INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460392 - I

SMR/107 - 14

WORKSHOP ON PATTERN RECOGNITION AND ANALYSIS OF SEISMICITY

(5 - 16 December 1983)

INTRODUCTION

V. KEILIS-BOROK

These are preliminary lecture notes, intended only for distribution to participants.
Missing copies are available from Room 230.

INTRODUCTORY NOTES

1. The scope of the Workshop: a know-how, which allows us to extract from the seismological data many more conclusions than are extracted at present. The data, already stored on the computer tapes, contain much information, which so far has remained undisclosed.

This is not an unfortunate trait of seismology. A similar situation is common for many fields, where a large volume of diverse data has to be analysed; among examples are almost all solid earth sciences; prospecting for mineral deposits, especially for oil fields; medicine; economy and social sciences, etc. Seismology shares with such fields the following difficulties:

(i) Existing theory does not yet provide a sufficient base for the data analysis. Specifically, seismology has no adequate physical model for the occurrence of the earthquakes.

(ii) Statistical methods do not provide such a base either, since the dimensionality of the data is too large for their volume. For example, suppose that we are interested in the occurrence of the earthquakes with magnitude ≥ 7. About 1500 of them are recorded instrumentally, and it seems a lot. But they have 3 basic types of mechanism - dip-slip, strike-slip and intermediate; 6 basic types of tectonic environment - within litospheric plates and on their margins: each within continents, within oceans and in ocean-continent transition zones; at least 6 basic types of tectonic history; 2 basic types of seismic background - high and intermediate - and occur during 3 types of seismic trend - stable, increase and decline. Altogether we have 3 x 6 x 6 x 2 x 6 = 648 different slots, to distribute our 1500 earthquakes - no straightforward statistics is available.

(iii) The last difficulty is that the data accumulation is traditionally introspective, if not addictive; it is not always oriented to the results.

2. We shall learn during this workshop, how these difficulties can be overcome (or avoided to a not inconsiderable extent) in two problems of earthquake prediction: recognition of earthquake-prone areas and long-term time prediction of the strongest earthquakes of a region.

In the second problem the research and the test of current hypotheses will be considered, since no reliable methods of routine prediction are developed as of now.

Table 1 shows the place of these problems in the general framework of the earthquake prediction. Our problems may sound disappointing, since prediction is sometimes associated with short-term warning, sufficient to get out of the city in a hurry. Actually, these two stages of prediction are of great importance. There is a general rule for the problems of applied mathematics: what kind of solution is looked for depends on how it will be used.

Earthquake prediction has to be used for disaster-and-relief preparedness, specifically - to activate a set of safety measures. Table 2 lists the types of disasters and Table 3 lists the types of safety measures. They are clearly of the civil defence type (actually, the civil defence is responsible for them in many countries). We see from Table 3 that our first two stages of prediction can be used independently for activation of important safety measures, and also constitute the necessary base for subsequent stages.

3. The practical importance of the development of earthquake prediction is already large and growing rapidly. According to UNESCO's estimation, the economic damage from the earthquakes per year in the whole world averages several billion U.S. dollars, and the number of people affected is tens of thousands. A single catastrophic earthquake may cause up to $ 100 billions of damage and several hundreds of thousands of casualities.

Quite recently, the danger of earthquakes has begun to increase rapidly. This is due to the explosion of the urban population in seismically active regions and to the proliferation of vulnerable danger-prone objects, such as life-lines, chemical and nuclear plants, etc. (seismicity, itself, remains about the same).

One of the lectures will show that by the end of this century, several tens of millions of people may experience a catastrophic earthquake, only in the largest cities of the world.

4. What can be done about this in our particular problems (which cover, of course, only a part of the whole spectrum of seismic risk)?

Right now, three major developments have brought about a revolution in seismology:

(i) A new approach to the theory of the occurrence of earthquakes is outlined, based on the general theory of dynamic systems.

(ii) Global observational systems have reached such a high resolving power that they provide sufficient data for many local studies; simultaneously, satellite-supported communications made these data quickly available and cheap.

(iii) Methods of exploratory data analysis proved that the data already accumulated, are sufficient for the solution of many key problems in earthquake prediction. A special software and the new, inexpensive microcomputers made these methods relatively easy to implement.

During this Workshop we shall discuss the mathematical and computational aspects of these three topics, and the new strategy which they bring to geophysics - especially in the developing countries. We shall focus attention on earthquake prediction. However, part of our studies are highly relevant to many other problems - particularly to the search of mineral deposits. I shall now outline briefly what we shall discuss with regard to each of these topics.

5. Theory. Tectonic earthquakes occur because the lithosphere of the Earth is divided into the hierarchy of blocks (volumes) which move relatively to each other. The largest blocks are the lithospheric plates themselves which are of the size of continents. They are consecutively divided into small blocks; after about 5 divisions we come to the blocks, like the hills around Trieste and continue down to tiny pieces of rocks, separated by microcracks. Some soild state physicists

tend to confirm this sequence up to crystal grid, but we will focus our attention on tectonic structures with linear dimension of hundreds to thousands km. The blocks are separated by relatively thin fault zones, from a few km. to several tens of km. wide. In the practical exercises we shall see the map of San-Andreas fault zone - between the Pacific and the North American plates in California. A rather thin fault zone can be seen on the slope of hills above the ICTP.

Until recently, the fault zones were considered as passive interfaces: they hold the blocks together by friction and cohesion, and they experience fracturing due to the relative displacement of the blocks. A significant part of the displacement is realized through the fast slips, i.e. through the earthquakes, Attempts to design a model of such a process were naturally directed towards traditional mechanics of elasto-viscous media, with finite-element method as a computational tool. However, complicated processes in the fault zones were discovered recently, so that these zones now look like active structures, which regulate the movement of the blocks. In other words, the energy of tectonic development is stored in the whole lithosphere, and possibly also below it, but the release of the energy is controlled within the fault zones by many factors which have recently come to our attention.

First of all, fault zones always contain migrating fluids - water, magmatic solutions, etc. These fluids control the friction by several types of lubrication and, possibly, by stress corrosion. Even a monomolecular layer of a lubrication can reduce the friction by factor $10 - 10^2$. Migration of the fluids is, in turn, a complicated, non-linear process, strongly dependent on the heat flow.

Another factor is the occurrence of the numerous phase and chemical transitions in the fault zones. Some of them, such as calcite-aragonite can lead to the loss of the volume, unlocking the fault for a short time; others - like dehydration of serpentine can release additional water, unlocking the fault by lubrication.

Yet another factor is the continuous relative slipping of the grains of the rocks. It is seen as "seismic emission", i.e., continuous emanation of a tremor on 50 - 100 hz by the whole body of the lithosphere, and by the fault zones in particular. All such factors make the fault zones unstable. The instability is illustrated by the fact that this emission is modulated by the vanishingly small (10^{-8}) tidal deformations.

Altogether, these factors can not be allowed for in a straightforward way - by the equations of the mechanics of continuous media. The only not impossible approach to the dynamics of the fault zones seems similar to the statistical physics: to look for some integral description. A fault zone is a complicated dynamic system, which in a seismic region is on the verge of instability.

One example to illustrate how consequentional is this approach is the following: The traditional and apparently natural view considers the development of the fault zones as the combination of fracturing and healing; the latter is necessary to explain why the fault zone has not been converted into dust by now. Now this view is questioned: an alternative hypothesis is that the fault zone is an ensemble of blocks, which merge in a different way during different episodes of the stick-slip process (ac. M.A. Sadovsky).

An important impact of this new approach is the correction of the traditional bias: to expect to find, and to try to recognize, in a natural phenomena, only three basic types of development - stability, trend and oscillation. Actually, other types exist in Nature and are well known in the studies of the dynamic systems - strange attractors, bifurcations and some others. This approach will be discussed in the lectures of Prof. L. Knopoff.

The more general aspects of the dynamical systems, relevant to a wide variety of phenomena, from economy to chemistry, will be discussed in the lectures of Prof. G. Puppi.

6. Exploratory data analysis. The earthquake prediction relates to the dynamics of the fault zone, as medicine to biology. They should, but still cannot, be merged, and something should be done, nevertheless, to save lives (also, to provide a phenomenological base for future theory). Research in this area is directed mainly to juxtaposition of the strong earthquakes (location or time), with potentially relevant phenomena, in the search for predictors.

The coat of arms for this line of research can be seen in Fig. 2. It is typical of the search for "precursors:, i.e. unusual phenomena before a strong earthquake. Such precursors were widely looked for among very diverse phenomena, from variations of geophysical fields to the behaviour of animals. A special trait of Fig. 1 is that it is obtained from two sets of random numbers. Exploratory data analysis is not only a safeguard against such self-deception. First of all, it allows the extraction of more conclusions from the data available - in muddy situation, when neither models nor statistics are adequate, but the decision has to be made.

Pattern recognition, for example, has allowed to increase by 30 - 70% the success score of diverse decisions: prediction of mineral deposits (ore bodies, ore fields, oil layers); prognosis of the outcome of surgical operations, etc.

Equally important are the rules of recognition, which are found in the process of pattern recognition. They often reveal the important regularities, obscured to our eye by the multi-dimentional nature of the data (on earthquake occurrence; on presidential elections in the U.S., etc.).

However, the results of the exploratory data analysis are never completely proven and they require fanatical testing. This spirit of testing, supported by computer software, will be the subject of the practical exercises.

Such know-how can be learned only by specific examples. The first week of the exercises will be dedicated to one of the premonitory seismicity patterns, and the second week to pattern recognition of the earthquake-prone areas.

Strategy of research and development. The idea of this Workshop was triggered by a real episode. A country decided to build a huge dam for a hydro-electric power plant and needed to study the seismicity around the proposed building site. It started by buying a rather expensive fully-automated seismic network to register microseismicity, and invited foreign experts to run it. The network outlined the strip of micro-earthquakes, i.e. the active fault, crossing the future dam. However, the same fault was outlined - with smaller, but sufficient accuracy - by Gutenberg-Richter in their catalogue of strong earthquakes of the world (1953).

Table 4 gives a brief outline of the data necessary for the first two stages of earthquake prediction. At least for magnitudes 6 or more, these data are already being routinely provided by global and some regional surveys. The auxiliary, more detailed, data can be obtained by a satellite-oriented local network of the new type, which have just appeared.

The most efficient strategy, both in time and expense , is not to repeat the history of geophysics, but to emphasize the analysis of the already available data; to enter the global or regional network; and, if and when necessary, to establish a specially oriented local network, using cheap microprocessors and satellite communications.

A similar strategy is applicable for many other geophysical problems. Example: preliminary reconnaissance of the oil-field-prone depressions, using global seismological observations (Fig. 2).

TABLE 1

GENERAL FRAMEWORK

EARTHQUAKE PREDICTION: INDICATION OF PLACE, TIME AND ENERGY OF FUTURE EARTHQUAKE WITHIN INFORMATIVE LIMITS ; CONSEQUTIVE NARROWING OF THESE LIMITS.

I. DETERMINATION OF EARTHQUAKE-PRONE AREAS. It is equivalent to the estimation of $\bar{E}(\lambda, \varphi)$, $\bar{E}$ - maximal possible energy of earthquakes for $\sim 10^2$ years; λ, φ - geographical coordinates.

ESTIMATION OF TIME AND ENERGY OF APPROACHING EARTHQUAKE IN EACH AREA. SO FAR - IN PROBABILISTIC TERMS, WITH PROBABILITIES OF FALSE ALARMS AND FAILURE-TO-PREDICT COUNTERBALANCED.

II. INITIAL PREDICTION IS POSSIBLE A LONG TIME $T_I(E)$ BEFORE AN EARTHQUAKE. SUCCESSIVE APPROXIMATIONS ARE POSSIBLE, WHILE EARTHQUAKE APPROACHES. CONVENIENT SUBDIVISION:

	INTERVAL BETWEEN THE PREDICTION AND THE EARTHQUAKE	$T_i(\bar{E})$
LONG - TERM	$T_1(E)$	YEARS
MIDDLE - TERM	$T_2 = T_i$: (3 - 10)	YEAR OR LESS
SHORT - TERM	$T_3 = T_i$: (3 - 10)	WEEKS
IMMEDIATE	T_4	DAY TO HOUR OR LESS

These estimations of $T_i(\bar{E})$ are relatively reliable for mountain areas; for others - best guess.

Note: each of the following statements may seem trivial. Attention is drawn to their set and consequences.

THE GOAL: PREDICTION SHOULD BE APPLIED TO REDUCING THE HAZARDS WHICH COME FROM:

COLLAPSE OF STRUCTURES

FIRES

FLOODS

AVALANCHES AND LANDSLIDES

RELEASE OF DANGEROUS MATERIALS - RADIOACTIVE, TOXIC, GENETICALLY ACTIVE, ETC.

TSUNAMIS

SOCIAL DISORDERS

DISRUPTIONS OF VITAL SERVICES - SUPPLY, MEDICAL, ETC.

EPIDEMICS

IMMEDIATE ECONOMIC LOSSES, LOSS OF EMPLOYMENT, ETC.

DISRUPTION OF ECONOMY

THE SECOND GROUP MAY ALSO BE GENERATED BY THE IMPROPER RELEASE OF A TIME-PREDICTION.

TABLE 3

MODES OF APPLICATIONS: (determine requirements to prediction)

I. FOLLOWING SAFETY MEASURES MAY BE INTRODUCED ACCORDING TO THE DETERMINATION OF EARTHQUAKE-PRONE AREAS.

RESTRICTION OF LAND - USE - ESPECIALLY FOR HIGH-RISK OBJECTS AND EARTHQUAKE - INDUCING ACTIVITIES.

REINFORCEMENT OF CONSTRUCTIONS.

TIGHTENING OF SAFETY REGULATIONS AND SERVICES (FIRE, MEDICAL, ETC.

IMPLEMENTATION OF RELEVANT OBSERVATIONS AND DATA - PROCESSING FACILITIES.

INSURANCE AND SPECIAL TAXATION.

PREPARATION OF THE RESPONSE TO TIME-PREDICTION, AND OF POST-DISASTER ACTIVITIES: Planning; legal background; accumulation of the stand-by resources; simulation alarms.

[illegible]

OF [illegible]

EMERGENCY (UP TO MARTIAL) LEGISLATION, TO FACILITATE THE RATIONAL RESPONSE TO PREDICTION.

MANDATORY REGULATION OF ECONOMY.

NEUTRALIZATION OF THE SOURCES OF HIGH RISK: LIFE-LINES; NUCLEAR PLANTS; CHEMICAL PLANTS; UNSAFE BUILDINGS, ETC. UP TO SUSPENSION OF OPERATION, PARTIAL DISASEMBLING, DEMOLITION ...

EVACUATION OF POPULATION AND HIGH-VULNERABLE OBJECTS (E.G. HOSPITALS).

MOBILIZATION OF POST-DISASTER EMERGENCY SERVICES.

PREPARING MEASURES FOR LONG-TERM POST-DISASTER RELIEF (RESTORATION OF JOBS, PRODUCTION, CREDIT, DWELLINGS, ETC.).

MONITORING OF PRECURSORS.

MONITORING OF SOCIO-ECONOMIC CHANGES, AND PREVENTION OF PREDICTION-INDUCED HAZARDS - SPONTANEOUS AND NOT: DISRUPTIVE ANXIETY; ECONOMICAL INSTABILITY, INFLUENCING PRICES, STOCK-MARKET, CREDIT, CURRENCY EXCHANGE, EMPLOYMENT, ETC.

Note: THESE MEASURES SHOULD BE, IN DIFFERENT FORMS, IMPLEMENTED ON INTERNATIONAL, NATIONAL, PROVINCIAL AND LOCAL LEVELS.

Details and references;

Intergovernmental conference on the assessment and mitigation of earthquake risk. Paris 10-19 February 1976. Final Report.

UNESCO SC/MD53, Paris, May 1976.

TABLE 4

MAGNITUDE RANGE	8	6	6	5
SIZE OF THE SOURCE, L km	10^3	10^2	10^1	1
SCALE OF THE MAPS AND SATELLITE PHOTOS (region ~ 10L)				
original	1 : 1 mln	1 : 1 mln	1 : 25 mlns	?
summary	1 : 5 mlns	1 : 2.5 mlns	1 : 1 mln- - .5 mlns	?
SENSITIVITY OF THE NETWORK, M ≥	~4.5	~4	~3	?

CAPTIONS TO FIGURES

Fig. 1. The danger of data-fitting. p - observed phenomena, t - time, vertical lines - the moments of strong earthquake; all three are taken from the table of random numbers.

Fig. 2. The detection of the potentially oil-prone depressions by the routine seismological data (on the dispersion of surface waves) from Levshin, A. and Berteussen, K.-A. Jeophys. J. R. ast. Soc. 1979, 56:97-118.

(a) Group velocities - along the normal shelf (top) and the path, crossing depression, Barents sea.

(b) The structures obtained from the group velocities: normal (dashed line) and in the depression.

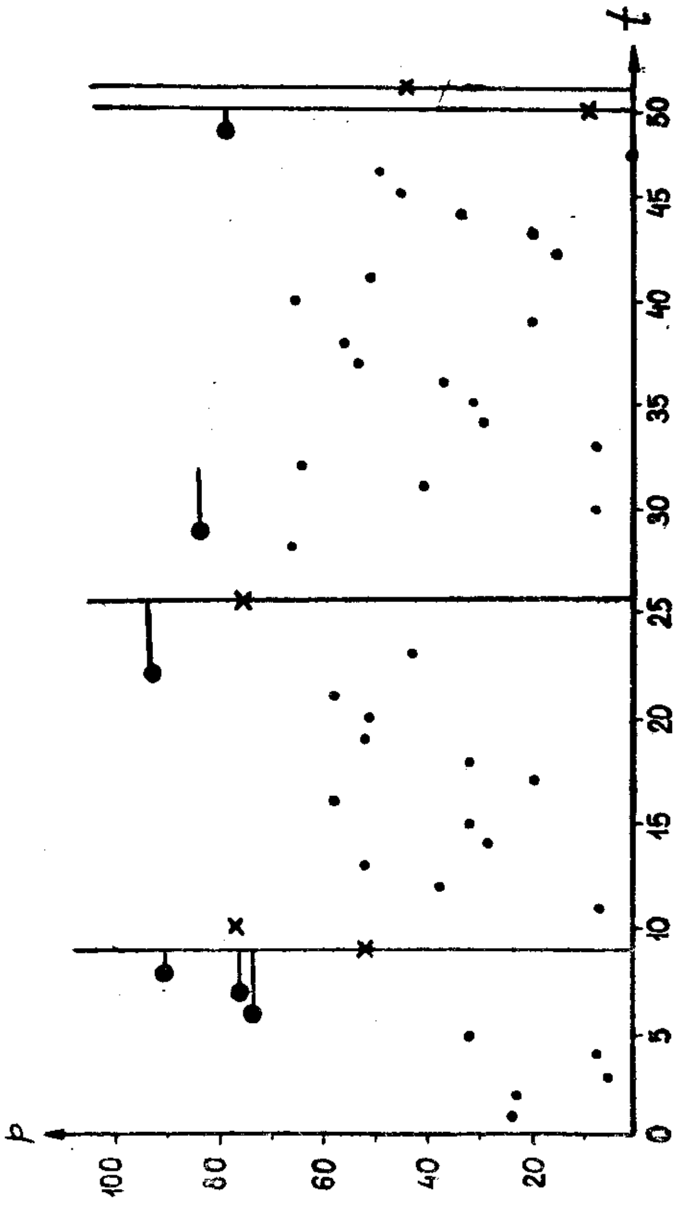

Figure 1

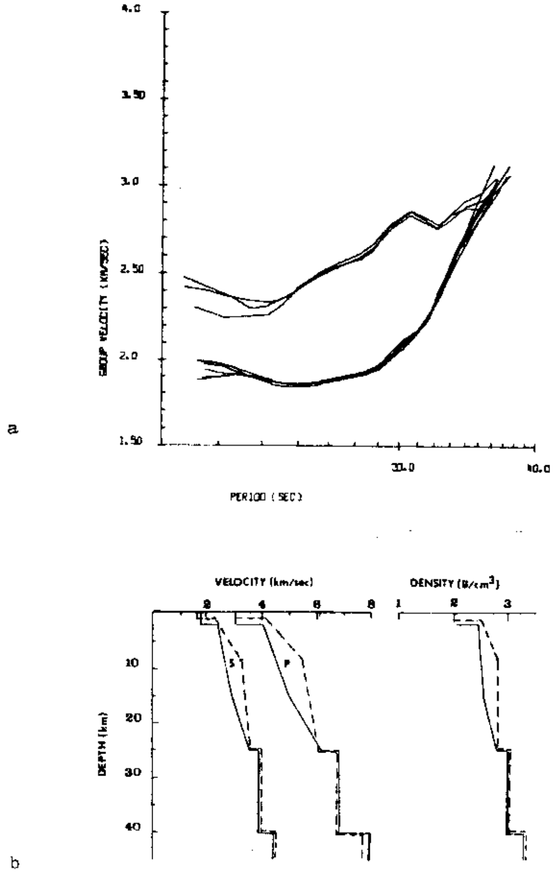

Figure 2.