ΣΥΝΕΔΡΙΑ ΤΗΣ 11ΗΣ ΝΟΕΜΒΡΙΟΥ 1993

ΠΡΟΕΔΡΙΑ ΚΩΝΣΤΑΝΤΙΝΟΥ ΔΕΣΠΟΤΟΠΟΥΛΟΥ

ΣΕΙΣΜΟΛΟΓΙΑ. — **Evaluation of the success of earthquake predictions beyond chance,** by *K. Eftaxias, F. Vallianatos, J. Polygiannakis**, διὰ τοῦ Ἀκαδημαϊκοῦ κ. Καίσαρος Ἀλεξοπούλου.

ABSTRACT

A 3 year continuous sample of official predictions based on the observation of Seismic Electric Signals in Greece was recently published by Varotsos and Lazaridou (1991). Four independent groups, Uyeda (1991), Hamada (1993), Shnirman et al (1993) and Nishizawa et al (1993), have analysed this sample and concluded that the success of the predictions is far beyond chance. However a fifth group, i.e. Mulargia and Gasperini (1992), claim that these predictions can be ascribed to chance. In the present paper we show that Mulargia and Gasperini's procedure: (i) violates general principles to such an extent that it «rejects» the results even of an ideally perfect earthquake prediction method, (ii) it leads to the paradox that the predictions can be more safely ascribed to chance when the ideal prediction method achieves a larger number of excellent predictions (although isolated in time and space), (iii) when applied to an ideally perfect earthquake prediction method it can lead to the wrong conclusion that the true precursors are «post-seismic effects», (iv) exactly for the same set of ideally perfect predictions (issued above a certain magnitude threshold) one can extract contradictory conclusions, i.e. that they either can be ascribed to chance **or** far beyond chance by just selecting different magnitude thresholds for the earthquakes. This lack of self-consistency is mainly due to the following two facts: (a) Mulargia and Gasperini do not select for their study a common magnitude range for the earthquakes and predictions but they take different magnitude threshold for these two sets of experimental data and (b) they use Poisson distribution when both mainshocks and (large number of) aftershocks are

* Κ. ΕΥΤΑΞΙΑΣ, Φ. ΒΑΛΛΙΑΝΑΤΟΣ καὶ Ι. ΠΟΛΥΓΙΑΝΝΑΚΗΣ, Βασικὲς ἀρχὲς γιὰ τὴν ἀξιολόγηση μιᾶς μεθόδου πρόγνωσης σεισμῶν.

involved in the calculation. Furthermore in the specific application of the MG-procedure to the SES predictions, Mulargia and Gasperini, except of the aforementioned points, also confused various kinds of electrical precursors that have different lead-times.

INTRODUCTION

Since a couple of years official predictions are issued by the VAN-group in Greece for earthquakes (EQ) with magnitude (M) larger (or equal to) 5-units that are based on the observation of the so called Seismic Electric Signals (hereafter called SES). They forecast the epicentral location and the magnitude of the impending EQ within a certain time-window Δt. A three year continuous sample of these predictions has been recently published by Varotsos and Lazaridou (1991) evaluated by five independent groups which followed different statistical methods. The conclusions drawn by these groups are as follows:

(i) Hamada (1993). «... for EQs with M_B (USGS) $\geqslant$ 5.0 the ratio of the predicted to the total number of EQs is 6/12 (50%) and the success rate of the prediction is also 6/12 (50%) with a probability gain of a factor of 4. With a *confidence level of 99.8% it is rejected that this success rate can be explained by a random model of EQ-occurrence* taking into account a regional factor which includes high seismicity in the prediction area...».

(ii) Shnirman, Shreider and Dmitrieva (1993). «... According to the test based on the independent earthquake catalogs (NOAA, ESMC), the earthquakes and the VAN prediction telegrams are in obvious correlation if we select both for strongest magnitudes...».

(iii) Nishizawa, Lei and Nagao (1993). «... Results show that the model assuming seismic electric signals as precursors of EQs gives the best fit to the data...».

(iv) Uyeda (1991). «... the actual success rate and alarm rate... are both estimated to be about 60%...».

(v) Mulargia and Gasperini (1992) (MG). «... the claimed success can be very confidently ascribed to chance; VAN-predictions show a much better association with the events which occurred before them...».

It is therefore obvious that the MG-claim contradicts the results of the other four groups the conclusions of which (although following quite different methodologies) practically *coincide*. It is the scope of this paper to examine the origin of this contradiction. It is organized into two parts as follows:

Part I contains three paragraphs that do not deal at all with the specific application of the procedure followed by Mulargia and Gasperini (MG) to the SES predictions but examines the validity of the MG-analysis from a general standpoint. More specifically in the first paragraph the attention of the reader is drawn to the fact that when the MG-procedure is applied to an **Ideally perfect prediction method** (i.e. to a method that by definition achieves the prediction of **all** earthquakes with an excellent accuracy as far as the time, epicentral coordinates and the magnitude are concerned) it leads to the «conclusion» that this IDEAL method should be rejected. In the second paragraph we recall the basic principles of the Poisson distribution and explain with precise examples that the MG-procedure (which is based on the Poisson distribution) cannot be used for the evaluation of an IDEALLY perfect prediction method when both mainshocks and aftershocks are involved in the calculation.

In the latter case we draw the attention of the reader that the MG-procedure can also lead to the wrong conclusion that true precursors are «post-seismic effects».

In the third paragraph we indicate the necessary precautions that should be taken in order to evaluate correctly (in the frame of Poisson distribution) experimental result of an earthquake prediction method.

Part II deals with the MG-application to the SES predictions. It clarifies that MG (except of the aforementioned points) have also misundestood some basic SES physical properties. Part II also shows that the MG-conclusions change dramatically when the points discussed in the Part I of the present paper are considered.

PART I. EXAMINATION OF THE VALIDITY OF THE MG-PROCEDURE FROM A GENERAL STANDPOINT

1. «Conclusion» of the MG-procedure when applied to an ideally perfect prediction method.

In this paragraph we shall «evaluate» an IDEALLY perfect prediction method by employing *different* magnitude thresholds for the EQs and for the predictions respectively in a way exactly similar to that followed by MG. We shall see that the result drastically depends on the thresholds chosen. For the sake of comparison we shall intentionally use numbers comparable to those used by MG, i.e. (almost) the same time-period of observation, the same time - window and comparable number of events.

Case a. Thresholds: M_S (EQ) $\geqslant$ 5.8 and M_S (pred) $\geqslant$ 5.1

Assume a total observation period (T) of approximately 3 years, e.g. 1100 days during which two strong EQs (M = 605) occurred at two remote areas A and B. Except of these two EQs a number (e.g. 18) of independent smaller shocks (NOT aftershocks) with M ranging from 5.1 to 5.5 (see Fig. 1) occurred

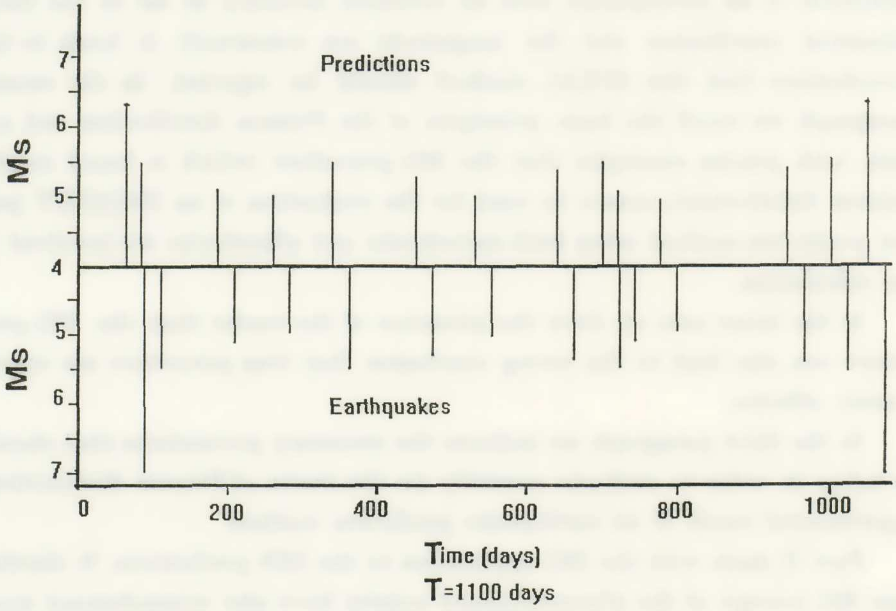

Fig. 1. Example of the results of an Ideally Perfect Earthquake Prediction Method (IPEPM) used to check the correctness of Mulargia and Gasperini's (MG) procedure. It is assumed that within almost 3 years (i.e. T = 1100 days) this IPEPM achieved to issue 20 *excellent* predictions (upper time chart), i.e. r = 0, M = 0, preceding (each one with t s 22 days) 20 independent EQs (lower time chart) with *different* (remote) epicenters and with magnitudes from 5.1 to 6.5. By applying the MG-procedure to this *ideal case* one finds s.l. = 0.2, i.e. that these excellent predictions could be ascribed to chance. The origin of this paradox is discussed in the text.

at various seismic regions. Assume now that for ALL these 20 EQs the IDEALLY Perfect Earthquake Prediction Method (**IPEPM**) achieved to issue predictions with an excellent accuracy in the epicentral location (i.e. $\Delta r = 0$) and in the magnitude determination (i.e. $\Delta M = 0$). The time-window, Δt, is quite small, e.g. it lies between a couple of hours and 22 days (and hence $\Delta t \leqslant$ 22 days).

Let us apply now the *same procedure and thresholds as in the MG-publi-*

cation in order to evaluate the above results. They selected a magnitude threshold of $M_S \geqslant 5.8$ for the EQs (EQs) and a different threshold of $M_S \geqslant 5.1$ for the predictions (pred). With in these thresholds one has 2 EQs and 20 predictions so that the two «mean» probabilities P_{EQS} and P_{pred} are:

$P_{EQS} = 2$ EQs / 1100 days and $P_{pred} = 20$ pred / 1100 days

Therefore the quantity $(\Delta t\ P_{EQS}\ P_{pred}\ T)$ calculated by MG is:

$(\Delta t\ P_{EQS}\ P_{pred} T) = 22$ days . (2EQs/1100 days) . (20 pred/1100 days) .
1100days $= 0.8$

The significance level (s.l.) is calculated from the (upper part of the) cummulative Poisson expression :

$$(\text{S.l.}) = \sum_{x=n}^{\infty} \frac{(\Delta t\ P_{EQS}\ P_{pred}\ T)^{x}}{x\,!}\ e^{-(\Delta t\ P_{EQS}\ P_{pred}\ T)}$$

By inserting into Eq (1) the above value of $(\Delta t\ P_{EQS}\ P_{pred}\ T) = 0.8$, one finds for $n = 2$ a result of around 0.19. For the convenience of the reader we give in Fig. 2 the value of the complicated expression (1) (Abramowitz and Stegun 1970) for various values of n, where n denotes the number of the EQs successfully predicted; the horizontal axis represents the aforementioned quantity $(\Delta t\ P_{EQS}\ P_{pred}\ T)$ and the vertical axis denotes the so called significance level, s.l., which should be smaller than 0.05.

As the above value of (s.l.) $= 0.19$ is much larger than 0.05, the MG-procedure «concludes» that *the excellent predictions depicted in Fig. 1 could be achieved by chance and hence the IDEALLY perfect prediction method* (cf. that it has predicted *all* the 20 EQs —isolated in time and space— with excellent accuracy) is *questionable.*

The paradox of the above result becomes more evident if we consider a larger number of EQ in the same time period e.g. if we assume that in the example depicted in Fig. 1, a *third* EQ with $M \simeq 6.5$ occurred at *another remote* area C and an additional number of 9 (independent) smaller EQs at various areas with M ranging between 5.1 and 5.5. The IDEALLY perfect prediction method would have predicted with excellent accuracy *all these* additional EQs (and hence we have now totally 30 EQs and 30 predictions) but the MG - calculation now gives:

$(\Delta t\ P_{EQS}\ P_{pred}\ T) =$ 22 days . (3EQs / 1100 days) . (30 pred/1100 days) .
. 1100days $= 1.8$

The value —see Fig. 2— corresponds (for $n = 3$) to a significance level =

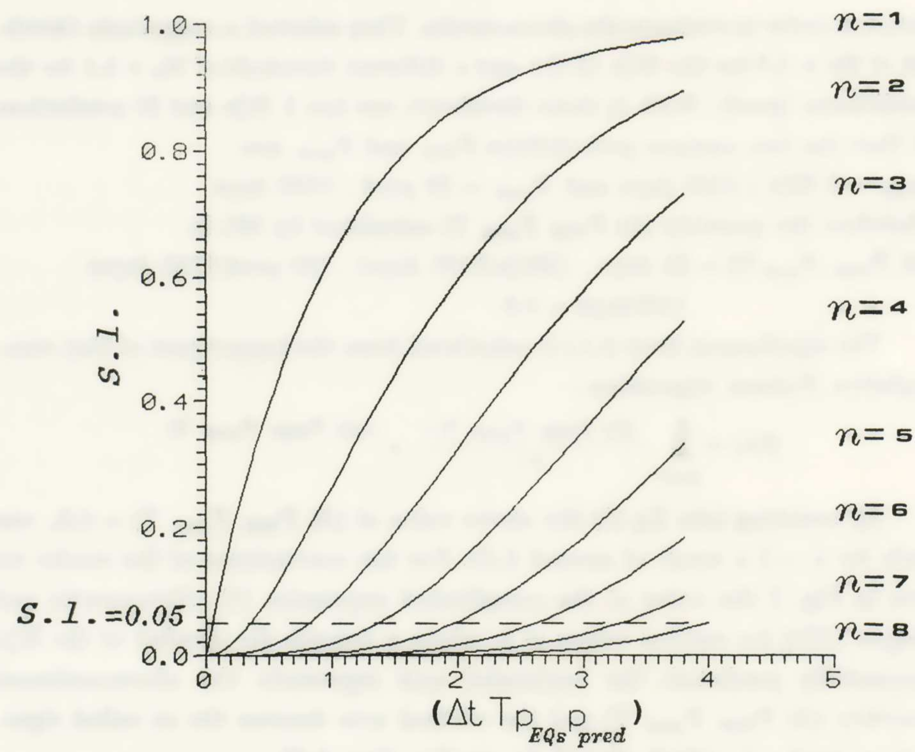

Fig. 2. The significance level (s.l.) as resulting from the calculation of the quantity.

$$(\text{s.l.}) = \sum_{x=n}^{\infty} \frac{(\Delta t\ T\ P_{EQS}\ P_{pred})^{x}}{x!} \exp\ (-\Delta t\ T\ P_{EQS}\ P_{pred})$$

versus the exponent ($\Delta t . T\ P_{EQS}\ P_{pred}$) for various values of n (= number of successful predictions). For the help of the reader the s.l. -value of 0.05 is also depicted (see the text).

0.27 *i.e. even worse than in the previous case.* In other words when *the IDEALLY perfect prediction method achieves a larger number of successful predictions, the MG-calculation leads to the conclusion that the results can be ascribed to chance to a higher degree.*

If we consider a *smaller* number of events than that depicted in Fig. 1, e.g. 14 excellent predictions and 14 EQs (two of them with Ms = 6.5 and 12 with Ms between 5.1 and 5.5), the MG-calculation leads to s.l. = 0.1. Therefore these predictions can be again ascribed to chance but with a *smaller s.l. value* than that resulted from Fig. 1.

Case b : M_S (EQs) $\geqslant$ 5.5 and M_S (pred) $\geqslant$ 5.1

Let us repeat the MG-calculation of s.l. by taking exactly the same data (i.e. those of Fig. 1) as in case a, but now changing only the magnitude threshold of the EQs from 5.8 to 5.5. In this case Fig. 1 shows that the number of EQs with $M_S \geqslant 5.5$ is 6 and hence the exponent becomes:

$$(\Delta t \; P_{EQS} \; P_{pred} \; T) = 22 \text{ days} \; . \; (6\text{EQs}/1100 \text{ days}) \; . \; (20 \text{ pred}/1100 \text{ days}) \; . \; 1100 \text{ days} = 2.4.$$

and then the (s.l.) is calculated to be: (s.l.) = 0.036 which means that the predictions of the EQs depicted in Fig. 1 *CANNOT be attributed to chance.* This result however is in obvious disagreement with that obtained in case (a). In other words when evaluating the IDEALLY perfect prediction method (the forecastings of which are depicted in Fig. 1) we can: (i) either «reject» it (because s.l. $= 0.19 > 0.05$) when we select magnitude thresholds: M_S (EQs) $\geqslant$ 5.8 and M_S (pred) $\geqslant$ 5.1 *or* (ii) to accept it as reliable (i.e. s.l. $= 0.036 < 0.05$) when we select: M_S (EQs) $\geqslant$ 5.5 and M_S (pred) $\geqslant$ 5.1. As in *both* cases (a) and (b) we have analysed the *same set of experimental data*, the contradictory results come from the selection of different magnitude thresholds for EQs and for predictions when checking their association.

Attention is drawn to the point that when we select in the example of Fig. 1 the *same magnitude thresholds* both for EQs and predictions we *always find* a s.l. -value appreciably smaller than 0.05; for example by taking either M_S (EQs) $\geqslant$ 5.8 and M_S (pred) $\geqslant$ 5.8 *or* M_S (EQs) $\geqslant$ 5.5 and M_S (pred) $\geqslant$ 5.5 we find s.l. $\ll 0.01$.

2. Does the MG-procedure follow the restrictions under which Poisson distribution is applicable?

The Poisson distribution gives the probability p(n) that 0, 1, 2, ..., n (n = integer) *unlikely* events will occur in a given period. The formula for the probability p(n) that a *rare* event will occur n times in a given period is:

$$p(n) = \frac{\mu^n}{n!} e^{-\mu}$$

where μ is both the *mean* and the variance of the frequency distribution; the value of μ is estimated from an earlier sample of data by computing X, where X is the average number of times the event occurred in the period of the same length in the past.

Attention is drawn to the point that the Poisson distribution is derived under the fundamental assumption that the events occurred: *independently of each other* (and that the probability did not change with time). It is therefore obvious that in the seismological data, the Poisson distribution can be used ONLY when mainshocks (but NOT aftershocks) are considered in the calculation. In other words, general principles demand that the evaluation of the IPEPM that was able to predict *both, all* the mainshocks and *all* the aftershocks CANNOT be done by using Poisson distribution in the way used by MG.

We proceed below to a few remarks indicating how one can reach non - acceptable results (from physical points of view) when violating the aforementioned general principles emphasized in several text books and in early papers (e.g. Aki, 1956).

(a). Let us take the example depicted in Fig. 3. In this example we assume that within a total period of T = 1100 days two major EQs with $M_S = 6.5$ occurred at two quite different areas A and B and that each of them was followed by number, e.g. 6, aftershocks. We assume that for all these EQs excellent predictions ($\Delta r = 0$, $\Delta M = 0$) have been issued with $\Delta t = 22$ days and that each aftershock occurred 12 days after the previous EQ. By following the MG-calculation of P_{pred} we have the value of $P_{pred} = 14$ pred/1100 days that considers both, the 2 mainshocks and the 12 aftershocks. This number of P_{pred} «means» that during a continuous period of around 3 years, on the average one prediction (for $M \geqslant 5.1$) was issued almost every 2.5 months. This is *not so* because Fig. 3 shows that the rate of issuing predictions was *very large* (i.e. 1 prediction/12 days) only during the two time-periods (each one lasting only 2.5 months approximately) of high seismic activity at the areas A and B respectively *but very small, i.e. zero*, during the intermediate period of around 2.5 years between the two seismic activities. Therefore when considering (a large number of) aftershocks, the so called «mean» probability does not represent the real picture.

(b) When we select the same magnitude threshold for both EQs and predictions, the Poisson distribution should also not be used when we have a significant number of aftershocks. This is obvious from the following simple example: Assume that during a period of T = 1100 days, 5 EQs occurred in two scenarios: (i) one 6.5 mainshock accompanied by 4 smaller aftershocks with M ranging from 5.1 to 5.5 (that occurred e.g. within 2 months after the mainshock), (ii) *five* EQs with M between 5.1 and 6.5 occurred at remote areas and

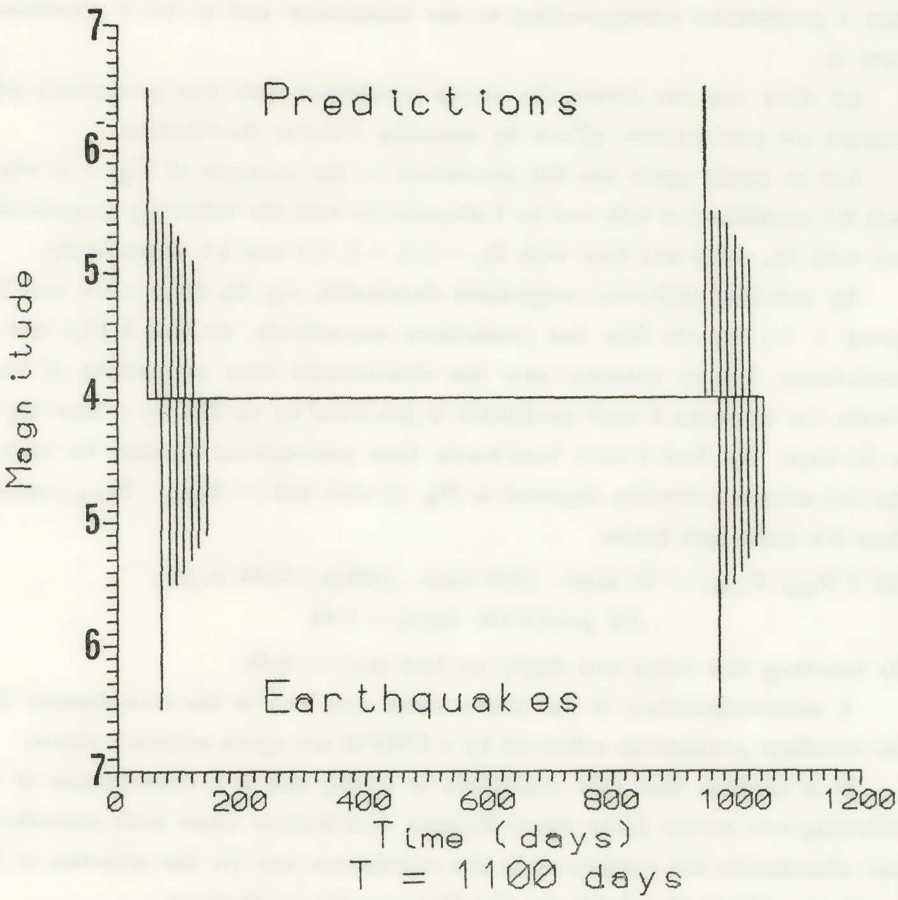

Fig. 3. Another example of the results of an Ideally Perfect Earthquake Prediction Method (IPEPM) in order to check the correctness of MG-procedure. It is assumed that within T = 1100 days the IPEPM achieved fourteen excellent (i.e. $\Delta r = 0$, $\Delta M = 0$) predictions preceding (each one by $\Delta t = 22$ days) fourteen EQs with magnitudes from 5.1 to 6.5. It is assumed that the two 6.5 EQs are *mainshocks* that occurred at two remote seismic areas A and B and that each of them was followed by 6 *aftershocks* (e.g. each aftershock occurs 12 days after the previous EQ).

at different time periods and hence completely independent. If we assume that for all these EQs correct predictions were issued (with the same time-window Δt) MG-procedure gives in *both* cases (i) and ((ii) the values: P_{EQS} = 5 EQs / 1100 days and P_{pred} = 5 pred / 1100 days and hence the *same s.l.-values.* This is not physically acceptable because achieving correctly the predictions of the five independent EQs (i.e. case ii) is appreciably more difficult than 5 predictions corresponding to one mainshock and to the 4 aftershocks (case i).

(c) *How can one derive the wrong conclusion that true precursory phenomena are post-seismic effects by misusing Poisson distribution.*

Let us apply again the MG-procedure to the example of Fig. 3 in which each 6.5 mainshock is followed by 6 aftershocks with the following magnitudes: two with $M_S = 5.5$ and four with $M_S = 5.4$, 5.3, 5.2 and 5.1 respectively.

By selecting *different* magnitude-thresholds, e.g. M_S (EQ) ≥ 5.5 and M_S (pred) ≥ 5.1, for the EQs and predictions respectively we find 6 EQs and 14 predictions. Let us examine now the «backward» time association of these events, i.e. examine if each prediction is *preceded* by an EQ, by a time-lag up to 22 days. We find 8 such backwards time associations (4 cases for each of the two seismic activities depicted in Fig. 3) with ΔM ($= M_{EQS} - M_{pred}$) smaller than 0.4-units and hence:

$$(\Delta t \; T \; P_{EQS} \; P_{pred}) = 22 \text{ days} \,.\, 1100 \text{ days} \,.\, (6\text{EQs} \,/\, 1100 \text{ days}) \,.$$
$$.\, (14 \text{ pred}/1100 \text{ days}) = 1.68$$

By inserting this value into Eq(1) we find (s.l.) << 0.05.

A misinterpetation of the above value can lead to the «conclusion» that the excellent predictions achieved by a IPEPM are «post-seismic» effects.

It is obvious that this conclusion is wrong and is a consequence of the following two errors: (i)the use of Poisson distribution when both mainshocks and aftershocks are considered in the calculation and (ii) the selection of different magnitude thresholds for the EQs and the predictions.

It is clear that when we have an aftershock sequence and a corresponding number of predictions, it is expected to find a «backwards» association with a small s.l.-value depending on the density of the (non-independent) events. This emphasizes again the critical importance of considering independent events when we study the association of two time-series.

3. Necessary precautions when evaluating the results of an EQ-prediction method in the frame of Poisson distribution.

It is far from the scope of this paper to suggest a method for the evaluation of a precursor beyond chance. As already mentioned such detailed methods hence already been suggested e.g. by Hamada (1993). We just indicate that if one wishes to use Poisson distribution to check the true correlation between EQs and predictions, the following precautions should be taken : As a first step we take only statistically *independent* EQs and predictions that belong to a (COMMON) *magnitude range* e.g. for M ranging from 5 to 7. As a second step we investigate which of these events are correlated *both* in time and space (e.g. $\Delta t \leqslant 22$ days and $\Delta r \leqslant 120$ km). Among these correlated «pairs» we select only those that have a magnitude difference ΔM smaller than a certain value, e.g. $\Delta M = 0.7$ which we accept as experimental error. As a third step we calculate the (s.l.) —by means of Eq. 1— that corresponds to the latter number of correlated events (i.e. those correlated in time, space and magnitude) and examine whether the resulting (s.l.) - value is smaller than 0.05.

We should clarify once again that the above precautions are necessary but not be sufficient. The procedure followed by Hamada (1993) —in which he also considers a regional factor including high seismicity in prediction area— is the most complete up to date.

PART II. COMMENTS ON THE SES-EVALUATION BY MG

SES Physical properties misunderstood by MG

Varotsos and Alexopoulos (1984, 1986), Varotsos and Lazaridou (1991) have discriminated two *different* precursory phenomena, i: the Gradual Variation of the Electric Field (GVEF) and the Seismic Electric Signals (SES) as follows :

(i) The GVEF starts *a couple of weeks before EQs with* $M_S \geqslant 5.5$ and lasts almost until the time of the EQ, i.e. it is a gradual variation of the electric field of the earth that has a long duration of the order of 1 month or so. It was emphasized (Varostos and Alexopoulos 1986) that when GVEF is observed then it is always followed by SES at the *same station* with the *same polarity.* (It is therefore abvious that when a prediction is issued on the basic of a GVEF, it is not necessary to issue additional prediction after the SES collection that is recorded later on).

(ii) A SES is a transient change of the electric field of the earth with *small duration* i.e. from 1/2 min up to a couple of hours (cf SES up to 20 hours have been observed). Concerning the time-window Δt one should pay attention to the following distinction emphasized by Varotsos and Lazaridou (1991) :

Single SES: Δt in most cases is smaller than 11 days.

SES electrical activity (i.e. series of SESs within a short time): although the seismic activity may start with small EQs within 11 days, however the first significant EQ (e.g. with M $\simeq$ 5.0) occurs after a few weeks (i.e. $\Delta t = 3 \sim 4$ weeks) after the initiation of the SES activity.

When a prediction is issued by Varotsos and coworkers, they always state on which type of precursor anomaly (GVEF, single SES, SES electrical activity) the estimation of the impending parameters was based on. Therefore for each of these three cases, MG should have considered the appropriate time windows and NOT mix them. As a characteristic example we state the following prediction: On April 26, 1987 a GVEF started at Pirgos station (PIR) and then a prediction was issued clarifying that EQ(s) with $M_S \simeq 5.5$ will occur at an epicentral distance of 50 km from PIR. Actually, almost after one month i.e. on May 29, 1987 (and later on June 10, 1987), seismic activity occurred at a distance only a few tens of kilometers from PIR with $M_S = 5.5$. This obvious successful prediction was misinterpreted by MG in two ways: a) they consider the VAN prediction of April 26, 1987 as unsuccessful (because they put as an upper limit for the corresponding expected EQ the value of 22 days, which is NOT correct for the GVEF) and b) they considered the 5.5 EQs of May 29, 1987 and June 10, 1987 as «missed» EQs as they did not find any prediction 22 days before their occurrence.

It is interesting to note that if the above misinterpetation by MG is corrected then the MG-calculation leads to quite different results (see Appendix 1)·

Theoretical models for the SES generation

As for the MG-claim that there is a lack of a convincing physical explanation for the phenomenon of precursory electric signals reported by Varotsos and coworkers (1984, 1991) we emphasize the following : Varotsos, Alexopoulos and Nomicos (1982) and later Varotsos and Alexopoulos (1986) have published detailed thermodynamical aspects on the possibility of the emission of the so called piezostimulated currents from the focal area. Furthermore other workers, e.g. Dobrovolsky et al (1989), Slifkin (1993), etc., Lazarus (1993), have already published a series of reliable physical models that can explain the SES generation and transmission. These models have been discussed in detail in a recent Conference (Park 1992, Park et al 1993) which also concluded that the SES are actually generated in the earth.

SUMMARY OF RESULTS AND CONCLUSIONS

I. We start from the general concept that in order to judge the correctness of a statistical procedure (which claims that is able to test the effectiveness of a prediction method **beyond chance**), an unambiguous way is the following: to apply this statistical procedure to the results of an Ideally Perfect Earthquake Prediction Method (IPEPM) —which of course does not still exist— and check the results.

In the present paper we have shown that:

a. The MG-procedure when applied to an IPEPM leads to the unacceptable «conclusion» that a series of repeated ideal predictions (i.e. those with $\Delta r = 0$ and $\Delta M = 0$) achieved in a period (e.g. of 3 years) appreciably larger than the time-window Δt (e.g. $\Delta t \leqslant 22$ days) for *various* (independent) seismic regions can be ascribed to chance.

b. If the IPEPM happened to make a larger number of *excellent* predictions (*isolated in time and space*) then the MG-procedure calculates a larger value of the significance level, i.e. it «concludes» the paradox that these predictions could be ascribed to chance to a higher degree.

c. Mulargia and Gasperini do not consider *a common magnitude range (e.g. from 5 to 7-units) for both sets of the earthquakes and the predictions*; as a consequence one finds contradictory results for the *same* set of «experimental data» (resulted from an IPEPM) when we just change the (different) magnitude thresholds for the EQs and predictions. The lack of any self-consistency becomes obvious from the following example discussed in the text: when considering EQs with $M_S \geqslant 5.8$ (with $M_{pred} \geqslant 5.1$) we find that the predictions can be ascribed to chance; on the other hand the consideration of all EQS with $M_S \geqslant 5.5$ (with $M_{pred} \geqslant 5.1$) leads to the *opposite conclusion*, i.e. that the predictions achieved are far beyond chance.

d. The MG-procedure when applied to an IPEPM can lead to the wrong conclusion that true precursors are «post-seismic» effects.

Note that the above points a, b, c and d, *are general considerations and do not depend on the SES predictions.*

II. Concerning the specific application of the MG-procedure to the SES predictions the following three points should be also emphasized:

a. By just repeating the MG-calculation after considering a *common magnitude range* for the EQs and SES predictions we get values of the significance level lying between 0.016 and 0.044 *which show (contrary to the MG-conclusions) that the predictions were issued far beyond chance.*

b. Although using the Poisson distribution, Mulargia and Gasperini consider both the statistically independent mainshocks together with (a large number of) aftershocks that are not independent events.

c. Mulargia and Gasperini confused various kinds of electrical precursors (e.g. GVEF with single SES) that have different time-lags.

APPENDIX 1

5. Results of MG-calculation when considering common magnitude range for earthquakes and SES predictions.

It is far from the scope of this paragraph to suggest a method for the evaluation of the SES predictions (because, as mentioned in the Introduction, it has already been done by four independent groups) *or* to *correct* the MG-procedure (because the latter suffers in general principles). We would like only to repeat here the MG-calculation after taking *common* magnitude range (e.g. from 5.3 to 6.5) for the EQs and the SES predictions in order to show that the resulting s.l. value is *quite different.* Towards this scope we give in Tables 1 and 2 the necessary data for *all* EQs (i.e. 11 EQs) with Ms $\geqslant$ 5.3 and *all* predictions (i.e. 14) also with Ms $\geqslant$ 5.3 that occurred during the period January 1, 1987 to November 30, 1989. In Table 3 we give only those EQs (i.e. 8 EQs) with Ms $\geqslant$ 5.3 which are successfully correlated with predictions (i.e. 8 predictions) with Ms $\geqslant$ 5.3 along with the corresponding deviations Δr and ΔM for each case. (Note that the 5.8 EQ of March 19, 1989 is not included as, during the preseismic period of this EQ, Varotsos and Lazaridou 1991 reported that they were absent due to the participation at the NATO International Conference in France). We recall that the prediction of April 27, 1987 refers to GVEF that is recorded, as clarified above, a couple of weeks before the EQs. By repeating now the MG-calculation we find the values summarized in Table 4. An inspection of this Table leads to the following remarks:

For $\Delta t \leqslant 22$ days, $\Delta r \leqslant 120$ km, the 8 predictions described in Table 3 leads to a s.l. = 0.016. If the prediction accuracy is increased to $\Delta r \leqslant 55$ km (and $\Delta t \leqslant 22$ days) the s.l. turns to 9.044 *Note that in both cases the s.l. - value is smaller than 0.05 and hence it is precluded that the predictions can be ascribed to chance.* We draw the attention of the reader that the omission of

TABLE 1

All EQs with Ms $\geqslant 5.3$ within N_{36}^{41} E_{19}^{25} during the period January 1987 to November 30, 1989

DATE YY MM DD	EPICENTER N E	MAGNITUDE M
87 02 27	38.37 20.42	5.9
87 05 29	37.53 21.60	5.5
87 06 10	37.17 21.46	5.5
87 08 27	38.93 23.81	5.3
88 05 18	38.35 20.47	5.8
88 05 22	38.35 20.54	5.5
88 09 22	37.99 21.11	5.5
88 10 16	37.90 20.97	6.0
89 06 07	37.99 21.65	5.3
89 08 20	37.22 21.08	5.9
89 08 24	37.89 20.11	5.7

* by excluding the EQs in Albania and the EQ on March 19, 1989.

TABLE 2

All predictions with Ms $\geqslant 5.3$; Period January 1, 1987 to November 30, 1989.

DATE YY MM DD		EPICENTER N E	MAGNITUDE M
87 02 26		37.94 20.32	6.5
87 04 27		37.67 21.46	5.5*
88 04 02		36.06 21.39	5.5
88 05 15		37.94 20.32	5.3
88 05 21		37.94 20.32	5.3
88 05 30		37.94 20.32	5.4
88 09 01		37.96 21.01	5.8
	or	39.87 21.01	5.3
88 09 30		37.96 21.01	5.3
88 10 03		37.96 21.01	5.3
88 10 21		37.96 21.01	6.4
	or	40.48 20.40	5.5
89 03 02		37.94 20.32	5.4
89 06 03		37.94 20.32	5.5
89 09 11		37.13 21.24	5.8
	or	38.62 21.73	5.2
89 10 18		37.96 21.01	5.5

* It refers to a GVEF (see text) that is recorded a couple of weeks before an EQ.

TABLE 3

All EQs with Ms ⩾ 5.3 and the corresponding predictions with Ms ⩾ 5.3.

Predictions						EQ's						
DATE			EPICENTER		MAGNITUDE	DATE			EPICENTER		MAGNITUDE	
YY	MM	DD	N	E	M	YY	MM	DD	N	E	M	Δr
87	02	26	37.94	220.32	6.5	87	02	27	38.37	20.42	5.9	48
87	04	27	37.67	21.46	5.5*	87	05	29	37.53	21.60	5.5	20
						87	06	10	37.17	21.46	5.5	55
88	05	15	37.94	20.32	5.3	88	05	18	38.35	20.47	5.8	47
88	05	21	37.94	20.32	5.3	88	05	22	38.35	20.54	5.5	49
88	09	01	37.96	21.01	5.8	88	09	22	37.99	21.11	5.5	10
88	09	30	37.96	21.01	5.3**							
88	10	03	37.96	21.01	5.3**	88	10	16	37.90	20.97	6.0	8
89	06	03	37.94	20.32	5.5	89	06	07	37.99	21.65	5.3	120

* The case of April 2, 1987 was reported as a GVEF (Varotsos and Lazaridou, 1991).

** The predictions of September 30 and October 3, 1988 corespond to SES activities, as reported by Varotsos and Lazaridou, 1991.

TABLE 4

Repetition of the MG calculation of the (s.l.) by considering a common magnitude range e.g. 5.3 ⩽ Ms ⩽ 6.5, both for earthquakes and SES predictions (the data are given in Tables 1, 2, 3).

T (days)	Δt (days)	P_{EQS}	P_{pred}	$(\Delta t . P_{EQS} . P_{pred} . T)$	Δr (Km)	n	(s.l.)
1064	22	11/1064	14/1064	3.1842	120	8	0.016
1064	22	11/1064	14/1064	3.1842	120	6*	0.104
1064	22	11/1064	14/1064	3.1842	55	7	0.044

* If GVEF is excluded.

the prediction on April 27 is critical as it turns the s.l. (see Table 4) at a value larger than 0.05. However we emphasize again that this omission made by MG is not justified in view of the obvious success of this prediction as it was followed by two EQs with excellent accuracy in the magnitude (ΔM = 0) and in the epicentral location (Δr = 20 km and 55 km respectively, see Table 3).

APPENDIX 2

Backwards correlation obtained by MG between VAN-predictions and Earthquakes

As already mentioned, MG claim that: «there is a little doubt that the occurrence of a large event (Ms ⩾ 5.8) has been followed by a VAN prediction with essentially identical epicenter and magnitude with a probability too large to be ascribed to chance».

In the main text we have drawn the attention of the reader to the fact that the MG-procedure when applied to an *ideally perfect earthquake prediction method* it can lead to the wrong conclusion that true precursors are «post-seismic effects». Therefore the answer to the aforementioned MG-claim is obvious. However we proceed to the following clarifications for the help of the reader:

The aforementioned MG-«conclusion» was based *ONLY* on three predictions issued after the following *TWO* EQs : (a) EQ on Oct. 16, 1988 and (b) EQ on Aug. 20, 1989. As the reader may get a wrong impression that the aforementioned two EQs were not preceded (but ONLY followed) by predictions we discuss each case separately:

Case of 6.0 EQ on Oct. 16. 1988. Thirteen days before this EQ, i.e. on Oct. 3, 1988, a prediction was issued (accompanied by a public warning on Oct. 5, 1988) forecasting a 5.3 EQ(s) with epicenter(s) that practically coincided with the actual one of the EQ on Oct. 16, 1988. This prediction however is *NOT* considered in Table 3 of MG although it is reported in their Appendix C that listed our predictions.

Concerning now the prediction of Oct. 21 which is correlated by MG backwards with the EQ of Oct. 16, we emphasize that the text as presented by MG is NOT correct but *differs drastically* than the real one; more precisely the text presented by MG (see Appendix C of their paper) reads: «predicted epicenter 240 km west of other with Ms = 6. 3-6 . 5 (or 400 km NW of Athens with Ms = 5.5)». On the other hand the correct text as published by Varotsos and Lazaridou (1991) reads: «predicted epicenter, several tens of km *away* from W 140 with Ms = 6.3 - 6.5...». It is therefore obvious that the correct text did *not* allow (due to the SES physical properties e.g. polarity, etc.) any correlation of this prediction with an EQ from the already active area lying 240 km west of Athens (this is the one correlation on which MG-conclusion is based). Actually on Nov. 8, 1988 a 5.3 EQ occurred with an epicenter 170 km SW of Athens, i.e. this epicenter was *indeed* several tens of km away from the previous one. Varotsos and Lazaridou (1991) has considered this prediction as unsuccessful due to the large deviation between the predicted magnitude from the actual one.

Case of the 5.9 EQ on August 20, 1989. This EQ was actually *preceded* by a prediction issued on August 16, 1989 that predicted a $M_S = 5.0$ EQ with an epicenter at 200 km WNW of Athens. By disregarding this correlation (may be due to the deviation $\Delta M = 0.9$), MG correlate backwards the prediction of August 24, 1989 with the EQ on August 20. We clarify however that the prediction of August 24, 1989 (epicenter within the zone WNW 190 — WSW 240 with $M_S = 5.2$ to 5.8) was actually followed by a $M_S \simeq 5.0$ EQ with an epicenter 170 km west of Athens i.e. only a few tens of km far from the predicted zone.

As for the prediction issued on Sept. 11, 1989 (the content of which was similar to that of Aug. 24, 1989) it was again correlated «backwards» by MG with the EQ on Aug. 20, 1989 (this is the «third backwards correlation» made by MG). We clarify however that this prediction was actually followed by EQs with Ms between 4.7 and 5.0 with epicenter(s) just in the predicted zone.

ΠΕΡΙΛΗΨΗ

Βασικὲς ἀρχὲς γιὰ τὴν ἀξιολόγηση μιᾶς μεθόδου πρόγνωσης σεισμῶν

Συνεχὲς δεῖγμα ἐπισήμων προγνώσεων, βασιζομένων ἐπὶ σεισμικῶν ἠλεκτρικῶν σημάτων ἐδημοσιεύθη τελευταίως ὑπὸ τῶν Βαρώτσου καὶ Λαζαρίδου (1991). Τέσσαρες ἀνεξάρτητες ὁμάδες (Uyeda (1991), Hamada (1993), Shnirman κ.ἄ. (1993) καὶ Nishizawa κ.ἄ.) ἀνέλυσαν τὸ δεῖγμα καὶ συνεπέραναν ὅτι ἡ ἐπιτυχία εἶναι πολὺ πάνω ἀπὸ τυχαίας προγνώσεις. Ἐν τούτοις μία πέμπτη ὁμὰς (Mulargia καὶ Gasperini 1992) ἰσχυρίζεται ὅτι αὐτὲς οἱ προγνώσεις μπορεῖ νὰ ἀποδοθοῦν στὴν τύχη. Εἰς τὸ παρὸν δημοσίευμα δείχνουμε ὅτι ἡ διεργασία τῶν Mulargia καὶ Gasperini: 1) παραβιάζει γενικὲς ἀρχὲς σὲ τέτοιο βαθμὸ ὥστε νὰ ἀπορρίπτονται ἀκόμη καὶ τὰ ἀποτελέσματα μιᾶς ἰδανικῶς τελείας μεθόδου προγνώσεως σεισμῶν, 2) ὁδηγεῖ εἰς τὸ παράδοξον ὅτι οἱ προγνώσεις νὰ μποροῦν νὰ ἀποδοθοῦν στὴν τύχη διαρκῶς σὲ μεγαλύτερο βαθμό, ὅσον ὁ ἀριθμὸς τῶν ἐπιτυχιῶν τῆς ἰδανικῶς τελείας μεθόδου αὐξάνεται (καίτοι ἀπομονωμένες σὲ χρόνο καὶ χῶρο), 3) ὁδηγεῖ στὸ ἐσφαλμένο συμπέρασμα ὅτι πραγματικὰ πρόδρομα σήματα εἶναι μετασεισμικὰ φαινόμενα ἀκόμη καὶ ὅταν ἐφαρμόζονται σὲ ἰδανικῶς τελεία μέθοδο καὶ 4) ἐπιτρέπει τὴν ἐξαγωγὴ ἀντιφατικῶν συμπερασμάτων διὰ τὸ ἴδιο δεῖγμα ἰδανικῶς ὀρθῶν προγνώσεων (ἐφ' ὅσον περιορίζονται σὲ σεισμοὺς πέραν δεδομένου κατωφλίου) ὥστε νὰ μποροῦν νὰ ἀποδοθοῦν στὴν τύχη ἢ ὄχι δι' ἐκλογῆς διαφόρων κατωφλίων μεγέθους. Αὐτὴ ἡ ἔλλειψη αὐτοσυνεπείας ὀφείλεται στὰ ἑξῆς δύο γεγονότα: α) Οἱ Mulargia καὶ Gasperini ἀντὶ νὰ ἐπιλέξουν διὰ τὴν μελέτη τους μία περιοχὴ μεγεθῶν κοινὴ διὰ σεισμοὺς καὶ προγνώσεις ἐχρησιμοποίησαν διαφορετικὰ κατώφλια διὰ τὶς δύο σειρὲς πειραματικῶν δεδομένων καὶ β) χρησιμοποιοῦν κατανομὴν Poisson προκειμένου διὰ κυρίους σεισμοὺς καὶ πολλαπλοὺς μετασεισμούς. Πέραν αὐτῶν, στὴν ἐφαρμογὴ τῆς διεργασίας στὶς προγνώσεις οἱ M.G. συγχέουν τὰ διάφορα εἴδη ἠλεκτρικῶν προδρόμων σημάτων, τὰ ὁποῖα ἔχουν διαφορετικοὺς προδρόμους χρόνους.