Investigating instrumental seismicity data sources for the Mediterranean region (IBCM-S)

Y. LEGROS[(1)] and J. BONNIN[(2)]

[(1)]*Formeoly with European-Mediterranean Seismological Centre, Institut de Physique du Globe, Strasbourg, France*
[(2)]*Institut de Physique du globe, Strasbourg, France*

(Received May 7, 1998; accepted May 6, 1999)

Abstract. Within the “International Bathymetric Chart of the Mediterranean (IBCM)” project of the Intergovernmental Oceanographic Commission (IOC) of UNESCO, thematic maps have been constructed; one theme selected is the seismicity within and around the Mediterranean Sea. To construct a map of seismicity, one has to report, on a map with a geographic background, the symbols representing occurrences of earthquakes. Current practice is to assign “attributes” to the epicenter as a function of the earthquake's parameters, such as focal depth, magnitude, etc., and to draw a symbol centered on the epicenter's projection of each earthquake onto the map; then the reported symbol itself varies (in size, shape, ornaments, color, etc.) according to the epicenter's attributes. The natural choice for background map was the bathymetric base map of the IBCM project (30° - 46° N, 6° W - 36.5° E + an insert for the Black Sea: 40° - 47.5° N , 26.5° - 42.5° E). Ideally the dataset reported on the map, should have been worked out for this purpose. It should have included one record per physical seismic event, with the parameters estimated/computed through a homogeneous procedure; this was clearly out of reach, as no general earthquake database has yet been established for the Mediterranean and surroundings areas. The decision to rely on datasets easily accessible to users was taken; this ruled out non-instrumental seismicity data (most often called “historical seismicity”). The European-Mediterranean Seismological Centre (EMSC) has been computing and disseminating earthquake locations (epicenters and focal depths) in a homogeneous manner, since 1976; although EMSC has not determined magnitudes, but has simply reported estimations by other institutions, it was considered that, because earthquake location is the predominant parameter here, the EMSC sub-dataset's homogeneity overpasses the other defects. For the instrumental period before 1976, the sub-

Corresponding author: J. Bonnin; Institut de Physique du Globe, Université Louis Pasteur, 5 rue René Descartes, F-67084 Strasbourg Cédex, France; phone: +33 (0) 388 416368; fax: +33 (0) 388 416477; e-mail: bonnin@selene.u-strasbg.fr

dataset chosen to start is the "earthquake catalogue" published by the International Seismological Centre (ISC); this "catalogue" actually often reports several determinations of earthquake parameters issued by different institutions, together with the ISC's proper determination when it has been computed (from 1964 on). A careful analysis of the ISC's sub- dataset has been conducted to unravel its intimate characteristics. A rationale has been developed to select the determination of parameters to adopt, so as to end up with a final sub-dataset showing one parameter determination for each physical seismic event. The details of the selection procedure are set down together with the statistical properties on which they are based. This procedure is not expected to be the only reasonable one to construct a map of seismicity; it is simply one way of solving the problems posed. Users will find the basic information necessary to check the procedure and its implementation, or even to design their own way of handling the datasets. Limitations on the reasoning are also stressed; they have, in turn, an influence on how to display the final results graphically.

1. Foreword

The investigations reported below were begun under the International Bathymetric Chart of the Mediterranean [IBCM] Project, launched by the Intergovernmental Oceanographic Commission. It was intended to report geological/geophysical information, in addition to bathymetric data, on transparent material which could be superimposed onto the topographic base sheets. For practical reasons, the technical aspects have evolved, but the fundamental ideas have remained: to collect available information on geological/geophysical features of the Mediterranean region, and put it at the disposal of potential users, under the form of a series of cartographic documents.

Present-day tectonic activity seemed, at first glance, one of the easiest parameters to report on a map. Actually, after preliminary investigations, it soon appeared that available data was too scarce in many areas, and/or too disparate over the whole area under consideration, precluding any publication within a reasonable time-frame; the so-called "seismotectonic" map of the Mediterranean was postponed to longer-term issues.

Subequently, it was thought that it would be easier to simply report the symbols representing the seismic activity in the domain as a first step. Once again, after a rapid overlook, it appeared that this limited project was, by far, too ambitious: it is well known that seismic activity is extremely diverse over the whole domain; if one wants to make sense out of the activity to illustrate how the lithosphere deforms in that region of the globe, in the brittle deformation mode, then a sufficient time span should be considered which significantly compares with the large scale deformation's time constants: the recurrence time of large events in many parts of the domain measures in centuries, if not more. The conclusion is obviously that, in terms of significant (brittle) deformation of the lithosphere, data displayed should encompass several hundreds of years at least. This is totally out of reach, for the entire geographic domain with a minimum homogeneity, unless it is carried out through long-term studies. On the other hand, things are much easier to handle when addressing that part of seismic activity recorded by instruments; fortunately, the whole domain is covered by data relatively easy to access, even though inhomogeneities remain. Hence, the decision to first design a map of "instrumental sesmicity" of the region, taking advantage of information available at that time.

Yet, the authors have faced a series of unexpected shortcomings in exploiting earthquake catalogues and bulletins which led them to dig up the data sets at their disposal. Their finds are presented here. The main outcome is that earthquake data users are strongly urged to investigate carefully the intrinsic properties of their data sets, before drawing any conclusion; the following considerations illustrate, in particular, how difficult it is (if possible at all) to design a general-purpose earthquake catalogue for any given region.

The authors are very much indebted to the IBCM Editorial Board, its chairman, members and associate members. Prof. C. Morelli, Dr. J. K. Hall, Prof. I. Makris, Commander D. P. D. Scott, were instrumental in encouraging the authors, and convincing them not to give up at various difficult moments of the project; their patience was several times almost brought to on end, but they resisted and maintained quite a remarkable trust in the authors, who, an their side, demonstrated a marked inclination to procrastination, up to the very end. Sylvie Couteaux played the role of a pioneer, opening the first paths through the jungle of published catalogues, magnetic tapes with strange encoding and/or formating, etc.

2. Introduction

Attempts at representing the seismic activity of the Mediterranean region on a map suppose that one has at hand some sort of a file, containing the relevant information on the occurrence of seismic events within the area. The seismic activity's map representation is only one example (that we will trace all along in the following) among many other situations where an overview of the seismic activity for a given territory is sought: seismic hazard assessment, seismotectonic studies, etc. Such a trivial requirement actually raises a lot of rather trite, although complex, questions that anyone confronted with the same kind of problem has had to face. Among these questions, one can tentatively list, without being in the least exhaustive:

- what has already been done in that field in the past?
- are the corresponding data sets readily accessible?
- is their intrinsic quality adequate for the newly-born purpose?
- how can they be complemented, in the sense of time series, in order to make them up-to-date?
- what are the data sets available for complementing the previously called-upon ones?
- does the former, as well as the latter, need some cleaning-up or refreshing?
- how to proceed with this up-dating?
- how to assess the comparative quality of different data sets?
- what rationale to adopt for comparing different determinations of parameters of a suspected unique physical event?
- how to decide that different parameter determinations relate to the same unique physical event?

- what rationale to adopt in choosing among the various parameter determinations supposedly relating to the same physical event?
- how to ensure a minimum geographic homogeneity in the representativeness of the final data set reported on the map?
- what lower threshold in magnitude is of significance?
- what are the pieces of information (magnitude, depth, location uncertainty, etc.) to be reported on a map?

Many other more specific questions which will be adressed in the bulk of the parameter selection procedure described below.

Finally, much more basic questions arise, like:

- how good is any parameter determination as an estimator of the corresponding feature of the actual event?
- is the parameterization in usage adequate to depict the actual events correctly?
- does it make any sense to print thousands of copies of a seismicity map, which could be used and quoted for years, while the data on which it is based is under constant questioning?

Most of the above questions, and especially the most basic ones, have probably no definite answer, if any at all. The only scientifically sound approach, in our view, would be:

- to re-determine the parameters of all the events under consideration through a unique procedure,
- to find a way of assigning a homogeneously-defined quality factor to each parameter set,
- and to try and convey the relevant information to the user/reader, including a detailed description of the entire procedure adopted.

This was obviously out of our reach: collecting the necessary primary data would keep a large group of workers busy for a while, also assuming that all the “owners” of seismological data are fully cooperative.

Our approach has been mid-way between the out-of-reach, but only fully reasonable, way of carrying out the task (re-estimating every parameter, with complete control over each step), and the method too-often used of grasping one data set and cooking it according to a by all means excellent recipe , which is however kept undisclosed.

It was clearly not feasible to collect all the relevant, most updated national and regional information, even though the above-mentioned questions were assumed to have been solved (not the case by far); to do so, we would have needed tools (like a purposely designed database) which were not available at the time the project started, and enough material, from all the places, to feed the database: altogether, a task which needed years to be completed; a priori not compatible with the time allotted to the project. In addition, we consider it of primary importance that any user should have at hand the necessary material to conduct his own tests on the data set reported on the map; this could not have been achieved in a reasonable delay, had the above-suggested approach been applied. An alternate solution was to rely on an existing data set, easy to access, and of reasonably good quality. We then started with a rapid analysis of the main sources of earthquake parameters. The first and immediate conclusion was that is was unrealistic to think of taking into consideration the pre-instrumental seismic activity, in so far as an historical seismicity database had not been built up for the area considered. Thus, the map to be constructed

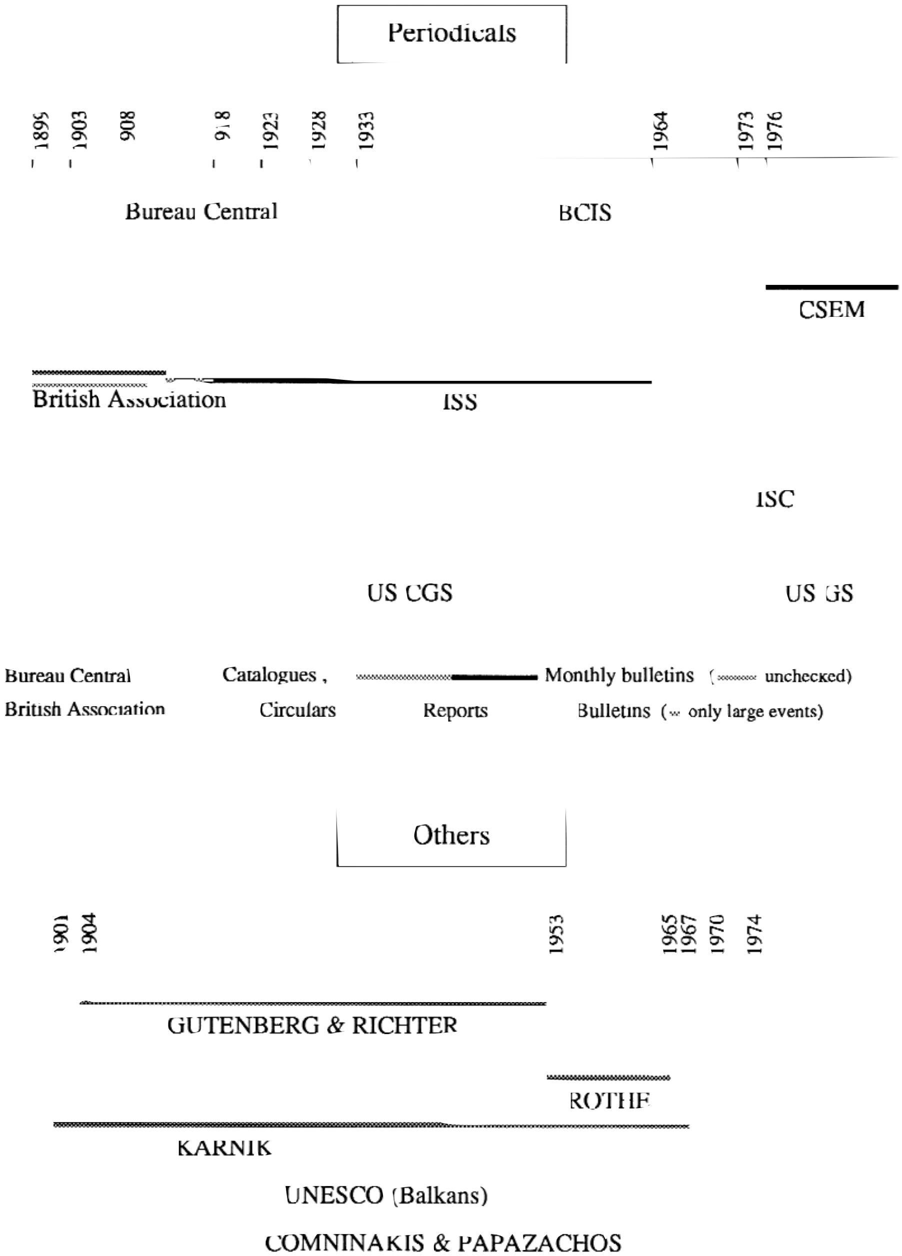

Fig. 1 - Bar-chart of the most important sources of instrumental earthquake parameters, relevant to the Mediterranean region.

would show only events which had been located using instrumentally recorded data; in practice, the so-called "instrumental" period began in the early 1900s for the Mediterranean region. It should be noticed that to date, as regards the recording of seismic activity, it is far from being uniform.

A rapid survey of the available data sets for the instrumental period (Fig. 1) shows that the main event has been the setting up of the International Seismological Centre [ISC], which started issuing hypocentral determination lists in 1964: this date marks a clear break between the previous situation and the period which follows. Actually, the European-Mediterranean Seismological Centre [EMSC] somewhat complicated the landscape in 1976: in practice, it led to another data set with very specific features which are described below.

After deciding to make use of the ISC's “Historical File” (up to and inclusive of 1963), of their “Catalogue” from 1964 to 1975 (both end years included), and then of the EMSC's monthly lists from 1976 onwards, a careful analysis of the first two files, which occasionally report several determinations of some events’ parameters, has been carried out in an attempt at defining “the best determination available” for final consideration (the third data set includes one location per event reported). Both the criteria and details of the procedure are described at length below: the purpose is to provide the user/reader with as much material as possible for him/her to evaluate precisely, according to his/her own criteria, the data set finally constructed.

The reader should not overlook the initial purpose of the investigations reported: the graphic design of the map depicting the seismic activity. It will hopefully be apparent how much the graphic representation must be tightly constrained by the fundamental philosophy of our approach.

3. Data sources

As it was not practicable to go back to the individual, best and most updated, lists of earthquake parameters (assuming that “best” can be properly defined) for each region and/or country included in the geographic domain involved, and to keep the whole study within reasonable time limits, we had to look for more general sources of information. To our knowledge, apart from specific chapters within studies extending over larger regions, most often the entire globe (see e.g. Gutenberg and Richter, 1941, 1949, Rothé, 1969; Lomnitz, 1974), the only study dedicated to the seismicity of the European area is the one by Karnik (1971); it is, no doubt, a master work in the sense that the author has examined the data he had at hand for each individual event, and has proposed a parameter set for each of them, following his own criteria; unfortunately, basic data is available with difficulty to the user, while the detailed procedure to applying the criteria adopted for accepting or rejecting a given parameter determination is not readily accessible. Karnik's entire masterwork is one more version, a major and sound one indeed, of the earthquake catalogue for the region. We could have simply drawn a map, according to the IBCM format, with Karnik’s catalogue data.

But we wanted to go a good step further, and to make it clearly understood to the user/reader that, however much care is taken, the parameters determined remain no more than one attempt at describing the actual features of the physical event under consideration. To make it very precise, we consider these present investigation notes much more significant than (but not independent from) the objective for which they were developed, although we are quite aware that, most probably, the user/reader will have the opposite approach, at least at the first glance.

3.1. Considerations on the selection of data sources

A few national and international organizations have produced more or less complete lists of parameters (Fig. 1) (established according to not always clearly stated criteria; most often,the availability of data for the first decades of this century was a limitation; at the very beginning, some organizations even used to publish lists of arrival times, "associated" or not, as we would say in our modern technical language, and eventually publish parameters derived from them, if possible): at the very end of the XIXth Century, the British Association for the Advancement of Sciences, relayed by the International Seismological Summary [ISS], started first; the U.S. Coast and Geodetic Survey [US CGS], relayed by the U.S. Geological Survey [US GS]; the Bureau Central International de Séismologie [BCIS] (Central International Bureau for Seismology), more or less relayed by the Centre Sismologique Euro-Méditerranéen [CSEM] (European-Mediterranean Seismological Centre [EMSC]), restricting the geographic domain of concern to Europe and the Mediterranean (in a broad sense). In addition, one should include in the list, the famous work of Gutenberg and Richter (1954), founding stone of global views on seismic activity on the Earth. And last, in 1964, the International SeismologicaI Centre [ISC] initiated its ongoing service of collecting data all over the world and regularly issuing earthquake determinations. It should be stressed that:

1. apart from ISC, US GS (partially ?), and EMSC, the corresponding data sets are not readily available in full digital form;
2. excepting the US CGS, the ISC was the first organization to systematically draw advantage from the digital computing facilities.

While collecting data, ISC gets both phase data and (more or less) preliminary evaluations of earthquake parameters. Phase data are the basic input data for ISC's computations of parameters, while preliminary evaluations are used eventually to initiate the iterative computational process; in addition to its own parameter determination, ISC reports on the evaluations provided by other agencies in its "Catalogue" [ISC-C] as well. In other words, ISC associates different estimations of parameters, supposedly related to the same physical event, and its bulletin comes out with, for any event (re)located by the Centre, a "principal determination" (i.e. the parameter determination computed by ISC), and a list of other determinations understood to depict the same unique physical event. Obviously, if the parameters of an event (usually a weak one) are reported to ISC, and the latter has no additional phase data to associate it to, then the parameters originally reported show up alone (as principal determination) on the final list (for a full account of the work performed by ISC, see the January or July issues of any year of the ISC Bulletin or refer to Adams et al., 1982).

ISC has built up a back file for the years before it started operations. This so-called "Historical File" [ISC-HF] (the name is somewhat misleading, as it causes confusion with "historical seismicity" data which refers to pre-instrumental observations of earthquake effects, which is by no means the background of ISC-HF) follows a scheme similar to the one adopted for ISC's Catalogue [ISC-C]: several determinations are eventually associated, by being related, in principle, to a unique physical event, and one of which is declared as the "principal" one. As explained above (see "Introduction"), and touched upon again just below, associating different determinations to a unique physical event is a very hard decision, and we have deliberately ignored the grounds on which associations were made by ISC, as well as how one determination out of the whole set is declared as the "principal" one.

Our primary interest is that such a file exists, and whatever the rationale adopted to build it up, it should be easily accessible to any user/reader and provide a data set to work on. This is equally true of the ISC-C. However, it is quite manifest that the determinations, grouped about a supposedly unique physical event, come from rather disparate sources; once one determination has been singled out of the determination set, following a certain procedure, there is little chance that the resulting overall set of singularized determinations show a large degree of homogeneity: this is a simple "mechanical" effect of using varied sources of parameters, which in turn is imposed by the deficiencies of each individual source.

For the period starting in January 1976 to date, EMSC has produced monthly lists of parameters which largely cover the Mediterranean area. From the beginning, EMSC's determinations were computed using the same algorithm and models throughout, and by the same operator who applied the same (or, at least, slowly and smoothly changing) criteria in the computation process. We then have at hand, for that period, a data set for which a certain homogeneity is guaranteed, as far as practicable; this is one of the very few data sets which displays such a feature over a relatively long period of time (13.5 years, in the present case) and for a rather vast area (the entire Mediterranean, for sure). It has thus been decided to rely on the EMSC's monthly lists from 1976 on; there is a unique determination per physical event considered, and thus no need to choose among several determinations during the corresponding period of time; the use of that data set is straightforward.

The initial Bathymetric Chart project had set geographic limits: 30° N to 46° N, and 6° W to 36.5° E (plus an insert for the Black Sea: 40° N to 47.5° N, and 26.5° E to 42.5° E). The seismicity map must match these limits, even though the tectonic processes causing the earthquakes affect a much larger area (i.e., from the large scale tectonics point of view, the territory is meaninglessly truncated). The earthquake parameter sources embrace a larger area as well: they should be run through a sharp spatial band-pass filter corresponding to the geographic limits stated above. In order to avoid undesirable side effects (and, in particular, so as not to eliminate abusively determinations falling just on the domain's borders), the whole analysis described below has been applied to an area extending beyond the nominal area (25° N to 48° N, and 15° W to 45° E): in the following, the words "[domain] restricted to the Mediterranean" refer to the territory included in the later limits.

Table 1 summarizes the overall situation regarding data sources.

Table 1 - Data sources.

• beginning of century - 1963 (included)	**ISC's Historical File**	eventually several determinations per physical events
• 1964 - 1975 (both ends included)	**ISC's Catalogue**	eventually several determinations per physical events
• 1976 (included) onwards	**EMSC's monthly lists**	1 determination per physical event

3.2. The ISC's Historical File

Considering the restriction of ISC-HF* to the Mediterranean area, the sources called upon by ISC to construct the file are:

- Gutenberg and Richter (1949)
- ISS
- BCIS
- US CG

(See Fig. 2)

(the other possible sources eventually used are not relevant to our zone of interest).

Gutenberg and Richter (1949), in their two successive editions, have compiled the data sets they could collect; the first edition (published in 1949) covers the period 1904-1946, while the second (published in 1954) includes a few addenda to the previous period and covers the additional 1947-1952 period. Their sources are much too numerous to be detailed here; nevertheless, those most often used are ISS (from 1918 to 1937), the bulletins of the International Union of Geodesy and Geophysics [IUGG] published by the Bureau Central [BCIS] (starting in 1937), and the bulletins of the US CGS (starting in 1937).

Prior to the formal BCIS bulletin, the Monthly Bulletins were actually published by the Bureau Central starting in 1923; they were issued then under the formal banner of BCIS. The British Association for the Advancement of Science has taken care of publishing the earliest data and parameter lists; the name ISS was introduced when the Association of Seismology of the IUGG endorsed the publication, in 1923.

Directly or through Gutenberg and Richter (1949), the ISC-HF has apparently exhaustively exploited Gutenberg and Richter (1949) itself, ISS, and partly the two other main sources: possibly, exploitation has been less exhaustive for the BCIS bulletins prior to 1937, and the US CGS bulletins after 1960.

In the ISC-HF, a determination appears as a list of parameters, which includes:
- reporting agency,
- origin time*,
- epicentral latitude and longitude,
- focal depth*,
- magnitude*,
- maximum intensity*,
- region (in the sense of Flinn and Engdahl, 1965).

Parameters marked by an * symbol are occasionally missing. Determinations are grouped in a cluster when they are considered relating to the same unique physical event; the first determina-

* It should be noticed that ISC also used to disseminate also a so-called "Extended Historical File", which was made up of the ISC-HF proper, the principal determinations (i.e. the ISC's prime estimate) for the period of availability (from 1964 up to the last ISC's Bulletin issue concatenated), and supplemented with other sources of preliminary determinations issued faster than ISC (in practice, the US GS's different types of rapid preliminary determinations).

tion in the list is then taken as the "principal determination". Naturally, when the cluster is made of only one determination, the latter is necessarily the "principal determination".

Table 2 shows the agencies which have most significantly contributed to ISC-HF. The left half-column reports the number of events for which the agency involved has contributed the only determination available for the corresponding event (= "singular determination events"), while the right half-column reports the number of events for which the agency involved has contributed one determination associated with other(s) (= "multiple determination events").

For the period 1904-1927, the ISC-HF restricted to the Mediterranean, shows only single determination events (see Table 2): this is because the only source involved (as is apparent on the bar-chart of Fig. 2) is Gutenberg and Richter (1949), who provide their own selection of determinations, or even their own computed determinations. From 1928 to 1934, most of the deter-

Table 2 - Contributions of agencies to the ISC Historical File [restricted to the Mediterranean; see Section 3. 1. "Considerations …", in fine] (in number of events per annum). Only those agencies which have the most significant contributions are shown; "minor" agencies are not included. The left half-column refers to the number of singular determination events; the right half-column refers to the number of multiple determination events. For the meaning of the acronyms designating the agencies, see the Appendix 2.

	GUTE	BCIS	ISS	CGS	ALG	ALI	ATH	BUC	MOS	PDE	PRA	ROM	STU	TRI	ZUR
1904...1927	56 :	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1928...1934	59 : 8	-	-	1 : 8	-	-	-	-	-	-	-	-	-	-	-
1935...1939	10 : 29	-	131 : 25	3 : 26	-	-	-	-	-	-	-	-	-	-	-
1940...1944	1 : 43	-	173 : 48	: 26	-	-	-	-	-	-	-	-	-	-	-
1945...1949	: 19	-	236 : 50	5 : 41	-	-	-	-	-	-	-	-	-	-	-
1950	: 1	12 : 44	25 : 52	: 14	-	-	-	-	-	: 4	: 5	1 : 12	-	1 : 4	-
1951	: 1	22 : 53	21 : 59	: 3	-	-	-	-	1 :	: 7	: 1	1 : 12	: 4	: 1	-
1952	-	28 : 79	23 : 84	: 2	-	1 : -	: 3	-	-	: 5	: 2	1 : 12	-	1 : 14	-
1953	-	111 : 55	2 : 48	: 3	-	: 1	-	-	: 9	3 : 31	: 2	7 : 9	-	-	-
1954	-	120 : 38	: 27	3 : 29	-	2 : 1	: 3	-	-	-	: 1	: 1	-	-	1 : 2
1955	-	61 : 23	: 19	1 : 19	1 :	5 :	: 1	-	-	-	: 1	4 : 3	-	1 : -	-
1956	-	82 : 45	: 35	: 29	4 : 2	7 : 3	: 1	4 : 1	1 : 16	-	1 : 5	1 : 1	: 1	: 1	-
1957	-	127 : 40	: 41	1 : 47	: 3	4 :	-	2 :	3 : 19	-	-	6 : 1	-	-	2 : -
1958	-	126 : 45	1 : 38	2 : 32	: 1	5 : 1	-	-	1 : 28	-	-	2 : 1	-	-	6 : -
1959	-	207 : 92	: 66	: 64	: 3	-	: 1	1 :	: 53	-	2 :	: 3	: 3	-	4 : 6
1960	-	271 : 62	: 44	10 : 62	2 : 1	3 :	-	1 : 1	1 : 38	-	-	2 : 2	: 1	-	: 2
1961	-	185 : 18	7 : 18	-	-	-	-	1 :	3 : 18	-	-	5 :	: 2	-	-
1962	-	252 : 57	2 : 38	-	-	-	: 1	-	26 : 49	-	-	3 : 3	1 : 4	-	-
1963	-	275 : 75	: 40	-	-	-	: 1	-	6 : 57	: 6	-	1 : 1	: 1	-	-

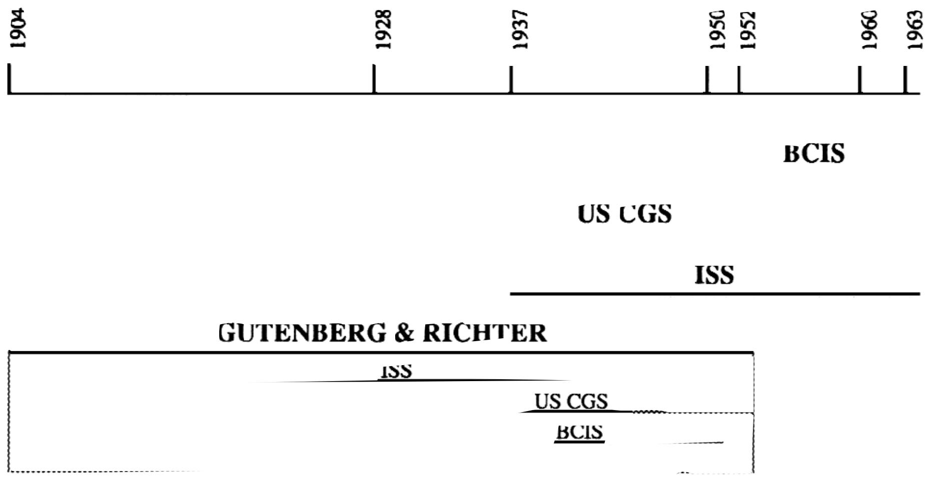

Fig. 2 - Bar-chart of sources making up the ISC Historical File [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine].

minations (about 80 %) are still contributed by Gutenberg and Richter (1949). This changes radically in 1935, when the number of determinations increases drastically (see Table 3): the average number of determinations per year is almost 60 over the period going from 1935 to 1949, while it was only 10 from 1928 to 1934. Over the same time span of 15 years (1935-1949), the number of determinations is, on the other hand, multiplied by a factor of 1.6; the increase affects, in about the same proportion, both the singular determination events and the multiple determination ones, although altogether the multiple determinations undergo an increase in the average number per year which is roughly twice as large as the increase in the same average number for the singular determinations; this comes from the fact that singular determination events largely outnumber the multiple determination ones. Nevertheless, the drastic change in 1935 is due to the influx of data from ISS. The same effect is again apparent, to a lesser extent, in 1950 when the BCIS's determinations started to be considered directly in the building up of ISC-HF; actually, the effect is more difficult to analyze because other, additional, agencies are considered and exploited starting at the same time (e.g. PRA, ROM, TRI) ; they are quoted from BCIS or ISS, and are very important because of their massive contribution to the creation of multiple determination events: these agencies (only the most important ones are reported in Table 2) have produced about 600 determinations between 1950 and 1963; the de facto filter, imposed on the data set by the selection process implied by using the Gutenberg and Richter (1949) source, is no longer operative as this latter source is practically exhausted. From 1953 onward, BCIS takes over ISS and is the source of most of the singular determinations. Furthermore, the number of determinations per year becomes large enough to allow a sensible statistical analysis on an annual basis, while preciously years had to be grouped together to make up a reasonably large number of events (in a statistical sense).

Tables 4 and 5 summarize the contribution of the different agencies in terms of percentage; including all the reporting agencies would have made the tables too big: the "minor" agencies which have contributed little data globally are not included. The two tables differ from each

Table 3 - ISC Historical File [restricted to the Mediterranean; see Section 3.l. "Considerations …", in fine]. Total number of events and determinations per annum.

	Number of determinations	Number of events	Number of singular determination events	Number of multiple determination events
1904 … 1927	56	56	56	0
1928 … 1934	76	68	60	8
1935 … 1939	224	177	144	33
1940 … 1944	291	222	174	48
1945 … 1949	351	291	241	50
1950	175	91	39	52
1951	203	108	47	61
1952	269	140	54	86
1953	305	182	123	59
1954	230	165	126	39
1955	141	96	73	23
1956	243	150	101	49
1957	299	199	146	53
1958	285	186	137	49
1959	512	307	214	93
1960	508	362	293	69
1961	260	226	202	24
1962	468	353	293	60
1963	487	361	288	73

other. Table 4 shows the origin of singular determinations: the total of each row of the table reaches 100 %, unless "minor" agencies have contributed some of the singular determinations, but these do not show up in the table. The predominance of Gutenberg and Richter (1949) before 1935, ISS from 1935 to 1950, and BCIS from 1953 onward, is clearly apparent. Table 5 shows the percentage of multiple determination events to which one given agency has contributed one of the associated determinations; each figure could ideally amount to 100 % if any agency shown had contributed one determination to each multiple determination event (which is, a long way from being the case); the agencies showing the largest percentages are US CGS, MOS and ROM.

Globally, the number of events for which parameters have been determined increases from 1950 onward, with two significant reverse occurrences, in 1955-1956 and in 1961 (it is less clearly significant in 1958) (Table 3). The proportion of singular determinations, as compared to the actual number of events, shows a marked low in 1950-1952 (see Fig. 3); this low corresponds to the short period during which all main sources contributed simultaneously according to the sketch in Fig. 2, which could a priori induce a larger number of the same events to be reported by different sources (inducing multiple determination events). However, things are, no doubt, more complex. Table 2 in fact shows that contributions to multiple determinations by Gutenberg and Richter (1949) virtually vanished after 1949; a more significant explanation seems, then, to be that the proportion of singular determinations effectively decreased at that time because of a lag between a certain loosening in the (relative) completeness of the survey of the actual seismic activity by ISS and the taking over by BCIS. Conversely, the relative number of multiple determination events against the total number of actual events, reaches a peak between in 1950-1952,

Table 4 - ISC Historical File [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Agencies showing the most significant contributions to singular determinations per annum (in %). A dash (—) in a box means that the source considered has not been used to build up the corresponding part of the ISC Historical File; an empty box stands for 0 % (the source considered has contributed no determination to the corresponding period of the ISC Historical File). For the meaning of the acronyms designating the agencies, see the Appendix 2.

	GUTE	BCIS	ISS	CGS	ALI	BUC	MOS	ROM	ZUR
1904 … 1927	100.0	—	—	—	—	—	—	—	—
1928 … 1934	98.3	—	—	1.7	—	—	—	—	—
1935 … 1939	6.9	—	91.0	2.1					
1940 … 1944	0.6	—	99.4						
1945 … 1949		—	97.9	2.1					
1950		30.8	64.1					2.6	
1951		46.8	44.7				2.1	2.1	
1952		51.9	42.6		1.8			1.8	
1953	—	90.2	1.6					5.7	
1954	—	95.2		2.4	1.6				0.8
1955	—	83.6		1.4	6.8			5.5	
1956	—	81.2			6.9	4.0	1.0	1.0	
1957	—	87.0		0.7	2.7	1.4	2.1	4.1	1.4
1958	—	92.0	0.7	1.5	3.6		0.7	1.5	4.4
1959	—	96.7				0.5			1.9
1960	—	92.5		3.4	1.0	0.3	0.3	0.7	
1961	—	91.6	3.5	—		0.5	1.5	2.5	
1962	—	86.0	0.7	—			8.9	1.0	
1963	—	95.5		—			2.1	0.3	

Table 5 - Historical File [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Agencies showing the most significant contributions to multiple determinations per annum (in % of events occurring during the year(s) considered, for which a certain agency has contributed one determination). A dash (—) in a box means that the source considered has not been used to build up the corresponding part of the ISC Historical File; an empty box stands for 0 % (the source considered has contributed no determination to the corresponding period of the ISC Historical File). For the meaning of the acronyms designating the agencies, see the Appendix 2.

	GUTE	BCIS	ISS	CGS	ATH	MOS	PRA	ROM	STU
1904 … 1927	—	—	—	—	—	—	—	—	—
1928 … 1934	100.0	—	—	100.0	—	—	—	—	—
1935 … 1939	87.9	—	75.8	78.8					
1940 … 1944	89.6	—	100.0	54.2					
1945 … 1949	38.0	—	100.0	82.0					
1950	1.9	84.6	100.0	26.9			9.6	23.1	
1951	1.6	86.9	96.7	4.9			1.6	19.7	6.6
1952		91.9	97.7	2.3	3.5		2.3	13.9	
1953	—	93.2	81.4	5.1		15.3	3.4	15.3	
1954	—	97.4	69.2	74.4	7.7		2.6	2.6	
1955	—	100.0	82.6	82.6	4.3		4.3	13.0	
1956	—	91.8	71.4	59.2	2.0	32.7	10.2	2.0	2.0
1957	—	75.5	77.4	88.7		35.8		1.9	
1958	—	91.8	77.6	65.3		57.1		2.0	
1959	—	98.9	71.0	68.8	1.1	57.0		3.2	3.2
1960	—	89.9	63.8	89.9		55.1		2.9	1.4
1961	—	75.0	75.0	—		75.0			8.3
1962	—	95.0	63.3		1.7	81.7		5.0	6.7
1963	—	100.0	54.8	—	1.4	78.1		1.4	1.4

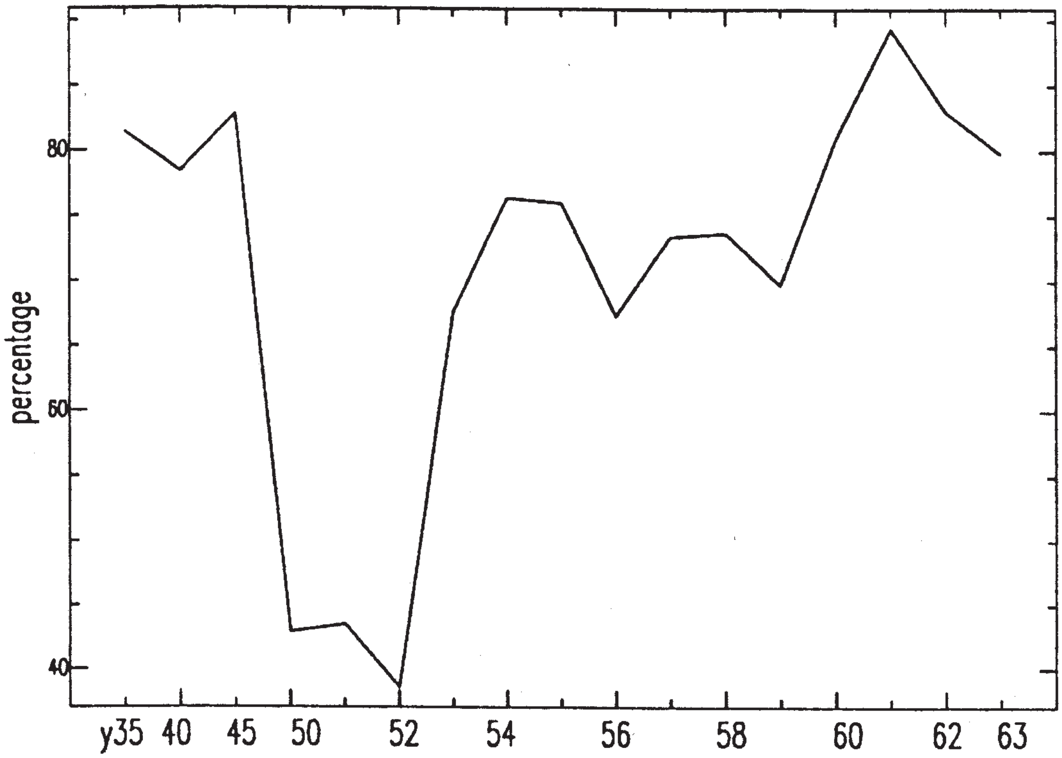

Fig. 3 - ISC Historical File [restricted to the Mediterranean; see Section 3.l. "Considerations …", in fine]. Percentage of singular events per annum.

before going roughly back to its prior 1950 level, displaying a minimum in 1961 (this is depicted in Fig. 3, if interpreted as the complement to 100 of the percentage of singular determinations).

Fig. 4 illustrates how multiple determinations tend to increase their multiplicity from about 1950 to about 1960, at least if one considers the fourfold multiple determinations; this is apparently still true with fivefold multiple determinations over a more limited period (1955-1960), but the curve is then strongly influenced by samples of little significance, in terms of statistics. It appears that there is a correlation between the above-noticed trend and contributions from BCIS: the latter becomes an important source of determinations from different origins in 1950. Before then, no event is described by four (or more) determinations. Similarly, in 1960-1963, the fourfold multiple determination proportion drops, most probably because of the non-consideration of determinations by US CGS, if any exist at all (see Fig. 2, Tables 2 and 5).

3.3. The ISC's Catalogue

As mentioned above (see Section 3. 1. "Considerations ..."), the ISC Catalogue [ISC-C]

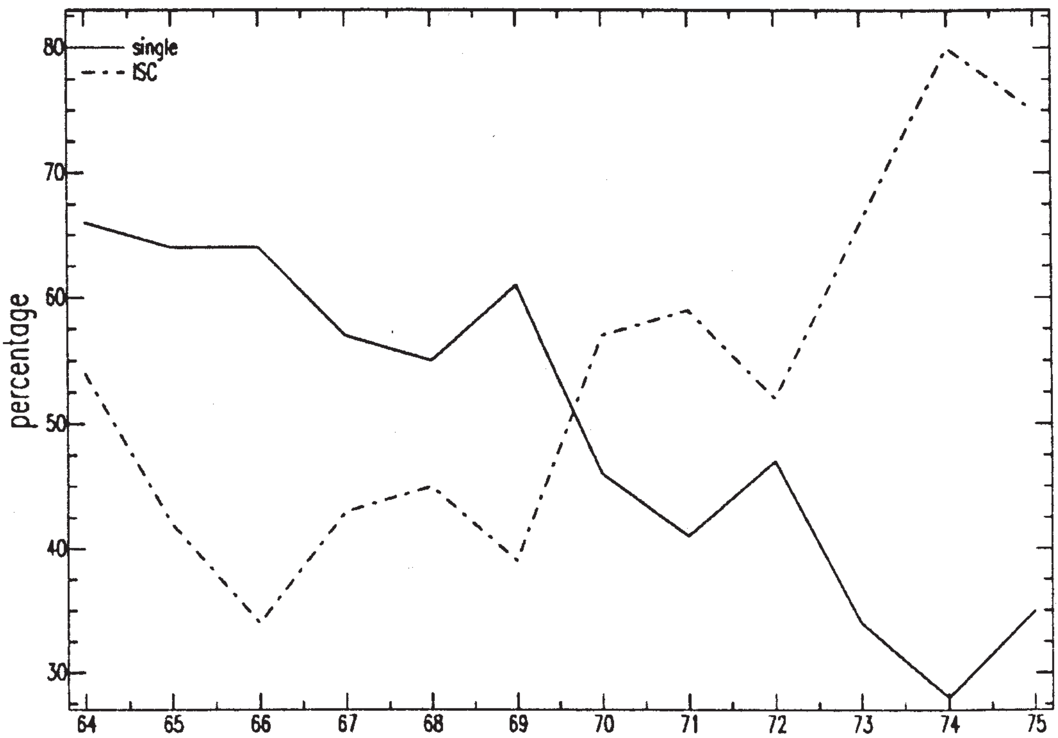

Fig. 4 - Historical File [restricted to the Mediterranean; see Section I.l. "Considerations …", *in fine*]. Percentage of n-fold determination events per annum, with n = 4 (solid line), and n = 5 (dashed-dotted line).

covers the period starting in January 1964. It includes, if appropriate, determinations reported by other agencies, associated together and to the determinations computed by ISC, if relevant (if it exists at all, the ISC's determination is "principal"; if there has not been any relocation computed by ISC, one of the reported determinations is declared "principal"). Parallel to what has been done on ISC-HF, some statistical analyses have been conducted on ISC-C. Actually, one could think of relying a priori on the "principal determinations" supplied by ISC from 1964 on, for the very reason that ISC reviews all the determinations issued by other agencies, and relocates the events for its own use when practicable. We have considered, a contrario, that

1. we had no indication that homogeneity in the processing and quality assessment by ISC, since 1964, could be guaranteed;
2. we could not decide a priori that no other agency could get "better results" (this term will be clarified in the following) than ISC, at least at the beginning of its operations;
3. it was interesting to pursue the statistical analysis applied to ISC-HF, and to compare it with the results obtained on both ISC-HF and ISC-C;
4. a posteriori, the approach has been somehow justified in the sense that the conclusion of the study is that ISC is recognized as one of the "best reporting" agencies for most of the period considered, as expected a priori; which, in a way, confirms the robustness of the

Table 6 - ISC Catalogue [restricted to thc Mediterranean; see Section 3.1. "Considerations …", in fine]. Agencies* showing the most significant contributions to singular deteminations *per annum* (upper row: number of events/determinations; lower row: percentage of contributions to singular determinations). For the meaning of the acronyms designating the agencies, see the Appendix 2.

	ISC	BCIS	KSA	MDD	VIE	LIS	ROM	ZUR	ATH	LAO	RBA SPGM	TIR	TIF
1964	134 33.8	71 17.9	7 1.8	65 16.4	16 4.0		30 7.6		47 11.9		11 2.8		
1965	74 19.6	71 18.8	16 4.2	29 7.7	26 3.4		10 2.7		118 31.3		28 7.4		
1966	2 0.3	53 7.2	20 2.7	33 4.5	2 0.3		1 0.1		568 76.7	3 0.4	48 6.5		
1967	30 5.3	31 5.4	7 1.2	27 4.7	5 0.9				383 67.3	14 2.5	60 10.5		
1968	13 1.8	41 5.6	4 0.5	29 4.0	6 0.8		2 0.3		473 64.4	12 1.6	152 20.7		
1969	14 1.4	7 0.7		19 1.9	7 0.7	1 0.1	2 0.2		226 22.6	14 1.4	693 69.2		
1970	49 8.8	9 1.6		5 0.9	5 0.9				160 28.8	71 12.8	247 44.5		
1971		12 2.1		12 2.1	2 0.4		1 0.2		107 18.8		209 15.8		
1972	3 0.4	50 6.3		14 1.8	12 1.5	4 0.5	9 1.1		320 40.2		126 15.8		254 31.9
1973	1 0.2	1 0.2		13 2.7	3 0.6				122 25.5		103 21.5		233 48.7
1974	156 31.3	2 0.4		10 2.0	6 1.2	4 0.8	11 2.2	31 6.2	153 30.7		112 22.5	13 2.6	
1975	254 32.1	1 0.1		12 1.5	5 0.6	14 1.8	3 0.4	105 13.3	134 16.9		91 11.5	146 11.5	

method applied.

For the restriction of ISC-C to the Mediterranean region (see Section 3.l. "Considerations …", in fine), Table 6 displays the number of singular determinations contributed by the most

* Actually, the following data is reported in Table 6 :
- if *n* is the number of determinations reported from 1964 to 1975 (included) by one agency,
- if *p* is the number of years during which this agency has been active contributing determinations (one should rather say, during which, results produced by this agency have been included in ISC-C), then the agency's contributions are reported in Table 6 in the following cases :
  - $n \geq 400$,
  - $p \geq 5$,
  - $n \geq 20$ and $p \geq 4$,
  - $n \geq 50$ and $p \geq 3$,
  - $n \geq 100$ and $p \geq 2$.

Table 7 - ISC Catalogue [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Agencies showing the most significant contributions to multiple determinations per annum (upper row: number of multiple determination events for which one determination has been issued by the corresponding agency; lower row: percentage of multiple determination events including a determination from the corresponding agency). For the meaning of the acronyms designating the agencies, see the Appendix 2.

	ISC	NEIC	BCIS	MDD	MOS	ROM	ATH	LAO	SPGM	HFS1	HFS2
1964	199 92.6	116 54.0	145 67.4	26 12.1	72 33.5	11 5.1	26 12.1				
1965	178 83.6	149 70.0	160 75.1	24 11.3	26 12.2	8 3.8	63 29.6				
1966	441 93.2	269 56.9	251 53.1	32 6.8	15 3.2		295 62.4				
1967	495 94.6	218 41.7	238 45.5	39 7.5	13 2.5	2 0.4	331 63.3		23 4.4		
1968	717 97.3	195 26.5	230 31.2	52 7.1	70 9.5	3 0.4	539 73.1		32 4.3		
1969	715 98.8	259 35.8	199 27.5	84 11.6	122 16.9	5 0.7	515 71.1	2 0.3	51 7.0		
1970	882 99.1	261 29.3	228 25.6	53 6.0	166 18.7	1 0.1	636 71.5	1 0.1	74 8.3		
1971	1059 99.6	213 20.0	214 20.1	48 4.5	134 12.6	2 0.2	519 48.8	235 22.1	40 3.8	17 1.6	
1972	1118 99.3	207 18.4	370 32.9	43 3.8	109 9.7	17 1.5	701 62.3	295 26.2	74 6.6		
1973	969 99.9	219 22.6	190 19.6	46 4.7	93 9.6	8 0.8	668 68.9	189 19.5	129 13.3	26 2.7	119 12.3
1974	1292 99.5	269 20.7	258 19.9	44 3.4	88 6.8	33 2.5	937 72.1	255 19.6	102 7.9	22 2.5	153 11.8
1975	1473 99.7	460 31.1	267 18.1	37 2.5	120 8.1	21 1.4	983 66.5	121 8.2	126 8.5	53 3.6	186 12.6

important agencies* (in terms of determinations contributed); because not all the contributing agencies show up in the table, the total of percentages for each row of the table does not amount to 100 % in the end. Many determinations were contributed by ATH, and RBA (1964-1966), followed by SPGM (1967 onward).

Table 7 displays the sources of multiple determinations. By inspecting of "ISC" column, it is evident that this Centre has relocated most of the multiple determination events, and thus contributed one determination for most of these events. Table 8 summarizes the overall picture for the period 1964 - 1975: the general trend is an increase in the number of actual events conside-

* In spite of the many successive changes in the administrative attachments of what has now become the National Earthquake Information Center [NEIC], under the US GS, the NEIC acronym only has been used in the following for the whole period 1964-1975, for simplicity. This is intended for convenience, and has absolntely no other implication of any kind.

Table 8 - ISC Catalogue [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Total number of events/determinations per annum.

	Number of determinations	Number of events	Number of singular determination events	Number of determinations related to multiple determination events	Number of multiple determination events
1964	1053	624	409	644	215
1965	1020	597	384	636	213
1966	2238	1308	836	1402	472
1967	2111	1230	707	1404	523
1968	2731	1621	899	1832	722
1969	3098	1851	1132	1966	719
1970	3098	1622	744	2354	878
1971	3636	1808	746	2890	1062
1972	4018	2136	1010	3008	1126
1973	3209	1477	509	2700	968
1974	4073	1807	508	3565	1299
1975	4920	2278	804	4116	1474

red per year, with some fluctuations and a marked relative minimum in 1973. Over the same period, there is a general decrease in the proportion of singular determinations, while Fig. 5 shows that ISC tends to compute parameters of more and more events. Comparing the two curves of Fig. 5 suggests that ISC either relocates events which are, otherwise, reported by only one non-ISC agency, and/or that ISC has been able to associate enough data to compute a singular determination (this probably happened in 1974-1975, where the percentage of singular determinations reported by ISC is unusually high; see Table 6).

The average number of determinations per multiple determination event (column 5 divided by column 6, Table 8) is close to 2.75 throughout the period considered. Fig. 6 shows a general decrease in the percentage of fourfold multiple determinations, while the fivefold multiple determinations proportion is roughly constant over the period under examination; as the percentage of singular determinations significantly decreased over the same time period, twofold and threefold multiple determination percentages must have increased.

3.4. EMSC's data file

In January 1976, the Europeo-Mediterranean Seismological Centre [EMSC] (which changed its name slightly, in 1983 to European-Mediterranean Seismological Centre) started its operations. It was basically the continuation of the BCIS operations, but restricted to Europe and the Mediterranean (in a broad sense: from the Azore Islands to the longitude of the shores of the Caspian Sea, and from North Africa to the Arctic Ocean). Based on phase data collected from a large number of stations (most of the stations located within the geographic domain of concern, and many others, depending on the magnitude of the event considered), hypocentral determinations are computed and the corresponding monthly lists issued with a 3 to 6 month delay; hypo-

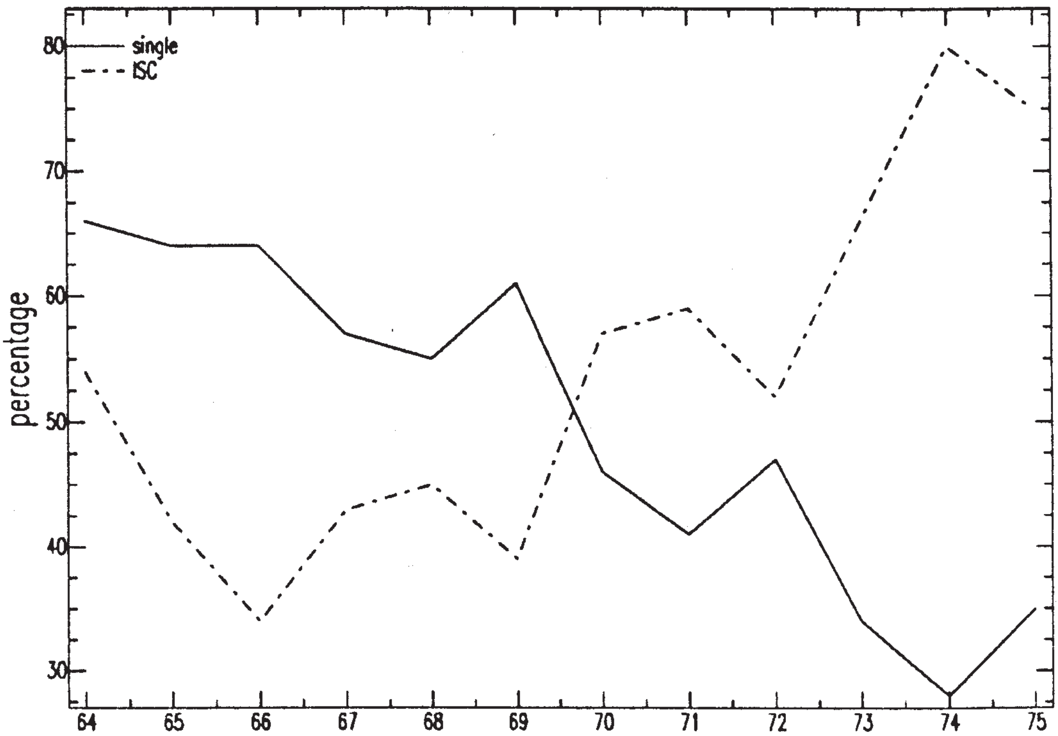

Fig. 5 - ISC Catalogue [restricted to the Mediterranean; see Section 3. 1. “Considerations …” in fine]. Global percentage of singular determination events *per annum* (solid line), and percentage of (singular and multiple determination) events determined by ISC per annum (dashed-dotted line).

central parameters are computed from the phase data available at the time calculations are done. Magnitudes are not computed, but simply reported from available estimations obtained from contributing agencies or stations. No magnitude threshold is considered for selecting the reported hypocenters; but only two criteria apply which are very different in nature. One criterium is a rule of thumb quality factor, evaluated by combining the usual computed error estimators (standard deviation on time of occurrence, error in distance in the [latitude, longitude] plane) with the azimuthal gap in station coverage, the number of station data used in the computations, and the number of phase readings discarded because yielding too large residuals. Another criterium is less “scientific” and aims at making sure that the Centre brings in a real international “added-value”: at least two national networks must have contributed data (exceptionally, if there is a relatively large number of data reported from one country, the hypocentral parameters are still computed; this is actually rather unlikely to occur in Europe, because of the parcelling out of space among many different countries). No event located with less than ten station data is included in the disseminated lists.

The procedure applied does not guarantee geographic homogeneity in the hypocenters determined, which is in any case primarily and strongly controlled by the distributon of stations reporting data; actually, the selection procedure has only a secondary influence. Nevertheless, the procedure described above has been applied, with no significant change, by the same analyst, since

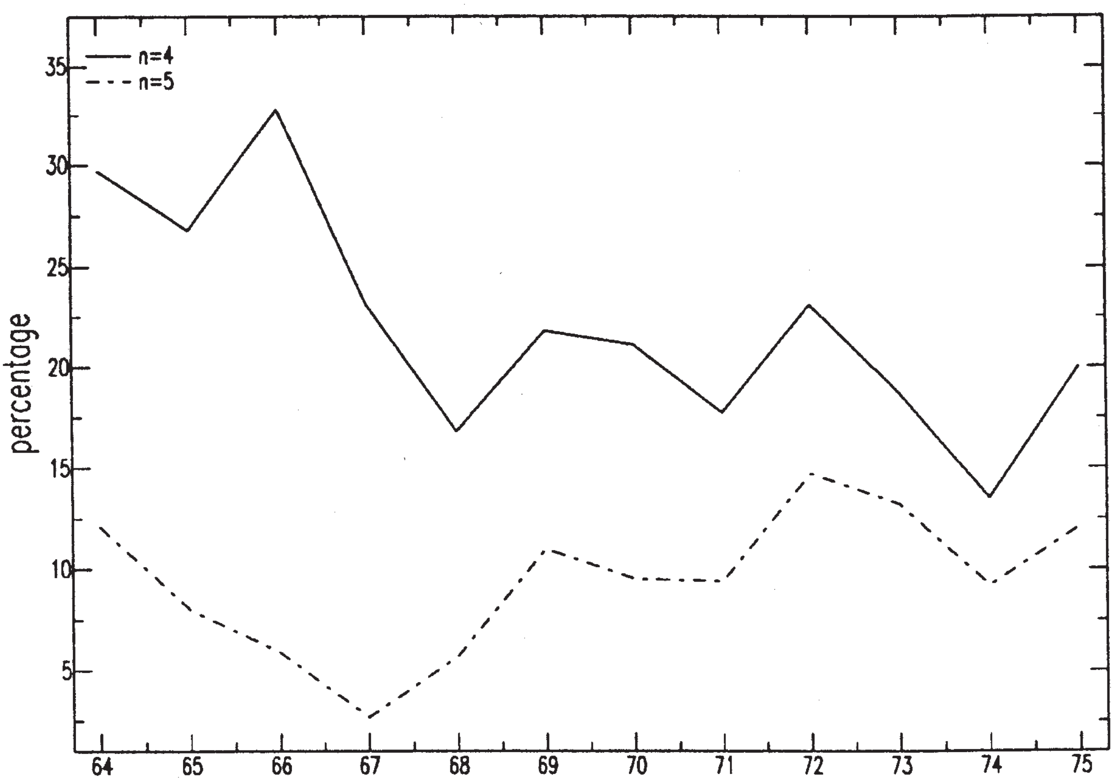

Fig. 6 - ISC Catalogue [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Percentage of n-fold determination events per annum, with n = 4 (solid line), and n = 5 (dashed-dotted line).

1976: this brings in some sort of homogeneity to the EMSC's data set which has, to our knowledge, no equivalent; it makes the data set very valuable in terms of statistics. This is the reason why it has been preferred to any other for the period 1976-1988.5. The EMSC's monthly lists are widely disseminated and readily accessible to any interested user.

However, in order to make sure that adopting the EMSC's lists as a primary data set from 1976 on does not corrupt the efforts made in choosing the primary data sets for the earlier periods, tests have been conducted. Anticipating on the analyses done concerning the epicentral locations issued by different agencies (see Sections 4.l. "Multiple determinations" and 5.l. "Epicentral locations in the final data set"), Fig. 7 presents a comparison between the proportion of large pairwise mismatch distances induced by EMSC's determinations for the multiple determination events from 1978 to 1981, and those induced either by ISC or NEIC; the two latter are reputed to be the "best reporting" agencies from 1964 to 1975 (see Section 5.l. "Epicentral locations in the final data set"). The pairs of determinations are formed by associating, for any given event, the determination issued either by EMSC, or ISC, or NEIC, to any other determination (represented by "ZZZ"). One can observe that the proportions of pairwise mismatch distances of more than 60 km roughly follow the same trends for the three above-mentioned agencies; the percentages do not deviate from each other by more than 4-5 units; percentages associated to

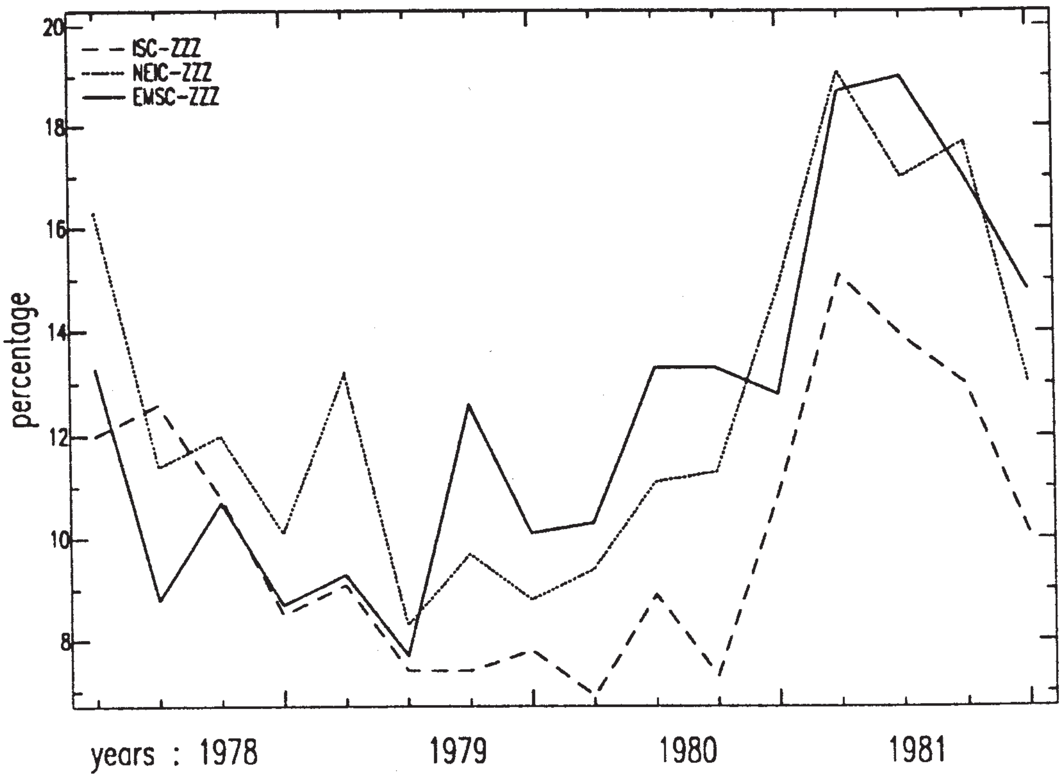

Fig. 7 - Proportions of pairwise mismatch distances larger than 60 km induced by the pairs involving ISC, NEIC or EMSC determinations for the period 1978 - 1981 (See Sections 3.l. "Considerations ...", 4.l. "Multiple determinations" and 5.1. "Epicentral locations in the final data set"). Solid line: pairs EMSC-ZZZ; dotted line: pairs NEIC-ZZZ; dashed line: pairs ISC-ZZZ. For the meaning of the acronyms designating the agencies, see the Appendix 2.

EMSC's determinations lie in between or close to the two others, except from 1979.75 to 1980.75 when it is slightly above them; on the other hand, up to 1979.5, percentages concerning ISC and EMSC's determinations are quite similar, and ISC is declared the "best reporting" agency for the immediately preceding period (1974-1975) (see Section 5.1. "Epicentral locations in the final data set"). The results of the test lead to the conclusion that EMSC data set does not behave significantly worse than the two other agencies, reputed the "best" ones for the previous twelve years. There is thus no obstacle to choosing EMSC's lists as the primary data set for 1976 on, on the ground of their intrinsic previously described properties.

4. About event parameters

4.1. Multiple determinations

The way the association of determinations from various sources has been made, has been

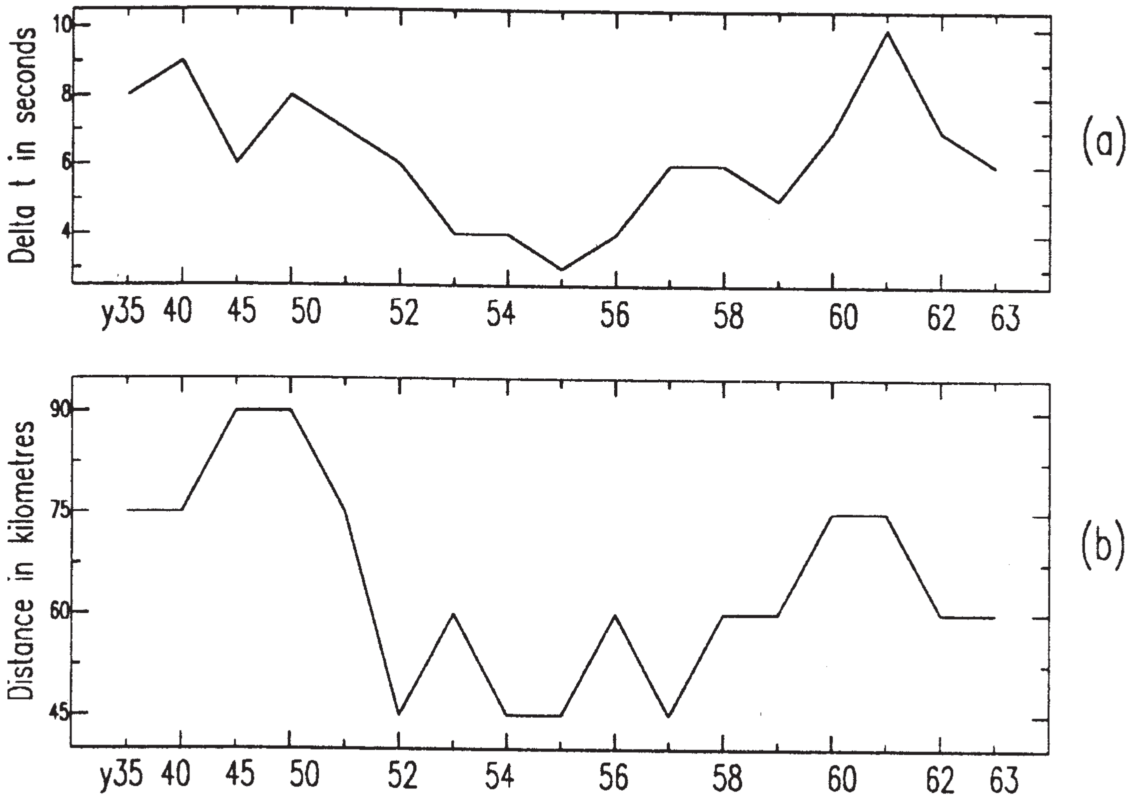

Fig. 8 - ISC Historical File [restricted to the Mediterranean; see Section 3.l. "Considerations …", in fine]. Multiple determination events. (a) Origintime pairwise difference threshold class (a class spans 1 sec) to be considered per annum in order to include 75 % of all origin-time pairwise differences. (b) Pairwise mismatch distance threshold class (a class spans 15 km) to be considered per annum in order to include 75 % of all pairwise mismatch distances.

ignored; firstly, because the rationale for it is not readily available (particularly for the ISC-HF); secondly, because it has been decided, as explained above, to rely on a data set that anyone could easily access and rework: the data set should be then taken as it is and further analyzed to get the best out of it. It follows that only one entry per physical event should show up in the final data set; hence, the problem of choosing one determination out of a set of several associated ones, comes up; and before setting any criterium for the choice, the properties of the set of associated determinations should be investigated.

In the following, we shall consider, for a certain earthquake (origin time, location), the distance (one- or two-dimensional) between any two values assigned to determinations related to one physical event: the so-called "pairwise differences" (in time or in location in the [latitude, longitude] plane). Within ISC-HF, all possible pairs in a set of associated determinations (n-fold multiple events) have been investigated. Pairwise differences in origin-times between the associated determinations have been sorted by classes of 1 second duration. Fig. 8a shows, as a function of time, which origin-time difference class should be considered in order to include at least 3/4 of the pairwise time differences: this threshold class has changed along the years (there is a marked minimum in 1955); nevertheless, as a rough estimate, one should go up to the 6 second

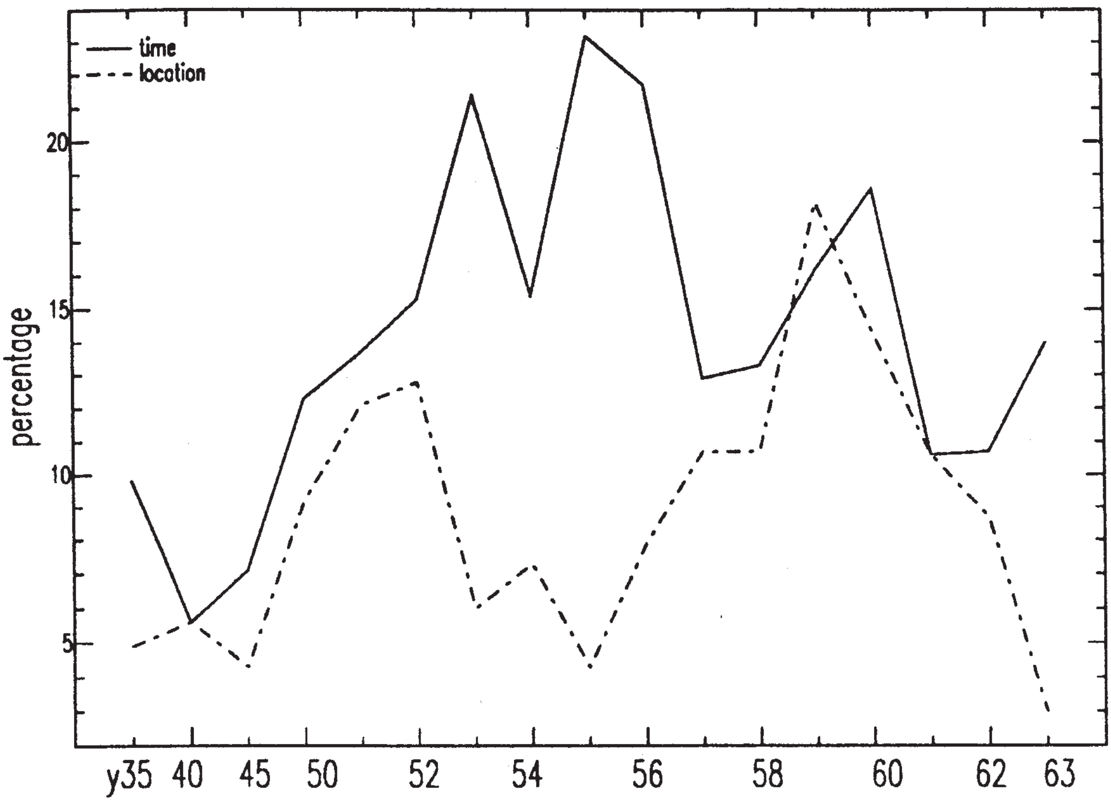

Fig. 9 - ISC Historical File [restricted to the Mediterranean; see Section 3.l. "Considerations …", in fine]. Multiple determination events. Percentage of pairwise origintime differences (solid line), and pairwise mismatch distances (dashed-dotted line) equal to zero.

difference class to include 3/4 of the pairwise origin-time differences; it reflects a rather large dispersion in the origin-times of the associated determinations. It is worth noticing that the threshold class, as defined above, does not show any decreasing trend as time goes on and seismological practice improves. However, as differences in origin-time, measured in seconds, do not influence the map representation, this parameter has not been investigated much further.

The same approach has been applied to pairwise distances (in the [latitude, longitude] plane) between the associated determinations, also coined "mismatch distances" in the following. Distances have again been sorted by classes of an amplitude of 15 km, and Fig. 8b shows, as a function of time, which class must be considered if one wants to include 3/4 of the pairwise mismatch distances. Surprisingly, one must go up to the 60 km class in most cases. Although one can notice that the average mismatch distance class is higher before the 50ies and lower from 1955 on, there is no clear tendency to a steady decrease with time.

Actually, the pairwise origin-time differences and mismatch distances could be seriously biased by various factors; one of them is that different associated determinations sometimes display identical values for some of the parameters; one must conclude that these determinations could well not be independent from each other, making any statistical analysis hazardous; it actually

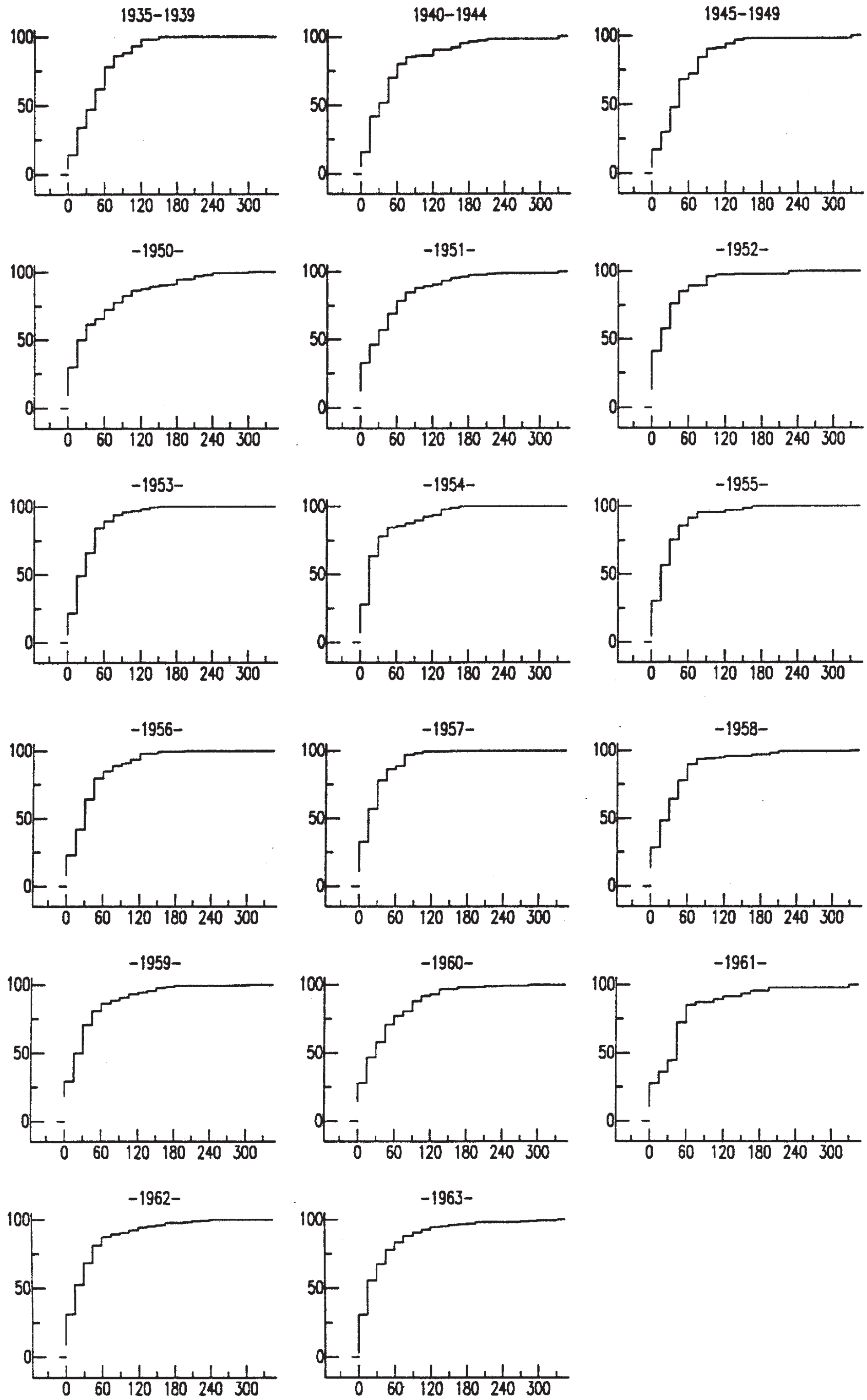

Fig. 10 - ISC Historical File [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Multiple determination events. Cumulative histogrammes of pairwise mismatch distances per annum.

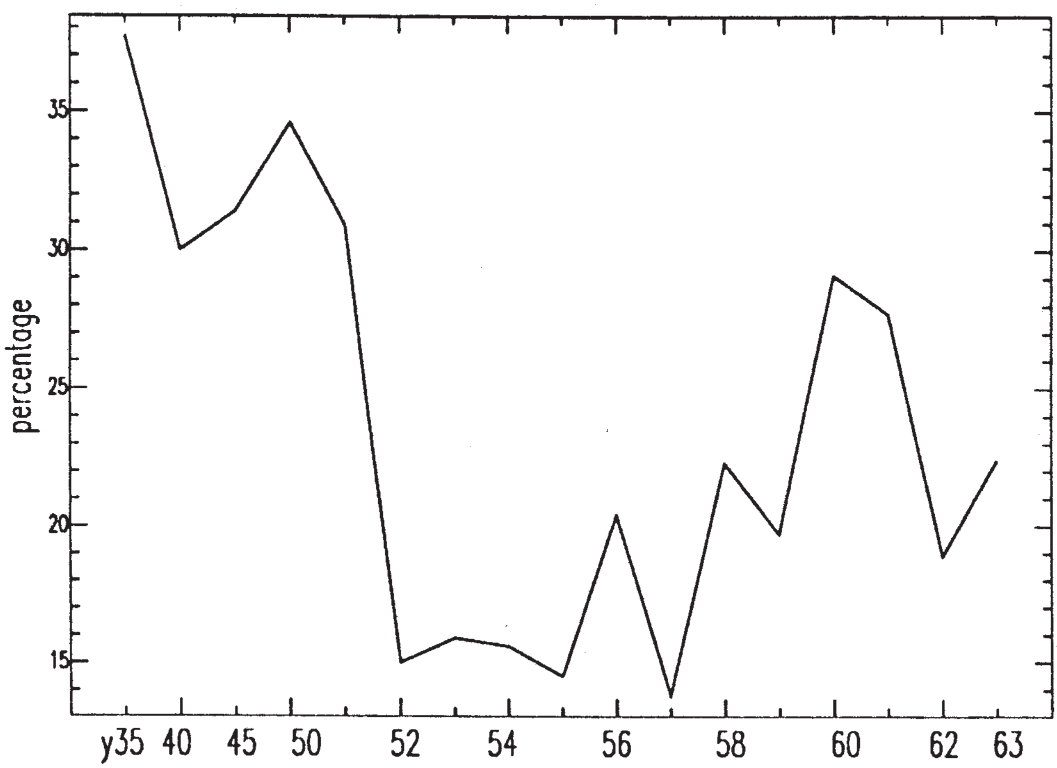

Fig. 11 - ISC Historical File [restricted to the Mediterranean; see Section 3.l. "Considerations …", in fine]. Multiple determination events. Percentage of large (> 60 km) pairwise mismatch distances per annum.

turns out that some parameters (the origin-time and/or the latitude-longitude location, for what concerns us here) happen to have been simply duplicated from one determination to another one, obviously resulting in a pairwise origin-time difference and/or a mismatch distance equal to zero; however, we had no way of deciding whether a zero pairwise origin-time difference and/or mismatch distance resulted from "chance" or from an actual duplication of parameters. In order to figure out how far data could be biased by this practice, Fig. 9 illustrates the percentage of pairwise time differences/mismatch distances equal to zero in the data sample under examination: it varies between 5 % and 25 %, with similar values for origin-time and location differences, except for 1953-1956, and 1963; for the latter periods, the percentage of zero difference in origin-time is rather large, while the opposite prevails for the percentage of zero difference in location; one possible interpretation is that the origin-time of one primary determination for each event has been steadily accepted as a reference, and the subsequent determinations have simply yielded locations which differ from the primary one. Of course, the relatively large percentage of zero difference in origin-time is reflected in the low of the curve of Fig. 8a, and explains part of it.

More details are shown in Fig. 10, which displays, for each year or group of years, the cumulative histogrammes of the mismatch distances: the *y*-coordinate of each step is proportional to the percentage of pairwise mismatch distances lower than the upper limit of the class considered

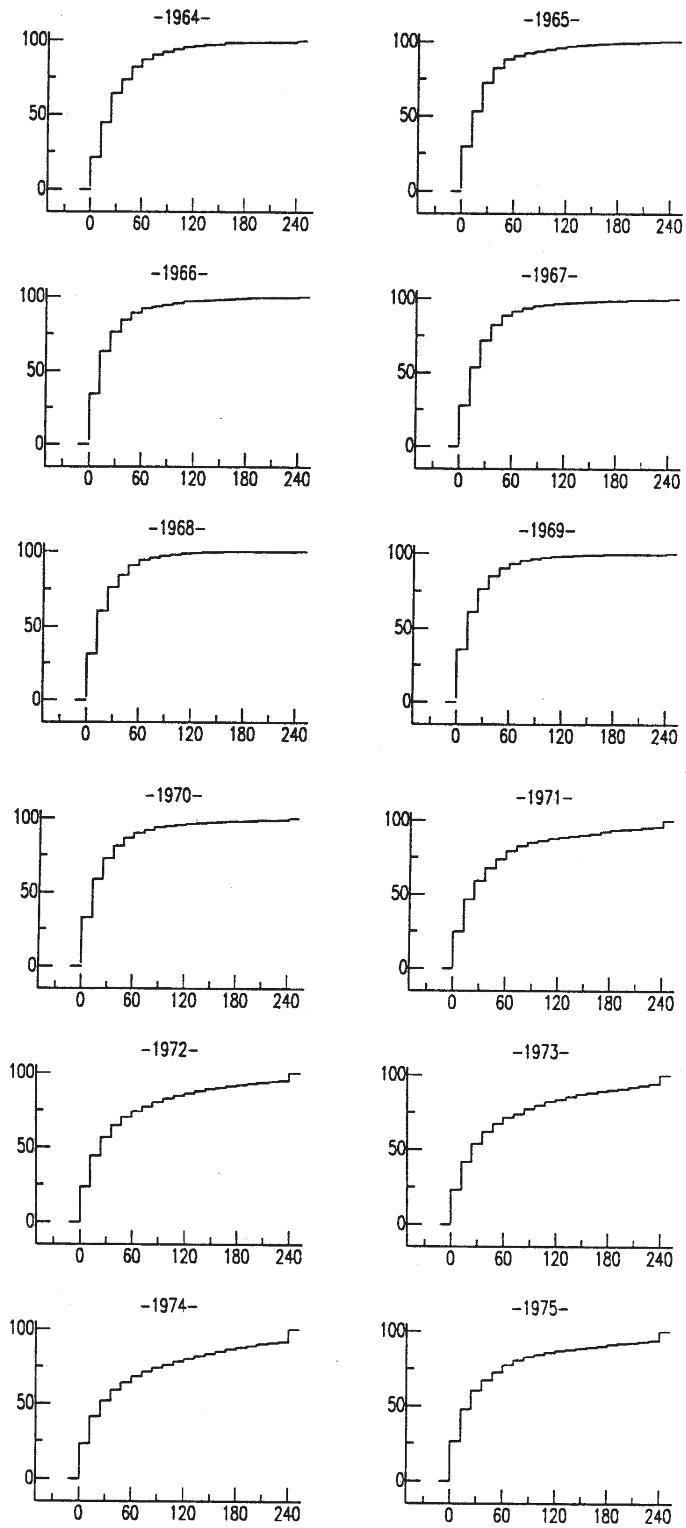

Fig. 12 - ISC Catalogue [restricted to the Mediterranean; see Section 3.1. “Considerations …”, in fine]. Multiple determination events. Cumulative histogrammes of pairwise mismatch distances per annum.

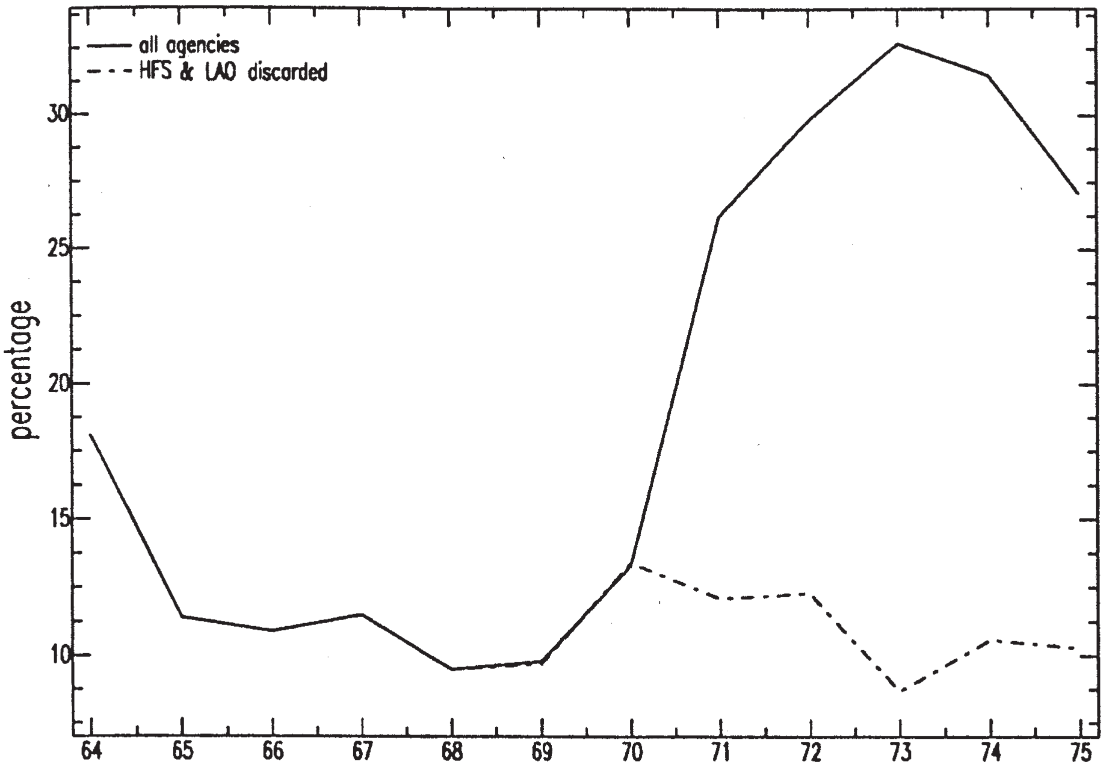

Fig. 13 - ISC Catalogue [restricted to the Mediterranean; see Section 3.l. "Considerations …", in fine]. Percentage of large (> 60 km) pairwise mismatch distances, considering all contributing agences (solid line), and discarding HFS and LAO contributions (dashed-dotted line).

(one class spans 15 km, except the last one which includes mismatch distances of more than 330 km; this means, by the way, that the largest mismatch distances cannot be read from Fig. 10, if, by chance, the maximum [100 %] value is reached only for the comprehensive last class). The often significant jump at 0 km accounts for the zero differences in location (see above). The curves closely approximate the graph of the probability density function obeyed by the distribution of mismatch distances.

The shape of the distribution graphs changes significantly from one year to another, and the larger the number of events involved, the smoother the curves; however, this does not account for all the shape's variations. As shown in Table 6, 1955 and 1961 both involve a small number of multiple determinations: yet, the curve's shapes for 1955 is smoother than the one for 1961 (and is close to the 1959 curve).

Nevertheless, one observes a rather high percentage of surprisingly large mismatch distances; we have put the lower limit for "large mismatch distances" at 60 km, the class level which includes approximately 3/4 of the pairwise mismatch distances and, conveniently, is about half a degree in geocentric angular distance. Fig. 11 displays the percentage of large mismatch distances as a function of the year. Although high values of this parameter were to be expected for the

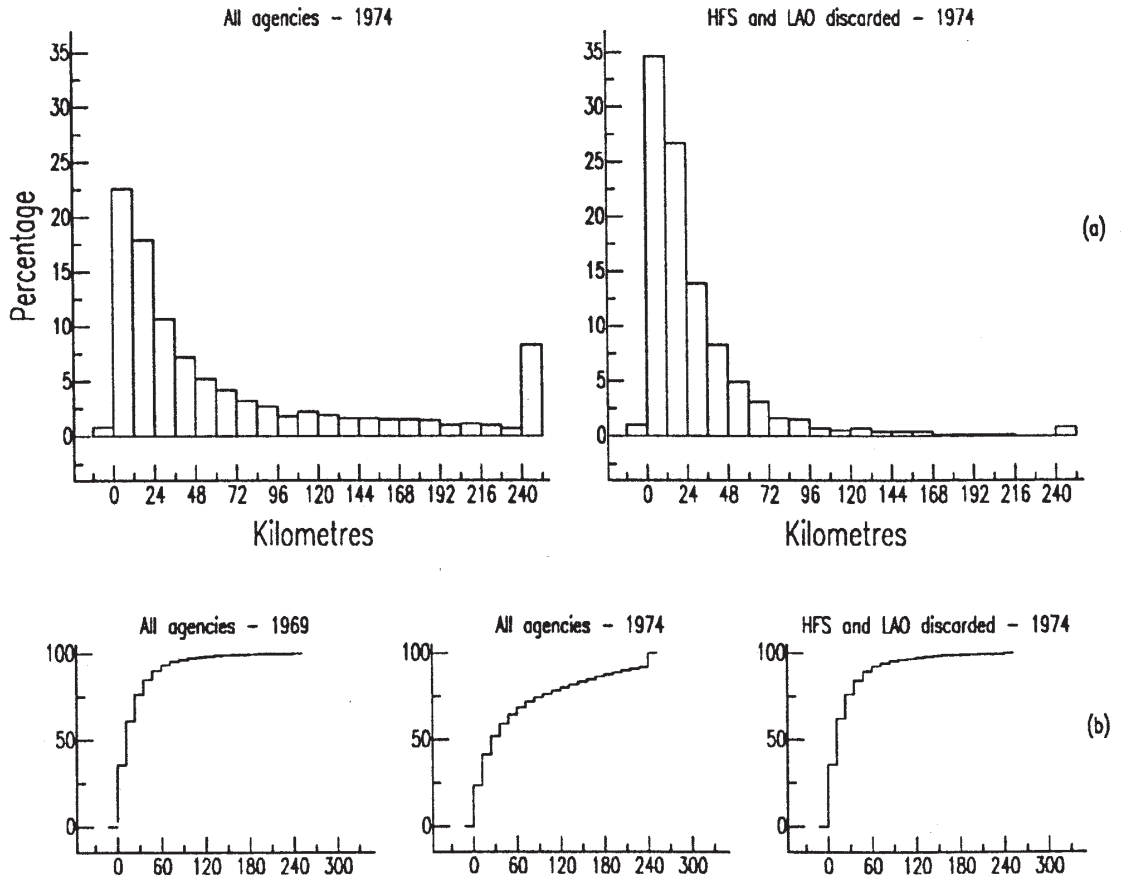

Fig. 14 - ISC Catalogue [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Multiple determination events. (a) Histogrammes of pairwise mismatch distances for 1974, considering all contributing agencies (left), and discarding HFS and LAO contributions (right). (b) Cumulative histogrammes of pairwise mismatch distances for 1974, considering all contributing agencies (middle), and discarding HFS and LAO contributions (right), as compared to the cumulative histogramme for 1969 (left).

first half of the century, because of the common practice of rounding off to the nearest full degree in latitude/longitude, it is still noticeable that it hardly goes down lower than 15 %, and even reaches almost 30 % by 1960. It is remarkable that the low in the curve of Fig. 11 corresponds roughly to the years when the lowest percentage in zero pairwise differences in location is observed, even though the reasons for this coincidence are not understood.

A similar analysis has been conducted on the ISC-Catalogue's data, with a slight change in the definition of the classes in location mismatch distances: because the samples were more numerous, it was possible to define classes spanning 12 km (instead of 15 km for the ISC-HF data); the last class includes mismatch distances larger than 240 km. Fig. 12 shows cumulative histogrammes for each year. The main features are:

- the jump corresponding to 0 km mismatch distance is small when it exists at all: very few determination pairs show up in the same latitude/longitude;

- the slope of the curves, for the first classes, are steeper and steeper from 1964 to 1969;
- the opposite holds for the years 1970 to 1974: it is flatter and flatter; the 1975 histogramme is close in shape to the 1972 one;
- the large mismatch distance class is well populated from 1971 on.

The last feature is very clearly illustrated in Fig. 13: obviously, something happens around 1970 which drastically disturbs the scenery. Actually, the large mismatch distances are mainly the result of locations issued by two agencies (LAO and HFS); if the processing is re-run ignoring the determinations computed by those two agencies, then the percentage of large mismatch distances in location does not display any solution of continuity (Fig. 13), and the corresponding cumulative histogrammes now resemble the ones prior to 1971 (Fig. 14). Once LAO and HFS's determinations are excluded, the average mismatch distances for the ISC-Catalogue's data are much lower than those for the ISC-HF's data: 3/4 of the mismatch distances for the former are less than 36 km.

*4.2. Focal depth**

While any event in the parameter lists considered, in the present study, is depicted at least by, sometimes rough, estimates of origin-time and latitude/longitude location, other parameter estimates are not always available. For our purpose, information on focal depth and strength of each event would be desirable, but is not readily accessible in the least, if at all it exists.

Depth is seldom reported in the determinations making up the ISC-HF. As shown in Fig. 15, it is most frequent that less than 1/10 of the determinations report an estimation of the focal depth; the percentage increased significantly during the last four years (1960-1963). The same holds for the events for which the focal depth is known (parameters of a physical event, indeed, can be computed by different agencies occasionally [“multiple determination events”]: each determination does or does not include a value for focal depth); the two curves are very much similar in Fig. 15. There is no apparent correlation between the small divergences of these two curves and the percentage of singular determinations (see Fig. 3) (nor is there one with its counterpart, the percentage of multiple determinations); it reflects the complex influence that singular determinations have on the global sample, in terms of percents. In Fig. 16, the percentile contributions of the different categories of determinations have been regrouped: one notes that, at both ends of the period under consideration, the singular determination contribution smooths down the contributions, in percentage, of multiple determinations and singularized multiple determinations; this still holds in between, from 1950 to 1959, but it is less apparent because the curves tend to come closer to each other; in addition, the few percentages of singular determinations with depth estimates in 1950-1952, when the global proportion of singular determinations drops (see Fig. 3), have no marked influence because they interact with a very small number of multiple determinations (singularized or not) associated to depth estimates. As will be

* The discussion in this section is strongly related to the way the final data set has been built up; see Section 5.2. "Focal depths in the final data set" for more details.

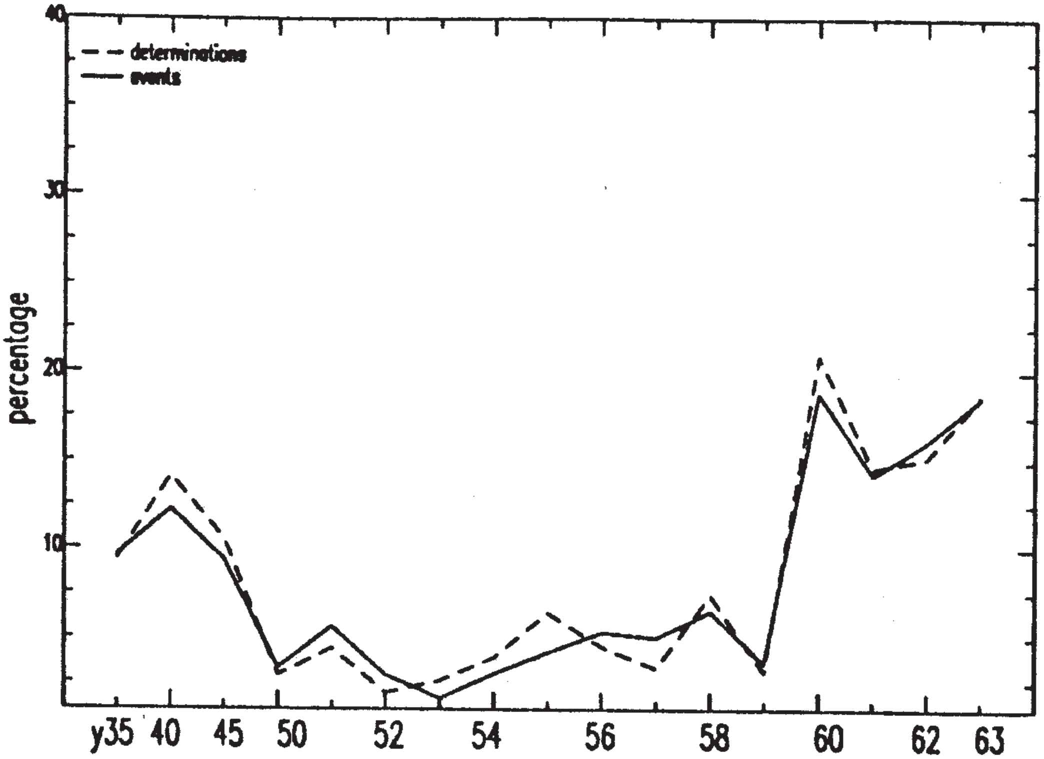

Fig. 15 - ISC Historical File [restricted to the Mediterranean; see Section 3. l. “Considerations …”, in fine]. Percentage of determinations (dashed line), and events (solid line) showing at least one depth estimate.

shown later (see Section 5.2. “Focal depths in the final data set”), it still helps to assign a depth to some of the events included in the final data set. Grouping multiple determination events into three periods for the total time-span of the ISC-HF, one observes that a sizeable proportion of those events show at least one depth estimation among the different determinations (see Table 9), except in 1952-1957 where the proportion drops to about 7 %.

Table 9 - ISC Historical File [restricted to the Mediterranean; see Section 3.l. “Considerations …”, in fine]. Multiple determination events and depth estimates per group of years.

	1904 - 1951	1952 - 1957	1958 - 1963
Number of multiple determination events	252	309	368
with at least one depth estimate	54 21.4 %	21 6.8 %	192 52.2 %
with depth estimate taken from a determination different from the determination selected according to the epicentral location	7 2.8 %	9 3.0 %	25 6.8 %

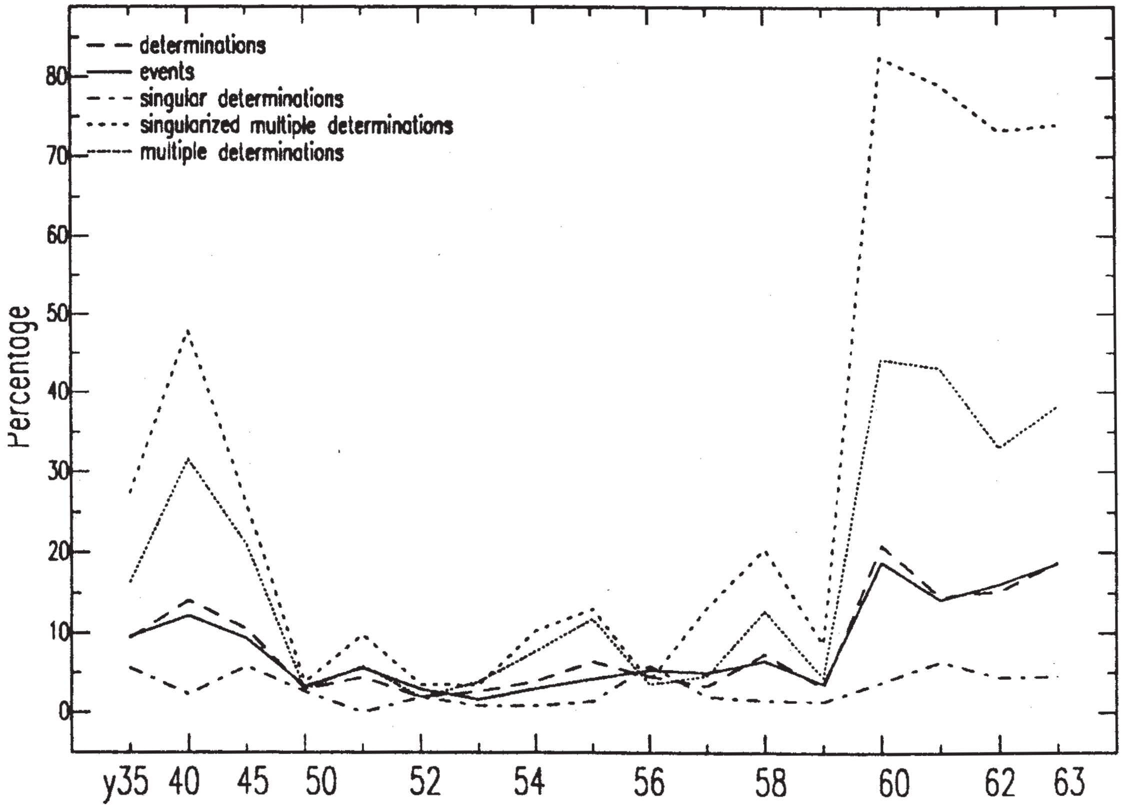

Fig. 16 - ISC Historical File [restricted to the Mediterranean; see Section 3.1. "Considerations…", in fine]. Percentages of determinations and events showing depth estimates. Solid line: all events (with at least one depth estimate; same as in Fig. 15); long-dashed line: all determinations (same as in Fig. 15); long-dashed/dotted line: singular determinations; short-dashed line: singularized multiple determinations; dotted line: multiple determinations.

In the ISC-C, most of the determinations appear, at first glance, to include an estimation of depth, at least for the period 1964-1970; in other words, the depth data field is filled with a value. A more careful examination shows, however, that the depth value is often put to zero, and we have no evidence to decide whether the zero value has been assigned to depth arbitrarily, because of some provision embedded in the algorithm (apart from the very few likely cases where the computations occasionally yielded a zero result), or it simply means that the depth value is missing. We have assumed that the latter was true, considering the large number of such occurrences; Fig. 17 displays, as a function of time, the percentage of determinations which include a non-zero estimation of depth: the decrease from more than half at the beginning of the period, down to less than 1/3 is striking; one must conclude that analysts have been more and more reluctant to let the algorithms yield a depth estimate, although an eventual bias could come from determinations of many weak events computed from a limited number of station data (insufficient to calculate depth with some confidence); surprisingly, the depth value classically imposed by the analyst or the algorithm when depth cannot be validly computed, does not seem to have been widely used in such cases.

All the determinations reported in the EMSC monthly lists include an estimation of focal

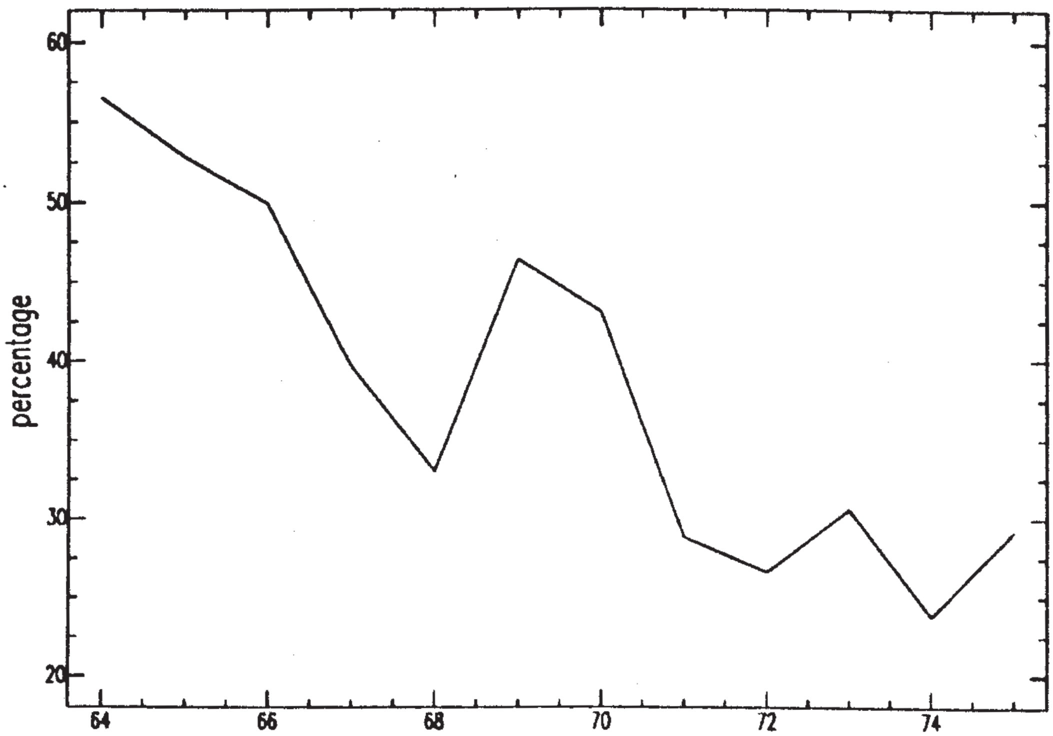

Fig. 17 - ISC Catalogue [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Multiple determination events. Percentage of determinations including a non-zero depth estimate.

depth, which is sometimes a value arbitrarily fixed by the analyst or by the provisions of the algorithm.

Considerations on depth estimates will be further developed below, once the data set finally adopted is described in detail.

4.3. Magnitude *

Paradoxically, from 1935 on, i.e. the year when Richter (1935) has first developed the concept of magnitude, the corresponding data field starts to be occasionally empty, while it is filled in the ISC-HF before that date. This is actually because the dominant data source prior to 1935 is precisely Gutenberg and Richter (1949): the authors re-worked their parameters and computed magnitude estimates according to the procedure designed, making use of the amplitude data they could collect. After 1935, magnitude estimates included in ISC-HF come from the various sources exploited; they are sometimes missing.

From 1904 to 1927, the 56 singular determinations excerpted from Gutenberg and Richter

* The discussion in this section is strongly related to the way the final data set has been built up; see Section 5.3. "Magnitudes in the final data set" for more details.

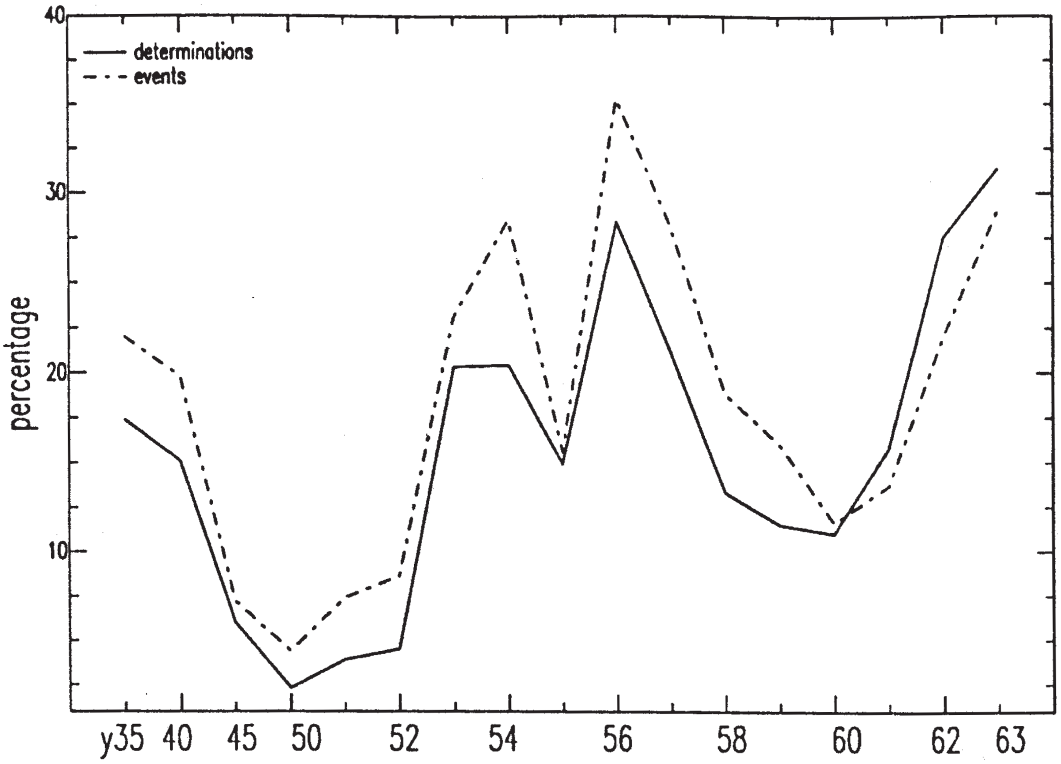

Fig. 18 - ISC Historical File [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Percentage of determinations (solid line) and events (dashed-dotted line) showing at least one magnitude estimate.

(1949), all have a magnitude estimate. From 1928 to 1935, only one event out of 68 is missing any magnitude estimate. From 1935 to 1963, the proportion of determinations or events which are assigned a magnitude estimate is displayed in Fig. 18: the two curves are close to each other; the latter (events with one magnitude estimate or more) is slightly above the former, and results from a complex combination of the no-magnitude class columns of Fig. 31a and b; multiple determination events are thus more often assigned at least one magnitude estimate rather than for the singular determination events.

No information is given in ISC-HF on the type of magnitude reported, except for the estimates computed by Gutenberg and Richter (1949) which are identified by a symbol (an "S", indicating an M_S magnitude).

The proportion of determinations reported with a magnitude estimate in the ISC-Catalogue is illustrated in Fig. 19a, where three curves are shown for all the determinations, the singular determinations, and the multiple determinations, respectively; the first one goes up (to 50 %) and down (to 40 %) from 1964 to 1971, and then shows a decreasing trend; the second displays large saw-teeth, ranging from 10 % to 40 %; while the third is much smoother, but shows a marked decreasing trend from 1970 on, which reflects, although attenuated, in the first (solid) curve. The cause for the decrease in the percentage of determinations of multiple events with magnitude estimates, lies in the output of ISC: starting in 1970, ISC contributes a large number of determi-

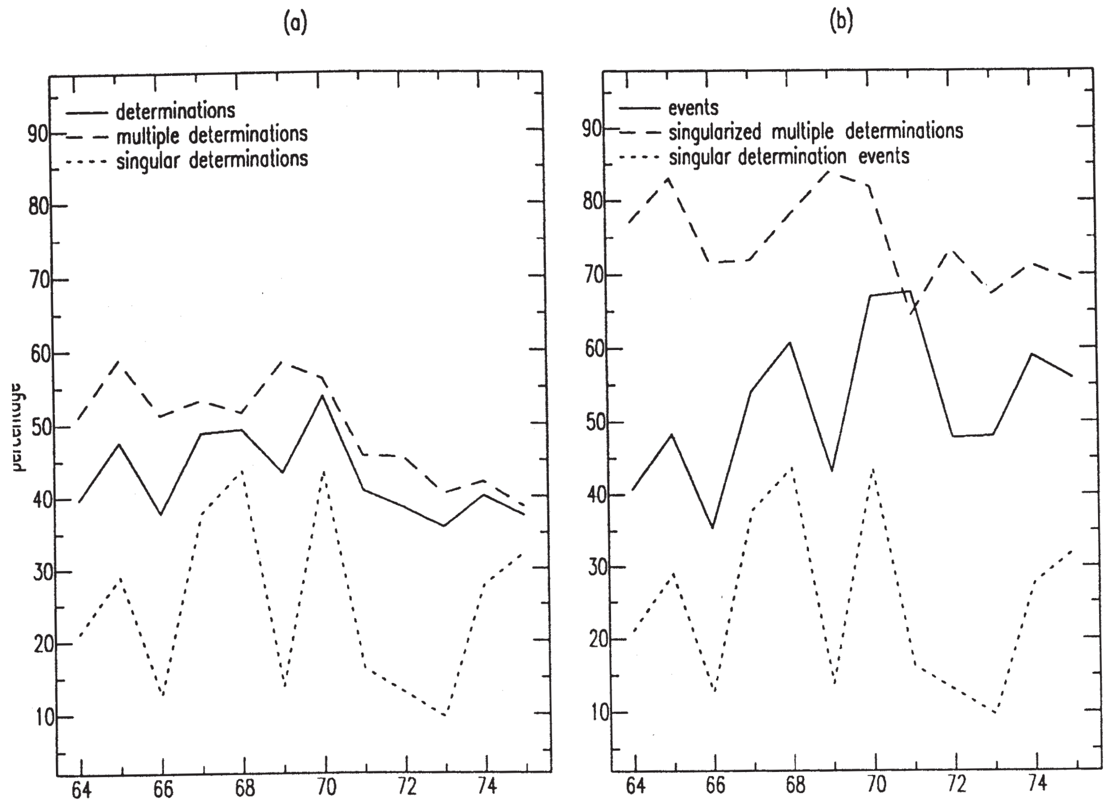

Fig. 19 - ISC Catalogue [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. (a) Percentage of all determinations (solid line), singular determinations (dotted line), and multiple determinations (dashed line), showing a magnitude estimate (the dotted line is identical to the one of Fig. 19b). (b) Percentage of all events (solid line), singular determination events (dotted line), and singularized multiple determinations (dashed line), showing at least a magnitude estimate (the dotted line is identical to the one of Fig. 19a).

nations of multiple determination events, but provides, in proportion, less and less magnitude estimates. On the other hand, the large variations of the (dotted) curve corresponding to singular determinations are more complex in origin: they come from a combination of changes of the dominant source of determinations, and the specific behaviour of each source vis-à-vis magnitude estimates; for example, SPGM is the most important contributor in 1969 (see Table 6), but does not supply magnitude estimates, causing the low observed on the dotted curve of Fig. 19a; in 1971, another agency, ISK (which does not even show up in Table 6, because its contribution is limited to that year), contributes almost half of the singular determinations, but does not supply magnitude estimates: again, there is a low in the corresponding curve; ATH plays an important role in 1966-1968 (see Table 6), but supplies a small number of magnitude estimates in 1966, this number increases in 1967 and becomes even larger afterwards (but ATH looses its importance in 1969): once again, this is reflected in the dotted curve.

Curves have also been drawn for the events with at least one magnitude estimate assigned (Fig. 19b; the dotted curve is identical to the one of Fig. 19a, naturally). The dashed curve shows that 70 %, or more, of the multiple determination events are assigned at least one magnitude esti-

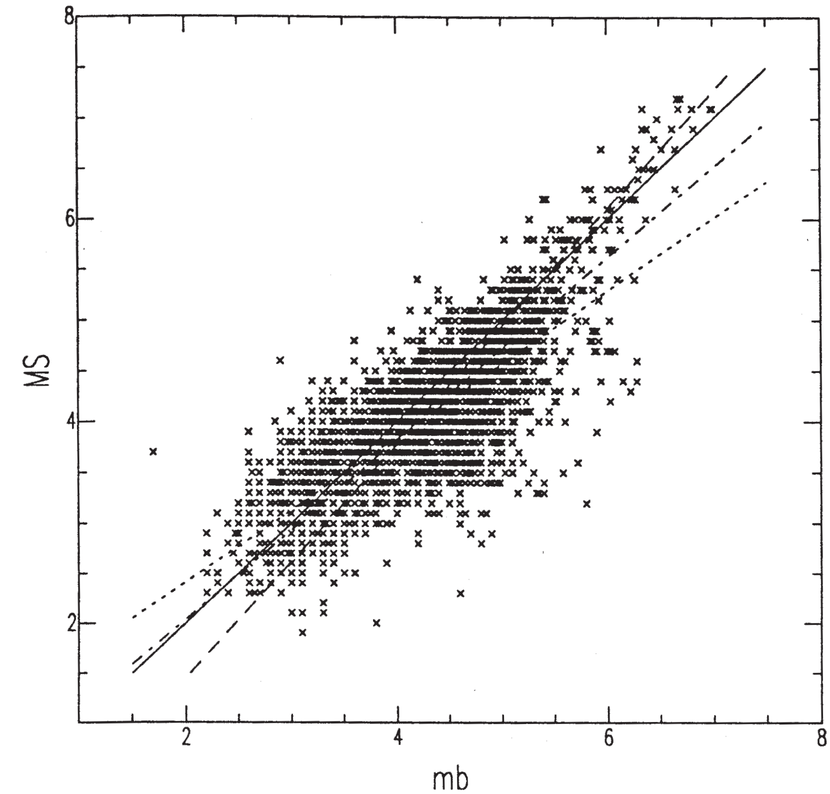

Fig. 20 - EMSC File [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. M_S versus m_b plot for all the events of the EMSC file showing both magnitude type estimates. Dotted line: M_S to m_b regression line; dashed line: m_b to M_S regression line; dashed-dotted line: orthogonal regression line M_S - m_b; solid line: bissector.

mate, although the percentage decreases toward the end of the period considered. However, as the relative importance of singular determination events tends to decrease (see Table 8), the solid curve, which represents all the events with at least one magnitude estimate, remains close to 50%, and even shows a tendency to increase. The marked difference between the dotted curve (singular determination events) and the dashed one (singularized multiple determinations) is also apparent when comparing the first columns (no-magnitude class) of Fig. 32a and b.

Although more often than for the ISC-HF, the type of magnitude estimated, when there is one, is seldom reported. Very few determinations (about 1 %) show up two types of estimated magnitudes (body-wave [m_b] and surface-wave [M_S] magnitudes); as, also, the magnitude range, for which estimates of both types are known, makes it impossible to convert estimates from a given type to another, when only one type is available. In other words, the magnitude estimates available have been taken as they are, with no attempt at converting them into another type.

In the EMSC's monthly lists, one magnitude estimate, and often two or more, is given for almost any event reported; in the files used in the present study, all the reported magnitudes belonging to one category (body-wave magnitudes) or the other (surface-wave magnitudes), have been combined to produce one averaged estimate for each category. About 4000 pairs have then

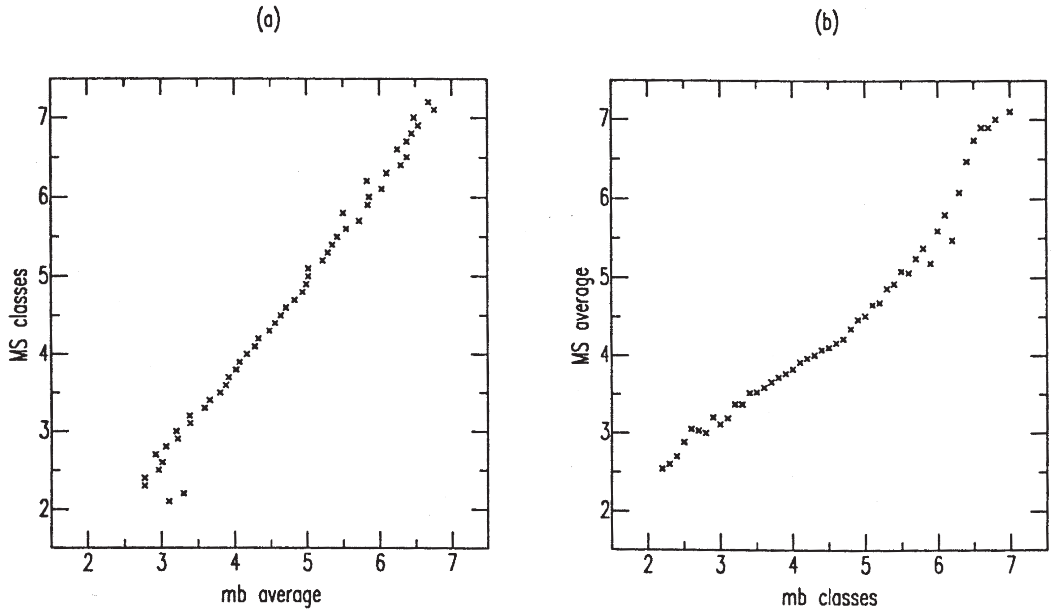

Fig. 21 - EMSC File [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. (a) m_b averages versus M_S classes (a class spans 1/10 of a magnitude unit). (b) M_S averages versus m_b classes (a class spans 1/10 of a magnitude unit).

been processed in the hope of deriving a formula for conversion of one magnitude type into the other. A plot of M_S *versus* m_b is given in Fig. 20, where, by the way, the rounding-off effects (to the nearest tenth of magnitude unit for M_S) are clearly apparent ("orchard" effect). Linear regression yields the three straight lines: the dotted one refers to the regression M_S to m_b; the dashed line to the regression m_b to M_S; and the dotted-dashed line to orthogonal regression; the three are close to the bissector (solid line). A better insight is gained by looking at the average magnitude of one type, computed from the values of that type for each pair pertaining to a certain class of the other type of magnitude; results are displayed in Fig. 21a and b: the reference classes span 1/10 of the magnitude unit of either type. If the m_b averages, as a function of M_S classes (Fig. 2la), fall close to a straight line similar to the dashed regression line of Fig. 20, the M_S averages, as a function of m_b classes, deviate significantly from any straight line: the dotted line of Fig. 20 should be discarded, and this is because, as it is well known, the m_b / M_S relationship is not uniform over the whole range of magnitude. Although the two approaches, m_b to M_S and M_S to m_b, are not equivalent, one can notice that the m_b averages versus M_S classes curve shows a larger dispersion in the same range (≥ 5) where the M_S averages versus m_b classes curve deviates from a straight trend. Of course, the orthogonal regression does not privilege anyone, but yields a roughly one-to-one relationship which simply conceals the difficulties.

Actually, this attempt at analyzing things in detail is based on a data sample (pairs m_b / M_S) the intrinsic quality of which is highly debatable: the values are already averages computed from statistically non-homogeneous data. Trying to go further in the analysis would have been hazardous if not fallacious. On the other hand, our final goal (i.e. to plot symbols on a map) does not necessarily require an extreme precision on magnitude: one simply wants to give an impression of where the seismically active zones are and how active they are, compared to each other, at the scale adopted (1/1 000 000). So, we have finally given up trying to convert all magnitude estimates into one unique type, and we have taken each estimate as it is when it is the only one available; we have made a simple average of the two estimates when they are both available. More details on magnitudes distribution will be found in Section 5.3. "Magnitudes in the final data set".

5. Building up the final data set

5.1. Epicentral locations in the final data set

As it has just been demonstrated, determinations related to one physical event, when there are several, could yield large differences in the epicentral locations; yet, the physical event itself is unique (as long as the assumption of a point source is accepted), and should show up in any final data set intended for application. Thus, one cannot escape from choosing among the different locations available; in other words, for multiple determination events, one of the determinations should be "singularized" and will conventionally depict the corresponding physical event; but still, nobody knows where the actual epicenter is located. As described before, extreme situations occur where it is relatively safe to tell that certain locations issued by an agency (or a group of agencies) are likely to be wrong, although it is difficult to definitely rule out the eventuality that all other locations are even more faulty. In the case of a manifold multiple determination event, one agency can be the originator of an epicentral location which causes very large mismatch distances between the different locations issued: this location will then be excluded from the set of likely locations. Assuming that such a situation is repeated for the given agency over a sizeable period of time, this agency will be regarded as unreliable for that period of time (and for the domain under consideration, of course). We have already met such a situation with HFS and LAO: it has been shown that they were the cause of large pairwise mismatch distances in the ISC-C, from 1970 onward and for the region of concern.

We would now like to reverse the reasoning and try to identify, if possible, the "best reporting" agencies for a certain time period. If such agencies exist, they are, by definition, likely to issue epicentral locations which fall in the closest vicinity of the actual epicenter, and to induce pairwise mismatch distances smaller than the other agencies do. Suppose that a collection of points M_n is uniformly distributed over a 2-D space; M_{n_m} is any point within a circle of radius r_i; M_{n_p} and M_{n_q} are taken within concentric circles of radii r_p and r_q, respectively, having the same centre; then the first moment (so called "mathematical expectation") of the pairwise distance distribution verifies the expression:

$$E\left\{d\left(M_{n_m}, M_{n_p}\right)\right\} < E\left\{d\left(M_{n_m}, M_{n_q}\right)\right\}, \text{ if } \quad r_p < r_q$$

where $d\ (P,Q)$ represents the distance between points P and Q.

This expression is of no help if applied tentatively to the set of determinations related to one physical event, as the number of determinations involved is never large enough for a sound statistical approach. What is important here is the distance between pairs of points, not their respective absolute location in the 2-D space. Let us then assume that, for a certain time period, all the physical epicenters are brought artificially together to the same point of the Earth's surface, and that the condition of uniform distribution is met by the set of epicentral locations from the different agencies, and for the whole set of events belonging to the period under consideration. If a pair of agencies can be identified which yields locations, statistically, closer to each other than the other pairs of agencies, these locations are likely to lie within the statistically smallest possible circle centered at the point representing the physical grouped epicenters; in other words, the locations issued by this pair of agencies are those which statistically best approximate globally the actual epicenters assumed to be all located at the same point. Naturally, the strongest limitation to this approach is the condition of uniform distribution, which has, a priori, no obvious reason to be met; moreover, the number of determinations is not necessarily large enough for statistical methods to apply validly: this is what the grouping of all physical epicenters in one very point tends to obviate; the determinations should be, in addition, independent from each other: this is certainly not the case, as, among other reasons, the different agencies make use, in their computations, of sub-sets of the set of all available station readings, and the sub-sets are not likely to be disjoint.

Once the pair of agencies yielding the epicentral locations statistically closest to the actual regrouped epicenters is identified, the second step is to decide which one of the pair is the "best" for the period under consideration. One then compares the mean of the distances between the epicentral location issued by each of the two agencies and all the other locations available. In other words, with the pair of agencies *a* and *b*, and A^j_l being the epicentral determination of event *j* issued by agency *l*, the two quantities

$$\frac{1}{\sum_{k=1}^{p_a} n_{a_k}} \sum_{j=1}^{p_a} \sum_{i_j=1}^{n_{a_j}} d\left(A_a^j, A_{i_j}^j\right)$$

and

$$\frac{1}{\sum_{k=1}^{p_b} n_{b_k}} \sum_{j=1}^{p_b} \sum_{i_j=1}^{n_{b_j}} d\left(A_b^j, A_{i_j}^j\right)$$

are compared, where p_a (resp. p_b) is the number of events for which agency *a* (resp. *b*) has issued a determination, and where n_a (resp. n_b) is the number of agencies (different from agencies *a* and *b*) which have issued a determination for event *j*.

Selecting the agency which yields the smallest value of the above expressions is equivalent to selecting the agency which issues epicentral locations, for each physical event, closest to the point which makes minimum the sum of mismatch distances to all the other epicentral locations available.

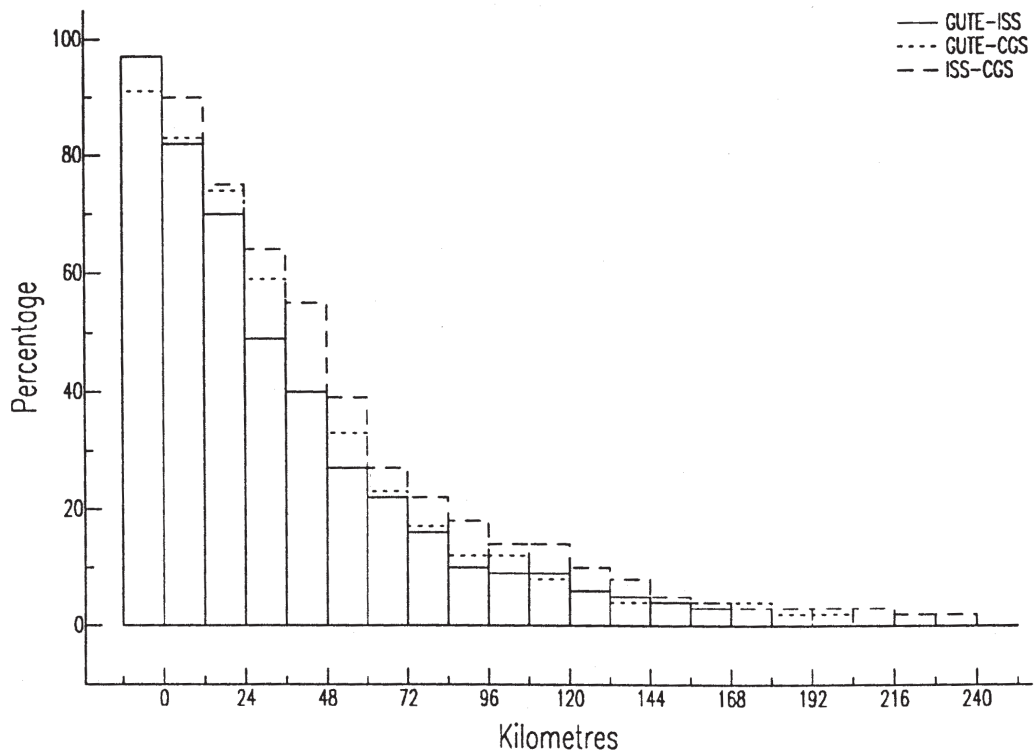

Fig. 22 - ISC Historical File, 1935-1949 [restricted to the Mediterranean; see Section 3.1. "Considerations …", *in fine*]. Histogrammes of pairwise mismatch distances between the epicentral locations issued by GUTE, ISS, and CGS. Solid line: GUTE-ISS pair; dotted line: GUTE-CGS pair; dashed line: ISS-CGS pair. For the meaning of the acronyms designating the agencies, see the Appendix 2.

Actually, this is only an approximation: to strictly apply the criterium derived from the properties of the mean distance between points uniformly distributed over concentric circles, one should have computed and compared, for each of the two "best reporting" agencies, the mean values of the means of the mismatch distances between different epicentral locations of the same physical event, while we compute and directly compare the means of all the mismatch distances induced by all the events at once. Doing it this way is much easier, and we have considered we did not violate more severely the conditions of validity of the overall method that it is done by assuming that the associated and grouped epicentral locations are uniformly distributed around the actual common-forced epicenter.

Other statistical approaches could have been thought of; in particular, one could have approximated the unknown actual epicenter by the centre of mass of the distribution of all the epicentral locations available. This would have led to assign, ipso facto, a heavy weight to the large mismatch distances which come from epicentral locations which are not statistically likely to lie close to the actual epicenter. The contradiction could be overcome by applying a weighting scheme to the mismatch distances, which in turn supposes that we have an a priori assessment of the quality of the corresponding determinations. To avoid this vicious circle, we have decided to apply the above-described method.

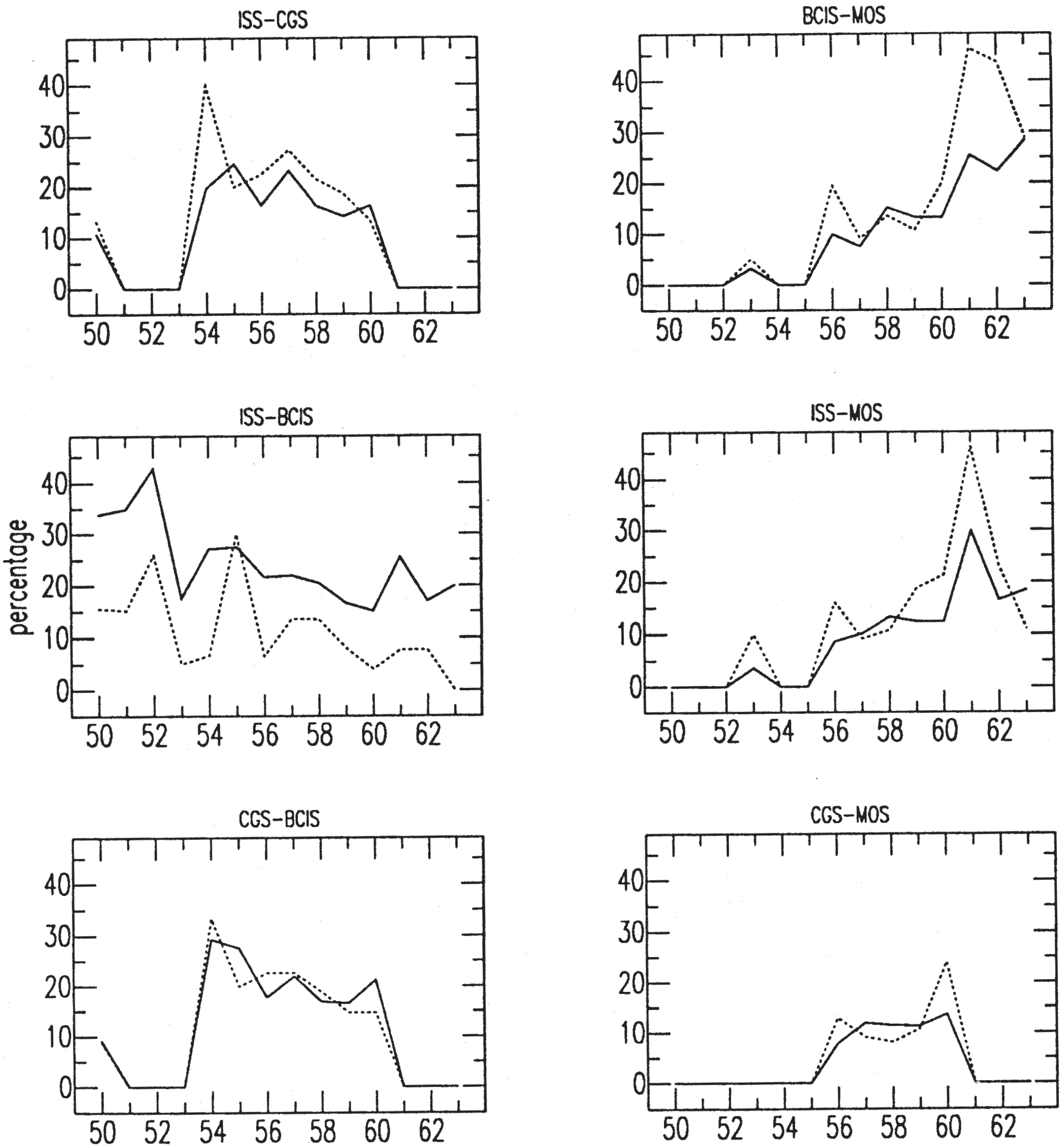

Fig. 23 - ISC Historical File, 1950-1963 [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Proportion of pairwise distances induced by each of six pairs of agencies among all pairwise mismatch distances (solid lines), and proportion of mismatch distances larger than 60 km contributed by the corresponding pair with respect to all mismatch distances larger than 60 km (dotted lines). See text for more explanation. For the meaning of the acronyms designating the agencies, see the Appendix 2.

Some care has been paid in using it. Naturally, the zero mismatch distances (see Section 4.1. "Multiple determinations") have been sorted out: they are essentially meaningless and could have seriously biased the already not-secured assumption of independence of the epicentral locations.

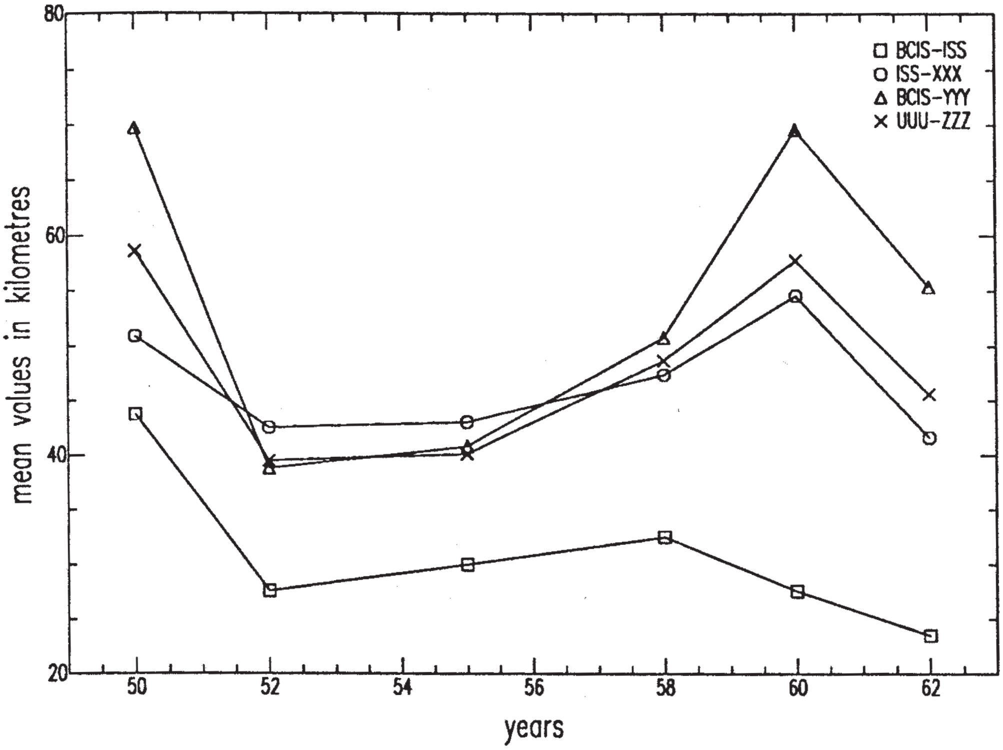

Fig. 24 - ISC Historical File, 1950-1963 [restricted to the Mediterranean; see Section 3.1. "Considerations ", in fine]. Mean mismatch distances for the pairs involving the two "best reporting" agencies for the period considered: BCIS and ISS. The mean mismatch distances of whatever pairs of agencies (UUU-ZZZ) are also displayed. The symbols correspond to the mean value over the time interval starting where the symbol is located on the plot. For the meaning of the acronyms designating the agencies, see the Appendix 2.

Very large mismatch distances have been eliminated as well, as they could correspond to determinations erroneously associated. The way the method has been implemented depends on the time period considered.

ISC-HF, 1904-1934. - Before 1935, there are only singular determinations (with eight exceptions): they have been included simply in the final data set.

ISC-HF, 1935-1949. - Three sources contributed: GUTE, ISS, and CGS. After identifying the pair of agencies yielding the statistically closest epicentral locations, the second step would consist in comparing each agency of the pair with the third one; actually, the mismatch distance histogrammes have been constructed for the three possible pairs (Fig. 22). The pair GUTE-ISS displays the most rapid decrease in class population, as the mismatch distances grows larger: it is well illustrated within the range 12-60 km. The pair GUTE-CGS remains below the third one (ISS-CGS), and tends to be very close to the first one for mismatch distances greater than 60 km (which, in any case do not amount to a very large proportion of the mismatch distance population). Averaging the mismatch distances (up to 480 km for the maximum, but not taking into account the zero mismatch distances), one comes up with:

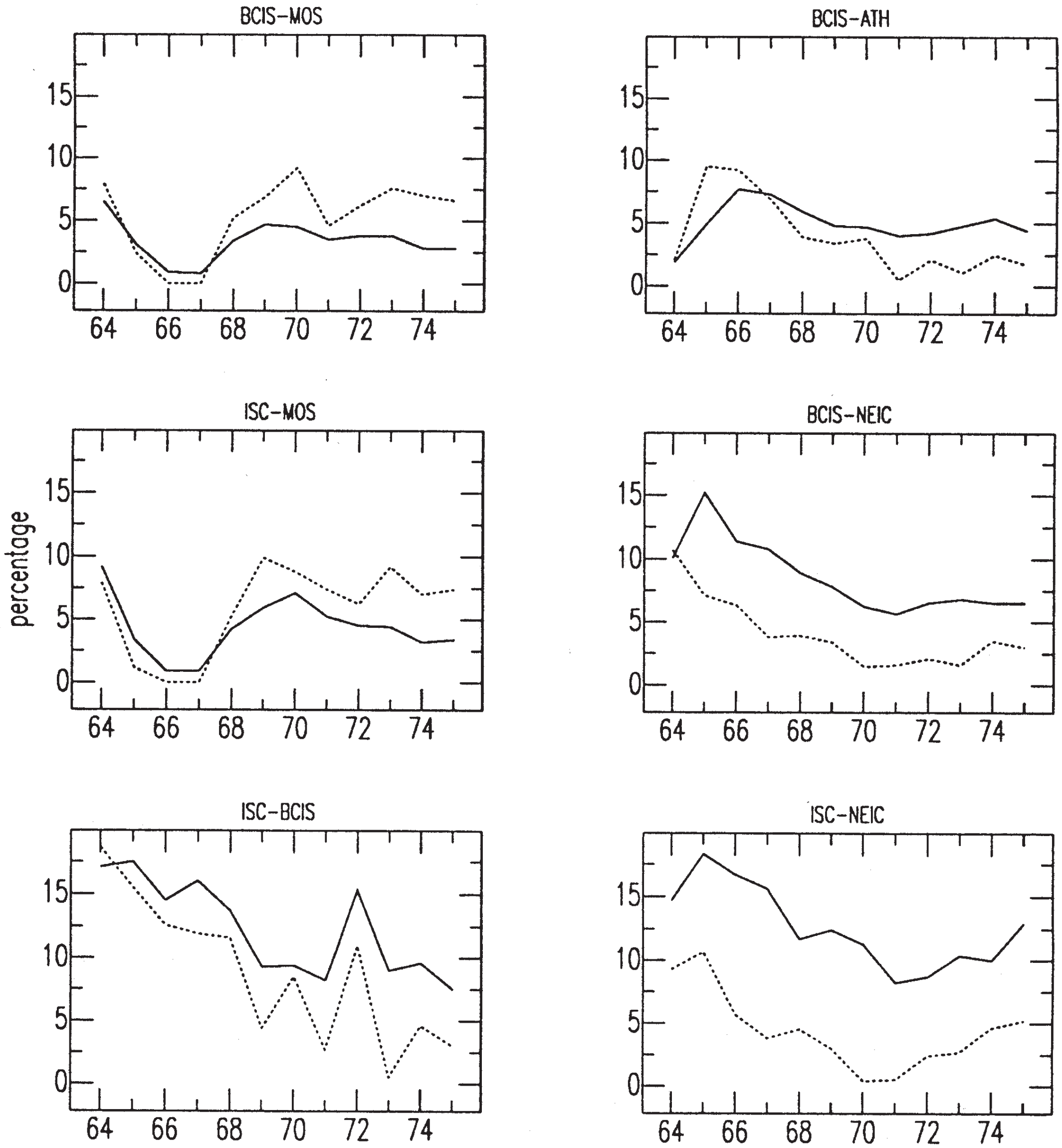

Fig. 25 - See caption on next page.

GUTE - ISS	50.3 km
GUTE - CGS	56.5 km
ISS - CGS	62.4 km

which is consistent with the histogrammes shown on Fig. 22.

According to our premises, GUTE determinations should be considered, as an average, as the "best" ones for the period 1935-1949. If it exists, a GUTE determination corresponding to a given physical event, will be included in the final data set; if the GUTE determination is missing, then the ISS one will be preferred as the latter agency appears to be statistically closest to the "best reporting" one. Actually, it turns out that there is no occurrence of a physical event with neither a GUTE

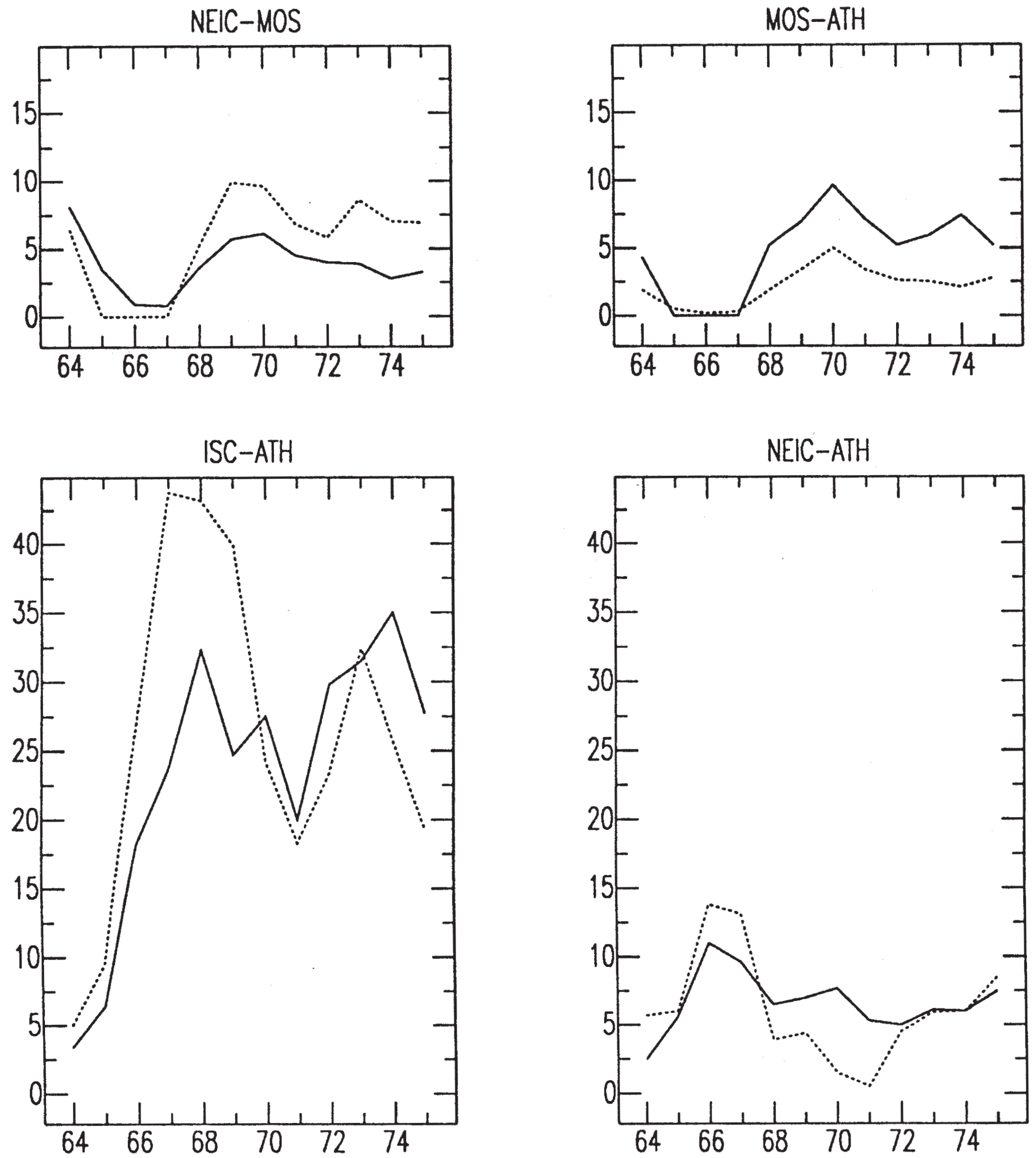

Fig. 25 - cont'd. ISC Catalogue, 1964-1975 [restricted to the Mediterranean, see Section 3.1. "Considerations …", i*n fine*]. Proportion of pairwise distances induced by each of ten pairs of agencies among all pairwise mismatch distances (solid lines), and proportion of mismatch distances larger than 60 km contributed by the corresponding pair with respect to all mismatch distances larger than 60 km (dotted lines). See text for more explanation. For the meaning of the acronyms designating the agencies, see the Appendix 2.

nor an ISS determination.

ISC-HF, 1950-1963. - During that period, there is enough data to apply the analysis scheme on a yearly basis. Table 5 shows that only BCIS, ISS, CGS, and MOS have significant contributions to multiple determination events over the whole period. Fig. 23 shows for each pair constructed on these four agencies, the proportion of mismatch distances induced by the pair considered (solid

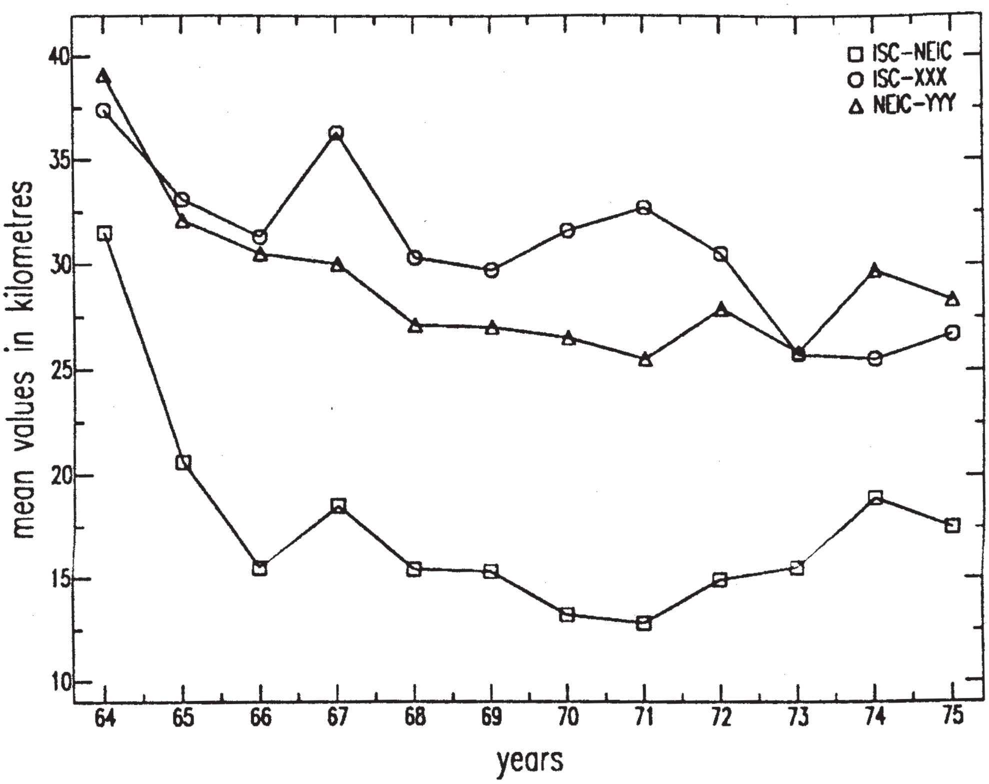

Fig. 26 - ISC Catalogue, 1964-1975 [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Mean mismatch distances for the pairs involving the two "best reporting" agencies for the period considered: ISC and NEIC. For the meaning of the acronyms designating the agencies, see the Appendix 2.

lines), and the percentage of mismatch distances larger than 60 km among the corresponding pair-induced mismatch distances greater than 60 km (dotted lines). Then, comparing the two curves pair by pair, one can observe the relative importance of large mismatch distances which the pair under consideration contributes (it would have been meaningless to compare the proportion of large mismatch distances induced by one pair against the total population of mismatch distances directly). The two "best reporting" agencies are obviously the ones which the pairs have made a significant contribution throughout the period and for which the dotted line is well below the solid line; clearly, the pair to be pointed out is the ISS-BCIS one (with an isolated accident in 1955).

The choice then between BCIS and ISS as the "best reporting" agency for the period 1950-1963 remains to be made. Mean mismatch distances between BCIS [resp. ISS] determinations and any other agency reporting a determination of the same event, denoted BCIS-YYY [resp. ISS-XXX], have been computed year by year; mismatch distances equal to zero or larger than 450 km have been ignored. Actually, trends observed allow us to draw Fig. 24, where mean mismatch distances for the pair BCIS-ISS, as well as those of whatever pair UUU-ZZZ (including BCIS and ISS) have been added. The mean distances between locations of the same events issued by BCIS and ISS are close to 30 km (except at the very beginning of the period considered), and show a decreasing tendency starting in 1958; they are systematically well below the means involving the pairs ISS-XXX and BCIS-YYY, which are grouped with the means involving the pairs UUU-ZZZ. At the beginning of

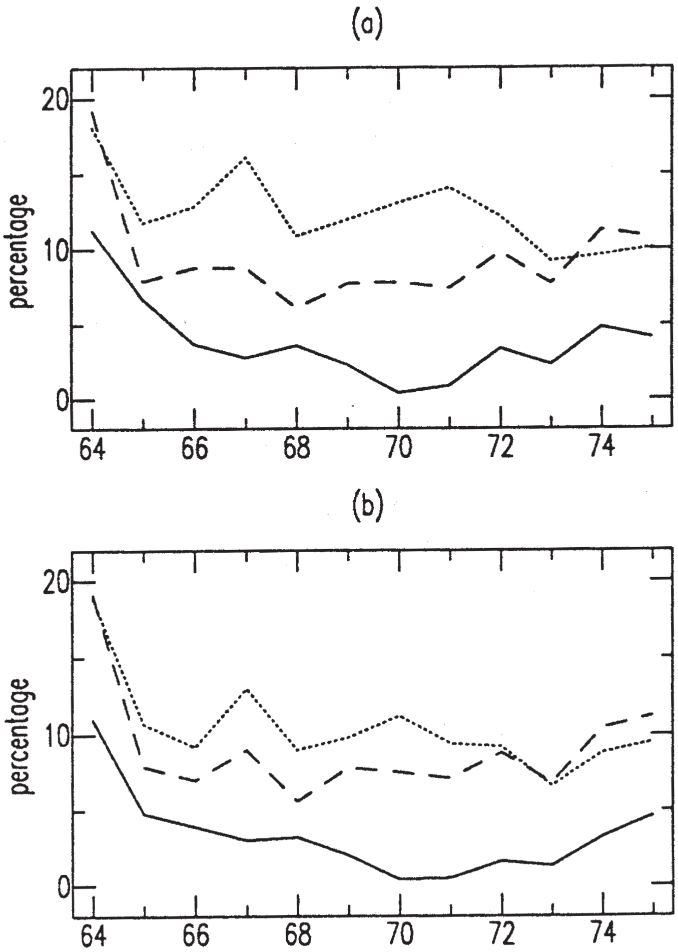

Fig. 27. - ISC Catalogue, 1964-1975 [restricted to the Mediterranean, see Section 3.1. "Considerations …", *in fine*]. Percentage of mismatch distances larger than 60 km within pairwise distances induced by the pairs involving the two "best reporting" agencies for the period considered (ISC and NEIC). Solid lines: pairs ISC-NEIC; dotted lines: pairs ISC-XXX; dashed lines: pairs NEIC-YYY. (a) no magnitude filter applied; (b) events with magnitude ≥ 4 only. For the meaning of the acronyms designating the agencies, see the Appendix 2.

the period, ISS-XXX shows a smaller mean, like the period from 1958 onwards; in between, BCIS-YYY appears just a little smaller.

Then, following our scheme we decided upon, the determinations to be included in the final data set as follows:

1950 - 1951	ISS
1952 - 1957	BCIS
1958 - 1963	ISS

If the first "best reporting" agency's determination is missing for an event, then the determination of the second "best reporting" agency is selected. If both are missing, then the principal determination of the event (as defined by ISC-HF) is selected, in order to avoid a further painstaking statistical analysis, meaningless in any issue due to the small number of cases involved.

ISC-C, 1964-1975. - An approach, similar to the one used for the latest period of ISC-HF, has been applied. Table 7 shows that the agencies to be compared are ISC, NEIC, BCIS, MOS, and ATH; the others have too small a contribution. Fig. 25 (equivalent to Fig. 23 for ISC-HF), together with Fig. 26 (equivalent to part of Fig. 24 for ISC-HF), allow us to conclude that the two “best reporting” agencies for the period 1964-1975 are ISC and NEIC. The mismatch distances equal to zero or larger than 240 km have been ignored in designing Fig. 25; the mean mismatch distances between determinations of the same events issued by ISC and NEIC are in the range of 12-20 km, except in 1964 where it reaches 31.5 km.

As for ISC-HF, separation between the two “best reporting” agencies is difficult and debatable. Our selection is as follows:

1964	ISC
1965 - 1970	NEIC (actually US CGS)
1971 - 1973	NEIC
1974 - 1975	ISC

As for ISC-HF again, the first default determination is the one issued by the second “best reporting” agency for the period considered, and the second default determination is the principal one (as defined by ISC-C).

As it will be discussed below (see Section 5.3. “Magnitudes in the final data set”, in fine*)*, only events with magnitude ≥ 4 have been included in the final data set, for the period 1964-1975. However, the selection procedure, as described above, has been conducted on the entire ISC-C, restricted to the Mediterranean (see Section 3.l. “Considerations …”, in fine); the question is, then, whether this difference in the samples introduces an undesirable bias in the conclusions drawn. The relative percentages of large mismatch distances induced by the pairs ISC-NEIC have been computed and reported on Fig. 27, for the entire sample on the one hand, and for the sample from which events with magnitude < 4 have been sorted out, on the other. Fig. 27 clearly demonstrates that the trends are the same in both situations, with a tendency for the curves ISC-XXX and NEIC-YYY to be closer to each other when the magnitude filter is applied.

Finally, the cumulated list of epicentral locations, after selection as described in the above sections, amounts to approximately 33 000 events. The sharp-cut space filter is yet to be applied in order to eliminate those events still on the list which fall outside the strict geographic limits imposed by the present study.

*5.2. Focal depths in the final data set**

Applying the selection procedure described in the preceding section has drastic consequen-

* The discussion which follows is closeley related to Section 4.2. "Focal depth".

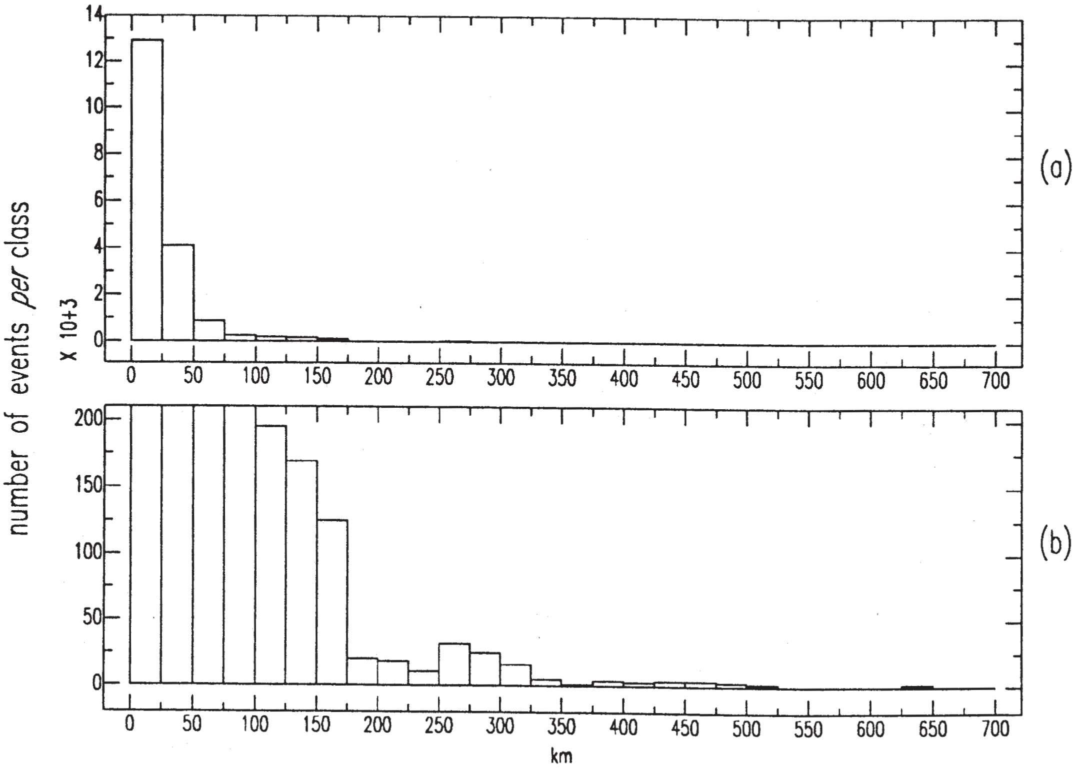

Fig. 28 - Histogrammes of non-zero depth distribution in the Mediterranean file (sec Section 5.1. "Epicentral locations in the final data set", in fine). The population scale is linear. (a) All classes are shown in full. (b) classes too largely populated are truncated.

ces on parameters which have not been taken in to consideration in the selection scheme: this holds for focal depth and magnitude.

Focal depth is not necessarily included in the parameter stream of the determination finally selected. When no depth estimate is included in any determination of a multiple determination event, the processing is unfortunately straightforward. If a depth estimate is included in at least one of the determinations of a multiple determination event which is not the one finally selected by considering the features of the corresponding epicentral locations, then one is confronted with a problem which has no ideal solution. Estimations of epicentral location and of focal depth are usually strongly interdependent; then, either the depth estimate available is ignored, in order to preserve the internal consistency of the data, but a piece of information will be lost, or the depth estimate corresponding to another epicentral location (of same event) is included in the final record at the expense of the internal data consistency. In view of the goals adopted for our project, we finally chose the latter solution: as concerns the ISC-HF derived data sub-set, a depth estimate has been kept whenever available (following the same hierarchy as for the epicentral location: second "best reporting" agency for the period under consideration, then the ISC's "principal determination"). It should be realized that such a deliberate inconsistency is limited to a rather small proportion of the records, as is shown in Table 9. Yet, one should be well aware that the solution adopted is a compromise and is strongly dependent on the goals pursued: circumstances certainly exist in which focal depth would be a crucial parameter; for a cartographic representation of seismicity, a rough separation into a few depth classes is acceptable.

In its final form, the data set obtained on an epicentral location basis, is about 33 000 records long; 56.8 % show a depth field with a non-zero value; of course, this relatively large proportion comes mainly from the contribution of the EMSC's data set.

In Fig. 28, histogrammes depict the distribution, as a function of depth, of events for which a focal depth estimate has been retained; the depth range considered goes down to 700 km, and each class spans 25 km. As expected, a wide majority of events are shallow ones. The large variation in class population makes it difficult to display; the upper part of Fig. 28 (Fig. 28a) has a linear scale for class populations: the most populated classes are easily visible while, very soon after (at around 150 km), the current value is well within the thickness of the drawing line; in the lower part of the figure (Fig. 28b), the linear scale has been preserved, but the most populated classes are truncated at 210 events, so that one can get an idea of how deeper events (focal depth ≥ 100 km) are represented in the Mediterranean area: actually, beyond 175 km, the number of events drops dramatically; there is a secondary accumulation of events at about a 250 - 300 km depth; residual activity (in terms of numbers of events) is yet observed between 350 and 525 km; two events are worth noticing, in the 625 - 650 km range, just below southern Spain.

Changing from a linear to a logarithmic scale for population classes, Fig. 29 (left) shows the histogrammes for each of the files (ISC-HF, ISC-C, and EMSC file); Fig. 29 (right) is constructed with a linear scale again, and large populations of shallow events are truncated. A rapid visual inspection of this figure leads to the following remarks:

- the 0-25 km class of ISC-HF is obviously underestimated when compared to the same class of the other two data sub-sets. Although there could be a complex influence of the distribution of the observation station network(s) in the first 60 years of the century (which, as it is well known, is crucial for the control of focal depth), it is highly probable that a large proportion of the ISC-HF events the focal depth of which is not known (and thus are not taken in account in to the charts) should populate the class under consideration; a less cautious way of saying it is that many such latter events are likely to be shallow;
- the differences in population of the two first classes (0-25 and 25-50 km) between 1964-1975 (ISC-C) on the one hand, and 1976-1988 (EMSC file) on the other, are most probably mainly due to the way focal depth is assigned arbitrarily (either by the analyst, or by the computation code itself) when it cannot be computed (i. e. when the computation convergence is bad if depth is left free to change): the arbitrary default value is 33 km (believed to be the mean depth of the base of the Earth's crust) for the first period, and 10 km (assumed to be representative of the seismogenic, brittle, part of the crust) for the second period;
- at a focal depth of beyond 50 km, the three histogrammes are largely different; however, similarities are still noticeable, like the low between 175 and 250 km for the last two periods; this suggests that the wide differences for the events deeper than 325 km, for these last two periods, are most probably due to non-stationarity of occurrence when looked at over a time period of about 25 years. A careful examination of the histogramme for ISC-HF suggests, in addition, that, for the events deeper than 50 km, the focal depth has often been rounded off to the nearest multiple of 50 km.

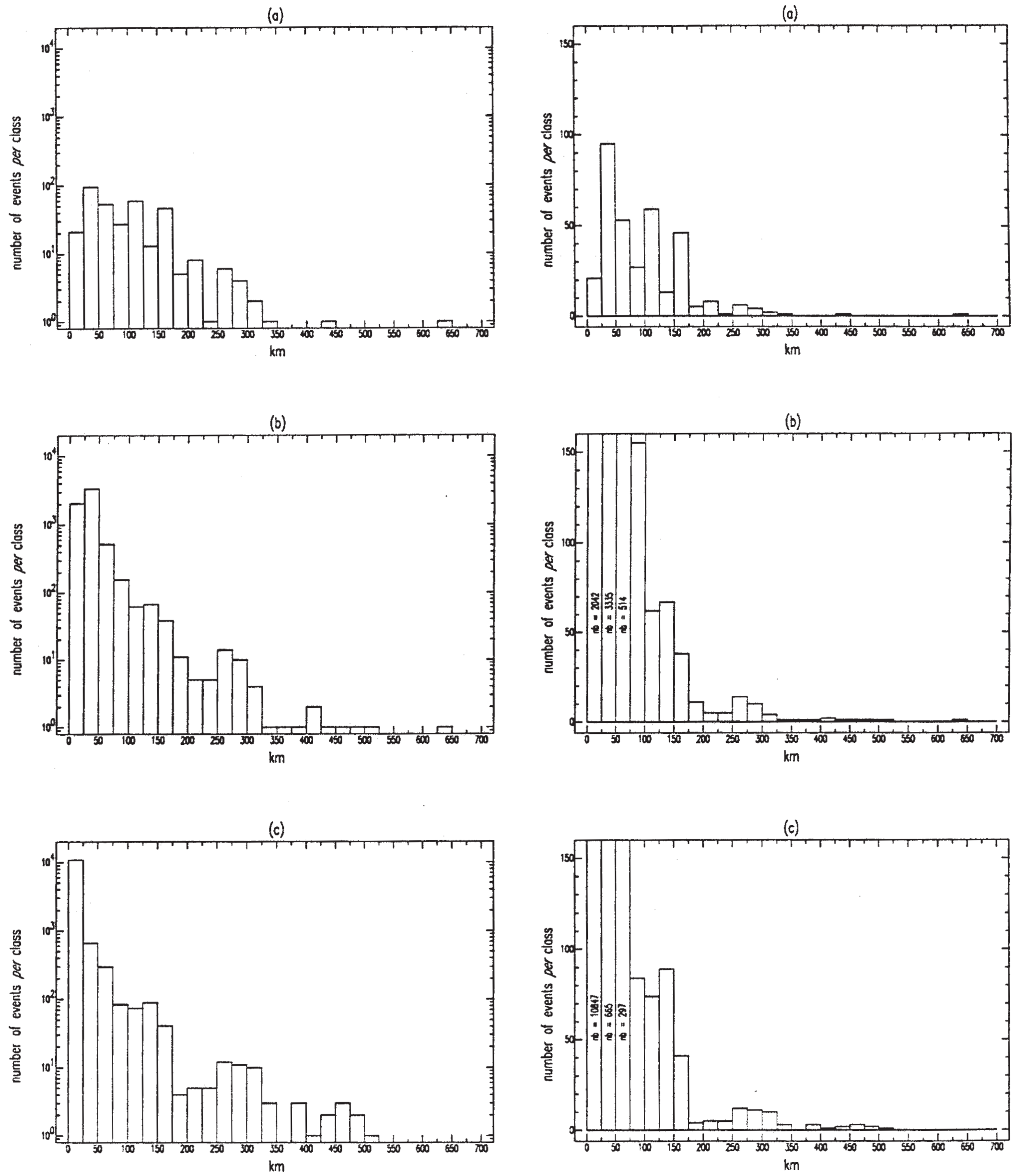

Fig. 29 - Histogrammes of non-zero depth distribution in the Mediterranean file (see Section 5.1. "Epicentral locations in the final data set", in fine). The population scales are logarithmic on the left-hand half of the figure, and linear on the right-hand half. (a) ISC-HF derived sub-set, 1904-1963. (b) ISC-C derived sub-set, 1964-1975. (c) EMSC derived sub-set, 1976-1988.

Classes with a large depth span have been considered and lead to Table 10, where the percentages of shocks deeper than 50 km are displayed according to the original file. The low decrease rate, as a function of depth for ISC-HF, differs strikingly from the other two (which differ from each other, as well). This is difficult to account for, unless as a bias effect, which in turn would be a by-product of the initial selection in constructing the ISC-HF (possibly and partially caused by the observation station network(s) at that time).

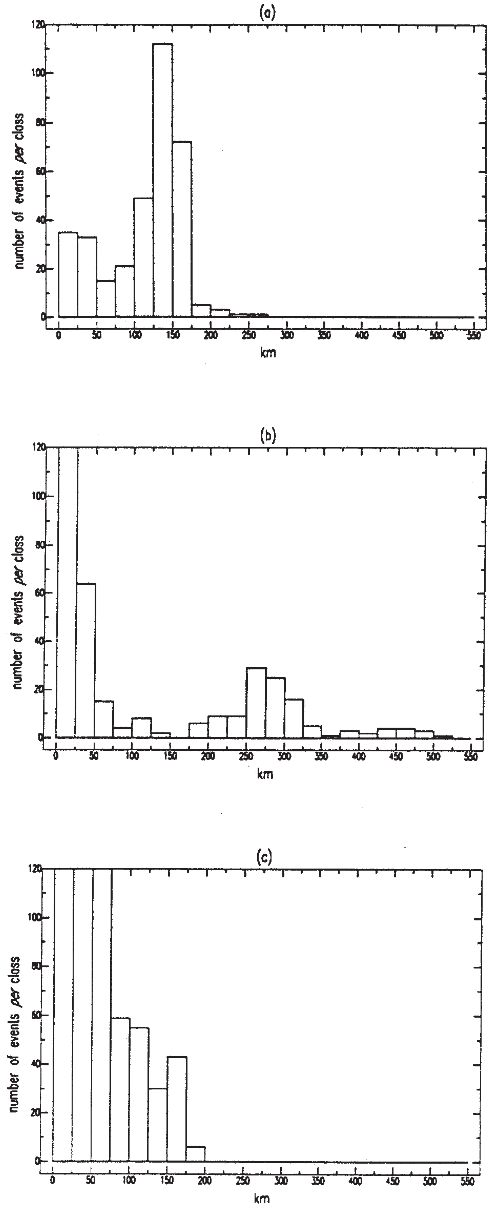

Fig. 30 - Histogrammes of non-zero depth distribution in three regions. The population scales are linear and truncation occurs. (a) Carpathians (44°- 48°N 24°- 28°E). (b) Tyrrhenian Sea (36°- 42°N 12°-17°E). (c) Aegean Sea (35°- 41°N 23°-27.5°E).

The traditional classification into shallow (from 0 to 50 [or 60, or 70, or 100] km), intermediate (from 50 [or 60, or 70, or 100] to 300 [or 350] km), and deep events (from 300 [or 350] to ...), does not really fit the histogrammes (with the exception of classical "deep events" (from 300 [or 350] to ...). A simple regionalization has been attempted, based on well-known peculiar tectonic features (Fig. 30). If one excludes the two exceptional shocks underneath southern Spain (about 625-650 km depth), the only region where earthquakes are occurring deeper than 275 km is the Tyrrhenian Sea region: the corresponding histogramme strongly suggests a significant

Table 10 - Percentages of events deeper than 50 km, for three depth classes, in the data sub-sets derived from the three initial files.

	50 - 100 km	100 - 150 km	150 - 200 km
ISC - HF	35.2 %	31.8%	22.5 %
ISC-C	74.7 %	14.4 %	5.4 %
EMSC	58.9 %	25.2 %	12.9 %

depth class between 150 and 350 km; the Aegean Sea and the Carpathians would rather suggest a single cut at 175/200 km. As for the upper limit of intermediate foci, attempts at establishing narrower classes (10 km) between 40 and 100 km for each file have failed showing any unique marked cut, which could be put at 70 km for ISC-HF, 60 km for ISC-C, and 50 km for EMSC sub-sets; no cut at all shows up for the global data set; regionalization could bring in some interesting trends but has not been attempted.

Finally, an arbitrary classification scheme has been adopted for the whole area:

$$H < 60 \text{ km}$$
$$60 \text{ km} \leq H < 150 \text{ km}$$
$$150 \text{ km} \leq H < 350 \text{ km}$$
$$350 \text{ km} \leq H$$

*5.3. Magnitudes in the final data set**

As in the case of focal depth (see preceding Section 5.2. “Focal depths in the final data set”), a magnitude estimate is not necessarily associated to the determination selected on the grounds of the “quality” of its epicentral location. If the determination is a singular one, or if one estimate is associated to any determination of a multiple determination event, then there is only one way to follow: either the only available magnitude estimate is included (whatever the type of magnitude) in the final data set, or none will be included. In the case of a multiple determination event, if no magnitude estimate is associated to the determination selected, then an eventual estimate is searched for in the determinations issued by other agencies and is included in the final data set, as well as the agency responsible for the magnitude estimate. The inconsistency introduced in the final data set by doing it this way, has been discussed at length in the case of the focal depth parameter. However, for the magnitude parameter, it is of much less importance, because magnitude evaluation has no influence on hypocentral location evaluation (the reverse is not true: magnitude evaluation strongly depends on hypocentral location and focal depth); one could say,

*The discussion which follows is closely related to Section 4.3. "Magnitude"; in particular, one should refer to that section for what concerns the type of magnitude

Table 11 - ISC Historical File [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Multiple determination events and magnitude estimates per group of years.

	1904 - 1951	1952 - 1957	1958 - 1963
Number of multiple determination events:	252	309	368
with at least one magnitude estimate	113	157	225
with magnitude estimate taken from a determination different from the determination selected according to the epicentral location	12	14	184

using deliberately poorly defined terms, that the inconsistency occasionally introduced affects only the homogeneity confidence one assigns to the agencies involved.

Table 11 shows that, up to 1957, most often the magnitude estimates come from the same determination as the epicentral location; this is no longer the case in 1958-1963, as during this period a large majority of magnitude estimates associated to multiple determination events have been taken from an earthquake determination different from the one contributing the epicentral location.

Figures 31 and 32 depict the distribution of events up to 1975 according to their magnitude (assigned as it is explained in Section 4.3. "Magnitude", and above in the present section), class by class, each class spanning one unit of magnitude. The column with no magnitude header gives the percentages of events for which the magnitude is not known.

Figure 31 relates to ISC-HF. The absence of shading in the first row of the lower "checker" reminds us that all the determinations are singular for the period 1904-1927 (see Section 3.2. "The ISC's Historical File"), and thus make up the upper "checker". In the case of events with an assigned magnitude, the situation is markedly different for the earlier periods than for the later ones: up to 1935, almost all the events are assigned a magnitude estimate (this is because of the Gutenberg and Richter (1949) source; see Section 4.3. "Magnitude"), but they belong to classes with a magnitude ≥ 5. From 1935 to 1950, the magnitude classes populated are more scattered than before, and the proportion of events with no magnitude assigned increases markedly; in addition, there are very few singular determination events larger than 5 in magnitude, and practically none larger than 6. From 1945 to 1952, no singular determination event is assigned a magnitude estimate. Generally speaking, the shading of the "checker's cells" is denser in the case of singularized multiple determination events (see Section 5.l. "Epicentral location"), and shifted to the right by about one magnitude unit, indicating a larger percentage than for the corresponding singular determination events. There is practically no multiple determination event with a magnitude estimate smaller than 4 (there are just a few in 1962).

Figure 32 illustrates the magnitude distribution within ISC-C. Comparing it with the ISC-HF (Fig. 31), the same features (larger magnitude estimates, shifted to higher values by about one magnitude unit, for singularized multiple determination events rather than for magnitude estimates of singular determination ones) superimpose themselves on a much broader and smoother

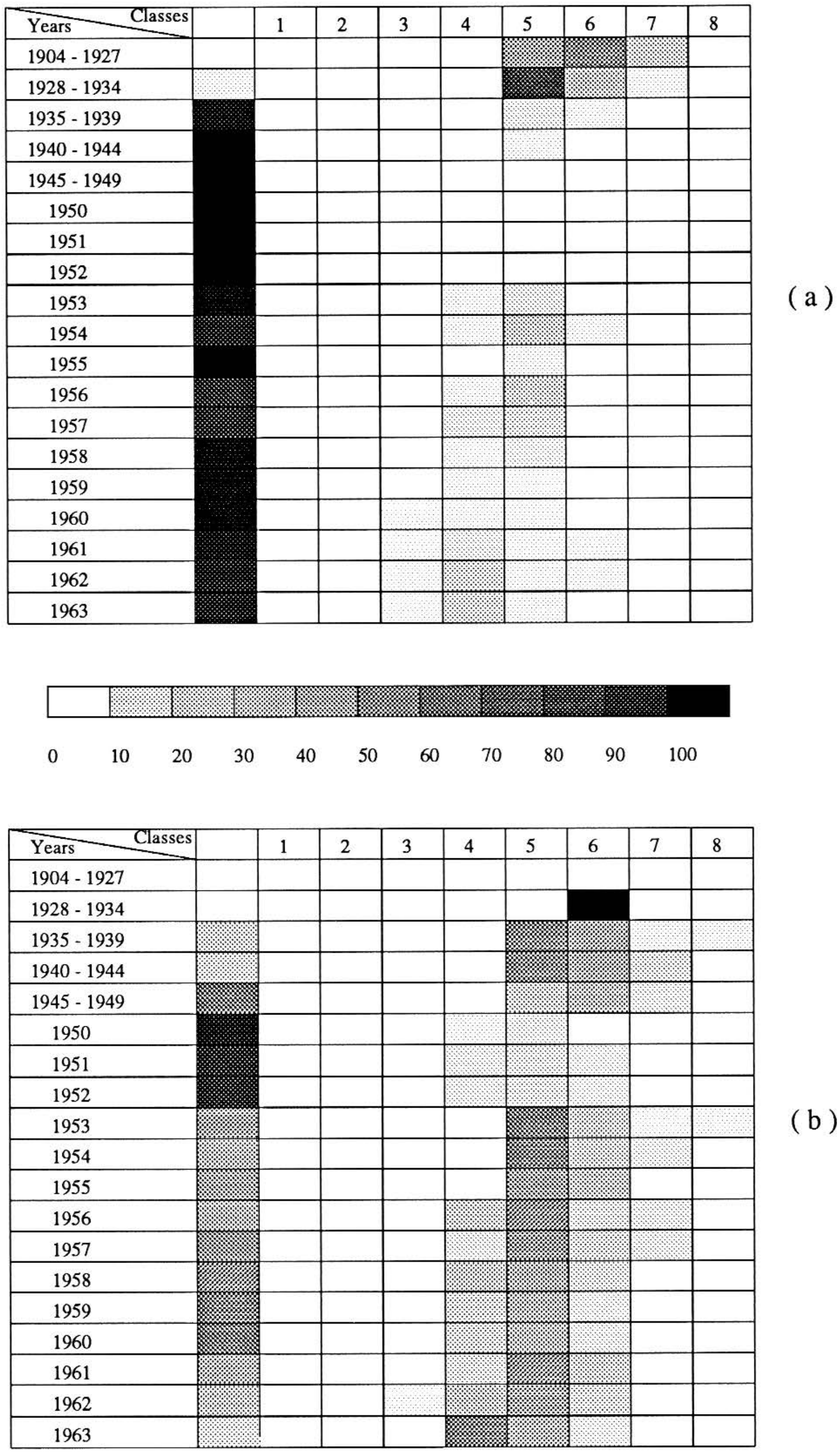

Fig. 31 - ISC Historical File [restricted to the Mediterranean; see Section 3.1. "Considerations …", in fine]. Distribution of events according to time and magnitude classes. The population of each cell is proportional to the density of shading, as shown on the shading scale. The (first) column with no magnitude header refers to events for which the magnitude is not known. (a) Singular determination events. (b) Singularized multiple determination events.

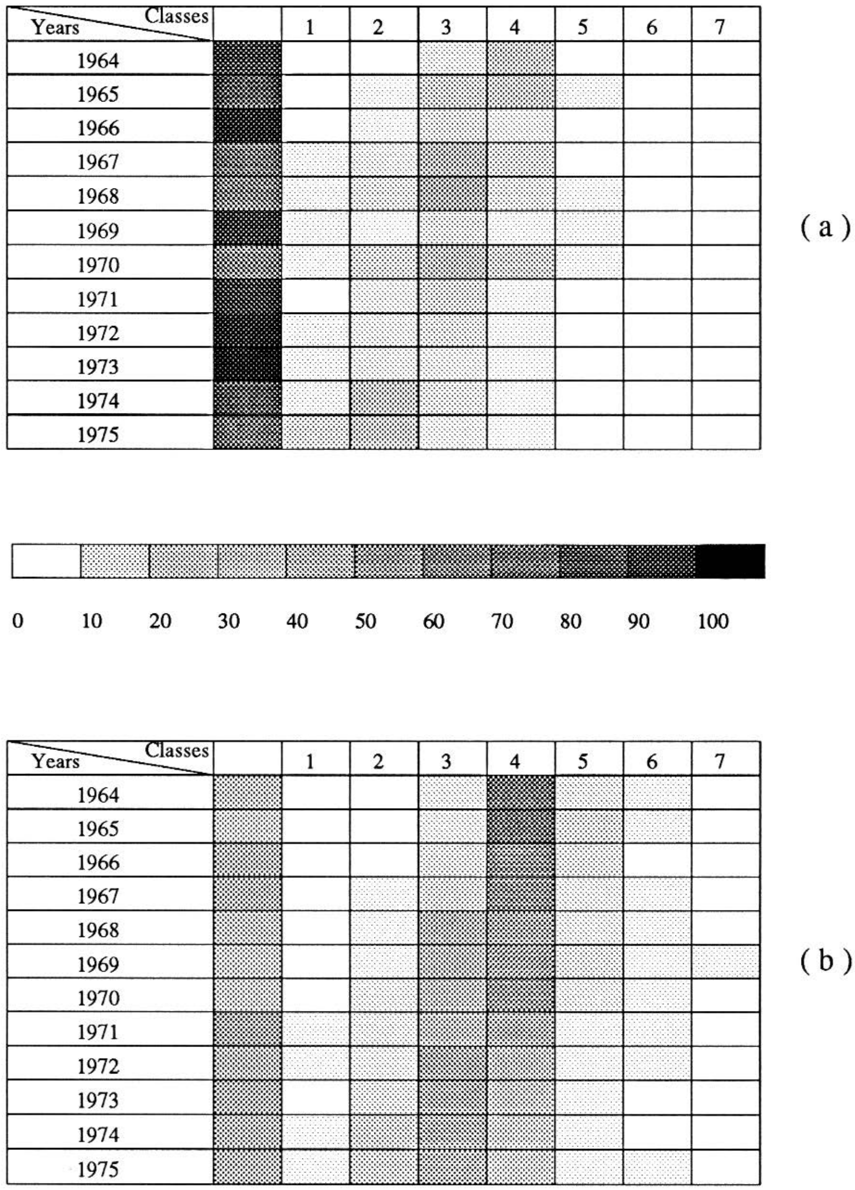

Fig. 32 - ISC Catalogue [restricted to the Mediterranean; see Section 3.l. "Considerations …", in fine]. Distribution of events according to time and magnitude classes. The population of each cell is proportional to the density of shading, as shown on the shading scale. The (first) column with no magnitude header refers to events for which the magnitude is not known. (a) Singular determination events. (b) Singularized multiple determination events.

distribution; whereas the very large magnitude values are not represented (no magnitude 8 event at all, and very few magnitude 7 ones), the low magnitude value cells are more densely populated.

In order to have a better insight into the magnitude distribution of the events of the Mediterranean region, data has been checked against the classical Gutenberg law. Magnitude

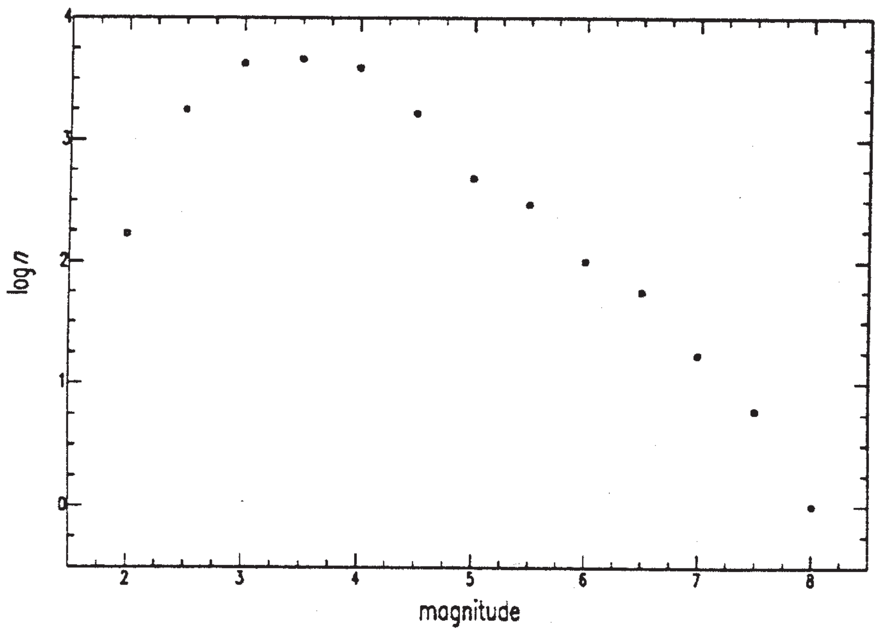

Fig. 33 - Final data set, 1904-1988.5. Plot log *n* versus magnitude; *n* is the number of events per class; a class spans half a unit of magnitude.

classes have been defined in such a way that two classes span one magnitude unit; this is naturally a compromise which takes into account the sparseness of the classes corresponding to too large values of magnitude, and the rounding-off effects (to the nearest unit or half-unit), especially in the case of ISC-HF (for which classes with a smaller magnitude span would have affected the smoothness of the curves). A census has been done on the data set, strictly restricted to 30°- 46°N, 6°W-36.5°E. Fig. 33 shows the results for the full-duration data set, from 1904 to 1988.5. The decreasing trend is approximately linear for magnitudes ≥ 4; but magnitudes < 4 clearly deviate from that trend. The reasons are that either the conditions for the Gutenberg law to be verified are not met (time period spanned by the data set; region covered; magnitude estimation procedures; etc.), or the data set is not "complete", or, obviously, any combination of the two preceding reasons.

Before trying to go further in the analysis of the magnitude distribution, it should be made clear that here the concept of "completeness" applies to that data sub-set to which magnitude estimates have been assigned; in other words, the "completeness" of the data set is probably screened or disturbed by the events to which no magnitude estimate has been assigned. Most often, one considers that the magnitude-unknown events are small size events: if this is plausible for the recent periods, it does not necessarily apply to the ISC-HF, for example.

Two approaches in evaluating the data set(s) completeness are followed below: a "linear approach" which assumes that the data set must be regarded as incomplete as soon as it deviates significantly from a decreasing linear trend; and the "monotonous approach" which simply supposes that the number of events is a smoothly decreasing function of magnitude. Applying these two procedures to the data set defined above (as for its geographic limits), the full-duration data set would be complete for magnitude 4 and above, with the "linear approach", while it will be so

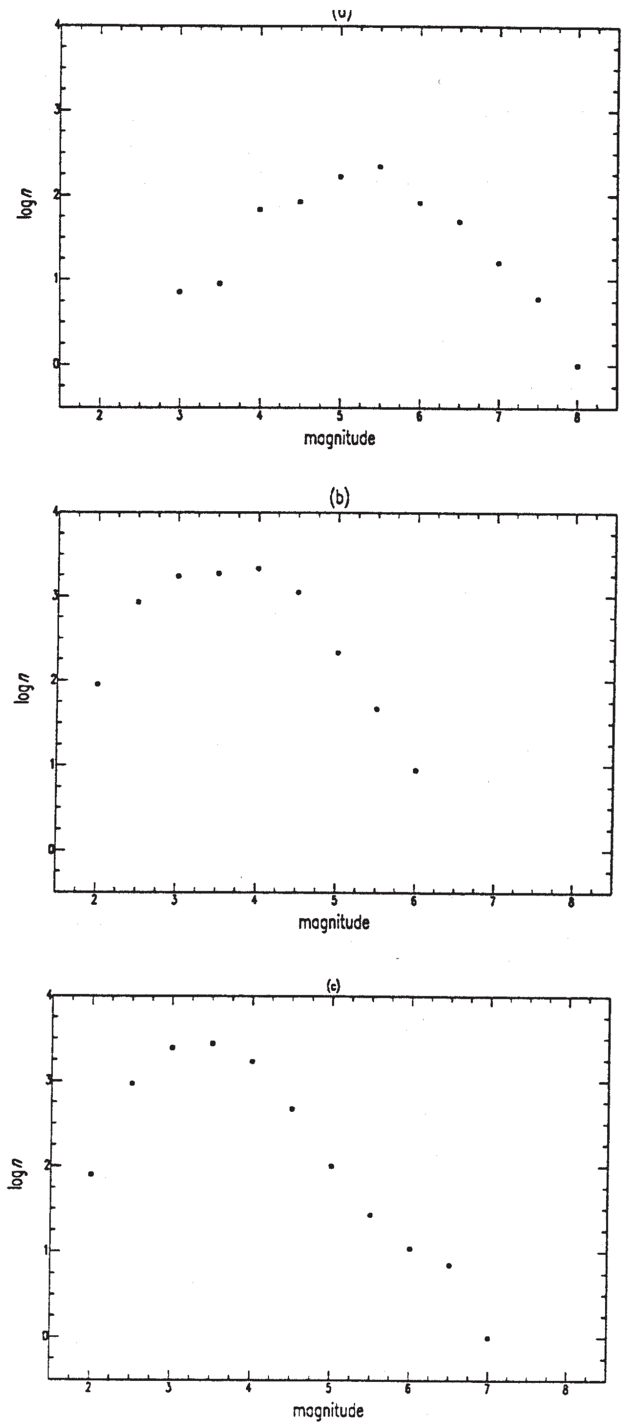

Fig. 34 - Final data sub-sets. Plots log *n* versus magnitude; *n* is the number of events per class; a class spans half a unit of magnitude. (a) ISC-HF derived sub-set, 1904-1963; (b) ISC-C derived sub-set, 1964-1975; (c) EMSC's files derived sub-set, 1976-1988.5.

for magnitude 3.5 and above according to the “monotonous approach” (Fig. 33). Naturally, looking at the full-duration data set conceals the varied behaviour of the different data sub-sets a great deal; in particular, the EMSC data sub-set is strongly dominant in the full-duration data set.

This is evidenced by splitting the full-duration data set into three sub-sets according to their respective files of origin (Fig. 34). Whatever approach is chosen, the ISC-HF derived data sub-set is not complete for magnitudes of < 5.5; the ISC-C derived sub-set would be complete down to magnitude 4.5 (for sure, and possibly to 4, with some degree of indulgence), but highest values are not represented; the EMSC's file derived sub-set would be complete for magnitude 4 and above (“linear approach”), and possibly 3.5 and above (“monotonous approach”). It is worth noticing that the time spans of both the ISC-C derived and the EMSC derived sub-sets are similar: 12 and 12.5 years, respectively. Yet, although the number of magnitude 6 events is similar in the two sub-sets, the number of events for magnitude classes from 4.5 to 5.5 is definitely higher in the former than in the latter; moreover, the slope of the linear trend is significantly steeper for the ISC-C derived sub-set, in spite of the fact that both correspond to the same geographic domain. Twelve years is simply non-significant when considering the time constant of the geological causes of earthquake occurrence: there could be differences between the two sub-sets simply because of the non-stationarity of seismic activity at the time-scale considered; this is evidenced, by the way, through the lack of shocks with a magnitude of over 6 in the ISC-C derived sub-set, as well as through the magnitude 6.5 number of events being well out of the linear trend in the case of the EMSC's file derived sub-set. Nevertheless, an influence of the non-uniform procedure for evaluating magnitude in the two sub-sets cannot be ruled out, and is on the contrary highly probable; actually, a more detailed analysis seems to lead to the conclusion that, using only m_b type magnitude estimates for the EMSC's file derived sub-set would tend to make the curve much more similar to the one of the ISC-C derived sub-set, than using only M_S magnitude estimates. Even though this tentative analysis cannot be taken as strongly conclusive, it suggests that the magnitude type, and thus the techniques used in deriving magnitude estimates, are responsible for (part of) the discrepancies observed.

Back to the ISC-HF derived sub-set, Fig. 35 separates the sub-set into two components, according to time. For the period 1904-1934, the rounding-off to the nearest half magnitude unit is clearly visible, and shortcuts any further comment. For the period 1935-1963, the general linear trend for magnitudes ≥ 5.5 is evident, even though the number of events in class 7.5 disturbs the scenery, while it was more or less compensated (Fig. 34a) by the number of events of the same class for the preceding period, when considering the ISC-HF derived sub-set as a whole. A combined examination of Figs. 31 and 35b suggests that the deviation from the Gutenberg law for the events of magnitude < 5.5 could be alleviated by working out magnitude estimates for the events presently reported without a magnitude estimate. In other words, the “middle-class” magnitude events lie, most probably, for that period, mainly in the first column of Fig. 31a and b.

The above analysis of magnitude distribution, when this parameter is known, has been conducted globally for the entire geographic domain considered. Quite obviously, seismic activity is not uniformly distributed over the whole domain, and a regionalized analysis should have been applied; however, to do so, a partition of the geographic area should have been adopted which identifies, tectonically, homogeneous sub-domains: this was considered going far beyond the goals pursued in this paper, as it would have implied collecting additional local seismicity data (see "Introduction" for a more detailed discussion).

It follows from the discussion on data set completeness, as based on magnitude parameter, that the ISC-C derived data sub-set is close to being complete for events with a magnitude of 4 and over; the other events from that source, weaker than magnitude 4, could be misleading and should be discarded for certain applications, as well as the events for which the magnitude is not known; this decision scheme surely leads to eliminating a certain number of events the strength of which would have justified their inclusion; the corresponding log *n* versus magnitude plot makes it likely that they are not very numerous, and, correlatively, that most of the magnitude-unknown events lie in the small magnitude range. It is probable that most of the events with a magnitude of 4-4.5 which are missing in the data sub-set for the period under consideration, would have fallen mainly in the badly covered regions, at that time, by seismographic stations; these are likely to be on the southern, eastern and northeastern fringes of the geographic domain under consideration.

In the case of the ISC-HF derived data sub-set, applying the same decision rules would have led to eliminating a relatively small but yet significant proportion of events over 2/3 of the time spanned by instrumental data; it thus appears unreasonable to discard magnitude-unknown events for the corresponding period: the log *n* versus magnitude plot does not allow any sound decision on the completeness before 1935, and would have led to putting the threshold at 5.5 for the second half of the period covered, resulting essentially in a collection of large events in the Mediterranean rather than a list of earthquakes which occurred.

For the EMSC's files derived data sub-set, the log n versus magnitude plot allows us to consider it as almost complete for magnitudes 3.5 and over. If one goes down to 3, the deviation from the linear trend is significant, but not very large yet; on the other hand, this data sub-set is the one which is globally the most reliable and the most "accurate" (if this word has any meaning here). It has then been decided to include all the events with magnitude 3 and over (the events with a magnitude lower than 3 [actually, a small number of events] have been discarded). Once more, this is to the expense of the homogeneity of the whole study, as deviation from completeness towards the weak activity primarily affects the geographically "marginal" zones; but, it has been demonstrated previously that this non-homogeneity could not be avoided, and that much more work was needed to alleviate it significantly.

Finally, as far as the magnitude parameter is concerned, we come up with the following data sub-sets:

1904 - 1963	no cut-off magnitude
1964 - 1975	magnitude ≥ 4
1976 - 1988.5	magnitude ≥ 3

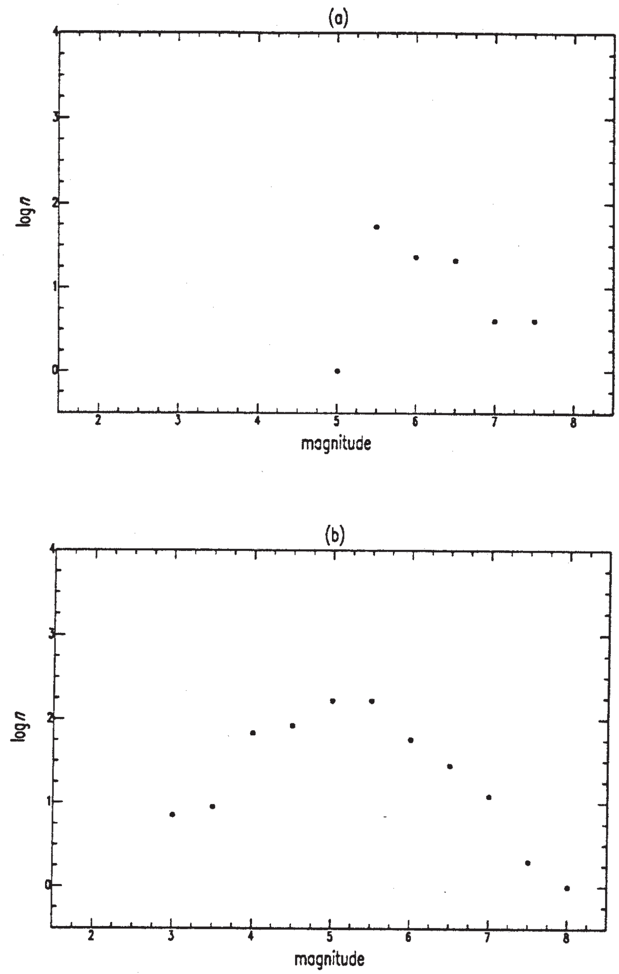

Fig. 35 - ISC-HF derived data sub-set, 1904-1963. Plots log *n* versus magnitude; *n* is the number of events per class; a class spans half a unit of magnitude. (a) 1904-1934; (b) 1935-1963.

6. Representing seismic activity on a map

The final product is a map at the mean scale of 1/1 000 000 (10 sheets) which can be obtained from UNESCO and which represents a data set, sorted as explained in the bulk of the paper (see its small size representation in Fig. 36). The graphics of the map has been designed in the following manner.

LOCATION. - Each event is, indeed, represented by a symbol centered on the location supplied by the determination selected. Except for the most recent period, the rounding-off of the location's coordinates results in an artificial alignment of symbols along straight lines ("orchard effect"), which is simply devoid of any tectonic significance; the older the data, the more apparent the effect: this is evidenced by events located at the crossing of integer latitude/longitude lines, which is obviously highly improbable.

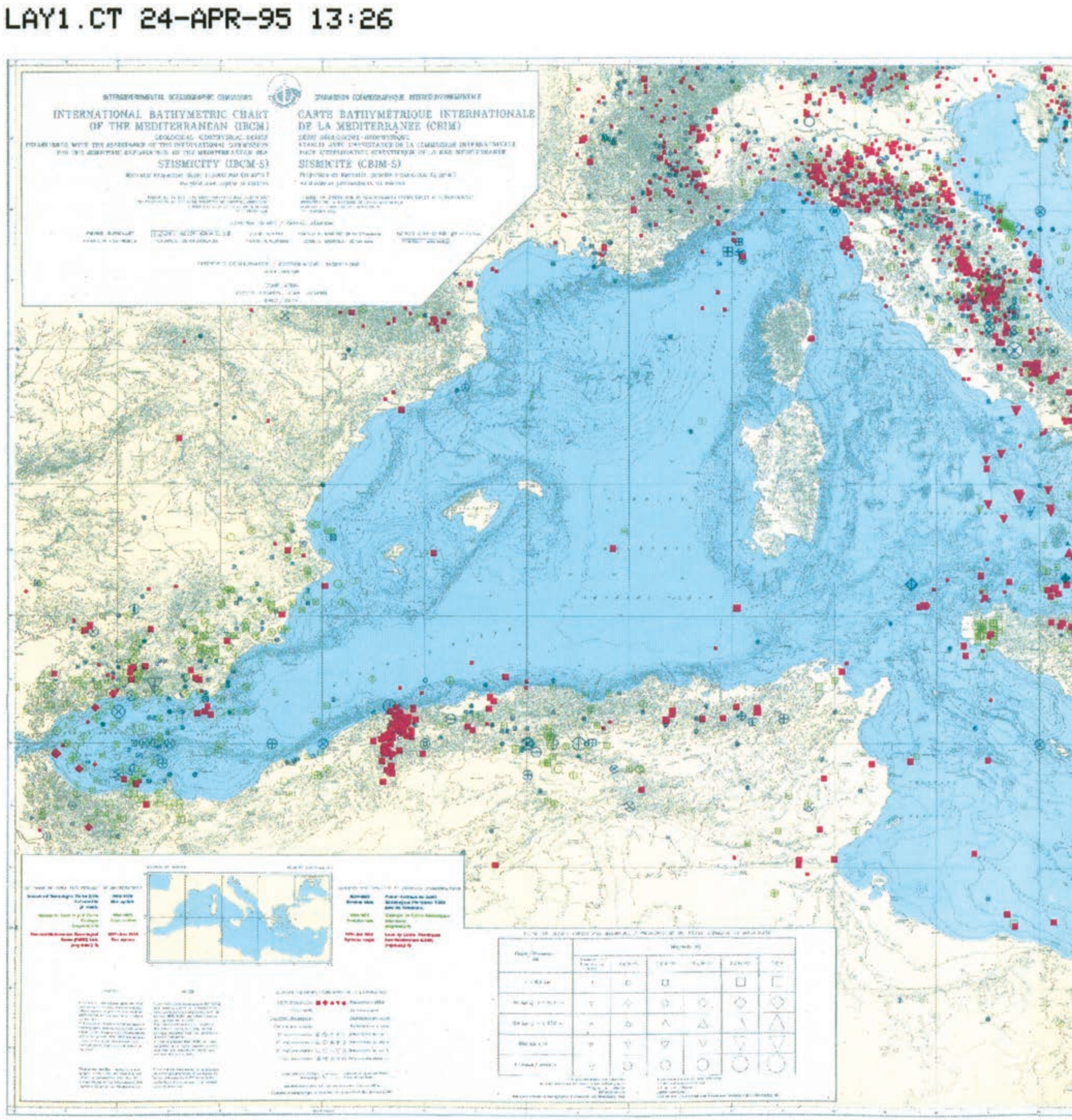

Fig. 36 - See caption on next page.

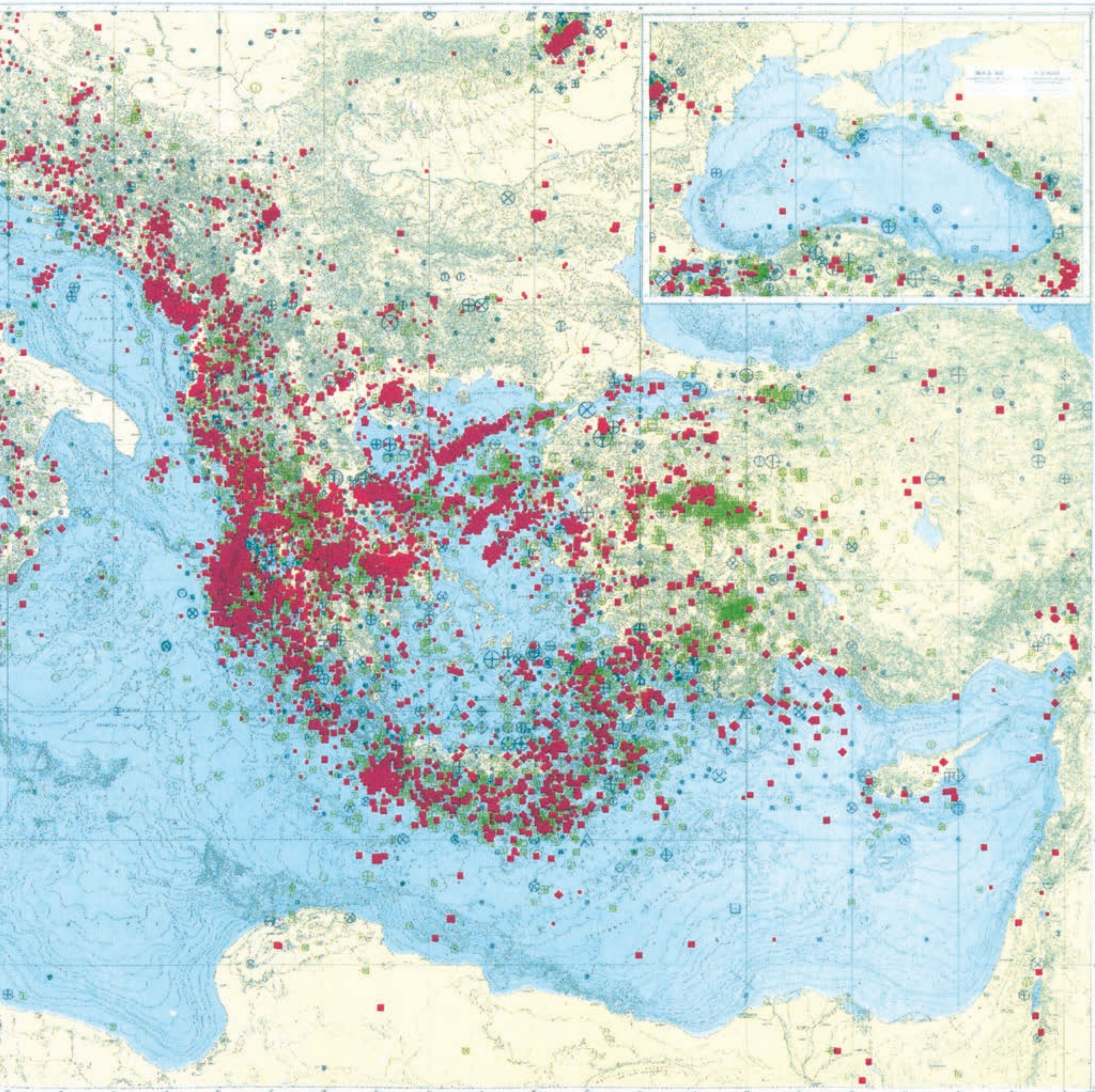

Fig. 36 - Seismicity map of the Mediterranean region.

Table 12 - Attributes of open symbols and ranks of selected determinations.

singular	0	singular	×
singularized, issued by first "best reporting" agency	1	1st rank	+
singularized, issued by second "best reporting" agency	2	2nd rank	|
singularized, "principal determination"	3	3rd rank	no ornament

In order to reinforce the visual impression according to the overall reliability of the data sub-sets, the most reliable one (EMSC's files derived sub-set) is represented by a solid symbol in bright red. The other two sub-sets are represented by open symbols, either in blue (ISC-HF derived sub-set) or green (ISC-C derived sub-set):

period	symbol	colour
1904 - 1963	open	blue
1964 - 1975	open	green
1976 - 1988.5	solid	red

The "quality" of the event's location is further illustrated by adding an attribute to the open symbols, depicting the "rank" of the determination plotted within the tentative classification applied to reporting agencies (see Section 5.l. "Epicentral locations in the final data set"). If the determination plotted is singular, a cross is superimposed on the symbol; if the determination plotted is the one issued by the first "best reporting" agency for the period considered, a + sign adorns the symbol; if the determination plotted is the one issued by the second "best reporting" agency (i.e. the first "best reporting" agency does not report any location for the event under consideration), a vertical bar is added to the symbol; if the determination plotted is simply the "principal" one (*i.e.* neither the first nor the second "best reporting" agencies report any location for the event under consideration), the symbol shows no special ornament. Table 12 summarizes the rank representation on the map.

FOCAL DEPTH. - The depth class for each event is reflected in the shape of the symbol, as displayed in Table 13. A depth-unknown class has been created (the corresponding symbol shape is a circle).

MAGNITUDE. - The magnitude class for the events plotted controls the size of the symbol. The events belonging to the period 1904-1963 (blue symbols) and the magnitude of which is unknown, are assigned a symbol with the same (smallest) size as the lowest magnitude class events belonging to the period 1976-1988.5; the smallest blue symbols are not very numerous. No green symbol of that size is, indeed, plotted.

As discussed at length previously (see Section 4.3. "Magnitude"), it would have been mislea-

Table 13 - Symbol shape according to focal depth class.

H unknown	○
H < 60 km	□
60 km ≤ H < 150 km	◇
150 km ≤ H < 350 km	△
350 km ≤ H	▽

ding to rely too much on the magnitude estimates finally included in the data set. This is why one has intentionally designed the symbol size scale in such a way as that a rapid visual inspection does not allow a precise identification of the corresponding magnitude class for a given event: an event assigned a magnitude class n could, almost as well, belong to magnitude classes (n - 1) or (n +1). Indeed, a more careful examination of the symbols' size compared to the scale reported in the legend, gives the precise magnitude class of the events represented.

In addition to the seismic events, the locations of the seismographic stations* falling into the map's limits have been plotted. Data has been taken from a file supplied in 1987 by the National Earthquake Information Service [NEIS] of the U.S. Geological Survey; the file has been supplemented with information collected here and there; there has been no attempt at making the file complete and accurate; in particular, international codes have not been checked, neither have their uniqueness; and the time period during which a station has been active, recording seismogrammes, has not been searched. Each seismographic station is plotted on the map as a star symbol (in black), associated to its (unchecked) alpha-numerical international code.

7. Conclusion

We have now come to an end of the analysis of the data sets we decided to work on, keeping in mind what we wanted to achieve. It should be recalled, at this point, that the above study was conducted with the final goal of designing a seismicity map of the Mediterranean domain; the path followed for the analysis is strongly dependent on this objective. It could not be otherwise. Parameterization of earthquakes has been so loose (and still is, by many aspects), that the events are (and still remain, even though parameterization will hopefully improve for events to come) poorly characterized. Nevertheless, applications require long time series of seismic activity: one has to take into account events "recorded" by observers (if there have been any at all) at a time when background knowledge did not allow accurate parameterization; yet those events, more or less poorly known, cannot be ignored. Does the final data set accurately depict the seismic acti-

* The reporting of the seismographic station locations results from a recommendation issued by the IBCM Editorial Board

vity of the area ? We feel unable to say more than:
1. it does not look worse than any other (the meaning of this sentence is kept intentionally fuzzy);
2. the potential users have at their disposal quite a number of estimators, which are more or less sound in terms of statistics, and which ideally permit them to accept or reject, step by step, our own results as regards quality assessment analysis;
3. other comparable attempts should be developed starting from other initial data sets, and/or with other basic assumptions, and their merits should be weighed against each other, including the present one.

By no means, do we claim that we have built the best seismicity data set for the Mediterranean; we simply intended to make one approach transparent enough so that any interested reader can understand what it contains. It would be very interesting for other similar studies to be conducted and the stability of the results assessed (not necessarily down to each very individual parameter, which would probably be beyond anyone's reach, and which would have no sense in view of the many ill-defined concepts called upon to be used).

Although we have carefully chosen the initial data sets with the aim of making things as simple as possible, the present study shows how complex they actually are. The initial data sets are well structured and documented; they have been stable in their structure over long periods of time; one cannot think of data sets more easily accessible and manageable. Yet, the results of the detailed analysis fully justify our circumspect attitude from the very beginning. In spite of the efforts developed, a bunch of unsolved problems about many aspects of earthquake parameters, as they were and still are handled by the various users, still remain. To list but a few:

- apart from when a distinct depth phase is detected, the interaction between epicentral location and depth estimation is mastered only when one has a good knowledge of the velocity structure of the medium, travelled through by seismic waves, and a well-adapted geometry of the monitoring network; those conditions are only met in very few cases of local detailed surveys, and there is little hope of retrieving the necessary ingredients and drastically improve the situation for 7/9 or 8/9 of the events instrumentally observed up to now; yet, there are a few lines (among them, a systematic use of joint epicentral location schemes), along which some improvement can be achieved, provided that basic data is available;
- the limited confidence one can give to magnitude estimations: with the exception of the last ten years or so, and especially when making use of data obtained with modern broad-band instruments, magnitude is, in practice, a poorly-defined parameter when it can be evaluated at all; going back to individual records and applying uniform reading and data reduction procedures is conceptually desirable, but in practice beyond the short term reach; working out a few events taken as references (paying attention to a good distribution, geographically and with respect to tectonic settings), and proceeding with some sort of comparison scheme, can be envisaged (and has been performed in special circumstances, for the largest shocks for example) if one has the necessary data at hand (feasibility would probably be limited to the most reliable data sets, coming from stations/agencies where instruments operating and record readings have been conducted in a stable manner over a sizeable period of time); high accuracy must be obviously forgotten with such a protocol, and it brings back to mind that good macroseismic observations

could be relied on and could allow, in certain circumstances, a reliable evaluation of the strength of the events;

- inhomogeneities in data processing could be easily alleviated by re-computing the parameters with a unique processing scheme, provided that the necessary basic data is available; it can simply be thought of in a relatively long-term perspective;
- more serious is the inhomogeneous geographic coverage by seismographic stations: there is no way of creating the basic data if it has not been recorded; fortunately, local networks have rapidly developed in many places in the last twenty years. Similar remarks hold, to a lower degree, for inhomogeneous station geometry with respect to the events to be located, and for location procedures.

We have concentrated our attention on considerations related to the data sets investigated. We should not forget that our ultimate aim was to design a map displaying the seismic activity within the Mediterranean domain. The reader have found a few considerations directly targeted to that specific purpose in the chapter 6. We shall then just conclude with a few remarks having a more general impact. A map of seismicity is mostly thought of as a picture of seismic hazard, even though one is aware that visual impression cannot replace a sound quantitative assessment. The few comments developed above substantiate the conclusion that seismic hazard assessment, at a regional scale, requires the design of methods adapted to inaccurate, or even fuzzy, data. Moreover, a map could constitute, to a certain extent, a pitfall, because it "freezes" a visual impression which is essentially based on questionable data. Then, one naturally concludes that what is badly needed is a database where the event parameters, and the basic data used in deriving the former, are stored and can be conveniently retrieved in such a way that parameter re-evaluation according to specific criteria can be easily made; a map becomes then a simple output, with a limited "life duration", which must disregarded of as knowledge and know-how progress.

A map has been drawn according to a certain graphic design, which is intended to be consistent with the information contained in the final data set. In spite of the strict limitations recalled in the general discussion, a number of features stand out from the map constructed, which were not all obvious so far. The first observation is that there is a marked difference between zones where seismic activity appears diffuse (southern Spain - Alboran Sea - northern Morocco; southeastern France and northern Italy), and others where activity is very much concentrated (axial "backbone" of peninsular Italy; the coastal parts of Albania and Montenegro; central Greece; the northern Aegean Sea; and the Hellenic trough from the Albanian seashore to the Island of Rhodes). Nesting of activity within geographically rather limited areas (sometimes close to "spots") is also observed (western Pyrenees, northwestern Algeria; western Sicily; several spots in western Anatolia; peculiar, but famous, is the case of the Vrancea region in Romania, with an intermediate-depth seismic nest); sometimes nests are more or less isolated within an almost inactive environment (nest lines across the Adriatic Sea, from north of Ancona towards the east), but they could just as well stand out in an active background (Navarra; offshore Livorno; southern Otranto Strait; northern Albania; northern Kosovo; northern Bulgaria; the Thessaloniki region; and the Erzincan - Erzurum region).

However, the visual impression created by the map could be misleading, if not enough attention is paid to the time dimension; seismic activity nesting is clearly apparent with the clustering of bright red symbols; epicentral locations are probably less accurate for green and (a fortiori) blue ones. There is, nevertheless, a strong suggestion that nesting existed before 1976 as well as later: a good example is the northern Aegean Sea where three very active "red" lineations run across, along a southwest - northeast direction; yet, a fourth band, less apparent because of the higher dispersion of epicenters, intercalated between the northernmost and middle red ones, seems to have been active before. Similarly, as already mentioned, a nest has been active offshore Pesaro, which appears quite comparable to the more recent one, south of Ancona. In central Greece and in western Anatolia, activity has been very high for the past 25 years at least; but it looks like a patchwork of places where activity clearly concentrated in episodes preceding 1975, and others which became active later: migration of seismic activity is manifest.

Other artifacts are caused by the shortness of the time-period spanned by the data set plotted: well-known features do not show up; a striking example is the northern Anatolian Fault which does not stand as a continuous lineament contrary to what one would have expected; this is certainly partly due to inaccuracy in epicentral locations, but the most important factor here is probably the too short time span. This remark brings us back to the role of pre-instrumental seismicity data in interpreting the tectonics of the region; as it seems quite reasonable to assume that brittle deformation proceeds, in the Mediterranean region, in the same manner as in other, more active, regions of the Globe, one should expect roughly recurrent activity, which is likely to show return periods of the order or longer than the time-span of instrumental data (from one to several centuries, roughly speaking): it demonstrates that no sound conclusion, on seismic hazard, can be drawn in the Mediterranean region without including "historical seismicity" data. This is the next compulsory, step to take together with the building-up of a comprehensive earthquake parameter database, if one wants to make progress in the domain, within the Mediterranean.

Acknowledgements. The authors are highly indebted to Helmut Aichele, Bertrand Delouis, Alexei D. Gvishiani, Piotr Shebalin, Dario Slejko, Dmitry Storchak, Mladen Živčić, who have reviewed the manuscript at various stages, and have made roborant remarks and constructive comments; they all deserve our warmest thanks, as well as anonymous reviewer who has significantly improved the English language. A special mention is made of Dario Slejko who has offered a timely opportunity to publish this study where circumstances had so far hampered it. Many thanks are due to John Hall, Geological Survey of Israel, for providing color separation of the Mediterranean Sismicity map. Unde the auspices of the IOC-FAO-CIESM and with data contributions from research institutions and private societies, after 20 years collaboration of an ad-hoc international group, the International Bathymetric Chart of the Mediterranean (scale 1:1000000; 10 sheets) was published in 1981. The importance of the physiogrphic results suggested their completion with the most pertinent geophysical and geological data. At the same scale maps of the Bouguer gravity anomalies (IBCM-G; 1989), of the seismicity (IBCM-S; 1991), of the Plio-Quaternary thickness (IBCM-P/Q; 1993) were also published, and maps of the unconsolidate sediments (IBCM-Sed.; 1997) and the magnetic anomalies (IBCM-M; 1998) are being printed. Each map is illustrated bu an explanatory brochure (IBCM-S also by catalogue) and can be obtained (when available, also in digital form) from: Ocean Mapping, IOC-UNESCO, 75732 Paris, Fax +33.1.4568.5812.

Appendix 1: Glossary

Determination. Any set of parameters supposed to describe (occasionally, partly) an earthquake source; usually, an instrumental determination assigns, at least, an origin time and an epicentral location to an earthquake source.

Earthquake catalogue. Any set of earthquake determinations describing the seismic activity of a certain territory over a certain period of time. In an earthquake catalogue, a physical event is described by one or several determination(s); ususally when several are present, one determination is shown as preferred by the author(s) of the catalogue.

Earthquake determination. Any set of parameter determinations supposed to describe (partly, sometimes) an earthquake source.

Earthquake parameter. Any characteristic feature used for assessing an earthquake source, e.g. origin time, epicentral location (in actual fact a pair of values to locate a point on a surface), focal depth, magnitude, *etc*.

Pairwise mismatch distance. Considering a pair of epicentral determinations of a physical event, pairwise mismatch distance is distance (in km) between the two epicenters.

Multiple determination event. In an earthquake catalogue, an event described through two or more determinations.

Parameter determination. Any specific value assigned to an earthquake parameter; most often, it is an numerical value coming out of measurements taken by dedicated instruments; sometimes, the parameter is assigned a class of values, because a more precise estimation of the parameter is not practicable (e.g. "crustal event" for an earthquake source located in the Earth's outer layer).

Singular determination event. In an earthquake catalogue, an event described through a single determination.

Singularized multiple determination event. In an earthquake catalogue, an event described by two or more determinations, one of which has been pointed out, following specific criteria, as best describing the physical event.

Appendix 2: Table of Acronyms

ALG	Algiers University, Algeria
ALI	Alicante, Spain
ATH	Athens, Greece
BCIS	Bureau Central International de Séismologie, Strasbourg, France
BUC	Bucharest, Romania
CGS	(US CGS) United States Coast and Geodetic Survey, Washington, D.C., USA
CSEM	Centre Sismologique Euro-Méditerranéen, Strasbourg, France
EMSC	European-Mediterranean Seismolocical Centre, Strasbourg, France
GUTE	Gutenberg and Richter, 1954, ed. 1954.
HFS	Hagfors Seismic Array Station, Sweden
HFS1	Hagfors Seismic Array Station, Sweden
HFS2	Hagfors Seismic Array Station, Sweden
IBCM	International Bathymetric Chart of the Mediterranean
ISC	International Seismolocical Centre, Newbury, UK
ISC-C	International Seismological Centre's Catalogue
ISC-HF	International Seismological Centre's "Historical File"
ISS	International Seismological Summary, Kew Observatory, Richmond, UK

KSA	Ksara, Lebanon
LAO	Montana Large Aperture Seismic Array, USA
LIS	Lisbon, Portugal
MDD	Madrid, Spain
MOS	Moscow, Russia
NEIC	National Earthquake Information Centre, Golden, Co., USA
NEIS	National Earthquake Information Service, Golden, Co., USA
PDE	Preliminary Determination of Epicentres (NEIS/CGS)
PRA	Prague, Czech Republic
RBA	Rabat, Morroco
ROM	Rome, Italy
SPGM	Service de Physique du Globe, Institut Scientifique, Rabat, Morocco
STU	Stuttgart, Germany
UNESCO	United Nations Educational, Scientific and Cultural Organization
US CGS	United States Coast and Geodetic Survey, Washington, D.C., USA
US GS	United States Geological Survey, Reston, Va., USA
TIF	Tiflis (= Tbilisi), Republic of Georgia
TIR	Tirana, Albania
TRI	Trieste, Italy
VIE	Vienna, Austria
ZUR	Eidgenössische Technische Hochschule (ETH), Zürich, Switzerland

References

Adams R. D., Hughes A.A. and Mc Gregor D. M.; 1982: *Analysis procedures at the International Seismological Centre.* Phys. Earth planet. Int., **30**, 85-93.

Comninakis P. E. and Papazachos B. C.; 1986: *A catalogue of earthquakes in Greece and the sourrounding area for the period 1901-1985.* University of Thessaloniki, Geophysical Laboratory, Publ. n° 1, 167 p.

Flinn E. A. and Engdahl E. R.; 1965: *A proposed basis for geographical and seismic regionalization.* Rev. Geophys., **3**, 123-149.

Gutenberg B. and Richter C. F. 1954: *Seismicity of the Earth and associated phenomena.* Princeton University Press, Princeton, 1949, 273 p.

Gutenberg B. and C .F. Richter C .F.; 1941: Seismicity of the Earth. Geol. Soc. Am., Sp. Pap. n° 34, 131 p.

Karnik V.; 1971: *Seismicity of the European area*, vol. I, Reidel, Dordrecht, 1969, 364 p.; vol. II, 320 p.

Lomnitz C.; 1974: *Global tectonics and earthquake risk.* Elsevier, 320 p.

Richter C. F.; 1935: *An instrumental earthquake scale.* Bull. Seism. Soc. Am., **25**, 1-32.

Rothé J. P; 1969: *The seismicity of the Earth 1953-1965 - La séismicité du globe 1953-1965.* UNESCO, 336 p.

Shebalin N. V.; 1974: *UNDP/UNESCO Survey of seismicity of the Balkan region*. UNESCO.