Δελτίο Ελληνικής Γεωλογικής Εταιρείας τομ. XXV/3, 205-216, 1991
Πρακτικά 5ου Επιστημονικού Συνεδρίου, Θεσσαλονίκη, Μάϊος 1990

Bulletin of the Geological Society of Greece vol. XXV/3, 205-216, 1991
Proceedings of the 5th Congress, Thessaloniki, May 1990

EVIDENCE FOR A STOCHASTIC MODEL OF GLOBAL SEISMICITY

B. C. Papazachos*, T. M. Tsapanos*, E. M. Scordilis*, C. B. Bagiatis** and CH. H. Koukouvinos**

A B S T R A C T

Four partially independent sets of complete data concerning main shocks (excluding foreshocks and aftershocks) which occurred in the whole Earth during the time period 1897-1986, have been used to test the hypothesis that the time difference of successive earthquakes follows a negative exponential distribution (Poisson process). For samples with large ($n \geq 29$) or small number ($n < 20$) of events the X^2 and the Kolmogorov-Smirnov tests were applied, respectively. All four sets of data show that this time difference of the large shocks ($M > 7.0$) follows a negative exponential distribution while this does not hold for the smaller ($M \leq 7.0$) main shocks. An interpretation of this observation is attempted.

Σ Υ Ν Ο Ψ Η

Τέσσερεις ομάδες από ανεξάρτητα και πλήρη δεδομένα κυρίων σεισμών, που έγιναν σε όλη τη Γη κατά τη διάρκεια της χρονικής περιόδου 1897-1986, χρησιμοποιήθηκαν στην εργασία αυτή για να εξεταστεί η υπόθεση ότι, η χρονική διαφορά μεταξύ διαδοχικών σεισμών, ακολουθεί την αρνητική εκθετική κατανομή (κατανομή Poisson). Τα δείγματα που είχαν μεγάλο ($N \geq 29$) ή μικρό ($n < 20$) αριθμό σεισμών εξετάστηκαν με εφαρμογή των μεθόδων X^2 - test και Kolmogorov-Smirnov, αντίστοιχα. Και οι τέσσερεις ομάδες δεδομένων έδειξαν ότι η χρονική αυτή διαφορά για τους μεγάλους σεισμούς ($M > 7.0$) ακολουθεί την αρνητική εκθετική κατανομή, ενώ αυτό δεν ισχύει για τους μικοότερους σεισμούς. Γίνεται μιά προσπάθεια για ερμηνεία αυτών των παρατηρήσεων.

Β.Κ.ΠΑΠΑΖΑΧΟΣ-Θ.Μ.ΤΣΑΠΑΝΟΣ-Ε.Μ.ΣΚΟΡΔΥΛΗΣ-Κ.Β.ΜΠΑΓΙΑΤΗΣ-Χ.Χ.ΚΟΥΚΟΥΒΙΝΟΣ. Ενδειξη για ένα στοχαστικό μοντέλο παγκόσμιας σεισμικότητας.
* University of Thessaloniki, Geophysical Laboratory, Thessaloniki 54006, GREECE
**University of Thessaloniki, Department of Mathematics, Thessaloniki 54006, GREECE

INTRODUCTION

Although the application of statistical techniques in studying the time distribution of earthquakes is of great importance in seismology, since such studies may bring up new ideas about the patterns of earthquake occurrence, few detailed stochastic models have been developed to study this distribution. This is due to difficulties in obtaining complete and reliable data and to the fact that statistical ideas needed to handle complex point processes are still under developement (Vere-Jones 1970). However, simple stochastic models have been applied by several seismologists for such investigations.

The simplest stochastic model for studying the frequency of earthquakes (number of shocks per time unit) is the stationary Poisson process, in which the probability of occurrence of an event is the same for any elementary interval along the time axis. Equivalent to the distribution of the frequency of shocks in accordance with the Poisson distribution, is the distribution of the time intervals, t, between the events in accordance with a negative exponential distribution.

Some seismologists believe that seismic events occur according to a Poisson process and that if seismic events do not seem to follow a Poisson process it is a result of observational error or because the sequence of shocks is a modified Poisson process (Mogi 1962, Lomnitz 1966). However, other seismologists believe that there are cases when the Poisson process is not followed (Singh and Sanford 1972). It seems that different types of earthquake data present very different features (Vere-Jones 1970). Evidence has been presented by some seismologists that on a global scale, the very large earthquakes occur according to a Poisson process (Gutenberg and Richter 1954, Shlanger 1960).

Because there is still a question whether or not the occurrence of earthquakes follows a Poisson process and because the answer to this question is very important for seismic hazard and earthquake prediction, this problem needs further study. An attempt is made in the present paper to test the Poisson distribution hypothesis by using four partial independent complete sets of data which concern globally occurred earthquakes. For samples with large number of events the X^2 test has been applied while for small number of events ($n < 20$) the Kolmogorov-Smirnov test is applied.

SOURCE OF DATA AND METHODS APPLIED

Tsapanos (1985) used most of the available catalogues (Gutenberg and Richter 1954, Duda 1965, Rothe 1969, Miyamura 1976,1978, Abe 1979, 1981, Kanamori and Abe 1979, Abe and Kanamori 1979, Geller et al. 1978) as well as data published in the Bulletin of the International Seismological Center to make homogeneous catalogue of earthquakes which occurred in the whole Earth during the period 1904-1980. The magnitudes of the earthquakes in this catalogue are surface-wave magnitudes and its completeness was securred by dividing the whole period into four subperiods (1904-1929, 1930-1951, 1952-1965, 1966-1980) and choosing for each of them a proper minimum earthquake magnitude (7.0, 6.5, 6.0, 5.5). For the present paper additional data were collected to make a similar catalogue which covers the period 1897-1986. The time periods and magnitude ranges for the four complete sets of data used in the present paper are shown on Table (I)

Table I.-Time periods and corresponding magnitude ranges for the four complete data sets used in the present paper.

Πίνακας I. Χρονικές περίοδοι και αντίστοιχες κλίμακες μεγεθών για τις τέσσερεις πλήρεις ομάδες δεδομένων που χρησιμοποιήθηκαν στην παρούσα εργασία.

Time periods	Magnitudes
1897-1986	$M \geq 7.0$
1930-1986	$M \geq 6.5$
1952-1986	$M \geq 6.0$
1966-1986	$M \geq 5.5$

Since foreshocks and aftershocks are events which clearly depend on the main shock and the X^2 test cannot be applied in clustering process (Vere-Jones 1970), an effort was made to distinguish foreshocks and aftershocks from main shocks. Any earthquake was considered as foreshock if its location was within 100Km of the epicenter of the main shock and it had occurred within 40 days period of the main shock (Jones and Molnar 1976). An earthquake was considered as aftershock if it has occurred within 100 days after the main shock and its location was within a distance L from the epicenter of the main shock, where L is the dimension of the aftershock area and is related to the surface-wave magnitude of the main shock by an empirical formula given by Utsu (1969). Thus, the final catalogue which was used included only main shocks and no foreshocks and aftershocks.

Each one of the data sets (Table I) was separated in several samples by using a magnitude step equal to 0.1. Thus the first sample of the first set of data (1897-1986) includes the shocks with $M \geq 7.0$, the second sample of the same set includes the shocks with $M \geq 7.1$, and so on. The time difference between successive events were calculated in years. The data were tested twice in order to check the null hypothesis Ho, that the events (shocks) follow the Poisson distribution. The first test examined the hypothesis that the annual number of shocks with magnitude $M \geq C$, where C=7.0, 7.1,...,8.3, has the Poisson distribution:

$$P(x) = e^{-\lambda} \frac{\lambda^x}{x!} \qquad x = 0,1,2,3,.... \qquad (1)$$

Our samples, consisted of 90 observations, i.e the number of shocks that had happened in the years 1897 through 1986. For every sample the parameter λ was estimated by the sample mean value $\bar{x}$, i.e

$$\hat{\lambda} = \bar{x} \qquad (2)$$

which is an unbiased estimate of λ.

In each sample the data were grouped, with respect to their expected frequency, so that each group would have at least a frequency of 5 shocks. In Table II the number of groups one denoted by NG and in the same Table, v denotes the degrees of freedom for the appropriate X^2-test, recall that $v = NG-2$. In Table II the P-column denotes the probability of a type I error when rejecting the null hypothesis H_0. From the results we conclude that the number of annual shocks having a magnitude $M \geq 7.2$ has the Poisson distribution while shocks with magnitude $M \geq 7.0$ do not. As for the sample of magnitude $M \geq 7.1$ the probability of a type I error is less than 0.05, so we would rather reject the null hypothesis.

We also examined the same data with a second test, to check the hypothesis that the time difference between two successive shocks has the negative exponential distribution which is implied by the fact that the number of shocks has the Poisson distribution

$$f(t) = \lambda e^{-\lambda t} \qquad (3)$$

For every sample the parameter was calculated by a maximum likelihood estimate, i.e. (4)

$$\hat{\lambda} = \frac{1}{\bar{t}}$$

Table II. Results of the application of the χ^2 test to check the hypothesis that the annual number of shocks has the Poisson distribution.

Πίνακας II. Αποτελέσματα της εφαρμογής της δοκιμασίας χ^2 για τον έλεγχο της υπόθεσης ότι ο ετήσιος αριθμός σεισμών ακολουθεί την κατανομή Poisson.

A/A	M	n	$\hat{\lambda}$	v	NG	χ^2	$\chi^2_{v;0.05}$	$\chi^2_{v;0.01}$	Hypothesis	P
1.	7.0	1208	13.42	10	12	41.53	18.31	23.21	Rejected	$P < 0.0005$
2.	7.1	904	10.04	8	10	18.81	15.51	20.09		$0.01 < P < 0.05$
3.	7.2	724	8.04	9	11	10.38	16.92	21.67	Accepted	$0.30 < P < 0.40$
4.	7.3	568	6.31	8	10	3.98	15.51	20.09	"	$0.80 < P < 0.90$
5.	7.4	461	5.12	8	10	7.81	15.51	20.09	"	$0.40 < P < 0.50$
6.	7.5	392	4.36	8	10	10.91	15.51	20.09	"	$0.20 < P < 0.30$
7.	7.6	294	3.26	6	8	10.50	12.59	16.81	"	$0.10 < P < 0.20$
8.	7.7	236	2.62	5	7	10.83	11.07	15.09	"	$0.05 < P < 0.10$
9.	7.8	167	1.86	3	5	6.93	7.81	11.34	"	$0.05 < P < 0.10$
10.	7.9	115	1.28	3	5	2.05	7.81	11.34	"	$0.50 < P < 0.60$
11.	8.0	82	0.91	2	4	2.74	5.99	9.21	"	$0.20 < P < 0.30$
12.	8.1	62	0.69	2	4	1.69	5.99	9.21	"	$0.40 < P < 0.50$
13.	8.2	47	0.52	2	4	3.74	5.99	9.21	"	$0.10 < P < 0.20$
14.	8.3	29	0.32	1	3	1.48	3.84	6.63	"	$0.20 < P < 0.30$

The time differences between successive events, defined as t, were classified with increasing order and the maximum, t_{max}, the minimum, t_{min}, the range $R = t_{max} - t_{min}$, the mean value t, and the standard deviation were calculated. The above mentioned range $R = t_{max} - t_{min}$ is divided to subspaces of equal lengths of time with a step between successive subspaces equal to 0.01 year. Thus, subspaces 0.00 - 0.01, 0.01 - 0.02, etc. were determined and the number of events, n, in each subspace was found. In the case when the number of events in a subspace was less than 5, this number is added to the number of events of the next subspace (or subspaces) until the number 5 or larger was obtained. In this way, the data were grouped in a number, NG, of groups and the degree of freedom, v, is NG - 2.

To check the validity of the hypothesis that the time difference between successive shocks follows a negative exponential distribution, the X^2 test was applied in all cases when the number of events were larger or equal to 29. This happens in all cases for which the earthquake magnitude was less or equal to 8.3. In the cases when the number of events was less than 20, and this happened when $M \geq 8.4$, the Kolmogorov-Smirnov test was applied.

APPLICATION OF THE X^2 TEST

This test was first applied to the subset of data which includes all main shocks with $M \geq 7.0$ which occurred during the time period 1897-1986. It was applied to test the hypothesis that the time difference between successive earthquakes of this period with $M \geq 7.0$ follows a negative exponential distribution, then to test the same hypothesis for earthquakes of the same time period but with $M \geq 7.1$, then for earthquakes with $M \geq 7.2$ and so on until $M \geq 8.3$. Table III shows the results for this subset of data, M is the smallest earthquake magnitude for each sample, n is the number of events (earthquakes) of each sample, NG is the number of groups, v is the degrees of freedom, λ, is the inverse of the mean value of the time difference, X^2 are observed values, $X^2_{v:0.01}$ are theoretical values at 0.01 level of significance and P is the probability that the observed X^2 value is excessed (type I error). Figure (1) shows plots of the values X^2 and X^2_v, given on Table III, versus the degrees of freedom, v, and the corresponding earthquake magnitude.

Table (III) and figure (1) show that with the exception of

Table III. Results of the application of the X^2 test to the worldwide data of earthquakes of the period 1897-1986.

Πίνακας III. Αποτελέσματα της εφαρμογής της δοκιμασίας X^2 σε παγκόσμια δεδομένα σεισμών της περιόδου 1897-1986.

A/A	M	n	NG	P	v	λ	X^2	X^2_v:0.01
1	7.0	1208	54	0.0001	52	13.4482	97.14	78.62
2	7.1	904	38	0.0103	36	10.0639	58.48	58.62
3	7.2	724	41	0.0142	39	8.0600	60.81	62.43
4	7.3	568	45	0.3820	43	6.3233	45.15	67.46
5	7.4	461	47	0.0542	45	5.1321	61.18	69.96
6	7.5	392	18	0.4518	16	4.3640	16.01	32.00
7	7.6	294	19	0.2705	17	3.2730	20.07	33.41
8	7.7	236	21	0.1330	19	2.6273	25.89	36.16
9	7.8	167	22	0.1780	20	1.8591	25.61	37.57
10	7.9	115	16	0.7253	14	1.2802	10.49	29.14
11	8.0	82	13	0.2409	11	0.9129	13.85	24.72
12	8.1	62	11	0.0603	9	0.6908	16.32	21.67
13	8.2	47	9	0.1896	7	0.6909	9.98	18.47
14	8.3	29	6	0.0273	4	0.4317	10.93	13.28

the case of main shocks with $M \geq 7.0$ in all other cases the observed X^2 values are less than the theoretical values X^2_v. The results of Table III, fit very well with the ones of Table II. One can easily realize that the time difference between successive shocks of magnitude $M \geq 7.2$ has the negative exponential distribution, which confirms the hypothesis that the number of annual shocks has the Poisson distribution.

Figure (2) shows the plots of the observed, X^2, and the theoretical $X^2_{v:0.01}$ values versus the degrees of freedom, v, and the magnitude, M, for the other three subsets of data. It is observed that for the main shocks with $M \geq 6.5$ of the period 1930-1986, the observed X^2 values are less than the theoretical values for $M \geq 7.1$ (or for $M \geq 7.3$) and not for smaller magnitudes (fig. 2a). For the main shocks with $M \geq 6.0$ of the period 1952-1986 (fig. 2b) as well as for the main shocks of the period 1966-1986 (fig. 2c) the observed X^2 values are less than the theoretical values for $M \geq 7.1$ and not for smaller magnitudes.

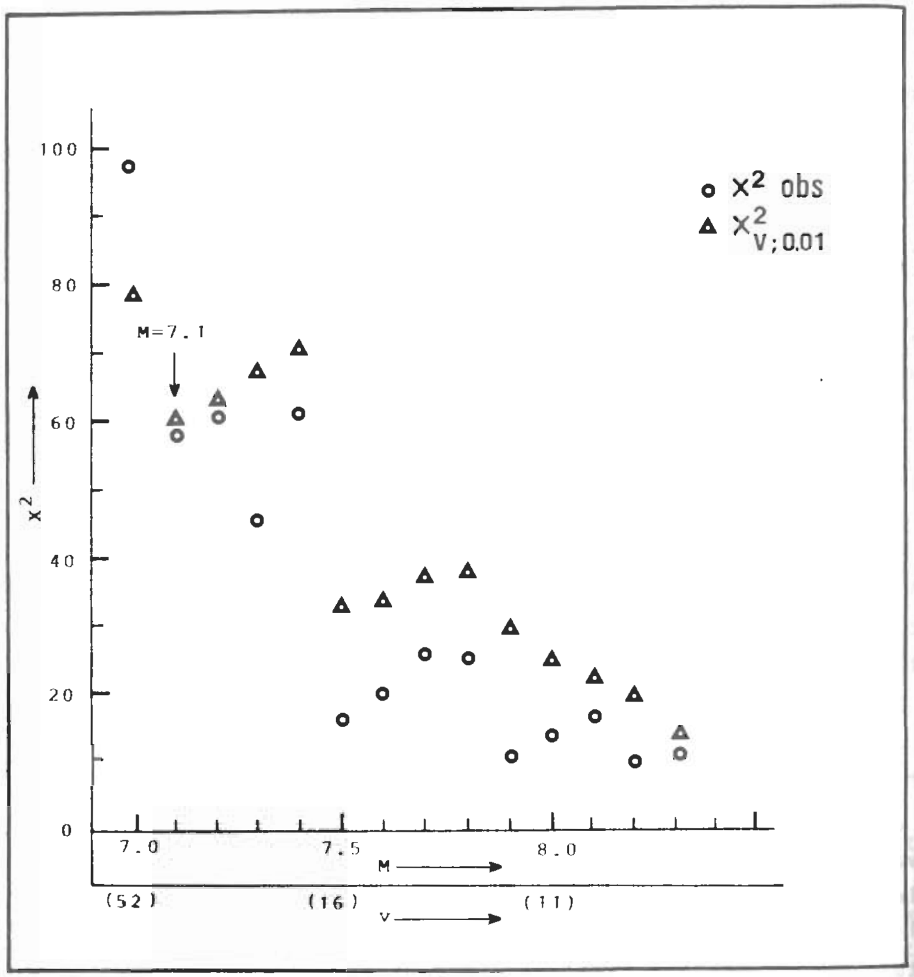

Fig. 1.- Plots of the observed X^2, and the theoretical X^2_v, values given on Table (III), versus the degrees of freedom, v and the corresponding earthquake magnitude for the period 1897-1986.

Σχ. 1.- Χαρτογράφηση των παρατηρούμενων, X^2, και των θεωρητικών, X^2_v, τιμών που δίνονται από τον πίνακα III σε συνάρτηση με τους βαθμούς ελευθερίας, v , και τα αντίστοιχα μεγέθη, M, των σεισμών για τη χρονική περίοδο 1897-1986.

The good agreement between the result of the application of the X^2 test to these four complete and to a considerable degree independent sets of data leads to the conclusion that the main shocks of the Earth of which the magnitudes lie in the interval (7.1, 8.3) follow a Poisson distribution.

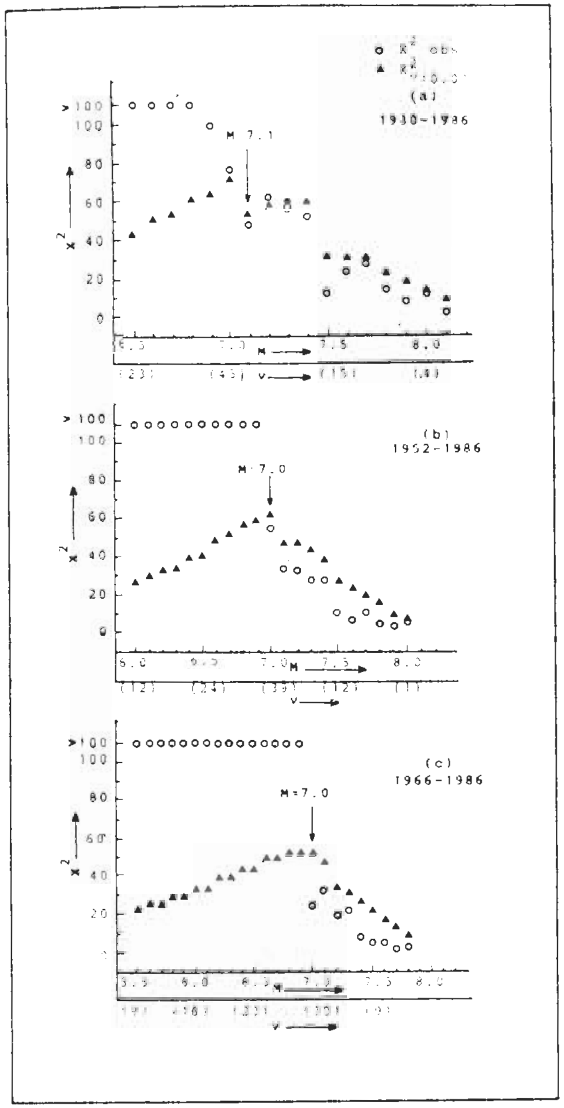

Fig. 2.- Plots of the observed, X^2, and the theoretical, $X^2_{v;0.01}$ values versus the degrees of freedom, v, and the magnitude, M, for the periods 1930-1986 (a), 1951-1986(b) and 1966-1986(c).

Σχ. 2.- Χαρτογράφηση των παρατηρούμενων, X^2, και των θεωρητικών $X^2_{v;0.01}$, τιμών σε συνάρτηση με τους βαθμούς ελευθερίας, v, και τα μεγέθη, M, για τις περιόδους 1930-1986(a), 1952-1986 (b) και 1966-1986 (c).

APPLICATION OF KOLMOGOROV-SMIRNOV TEST

The number of shocks with $M \geqslant 8.4$ which occurred during the period 1897-1986 is smaller than 20. Therefore, the X^2 test cannot be effectively applied in these cases. One of the proper tests in these cases is the Kolmogorov-Smirnov (Lehmann 1975, Kounias et al 1985) and for this reason this test has been applied for all main shocks with $M \geqslant 8.4$ which occurred during the period 1897-1986. As estimate of $\hat{\lambda}$ in the equation (3), the maximum likelihood estimate from equation (2) was also used in this case.

Table IV shows the results of this test for the four samples with $M \geqslant 8.4$, $M \geqslant 8.5$, $M \geqslant 8.6$, $M \geqslant 8.7$ and the corresponding number of observations (number of earthquakes) 19, 10, 7 and 4. D_n denotes the absolute difference between the sample $F_n(t)$ and the corresponding theoretical probability function $F_0(t)$. $D_{a,n}$ denotes the theoretical values at a level of significance a and number of observations n.

Table IV . Results of the application of the Kolmogorov-Smirnov test for the very large earthquakes of the period 1897-1986.

Πίνακας IV . Αποτελέσματα της εφαρμογής της δοκιμασίας Kolmogorov-Smirnov για τους πολύ μεγάλους σεισμούς της περιόδου 1897-1986.

n	M	D_n	$D_{.20,19}$ $D_{.20,10}$ $D_{.20,7}$ $D_{.20,4}$	$D_{.05,10}$	$D_{.15,7}$	$D_{.01,4}$		Hypothesis H_0
19	8.4	0.2238	0.237				$D_n < D_{.20,19}$	accepted
10	8.5	0.1789	0.322	0.409			$D_n < D_{.20,10}$	"
7	8.6	0.3839	0.381		0.405		$D_{.20,7} < D_n < D_{.15,7}$	"
4	8.7	0.4905	0.494			0.734	$D_n < D_{.20,4}$	"

It is observed that the hypothesis of the negative exponential distribution of t is accepted at a level of significance less than 0.15 in all cases. This is an evidence that even the very large earthquakes ($M \geqslant 8.4$) follow a Poisson process.

DISCUSSION

The observation that the main shocks of the Earth with $M > 7.0$ follow a Poisson process and that this does not hold for smaller shocks cannot be easily explained. This result, however, shows that the rate of seismic energy release remains almost constant (Vere-Jones 1970) since the seismic energy is mainly released by the large earthquakes. This result probably indicates that the large earthquakes ($M > 7.0$) are independed events occurring in separated regions of the world and that the smaller shocks are not indenpendent but are affected by the occurrence of the large ones. Although the sample for the very large earthquakes ($M \geqslant 8.4$) is small and the result that these earthquakes also follow a Poisson process cannot be considered as definite, this result do put a serious question about the relation between these earthquakes of the Earth which needs further investigation.

REFERENCES

ABE,K. (1979). Size of great earthquakes of 1837-1974 inferred from tsunami data, *Jour. Geophys. Res.*, 84, 1561-1568, U.S.A.

ABE,K. (1981). Magnitudes of large earthquakes from 1904 to 1980, *Phys. Earth.Plan. Inter.*, 27, 72-92, U.K.

ABE,K. and KANAMORI,H. (1979). Temporal variation of the activity of intermediate and deep focus earthquakes, *Jour. Geophys.Res.*, 84, 3589-3595, U.S.A.

DUDA,S.J. (1965). Secular seismic energy release in the circum-Pacific belt, *Tectonophysics*, 2, 409-452, The Netherlands.

GELLER,R.J.-KANAMORI,H. & ABE,K. (1978). Addenda and corrections to magnitude of great shallow earthquakes from 1904 to 1952, *Bull. Seismol.Soc. Am.*, 68, 1763-1764, U.S.A.

GUTENBERG,G. & RICHTER,C.F. (1954). Seismicity of the Earth and associated phenomena, *Princeton Univ.Press*, Princeton,N.Y., 310p.

JONES,L. and MOLNAR,P. (1976). Frequency of foreshocks, *Nature*, 262, 677-679, U.K.

KANAMORI,H. & ABE,K. (1979). Reevaluation of the turn-of-the-century seismic peak, *Jour.Geophys.Res.*, 84, 6131-6139, U.S.A.

KOUNIAS,S.-BAGIATIS,C.B.-KOLYVA-MACHERA,F. & BORA-SENTA,E. (1985). Introduction to statistics, (in Greek), *Univ. of Thessaloniki*, Thessaloniki, 417p, Greece.

LEHMANN,E.L. (1975). Nonparametrics: Statistical methods based on ranks, *Holden-Day Inc.*, San Francisco, 457p, U.S.A.

LOMNITZ,C. (1966). Statistical prediction of earthquakes, *Rev.Geophys.*, 4, 377-393, U.K.

MIYAMURA,S. (1976). Provisional magnitudes of Middle American earthquakes not listed in the magnitude catalogue of Gutenberg-Richter, *Bull. Intern.Inst.Seism.Earthq.Eng.*, 14, 41-46, Japan.

MIYAMURA,S. (1978). Magnitude of large earthquakes not included in the Gutenberg-Richter magnitude catalogue, *Tectonophysics*, 49, 171-184, The Netherlands.

MOGI,K. (1962). Study of elastic shocks caused by the fracture of heterogeneous materials and its relations to earthquake phenomenon, *Bull. Earthq.Res. Inst.*, 40, 125-173, Japan.

ROTHE,J.P. (1969). The seismicity of the Earth 1953-1965, *UNESCO*, Paris, 330 p, France.

SHLANGER,A. (1960). Some consequences of earthquake statistics for the years 1918-1955, *Gerl. Beitr.z. Geophys.*, 69, 68-72,Germany.
SINGH,S. and SANFORD,A.R. (1972). Statistical analysis of micro-earthquakes near Socorro,New Mexico, *Bull.Seismol.Soc.Am.*, 62, 917-926,U.S.A.
TSAPANOS,T.M. (1985). A contribution to the study of the seismicity of the Earth, Ph.D.Thesis (in Greek) , *Univ.of Thessaloniki*, 147p,Greece.
UTSU,T. (1969). Aftershocks and earthquake statistics (I),*Jour. of Fac. of Science*, Hokk. Univ.Serries VII, 3, 3, 130-195,Japan.
VERE-JONES,D. (1970). Stochastic models for earthquake occurrence, *J.R.Statistic, Soc.*, B32, 1-62,Australia.