INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/709-9

Second Workshop on Non-Linear Dynamics
and Earthquake Prediction

22 November - 10 December 1993

Analysis of Earthquake Catalogs

I. Rotwain
Russian Academy of Sciences
International Institute of Earthquake Prediction
Theory and Mathematical Geophysics
Moscow 113556
Russian Federation

A N A L Y S I S   O F   E A R T H Q U A K E   C A T A L O G S

I.Rotwain

The algorithms for intermediate-term earthquake prediction, which you will study on this workshop are based on seismicity data - earthquake catalogs.

Why do we use seismicity data for earthquake prediction?

-- We assume that an earthquake catalog contains enough information to diagnose incipient large earthquake.

-- Earthquake catalogs are the most available type of geophysical data: there are catalogs for practically all seismoactive regions in the world and compiled for a long period of time.

First of all - some common information about catalogs. As you know, the earthquake catalog is a file which contains the following information:

Time of the earthquake occurrence

Geographical coordinates of the earthquake epicenter

Focal depth

Magnitude

Catalog in original format and main parameters of catalog is considered in example 1.

There are global and local catalogs.

Most complete global catalogs:

-- NEIC - issued by US National Earthquake Information Center; the catalog, based mainly on the observations of the WorldWide Standard Seismological Network (WWSSN) which started operations in 1963. The catalog is complete since 1964 for magnitudes above 4.5.

-- ISC - issued by International Seismological Center, based on reportings of all national networks all over the world, it is more complete than NEIC, but compiled with a 2- years delay.

-- CSEM - issue by Centre Seismologique Euro-Mediterranean for the territory of Mediterranean Sea and adjacent areas.

There are local catalogs for many regions, which may be more relevant for usage in prediction problem, because usually they are more complete and representative than the global catalogs in the region.

Description of the parameters of the catalogs:

Origin time T

In global and most of the regional catalogs Greenwich Mean Time (GMT) is given, in regional, however in some regional catalogs local times are used instead of GMT. The accuracy of the origin time is about seconds. Events in catalogs are usually sorted by the increase of their origin time.

Geographical coordinates of epicenter LAT, LON

Usually the coordinates are given in degrees and decimal fractions of a degree. In some local catalogs fractions of a degree are given in minutes and seconds. The values of Southern latitude and Western longitude are negative by definition or marked by letters S and W respectively. For example: -25.00 = 25.00 S means 25 degrees of Southern latitude. The accuracy of epicenter determination vary in different catalogs: it is a few tens of kilometers for global catalogs and a few kilometers or even less than a kilometer for local ones. Usually accuracy of epicenter determination improves in time with the development of seismological network. Some catalogs explicitly indicate the accuracy of epicenter determination for the other ones the accuracy have to be found elsewhere in literature or in catalog descriptions.

Focal depth DEP

Focal depth is given in kilometers. Usually depth is determined less accurately than coordinates. When the exact value of focal depth is impossible to find out the standard value of depth is given. It concerns only shallow earthquakes and in NEIC catalog the standard value is equal to 33 km; in California Institute of Technology (CIT) catalog it is equal to 10 km.

Magnitude M

In some catalogs several definitions of magnitude are used. In NEIC catalog four values of magnitude are given:

mb - averaged body wave magnitude determined using global network;

Ms -averaged surface wave magnitude determined using global network;

ml - magnitude determined using local network;

mp - magnitude determined by other seismological agencies.

In local catalogs one or two magnitudes are usually given.

The accuracy of magnitude determination inside one magnitude scale is about several decimals.

However the magnitudes of the same earthquake determined by different magnitude scales or different agencies may differ by unit or even more.

Preliminary preparation of the catalog.

A catalog has to meet some special requirements to be analyzed by the software worked out in International Institute for Earthquake Prediction Theory and Mathematical Geophysics (IIEPT). They are as follows:

-- *Transformation into standard format.*

All the programs concerning the research and processing of

catalogs use standard format. It is a binary fil with all the records with equal length of 20 bytes.

The first record is not the same as the others, it contains the number of earthquakes +1, 4 bytes integer value in the internal ccde of IBM-PC computer.

The format of the other records is the following:

The first parameter:

1-4 Origin time, A.D., GMT, in minutes is the 4 bytes integer value.

All other parameters are two bytes integer values:

5-6 Latitude in degrees and decimals multiplied by 100 to transform it into an integer number,

7-8 Longitude in degrees and decimals multiplied by 100, (Southern latitude and Western longitude are negative by definition)

9-10 Focal depth in kilometers - 2 bytes integer,

11-12 Magnitude mb multiplied by 100,

13-14 Magnitude Ms multiplied by 100,

15-16 Magnitude ml multiplied by 100,

17-18 Magnitude mp multiplied by 100,

19-20 Blank positions.

Such format omits part of the information contained in the original catalog: seconds of the origin time parameter, accuracy of parameters determination, number of seismic stations registered the earthquake, etc. The omitted information is not used by our algorithms. The advantage of the standard format is the following:

- the considerable compression of data;
- we use common information, thus it is easy to work with

different catalogs.

-- *Time ordering.*

In some catalogs, events are not ordered by time. For example they ordered by latitude or longitude. Even if the catalog is supposedly ordered by time, some errors may occur. The correct processing of programs require time ordered catalogs.

-- *Removal of duplicate records.*

Some catalogs include the data from various sources, and may contain several records corresponding to the same event. Sometimes records are doubled in file by mistake.

-- *Elimination of* the *obvious errors.*

The errors of the type LAT = 99.00 or month = 13 are frequently occur in many catalogs.

General analysis of the catalog.

Catalog analysis is the most informal stage of the work. The quality of input data can affect the result of a study. How to analyze a catalog? We are going now to give you some recommendations which are based on our experience.

1. *Determination of the time period* when catalog contains sufficient annual number of events. Usually years at the beginning of the catalog should be ignored.

The histogram of number of earthquakes according to magnitude and time is shown on example 2. From the first glance it seems that we have two reasonable points to define the beginning of the time interval. The first one is from 1958, because since this time the earthquakes with magnitude above 3.5 are represented homogeneously. The second one is from 1964 because since this time the representativeness of catalog improves considerably. After

considering the table in more detail, we find out that earthquakes with magnitudes from 3.5 to 5 are presented very poorly before 1964. The values of annual number of events are much smaller before 1964 than after this year. We can suspect that this phenomenon is not a real seismic quiescence, because large earthquakes on the contrary to the small ones were more numerous before 1964 than after. Finally for the purposes of our study we choose a period from 1964. So, we recommend to ignore the seismic data from this catalog before 1964 in our analysis.

Let me show you another phenomenon in this table: there is a short period from 1980 to 1982 when the number of events decreases. It may be a real seismic quiescence, because there are no large earthquakes during this period. However, there is another possible cause for this gap. You can guess it, because the number of small events (below magnitude 3) increases considerably since 1982. The local network was reorganized in this period. This example shows that seismic gaps in catalogs may reflect manmade changes which are not connected with the behavior of seismicity.

It is impossible to correct such defects in the catalog, but the information about the reorganization of network is very useful to interpret the results obtained by earthquake prediction algorithms.

2.Choice of area. It is natural to choose regions according to their actual seismicity. A local network may cover only a part of seismoactive region. It may influence on the results if some part of the preparation zone of a strong earthquake is outside the area covered by the local network.

In the example 3 the space distribution of seismicity is shown. The first map shows the local network data, and the second one - the global network data for the same area. Many earthquakes occur to southeast and to northwest from the area represented by the local network.

Another difficulty is the time changes in the observed area if you use local catalogs. Therefore it is necessary to choose a part of territory represented uniformly during all time periods.

The example 4 shows the earthquake distribution for the same territory but for two different periods of time, according to local network data. As you see the southern part of the region is not presented in 1964-1969. This difference comes from local network changes. When you choose the region it is necessary to define not only the southern boundary, but other boundaries too, to exclude possible effects of further network expansion.

3. Definition of the maximum depth threshold for shallow earthquakes. For the purposes of a study it may be useful to distinguish shallow earthquakes from other ones, because in many cases we will use only shallow seismicity.

Usually there is some local minimum in event distribution by depth between shallow and other earthquakes. In the example 5 such a local minimum is around 70-75 km. We can choose this value as the threshold of depth in this region. In many regions there is no intermediate or deep seismicity at all. For such regions you can skip this step of analysis.

4 Choice of a magnitude scale. Most of IIETP algorithms require only one value of magnitude for each event. There are no formal rules in the choice of magnitudes. Following are several rules:

If several magnitude values are given in catalog you may:

- choose one kind of magnitude and ignore others. This way is

used if this type of magnitude is determined for most of the events.

- use maximum value from the magnitudes attributed to an event.

The last method allows you to use all events with known magnitudes. But be careful in this case. Magnitudes of different kinds may be related to different periods of time or parts of the region. This choice needs a special analysis.

- to define priority magnitudes. It means that you declare an order in magnitudes: m1,m2,m3,m4. If magnitude m1 is known for the earthquake, m1 is used; if, not m2 is used, if m1 and m2 unknown, m3 is used; etc. This method allows you to use the magnitude which seems to be the "best" if it is present in the catalog and use another magnitude when it is not.

- You can choose one kind of magnitude and when it is absent recalculate the other ones according to special formulae. As a rule we use Ms magnitude. The segment of catalog NEIC is represented in example 6. It is known that Ms magnitude is usually much higher than mb for large earthquakes. Practically mb never exceeds 7.0 but Ms can reach 8.5 value. However for small earthquakes Ms is usually unknown. For them magnitude mb only is given in the catalog. Therefore we have to recalculate magnitude mb to Ms via statistical regression formula.

To recalculate different magnitudes from one scale to another, you have to analyze preliminary relations existing between them.

The example 7 shows the relation between mb and Ms magnitudes. Here is the linear orthogonal regression formula to recalculate mb magnitude into Ms one. You can use other methods to determine the relationship.

It may be useful to compare magnitudes in local and global catalogs. Sometimes the method for magnitude determination may change in time.

The example 8 shows the changing of magnitude scale in time for CIT catalog. We compared this catalog and NEIC. The magnitude determination in CIT catalog had probably changed after 1980, because since that time magnitudes differ significantly from the NEIC magnitudes. In such a case it is reasonable to take into account the information from global catalog.

Determine a minimum representative value of magnitude at the chosen magnitude scale. One way to do it is the usage of the distribution of the number of shocks by magnitude and time (example 9). In this example the threshold magnitude 3.0 may be chosen, because it is the minimum magnitude presented homogeneously in the whole period of time.

The other way is to use the shape of the frequency-magnitude relation in semi-logarithmic scale or the so called Gutenberg-Richter law. According to this law lg(N) where N is the number of events should decrease linearly with increasing magnitude. Empirical points in example 10 deviates from linear regression in two ranges: in the range of strong magnitudes because such shocks occur very seldom, and in the range of low magnitudes due to the deficit of small earthquakes in a catalog. It seems that starting from magnitude 3.0 empirical points fit to the approximating straight line. It means that events with magnitude above 3 are presented completely enough in the catalog.

To merge catalogs. If catalog is not sufficiently complete, or if some parts of the region under study are presented in different

catalogs, you can try to merge them. When you merge the catalogs you should compare the magnitude scales in them, determine a relation between magnitudes and recalculate magnitudes before merging if it is necessary.

But there is no formal rule which can guarantee the correctness of the merging. As a rule, epicenters and magnitudes of the same events are differ in different catalogs. You should make sure that relatively large earthquakes are not duplicated or lost.

As a result you will have a complete, representative catalog for some region and time period, with homogeneous magnitude scale - it is the ideal situation. Practically, we never get such an ideal catalog. It will always have some defects for objective reasons. Area, period of time, and minimum representative magnitude are not independent. By decreasing the time period you probably can expand the area or decrease minimum magnitude, etc. It is an iterative process, and you have to stop it at a reasonable stage.

The requirements to catalog may seem too hard, but really they are not. To apply, for example algorithm CN, catalog should contain only several shocks annually. The earthquake prediction algorithms provide an automatical adjustment to data. In reality only one hard necessary requirement to data is - to be HOMOGENEOUS.

The catalogs are updated in time. Don't forget to test each new portion of data, to save the homogeneity of your catalog. The changes of catalog characteristics in time may cause some errors in forward prediction.

<u>Catalog of main shocks</u>.

Our algorithms use catalogs of main shocks. To get such catalog it is necessary to remove aftershocks from initial catalog.

We use the space-time windows to identify aftershocks.

Earthquake (j) is an aftershock of earthquake (i) if:

magnitude $M(j) \le M(i)$

origin time $t(j) \ge t(i)$

time difference $t(j) - t(i) \le T(M(i))$

distance between epicenters $r(i,j) \le R(M(i))$

depth difference $h(i,j) \le H(M(i))$

where T(M), R(M), H(M) are parameters of the algorithm dependent from the magnitude of the main shock (i).

The first earthquake in the catalog is declared as a main shock, then all its aftershocks are removed from the catalog. The first of the remaining earthquakes is declared as the next main shock, etc., until the end of the catalog.

Software

We have a problem-oriented software for catalog processing. It provides various facilities for data analysis.

- EDCAT - to edit catalog and transform it into standard format;
- CATAL - to select subcatalogs, merge and compare catalogs;
- HIST - to make histograms for any parameters of catalog;
- AFT - to create the catalog of main events (to exclude aftershocks);
- MAP - to create the map of epicenters and other data;
- CINEMA- to view the dynamics of seismic process in time.

At the end of my lecture I would like to stress that no formal algorithm provides automatic preparation of data. Software can facilitate analysis, but only the researcher himself should make the final decision.

EXAMPLE 1. Catalog in original format

	Year	d m h m sec	lat	long depth mb Ms	M author
PDE	1990	0805013457.50	29.513	137.596516.1D220.915.956	6.50mbBRK
PDE	1990	0805031610.20	27.503	141.593 47.7D220.745.430	
PDE	1990	0805033624.33	36.300	141.083 41.7D220.835.7506.0Z13	
PDE	1990	0805065213.60B&	36.855	-121.638 7.0 3.6 1	4.00MLBRK
PDE	1990	0805071301.83	35.088	139.042 37.0 0.655.4334.5Z 4	
PDE	1990	0805072436.67*	-5.258	129.596200.0? 1.164.8 7	

Main parameters of catalog

Origin time						Epicenter		Depth	Magnitudes			
Year	mo	da	hr	mi	sec	lat	lon		mb	Ms	Ml	mp
1990	08	05	01	34	57.50	29.513	137.596	516.1	5.9			6.50
1990	08	05	03	16	10.20	27.503	141.593	47.7	5.4			
1990	08	05	03	36	24.33	36.300	141.083	41.7	5.7	6.0		
1990	08	05	06	52	13.60	36.855	-121.638	7.0	3.6		4.00	
1990	08	05	07	13	01.83	35.088	139.042	37.0	5.4	4.5		
1990	08	05	07	24	36.67	-5.258	129.596	200.0	4.8			

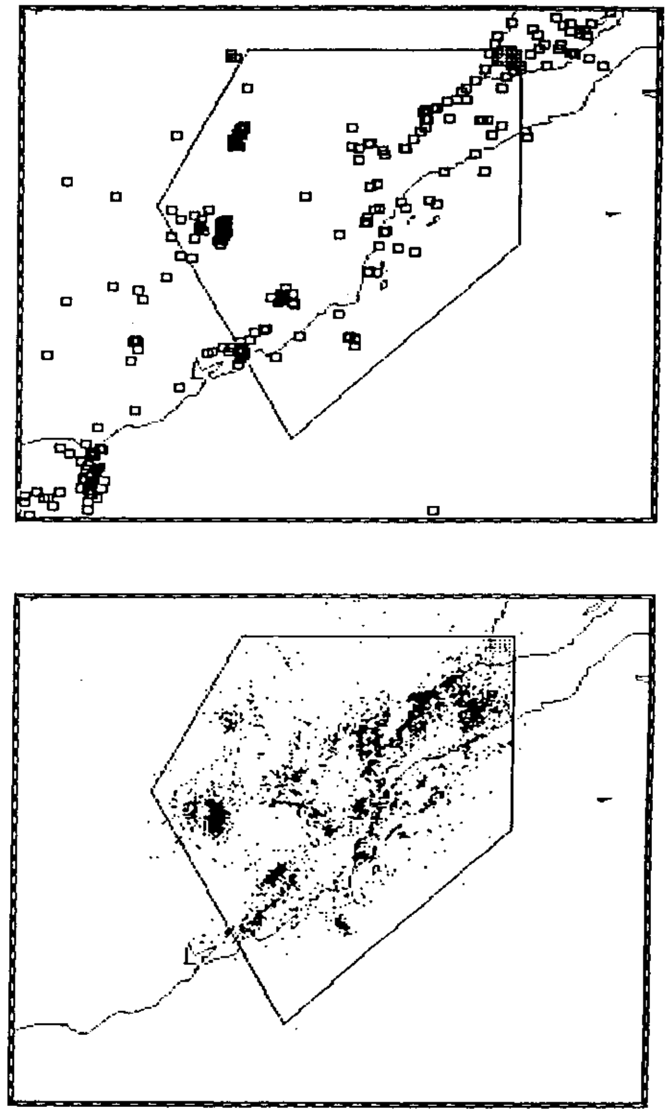
EXAMPLE 2.

CATALOG REPRESENTATIVENESS IN TIME

Year	Magnitudes 2.0–2.5	2.5–3.0	3.0–3.5	3.5–4.0	4.0–4.5	4.5–5.0	5.0–5.5	5.5–6.0	6.0–6.5	Total
1950	.	.	.	.	1	.	1	.	1	3
1952	.	.	.	1	1	.	.	1	1	4
1954	.	.	.	2	.	.	.	1	1	4
1956	.	.	5	1	.	1	.	2	1	10
1958	.	.	19	3	1	.	1	1	.	25
1960	.	.	26	2	.	.	.	.	1	29
1962	.	.	12	4	.	.	1	.	.	17
1964	6	7	19	8	6	8	.	.	.	54
1966	2	15	23	12	4	1	.	.	.	57
1968	.	4	10	14	3	1	.	.	1	33
1970	4	13	15	16	8	4	2	.	.	62
1972	4	9	21	16	.	1	1	.	.	52
1974	2	15	20	19	9	1	2	.	.	68
1976	15	17	13	13	6	.	.	.	.	64
1978	12	11	11	10	8	1	1	.	.	54
1980	4	18	1	2	3	.	.	.	.	28
1982	80	81	52	34	33	12	2	.	.	294
1984	108	103	43	23	13	5	3	.	.	298
1986	101	48	30	5	7	1	.	.	.	192
1988	138	68	36	7	15	2	1	.	.	267

Example 3

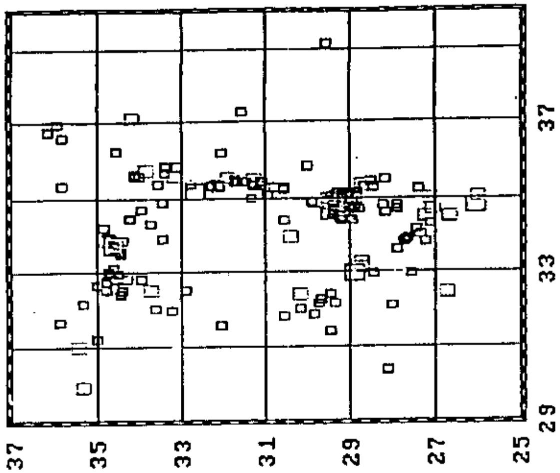
Local and global network data



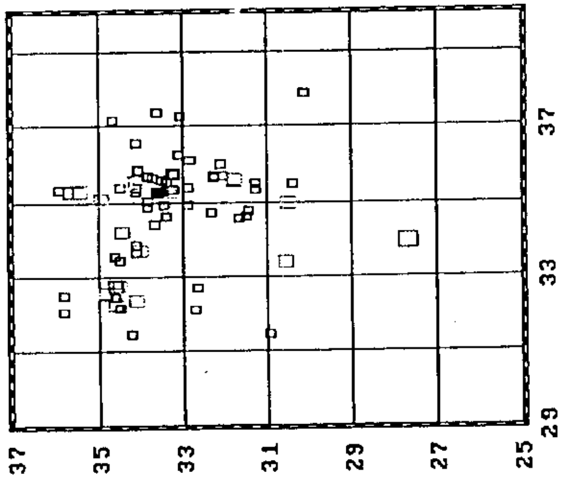
Example 4

Changes in local network affect the catalog

1981 - 1985

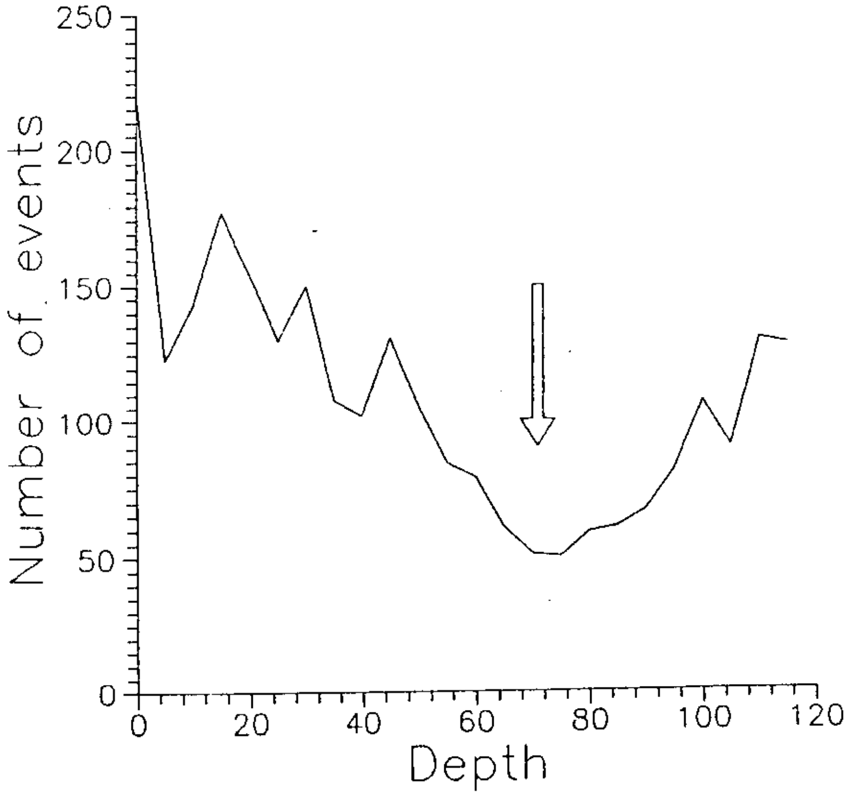


1964 - 1969



Example 5

Choice of depth threshold



EXAMPLE 6.

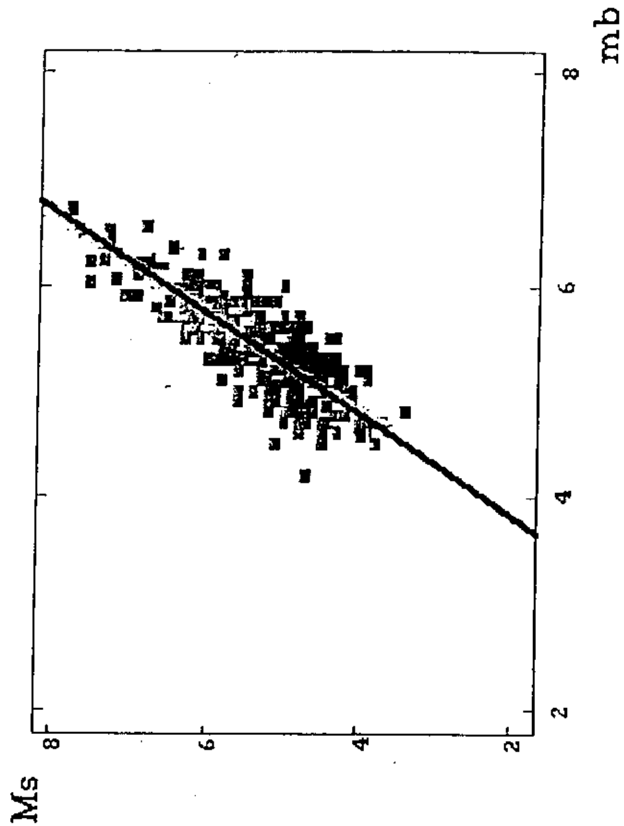
DIFFERENT MAGNITUDE SCALES: MB and MS

Date			Epicenter		Magnitudes	
year	mo	d	lat	long	mb	ms
1986	5	7	51.52	-174.77	6.40	7.70
1986	10	20	-28.11	-176.36	6.60	8.10
1986	11	14	23.90	121.57	6.30	7.80
1987	9	3	-58.89	158.51	5.90	7.30
1987	10	16	-6.26	149.06	5.90	7.40
1987	11	30	58.67	-142.78	6.70	7.60
1988	3	6	56.95	-143.03	6.80	7.60
1989	5	23	-52.48	160.20	6.50	8.30

Example 7

Relation between mb and Ms magnitudes

$$Ms = 2.11\ mb - 6.08$$



EXAMPLE 8.

CHANGING of MAGNITUDE SCALE in TIME

Date year	mo	d	Time h	mi	Magnitudes NEIS	CIT	DELTA
1966	8	7	17	36	6.30	6.30	0.0
1968	4	9	2	28	6.40	6.40	0.0
1971	2	9	14	0	6.50	6.40	0.1
1979	10	15	23	16	7.00	6.60	0.4
1980	5	25	16	33	6.50	6.40	0.1
1980	5	25	19	44	6.70	6.50	0.2
1980	5	27	14	50	6.30	6.30	0.0
1980	6	9	3	28	6.40	6.10	0.3
1981	4	26	12	9	6.30	5.70	0.6
1983	5	2	23	42	6.70	6.30	0.4
1986	7	21	14	42	6.50	5.90	0.6
1987	11	24	1	54	6.50	5.80	0.7
1987	11	24	13	15	6.70	6.00	0.7

EXAMPLE 9

CHOICE OF REPRESENTATIVE MAGNITUDE

Year	Magnitudes 2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0 6.5	Total
1964	6	6	18	7	6	8	.	.	.	51
1966	2	14	23	12	4	1	.	.	.	56
1968	.	3	10	14	3	1	.	.	1	32
1970	4	13	12	14	5	2	1	.	.	51
1972	4	9	21	14	.	1	.	.	.	49
1974	2	12	16	16	5	1	2	.	.	54
1976	15	16	12	10	2	.	.	.	.	55
1978	12	11	11	9	7	1	1	.	.	52
1980	4	18	.	2	3	.	.	.	.	27
1982	43	24	11	8	5	3	.	.	.	94
1984	99	78	24	16	6	4	1	.	.	228
1986	93	41	22	2	4	.	.	.	.	162
1988	121	52	22	5	8	.	1	.	.	209
Total	405	297	203	129	58	21	6	1	1	

Example 10

Choice of magnitude threshold according to frequency-of-occurence graph