reliable regression analysis for determining the magnitude-frequency relationship.

APPLICATION OF EXTREME EVENT ANALYSIS

Application of extreme event analysis method has also been proposed for earthquake occurrences (Nordquist, 1945; Gumbel, 1958; Epstein and Lomnitz, 1966; Lomnitz, 1976). According to this formulation, the annual highest earthquake magnitude in a specified region is distributed according to the cumulative distribution function $G(Y)$ given by:

$$G(Y) = \exp[-\alpha \exp(-\beta Y)] \tag{3}$$

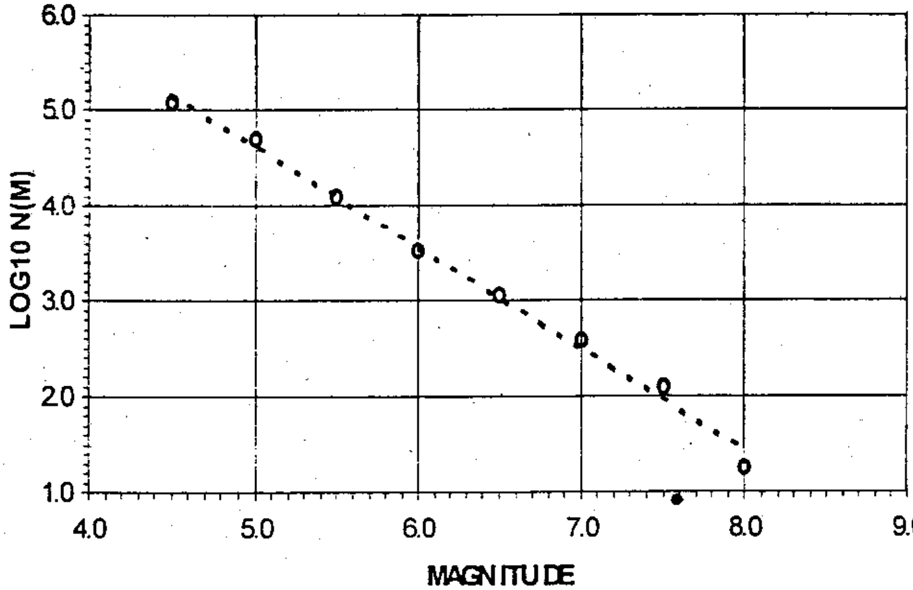

Fig. 1 Global $\log_{10} N(M)$ versus M values of 32 years data from Table 1

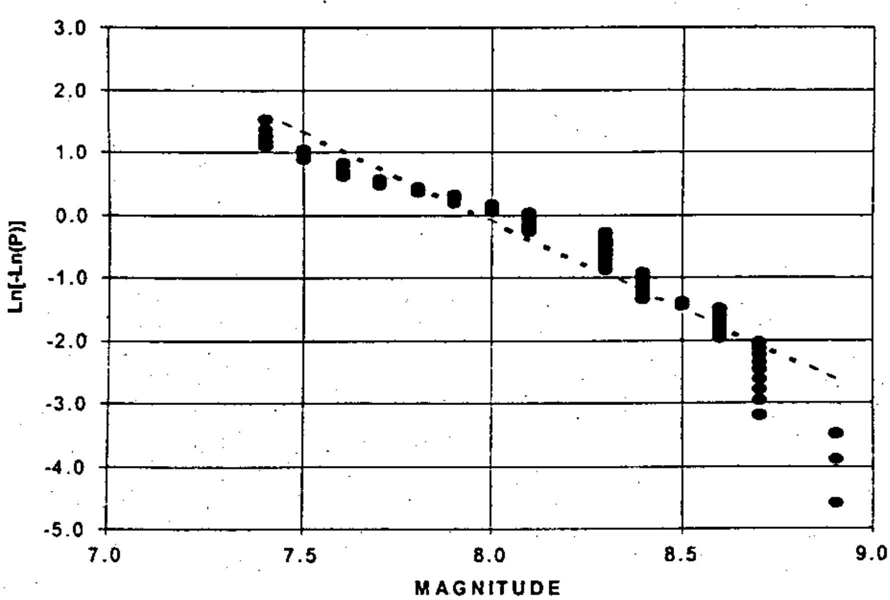

Fig. 2 Maximum yearly magnitude value versus Ln[-Ln(P)] from NOAA (1976,1996)

The parameters α and β are related to the constants a and b of Equation (1) as follows:

$$\alpha = \exp[a \log_e 10] \text{ and } \beta = \exp[b \log_e 10] \tag{3a}$$

For the annual maximum magnitudes Y_j $(j = 1, \ldots, N)$ arranged in an ascending order, the probability P_j of not exceeding the jth magnitude is given by:

$$P_j = G(Y_j) = j/(N+1) \tag{4}$$

The values of α and β are estimated from a least square fit of the $G(Y_j)$ and Y_j values to the equation:

$$\log_e[-\log_e\{G(Y)\}] = \log_e \alpha - \beta Y \tag{5}$$

The slope and the intercept, respectively, provide the estimates of β and $\log_e \alpha$. The left-hand side of Equation (5) has been plotted in Figure 2 against the global yearly maximum earthquake magnitudes between 1897 and 1995 (NOAA, 1976, 1996).

A straight line fitted to these data gives the following magnitude-frequency relationship for global earthquake occurrences:

$$\log_{10} N(M) = 9.86 - 1.23M \tag{6a}$$

The values of the parameters a and b in Equation (6a) are relatively higher compared to those of Equation (2). This results from the deviation (due to a systematic error component) of the data points from a single straight line. A straight line fitted to the points in the apparently linear portion (nearly 90 points) gives the following relationship:

$$\log_{10} N(M) = 8.24 - 1.03M \tag{6b}$$

These estimates are fairly close to those of Equation (2). It is apparent from Figure 2 that the slope of the fitted straight line as well as its intercept with the frequency (probability) axis varies in different parts of the data. Accordingly, the two parameters a and b also vary substantially. This behaviour of the data remains with any change in the length of estimation interval, the duration for which the maximum yearly magnitudes are analysed. One important observation, which comes out of the analysis, is that higher b values are associated with higher a values and vice-versa. This latter feature is, at least partly, the result of the formulation of the magnitude-frequency relationship itself, namely: if one of the parameters is underestimated or overestimated the estimation error is directly transmitted to the other parameter, some people call it correlation. There may also be a correlation between the values of the two parameters, as inferred by some authors earlier (Karnik, 1964; Chauhan et al., 1968; Ram and Rathor, 1970; Kaila and Narain, 1971; Kaila et al., 1972; Rao and Rao, 1984; Srivastava and Dattatrayam, 1986). For the three estimates given above (Equations (2), (6a) and (6b)), the a/b ratio comes close to 8.

In the global data, on which the estimates of Equations (6a) and (6b) are based, the yearly maximum magnitudes vary between 7.4 and 8.9. We synthesised 99 random numbers on a digital computer (using standard random number routine from the library) in the range, 7.4-8.9, and used these to represent the yearly maximum magnitudes for 99 years. The magnitude-frequency relation estimated from these data is

$$\log_{10} N(M) = 9.897 - 1.134M \tag{6c}$$

The a and b values were also, then, estimated for different estimation intervals. This was done by, first putting the consecutive first 60 magnitude values out of the 99 simulated ones (representing an estimation interval of 60 years) to extreme value analysis to determine the a and b values. The estimation window was then shifted by one year to exclude the first value and to include the 61st value, and again the values of a and b were estimated. The process was repeated until the 99th magnitude value was included in the estimation window. The maximum, minimum and mean values and the standard deviation of the 40 values of a and b were then estimated for the sixty year estimation interval. These values are listed in the first row of Table 2. The estimation interval was then increased to 61 years, thereby reducing number of a and b values in the set to 39 (second row of Table 2). The estimation interval was increased in this manner until the number of a and b values to be averaged remained above 10 (below this number, the computed statistics, mean and standard deviation, are likely to be rendered meaningless). Table 2 shows (from the left, next to the serial number) values of the estimation interval, number of observations in one set of calculations, maximum, minimum and average values of the a and b parameters, their standard deviations, and the a/b ratio for one set of calculations (characterised by the initialisation of the random number).

The trends described above come out quite clearly. The analysis is suggestive of the following:

1. It is not possible to get unique values of the seismicity parameters a and b using the extreme event analysis even if they exist. There may be several reasons for this, e.g., the tendency of the maximum yearly magnitudes to bunch near certain values (the effect being more pronounced towards the higher magnitudes), saturation of the magnitude scale, accuracy of magnitude determination and limitations of the method of analysis, which uses the least–square method and delivers biased estimates.
2. The a and b values estimated from a synthetic data set comprising randomly distributed numbers in the range of the observed maximum yearly magnitudes are very similar to those obtained for real

data. It is, therefore, logical to assume that the global highest annual magnitude is randomly distributed.

3. For a specified data set, the a and b values vary with estimation interval or starting points. However, the a/b ratio remains practically constant. This is because of the fact that an exponential distribution has been used in the derivation.

Table 2: Values of Seismicity Parameters Estimated from Yearly Maximum Earthquake Magnitude for Different Estimation Intervals and their Statistical Properties

S. No.	Est._ Int.	No._ Obs.	a_{max}	a_{min}	a_{ave}	a_{sigma}	b_{max}	b_{min}	b_{ave}	b_{sigma}	a/b
1	60	40	9.184	8.203	8.626	0.295	1.159	1.030	1.086	0.039	7.946
2	61	39	9.265	8.252	8.637	0.283	1.169	1.037	1.087	0.037	7.946
3	62	38	9.215	8.329	8.647	0.277	1.164	1.047	1.088	0.037	7.947
4	63	37	9.263	8.301	8.657	0.265	1.170	1.042	1.089	0.035	7.947
5	64	36	9.342	8.219	8.665	0.257	1.181	1.031	1.090	0.035	7.948
6	65	35	9.257	8.229	8.675	0.239	1.172	1.031	1.091	0.032	7.948
7	66	34	9.114	8.285	8.680	0.218	1.151	1.039	1.092	0.030	7.949
8	67	33	9.085	8.320	8.691	0.205	1.147	1.042	1.093	0.028	7.949
9	68	32	9.104	8.388	8.703	0.197	1.149	1.050	1.095	0.027	7.949
10	69	31	9.048	8.440	8.713	0.189	1.140	1.057	1.096	0.026	7.949
11	70	30	9.033	8.385	8.721	0.187	1.141	1.053	1.097	0.026	7.949
12	71	29	9.060	8.389	8.731	0.188	1.143	1.053	1.098	0.026	7.949
13	72	28	9.095	8.335	8.741	0.192	1.148	1.047	1.100	0.027	7.949
14	73	27	9.066	8.388	8.750	0.188	1.146	1.054	1.101	0.026	7.950
15	74	26	9.046	8.414	8.756	0.181	1.144	1.057	1.101	0.025	7.950
16	75	25	9.110	8.490	8.762	0.175	1.152	1.066	1.102	0.025	7.951
17	76	24	9.175	8.521	8.766	0.175	1.160	1.072	1.102	0.024	7.952
18	77	23	9.196	8.552	8.764	0.164	1.162	1.076	1.102	0.023	7.953
19	78	22	9.127	8.537	8.757	0.148	1.155	1.075	1.101	0.020	7.954
20	79	21	9.031	8.531	8.751	0.128	1.142	1.074	1.100	0.017	7.954
21	80	20	8.943	8.493	8.753	0.117	1.129	1.070	1.100	0.015	7.954
22	81	19	8.945	8.543	8.761	0.107	1.123	1.077	1.102	0.013	7.953
23	82	18	8.946	8.562	8.769	0.098	1.121	1.078	1.103	0.011	7.953
24	83	17	8.978	8.631	8.780	0.091	1.126	1.087	1.104	0.010	7.952
25	84	16	8.902	8.623	8.793	0.080	1.118	1.088	1.106	0.009	7.951
26	85	15	8.937	8.671	8.812	0.069	1.123	1.094	1.108	0.007	7.950
27	86	14	8.924	8.686	8.835	0.065	1.120	1.095	1.111	0.007	7.949
28	87	13	8.986	8.753	8.860	0.068	1.128	1.102	1.115	0.008	7.948
29	88	12	8.957	8.811	8.888	0.056	1.126	1.110	1.118	0.006	7.947

It turns out, therefore, that from a given data set the two seismicity parameters cannot be determined uniquely, unless the value of one of the parameters is fixed from other sources. It may, however, be noted here that we are attempting to derive an average magnitude–frequency relationship from a range of magnitudes, over which a single exact relationship of the type of Equation (1) may not exist. Under these circumstances, it is difficult and probably not required to pinpoint the cause of the deviation. Determination of b value is possible with a limited data set. If a sensitive network of seismographs is operated in an area for some time and the b value is estimated with some accuracy, the a value can, then, be estimated from the data on annual yearly maximum magnitudes, real or synthetic, by fixing the a/b value for a given situation.

AN EXAMPLE: THE CASE OF PENINSULAR INDIA

The peninsular region in India had often been thought by many as an aseismic region, though earthquakes of moderate intensity have occurred in this region from time to time. The observed maximum magnitude has remained 6.5 or lower. The seismic signal detection capability, which has existed for the area even in the post instrumental period, cannot be used very effectively to tabulate the maximum yearly magnitudes on yearly basis very reliably. Whereas an upper limit on the highest yearly maximum magnitude can be placed safely at 6.5, the lower limit cannot be stated with confidence. The magnitude of the Koyna earthquake of December 10, 1967, the Latur earthquake of September 30, 1993 and the Jabalpur earthquake of May 22, 1997 support the upper limit. We assumed, tentatively, the lower limit on the maximum yearly magnitude at 4.5, generated 100 random numbers in this range (on a digital computer) for yearly maximum magnitudes and estimated the a and b values in the manner described above (see Table 3 and Figure 3). The estimated magnitude-frequency relationship is given by:

$$\log_{10} N(M) = 4.35 - 0.83M \qquad (7)$$

The b value is reasonably close to observations for the Koyna region (Guha et al., 1968; Rao and Rao, 1984; Gupta et al., 1997) indicating that our assumption about the range of the yearly maximum magnitude is close to the acceptable one. The a value is applicable to the entire region for which the highest and lowest values of earthquake magnitude have been assumed (scaling of the a value is necessary during any application).

Table 3: Values of Seismicity Parameters Estimated from Synthesized Yearly Maximum Earthquake Magnitudes for Different Estimation Intervals and their Statistical Properties Applied to Peninsular India

S. No.	Est._ Int.	No._ Obs.	a_{max}	a_{min}	a_{ave}	a_{sigma}	b_{max}	b_{min}	b_{ave}	b_{sigma}	a/b
1	80	20	4.421	4.186	4.324	0.061	0.847	0.803	0.825	0.011	5.239
2	81	19	4.434	4.210	4.327	0.057	0.842	0.807	0.826	0.010	5.238
3	82	18	4.432	4.222	4.331	0.054	0.841	0.809	0.827	0.008	5.237
4	83	17	4.446	4.256	4.336	0.051	0.844	0.815	0.828	0.008	5.237
5	84	16	4.403	4.244	4.342	0.045	0.838	0.816	0.829	0.007	5.235
6	85	15	4.418	4.267	4.351	0.039	0.842	0.821	0.831	0.006	5.234
7	86	14	4.414	4.277	4.362	0.037	0.840	0.822	0.834	0.005	5.233
8	87	13	4.446	4.310	4.373	0.038	0.846	0.826	0.836	0.006	5.231
9	88	12	4.428	4.341	4.386	0.032	0.844	0.833	0.839	0.005	5.229
10	89	11	4.460	4.343	4.396	0.030	0.850	0.832	0.841	0.004	5.227
Average Values					**4.35**				**0.83**		**5.234**

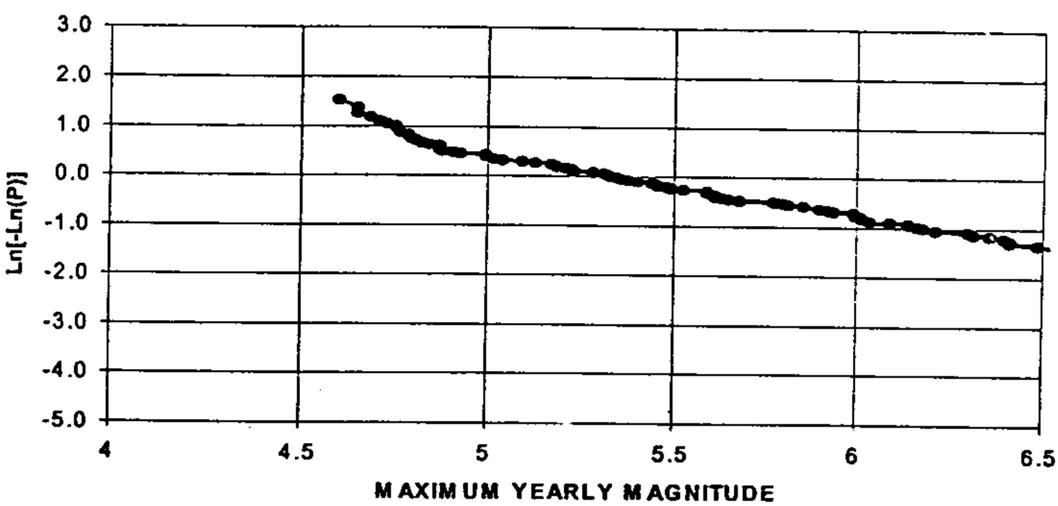

Fig. 3 Probability of yearly maximum magnitude value in peninsular India

These estimates appear far more stable as compared to those of Table 2 (see Figure 3). This is because of a relatively narrow range of magnitudes offering a better fit to the magnitude probability data. The values of the a, b parameters and the a/b ratio are remarkably close to those where the analysis has been based on actual data (Rao and Rao, 1984). If the range of the yearly maximum magnitudes is increased or decreased, the estimates deviate.

APPLICATIONS IN SEISMIC HAZARD ASSESSMENT

Earthquakes occurring at distances farther than a couple of hundred kilometres from a site are not likely to affect well-built structures adversely. For the purpose of estimating seismic hazard, which may be quantified in terms of the probabilities of different levels of the peak values of ground motion parameters (acceleration, velocity or displacement, though use of acceleration is common), a magnitude-frequency relationship derived in the manner described above can be assumed for a circular area of 300 km radius. The figure of 300 km is used in assessing seismic hazard at nuclear power plant sites (USNRC, 1980).

The probability of a specified value α of the peak ground acceleration (PGA) at a site being exceeded may be written as:

$$P(\alpha) = \int_{S=\alpha}^{S=S_{max}} \int_{M=M_o}^{M=M_{max}} P_1(M) P_2(S/M) dM\, dS \tag{8}$$

Here, $P_1(M)$ is the probability density function of M and $P_2(S/M)$ is the conditional probability density function of PGA for a given M. M_{max} is the upper bound on the earthquake magnitude in the region and S_{max} is the maximum value of PGA, which can be reached. The return period of the causative event, which produces PGA in excess of α is given by:

$$T(\alpha) = -L/\log_e[1.0 - P(\alpha)] \tag{9}$$

Here, L is the estimation period (e.g., the life of a structure under consideration, say 50 years). The probability of the event occurring during a period equal to the return period of the event is about 63% (Lomnitz, 1976). The level of PGA associated with an acceptably low probability (or high return period) event is a measure of seismic hazard in the area (or, at a construction site).

Evaluation of $P(\alpha)$ and $T(\alpha)$ requires the temporal and spatial distributions of earthquake magnitudes and a relationship to estimate the PGA value from an earthquake of specified magnitude and location. Unless the region has shown a definite observable seismicity pattern during historical times, the area is geologically well mapped (so that the possibility of discovering new faults can be ruled out) and association of earthquakes with geological faults is well understood, the probabilities and return periods of exceeding specified values of PGA can be estimated based on the assumption that earthquakes can occur anywhere and everywhere in the site region (Housner, 1975). Depending on the region and the knowledge of the seismicity distribution in the area under consideration, a more accurate model may be possible. However, if such an accuracy is feasible, we will not talk of the approximation discussed above. A regression analysis of the data should be used in that case. The PGA value from an earthquake may then be calculated using an applicable empirical relationship between earthquake magnitude, hypocentral parameters and PGA (Campbell, 1985). In calculations given below, we use the McGuire (1978) relationship, namely:

$$\alpha = 0.0306 \exp(0.89M) \exp(-0.2\beta) / R_h^{1.17} \tag{10}$$

for the purpose of illustration. Here α is the horizontal component of PGA (in g) from an earthquake of magnitude M at hypocentral distance R_h, $\beta = 1$ for soil sites and 0 for rock sites. Figure 4(a) shows the return periods of different levels of PGA for a typical rock site, based on Equations (7) and (9). An a value equal to 4.35 has been applied to an area of 300 km radius around the site. A fixed focal depth of 15 km has been assumed. Maintaining the average a/b ratio at 5.234 (see Table 3), the calculations were repeated with b value of 0.70 (giving a = 3.67). The corresponding return periods are shown in Figure 4(b). Results of similar calculations for the older a value, namely 4.35 and latter b value, i.e., 0.7 (in place of 0.83) are shown in Figure 4(c). Figure 4(a) has been plotted on two different scales to illustrate the behaviour of the curve in different ranges of PGA. The figures show how the return period of the causative event (which leads to a special value of PGA) increases, first gradually, and then rapidly. The plot in the upper figure brings out the concept of a low probability event: it is that event for which the

PGA value, α, is associated with a finite return period T, but for an event with a PGA value $\alpha + h$ (where h is small, tending to zero), the return periods tend to infinity. In engineering practice, an event having a return period of 10,000 years is considered as a low probability event. This is an over-simplification of the low probability concept. As long as the PGA value of 10,000 years return period remains significantly lower than any value associated with a higher return period (say 20,000 or 40,000 years), the choice of a 10,000 years return period PGA value for design offers no assurance towards safety. It turns out that more useful information in the figures lies where the return period begins to rise rapidly, reaching asymptotically a limiting value of PGA (represented by a broken line in the figure). For smaller values of PGA (which are associated with lower return periods), the curves show how the return periods are affected by the changes in the values of the seismicity parameters. For a given value of a, the return periods for specified value of PGA decrease with decrease in the b value implying increased seismic hazard (compare Figures 4(a) and 4(c)). For a given value of b, the return period for any specified value of PGA decreases with increase in the a value (compare Figures 4(a) and 4(b)), signifying again increased seismic hazard.

In the extreme event analysis approach, the values of the seismicity parameters, a, b and a/b are dependent on the lowest and highest values of the maximum yearly magnitude. The appropriate a value can be chosen only if the a/b and b values are known apriori or assumed, or the highest or lowest values of the annual maximum values are catalogued. Table 4 shows the results of some calculations to illustrate this point.

Table 4: Estimates A, B and a/b using Extreme Value Analysis for Different Values of Maximum Yearly Earthquake Magnitude over a Period of 99 Years

S. NO.	LOWEST VALUE OF YEARLY MAXIMUM MAGNITUDE	HIGHEST VALUE OF YEARLY MAXIMUM MAGNITUDE	$\bar{a}$	$\bar{b}$	a/b
1	3.0	6.5	2.0532	0.4805	4.2730
2	3.5	6.5	2.5738	0.4805	4.5912
3	4.0	6.5	3.3026	0.5606	4.9093
4	4.5	6.5	4.3957	0.6727	5.2274
5	5.0	6.5	6.2177	0.8409	5.5456
6	5.5	6.5	9.8600	1.1212	5.8637
7	6.0	6.5	20.7933	1.6818	6.1819
8	6.0	7.0	10.7025	3.3636	6.3637
9	6.0	7.5	7.3389	1.6818	6.5456
10	6.0	8.0	5.6571	1.1212	6.7274
11	6.0	8.5	4.6480	0.8409	6.9093
12	6.0	9.0	3.9753	0.6727	7.0912
13	6.5	9.0	4.9844	0.5606	7.4093
14	7.0	9.0	6.4980	0.6727	7.7274
15	7.5	9.0	9.0207	0.8409	8.0456
16	7.4	8.9	8.3900	1.1212	7.9820

It turns out, from the calculations, that the variation in synthesised magnitudes (introduced by changing the starting number for the random number generator) affects the values only marginally. One limitation of the analytical process arises from the arbitrariness in scaling the a value, which should be related to the size of the area under consideration. For instance, the a value of 8.43 in Equation (2) is applicable for the entire globe. This value normalised to a circular area of 300 km radius gives 3.26 for the a value of the area of this size and 3.07 for the ratio a/b. While considering seismicity for assessing seismic hazard on regional basis, an area of this size may be considered adequate because earthquakes having their sources farther than this are unlikely to have any adverse impact at a site. Equation (7) above may be considered valid for an area of this size.

DISCUSSION

The analysis in this paper is focussed on the earthquake problem in engineering design, where values of earthquake ground motion parameters (acceleration, velocity and displacement) are to be specified for a construction site. For areas of infrequent earthquake occurrences, adequate data on locations and magnitudes of past earthquakes are not available for applying deterministic procedures for fixing these parameters. It is for such areas that the formulation of the magnitude-frequency relationship on the basis of the extreme event analysis is considered. Applicability of a relationship of the type of Equation (1) is assumed, rather being proved. Such a formulation of seismicity firstly provides a baseline quantification of seismicity for estimating the earthquake design parameters and, secondly, helps in laying guidelines on any data collection exercise in the region since the collected data, to be useful, must lead to improvement over the baseline formulation. Peninsular India has been cited in the paper as an example of a region with inadequate data. Some estimates of a and b values are available, and these have been used to make some comparisons. The analysis does not, by any means, claim to give a prescription for the region. Nevertheless, any other values, if adopted for the region, must have reasons for becoming more acceptable.

In dealing with the earthquake problem in engineering design (or quantification of earthquake hazard assessment) seismological tools become the subject of review and discussion between specialists in different areas as well as generalists. Not infrequently, the discussions betray a communication gap. A linear (or non-linear) frequency-magnitude relationship means that $\log_{10} N(M)$ and M relationship is linear (or non-linear). The b value in Equation (1) determines the distribution of magnitude in the earthquake population and association of lower b values with one type of geological conditions (rock types) and higher b values with another type is in the realm of physics. Return period of the causative event (which, say, results in exceeding a certain specified value of ground motion) is helpful in understanding the data, and determining of this can be termed as a low probability event. In design, a 10,000 years return period event is often considered a low probability event. To a generalist, a return period of 10,000 years makes little sense when the used data set is of not more than a few hundred years. Figures 4(a) and 4(b) illustrate the role of return period. Firstly, return period is a property of the data set itself, and no extrapolation is implied. Secondly, a low probability event is associated with a high return period, tending to infinity for a finite value of the ground motion parameter in question. A 10,000 year return period can qualify as a low probability event, but only as long as the values of the ground motion parameter for the 10,000 year return period event are sufficiently close to the one for a 20,000 or a 40,000 year return period event. Also, choice of an 10,000 years return period event for design does not necessarily mean increased conservatism over the use of a 5,000 years return period event. The increased value of the design parameter, from one case to the other, may lead to only a marginal improvement in safety against earthquakes.

In this paper, an attempt has been made to demonstrate how to deal with the 'lack of data' situation in fixing the design parameters. A set of random numbers in the range of the highest and the lowest yearly maximum magnitudes and application of extreme event analysis to derive a magnitude-frequency relationship can be used to generate the baseline information for design, which is based on assumptions and approximations arising from the compulsions of the situation, and can be superseded when 'better' information becomes available.

CONCLUSIONS

1. In case of areas, where adequate data on magnitudes of past earthquakes to allow a regression analysis are not available, the extreme event analysis procedures can be applied to quantify seismicity in terms of magnitude–frequency relation.
2. For extreme event analysis to give reliable results, one of the parameters (the b value is a more likely candidate) must be known.
3. The ratio a/b is a better descriptor of seismicity of an area compared to a and/or b value.
4. Seismicity of an area can be quantified using the extreme event analysis method if the highest and the lowest yearly maximum magnitudes, and the b value can be fixed or assumed.

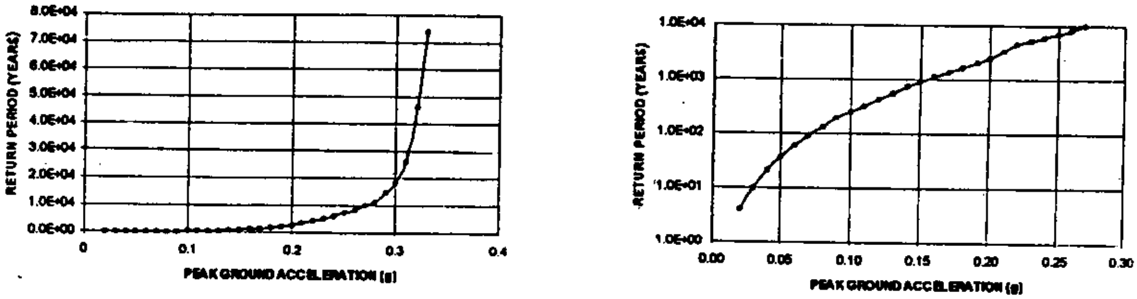

Fig. 4(a) Return period of exceeding PGA values ($a = 4.35$, $b = 0.83$)

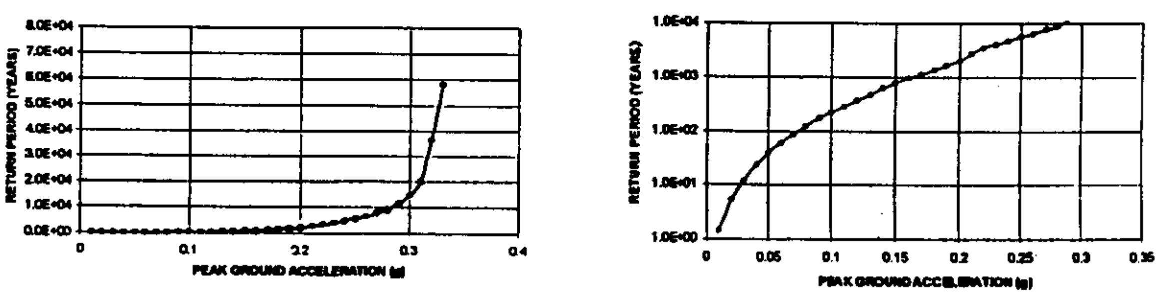

Fig. 4(b) Return period of exceeding PGA values ($a = 3.67$, $b = 0.70$)

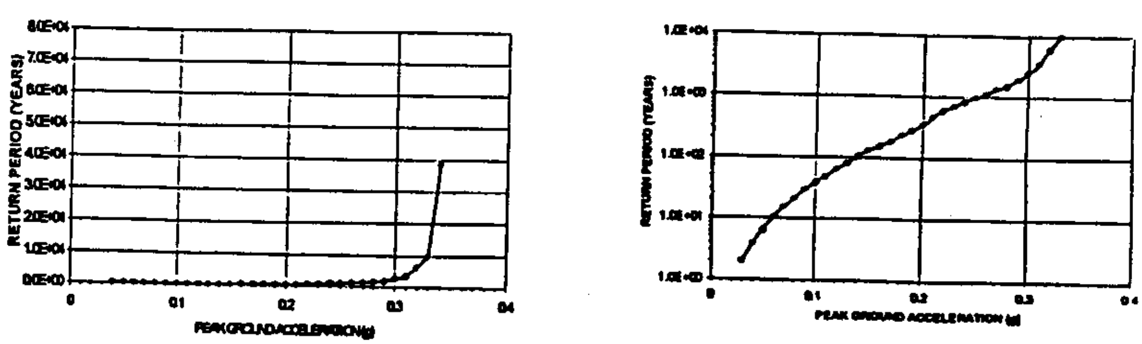

Fig. 4(c) Return period of exceeding PGA values ($a = 4.35$, $b = 0.70$)

ACKNOWLEDGEMENTS

We are grateful to Shri A. Sanatkumar, Director (Engg.), Nuclear Power Corporation for encouragement and permission to publish this paper. We thank Prof. C. Lomnitz for a detailed scrutiny of the paper, comments and suggestions and anonymous reviewers for their comments and suggestions to improve the manuscript.

REFERENCES

1. Campbell, K.W. (1985). "Strong Motion Attenuation Relations: A Ten-Year Perspective", Earthquake Spectra, Vol. 1, pp. 759-804.
2. Chauhan, R.K.S., Singh, C.L. and Singh, R.D. (1968). "A New Measure of Seismicity", Pure Appl. Geophys., Vol. 70, pp. 47-60.
3. Epstein, B. and Lomnitz, C. (1966). "A Model for the Occurrence of Large Earthquakes", Nature, Vol. 211, pp. 954-955.
4. Guha, S.K., Gosavi, P.D., Varma, M.M., Agarwal, S.P., Padale, J.G. and Marwadi, S.C. (1968). "Recent Seismic Disturbances in the Sivajisagar Lake Area of the Koyna Hydroelectric Project, Maharastra, India", Central Water and Power Research Station, Khadakwasla, Poona.
5. Gumbel, E.J. (1958). "Statistics of Extremes", Columbia University Press, N.Y., U.S.A.
6. Gupta, S., Bansal, B.K. and Kumar, S. (1997). "Seismic Activity of Koyna Region for the Period Sept., 1993 – Aug., 1996", Bull. Indian Soc. of Earthq. Tech., Vol. 34, pp. 209-218.
7. Housner, G.W. (1975). "Strong Ground Motion, Chapter-4 in Earthquake Engineering" : R.L. Weigel (Ed.), Prentice Hall, N.J., U.S.A.
8. Kaila, K.L. and Narain, H. (1971). "A New Approach for Preparation of Quantitative Seismicity Maps as Applied to Alpide Belt Sunda Arc and Adjoining Areas", Bull. Seism. Soc. Am., Vol. 61, pp. 1275-1291.
9. Kaila, K.L., Gaur, V.K. and Narain, H. (1972). "Quantitative Seismicity Maps of India", Bull. Seism. Soc. Am., Vol. 62, pp. 1119-1132.
10. Karnik, V. (1964). "Magnitude-Frequency Relations and Seismic Activity in Different Regions of the European Area", Bull. IISEE, Tokyo, Vol. I, p. 9.
11. Lomnitz, C. (1976). "Global Tectonics and Earthquake Risk", Developments in Geotectonics, Chapter 5, Elsevier Scientific Publishing Company, U.S.A.
12. McGuire, R.K. (1978). "Seismic Ground Motion Parameter Relations", Journal Geotech. Engg. Div., Am. Soc. Civ. Engs., Vol. 104, pp. 481-490.
13. Merz, K. and Cornell, C.A. (1973). "Seismic Risk Analysis Based on a Quadratic Magnitude–Frequency Law", Bull. Seism. Soc. Am., Vol. 63, pp. 1999-2006.
14. NOAA (1976). "Earthquake Data File Summary", US Deptt. of Commerce, National Oceanic and Atmospheric Administration Environmental Data Services, National Geophysical Data Centre, Boulder, Colorado, U.S.A.
15. NOAA (1996). "Seismicity Catalogs, Volume-2: Global and Regional, 2150 B.C.–1996 A.D.", National Geophysical Data Centre, Boulder, Colorado, U.S.A.
16. Nordquist, J.M. (1945). "Theory of the Largest Values Applied to Earthquake Magnitudes", Trans. Am. Geophy. Union., Vol. 26, pp. 29.
17. Oliver, J., Ryall, A., Brune, J.N. and Slemmons, D.B. (1966). "Microearthquake Activity Recorded by Portable Seismographs of High Sensitivity", Bull. Seis. Soc. Am., Vol. 56, pp. 899-924.
18. Ram, A. and Rathor, H.S. (1970). "On Frequency–Magnitude and Energy of Significant Indian Earthquakes", Pure and Applied Geophysics, Vol. 79, pp. 26-32.
19. Rao, B.R. and Rao, P.S.(1984). "Historical Seismicity of Peninsular India", Bull. Seis. Soc. Am., Vol. 74, pp. 2519-2533.
20. Richter, C.F. (1958). "Elementary Seismology", W.H. Freeman and Co., San Francisco, California, U.S.A.

21. Sharma, R.D. (1989). "Magnitude and Frequencies of Earthquakes in Relation to Seismic Risk", Bhabha Atomic Research Centre, Trombay, Mumbai, Publication BARC-1485.
22. Shlien, S. and Toksoz, M.N. (1970). "A Clustering Model for Earthquake Occurrence", Bull. Seis. Soc. Am., Vol. 60, pp. 1765-1787.
23. Srivastava, H.N. and Dattatrayam (1986). "Study of Return Periods of Earthquakes in Some Selected Indian and Adjoining Regions", Mausam, Vol. 37, pp. 333-340.
24. USNRC (1980). "Seismic and Geologic Siting Criteria for Nuclear Power Plants, Rules and Regulations, Title 10; Chapter 1", Code of Federal Regulation Energy, Part 100, Appendix–A.