Basic principles for evaluating an earthquake prediction method

P. Varotsos[1], K.Eftaxias[1], F.Vallianatos[2] and M.Lazaridou[1]

Abstract. A three year continuous sample of earthquake predictions based on the observation of Seismic Electric Signals in Greece was published by *Varotsos and Lazaridou* [1991]. Four independent studies analyzed this sample and concluded that the success rate of the predictions is far beyond chance. On the other hand, *Mulargia and Gasperini* [1992] (hereafter cited as MG) claim that these predictions can be ascribed to chance. In the present paper we examine the origin of this disagreement. Several serious problems in the study of MG are pointed out, such as: 1. The probability of a prediction's being successful by chance should be approximately considered as the product of three probabilities, P_T, P_E and P_M, i.e., the probabilities with respect to time, epicenter and magnitude. In spite of their major importance, P_E and P_M were ignored by MG. The incorporation of P_E decreases the probability for chancy success by more than a factor of 10 (when P_E is taken into account it can be shown that the VAN predictions cannot be ascribed to chance). 2. MG grossly overestimated the number of earthquakes that should have been predicted, by taking different thresholds for earthquakes and predictions. With such an overestimation, MG's procedure can "reject" even an ideally perfect earthquake prediction method. 3. MG's procedure did not take into account that the predictions were based on three different types of electrical precursors with different lead-times. 4. MG applied a Poisson distribution to the time series of earthquakes but included a large number of aftershocks. 5. The backward time correlation between predictions and earthquakes claimed by MG is due to misinterpretation of the text of some predictions and an incorrect use of aftershocks. Although even the discussion of the first problem alone is enough to invalidate the claims of MG, we also discuss the other four problems because MG violated some basic principles even in the time domain alone. The results derived in this paper are of general use when examining whether a correlation between earthquakes and various geophysical phenomena is beyond chance or not.

Introduction

During the last 10 years predictions have been issued in an officially documented form by the VAN-group in Greece for earthquakes (EQ) with (expected) magnitude (M_s) larger or equal to 5-units. These predictions were based on the observation of electrical precursors. The most usual form of these precursors are the so-called Seismic Electric Signals (SES) [*Varotsos and Alexopoulos*, 1984a,b,c] the generation of

1. Solid State Section, Department of Physics, University of Athens

2. Technical University of Creta, Chania, Greece.

which can be explained by various physical models [e.g., *Lazarus*, 1993; *Slifkin*, 1993; *Gershenzon and Gokhberg*, 1993; *Varotsos and Alexopoulos*, 1986]. Each prediction forecast the epicentral location and the magnitude of the impending EQ with a certain lead-time Δt; for the two most destructive events even public warning was also made [*Varotsos and Lazaridou*, 1991; *Varotsos et al.*, 1993]. A three year continuous sample of these predictions has been recently published by *Varotsos and Lazaridou* [1991] and evaluated by five independent groups. The conclusions drawn by these groups are as follows:

(i) *Hamada* [1993]: "... With a confidence level of 99.8%, the possibility of this success rate being explained by a random model of EQ-occurrence, taking into account a regional factor which includes high seismicity in the prediction area, can be rejected."

(ii) *Shnirman et al.* [1993]: "... the earthquakes and the VAN prediction telegrams are in obvious correlation if we select both for strongest magnitudes..."

(iii) *Nishizawa et al.* [1993]: "...The results show that SES are not postseismic events... We conclude that there is a causal relationship between SES and EQ in the selected interval of time and selected area of space..."

(iv) *Uyeda* [1991]: "... the actual success rate and alarm rate... are both estimated to be about 60%..."

(v) MG: "...this claimed success can be very confidently ascribed to chance; ... VAN predictions show... a much better association with the events which occurred before them..."

MG's claims disagree with the conclusions of the other four groups. In view of this disagreement, a recent comprehensive review on low frequency electrical precursors [*Park et al.*, 1993] which generally supported VAN-observations stated that no consensus has yet been reached among researchers concerning the SES-success rate beyond chance. We examine the origin of this disagreement below.

Before proceeding further we summarize MG's procedure as follows: The significance level (s.l.), which indicates the risk that one takes when the predictions are regarded as "non-chancy," is expressed by Eq.(1). This equation indicates that s.l. is evaluated on the (upper part of the cumulative) Poisson curve with a mean value of $\mu = N_{EQs} N_{pred} \Delta t/T$ where N_{EQs} and N_{pred} represent the number of earthquakes (EQs) and predictions (pred) respectively, n the number of successful predictions, T the total observation period and Δt the precursory time-lag. The mean value μ can be alternatively written as $\mu = \Delta t \, P_{EQs} P_{pred} T$ where $P_{EQs} = N_{EQs}/T$ and $P_{pred} = N_{pred}/T$.

Therefore, the s.l. is calculated from the usual expression:

$$s.l. = \sum_{r=n}^{\infty} \frac{\mu^r}{r!} e^{-\mu} \qquad (1)$$

which is just Eq.(A.9) of MG. Note that when μ increases the s.l. value also increases (for the same n). (More precisely, s.l. rises from "near zero" to "near unity" in a range of μ centered on about $n-1$ if $n>1$.) When s.l. is larger than 0.05, MG conclude that the predictions can be ascribed to chance. In other words MG estimate the probability of an EQ to occur in the

predicted domain by chance to be equal to

$$N_{pred} \, \Delta t/T. \qquad (2)$$

This approach would be justified in the time domain if applied properly. MG's mistakes are in their wrong use of "the rules of the game" and in the violation of some basic principles, as will be shown in the following.

1. Probability of a prediction's being successful by chance

Varotsos and Alexopoulos [1984a,b,c] (see p.119 of *Varotsos and Alexopoulos* [1984b], and the section entitled "Probability of a prediction being made by chance" of *Varotsos and Alexopoulos* [1984c]) indicated that when EQs occur in various seismic regions, and the predictions determine the time, epicenter and magnitude in advance, the probability of achieving a successful prediction by chance is approximately given by the product of three probabilities (i.e., for time, space and magnitude). For example when we issue predictions within an area of 500 km x 600 km (which is comparable to the area within which the VAN predictions are issued) with an epicentral accuracy Δr=50 km, the probability P_E of predicting the epicenter (of an independent EQ) by chance is approximately given by: $P_E = \pi \, (50 \text{ km})^2 / (500 \text{ km x } 600 \text{ km}) = 2.6 \text{ x } 10^{-2}$ if the seismicity is distributed roughly "homogeneously" (a more rigorous calculation should consider, of course, the inhomogeneity of seismic activity).

In spite of their major importance, P_E and P_M were not considered at all by MG. The importance of P_E has also been recently pointed out by *Takayama* [1993]. For example, for the case of $M_{pred} \geq 5.8$ (and $M_{EQs} \geq 5.1$), $\Delta t \leq 22$ days, $\Delta r \leq 30$ km (Table 1 of MG) Takayama found s.l.=0.020 to 0.021 in contrast to MG's value of 0.859. Similarly, for $M_{pred} \geq 5.3$ (and $M_{EQs} \geq 4.6$) for $\Delta t \leq 22$ days, $\Delta r \leq 30$ km, Takayama obtained s.l.=0.035 to 0.028, whereas MG's value is 1.00. We emphasize that the role of P_E in evaluating the set of predictions published by *Varotsos and Lazaridou* [1991] becomes of primary importance because the predictions for most of the main shocks (e.g., 5.9 EQ on Feb. 27, 1987; 5.8 EQ on May 18, 1988; 6.0 EQ on Oct. 16, 1988, etc.) had small actual Δr (i.e., $\Delta r \approx 30$ km or so).

2. The appropriate number of the earthquakes that should have been predicted

Varotsos and Lazaridou [1991] have clearly stated that predictions are issued only when the expected magnitude is larger than (or equal to) 5.0 - units. They issued in total 29 predictions (with Ms≥5.0) and stated that the large majority of them "...showed differences ΔM between predicted and true magnitudes of between 0 and 0.7-units." However, this does not mean that all 547 EQs with Ms≥4.0 or 204 EQs with Ms≥4.3 should have been predicted as MG stated in their Appendix D and Table 1. The number of EQs increases drastically with decreasing M, following the Gutenberg-Richter relation. According to the detailed analysis of *Hamada* [1993]: "...The higher value of the SRP [success rate of predictions] when M$_B$(USGS) ≥5.0 and the dramatic increase in RPE [the ratio of predicted earthquakes to all earthquakes] from 6.6% to 50%, as magnitude is increased [i.e., from M$_B$≥4.0 to M$_B$≥5.0], again suggest the existence of a physical relationship between the SES signal and subsequent earthquakes..." and "... It should

be noted that there are neither predictions (expected Ms≥5.3) nor EQs (recorded M$_B$≥5.0) for approximately 10 months..."

In other words, MG selected different magnitude thresholds for the predictions and EQs. We shall show in the Appendix that when such a selection of different magnitude ranges for earthquakes and predictions is applied, even an ideally perfect earthquake prediction method can be "rejected."

We emphasize that the same magnitude range (of course appreciably larger than $|\Delta M|$) must be taken for both predictions and EQs. For example, by taking both threshold values Ms≥5.3 (i.e., M$_B$≥5.0 [*Hamada*, 1993]) and the appropriate Δt-value for each prediction (see next subsection) the repetition of their calculation gives, for $\Delta r \leq 120$ km, a s.l.value ≈ 0.04, which is far smaller than MG's. (This value becomes, of course, appreciably smaller if we consider P_E, as discussed above.)

3. MG did not consider the appropriate lead times in their calculation

The predictions in *Varotsos and Lazaridou* [1991] were based on the observation of the following three kinds of electrical precursors: (a) gradual variation of the electric field of the earth (GVEF) for which Δt is of the order of 1 month; (b) single SES with $\Delta t \leq 11$ days: and, (c) SES electrical activities (i.e., many SESs within a relatively small time, e.g., 1 hr) with $\Delta t \leq$ a few weeks (e.g., ≤ 22 days but see also *Varotsos et al.* [1996]).

The announcement of each prediction states the phenomenon on which it was based. Therefore a correct statistical procedure should necessarily consider the appropriate Δt-value for each prediction. On the other hand, all of MG's calculations were made by considering for all predictions the same Δt, i.e., $\Delta t \leq 11$ days or $\Delta t \leq 22$ days. When they consider $\Delta t \leq 11$ days they immediately "reject" a number of successful predictions associated with SES-electrical activities and with GVEF Alternatively when they assume $\Delta t \leq 22$ days they allow for the predictions based on single SESs, which is the majority of the predictions studied by MG, a significantly larger (i.e., by a factor of 2) probability $N_{pred} \, \Delta t/T$ (see Eq.2) that an EQ will occur in the predicted domain "by chance." We present two characteristic examples:

1st example: On April 27, 1987, a prediction was issued stating that a GVEF was detected at Pirgos (PIR) station and hence "EQ(s) with (expected) Ms=5.5 should occur at a distance 50 km from that station" (Table 2 of *Varotsos and Lazaridou* [1991]). Actually 4 weeks later, i.e., on May 29, 1987, a 5.5 EQ occurred just in the predicted area, i.e., a distance only a few tens of kilometers from PIR. The MG-calculation "rejected" this successful prediction because they only considered either $\Delta t \leq 11$ days or $\Delta t \leq 22$ days.

2nd example: In their Table 3 where MG consider $\Delta t \leq 11$ days, the Ms=5.3 successful prediction issued on Sept. 30, 1988 based on an SES electrical activity (Table 1 *of Varotsos and Lazaridou* [1991]) was rejected because the catastrophic 6.0 EQ (with an epicenter just at the predicted "point") occurred 16 days later.

4. Restrictions under which Poisson distribution is applicable

The Poisson distribution is derived under the fundamental constraint that the events occurred independently of each other

Table 1. Evaluation of IPEPM by MG-procedure

Prediction range	Total pred.	Correct pred.	No. of events	μ	s.l.
All	29	29	547	317.3	1.00
M≥5.0	29*	29	204**	118.3	1.00
M≥5.3	14*	14	80**	22.4	0.98
M≥5.5	9*	9	44**	7.9	0.40
M≥5.8	3*	3	19**	1.1	0.11

Rules of the game: Magnitude of predictions ±0.7, Δt≤22 days, Δr≈0 km

* All predictions with Ms≥M, where M denotes the values of the first column. All predictions are assumed correct (3rd column).

** All EQs with Ms ≥ M - 0.7.

The number 547 (first row) refers to the totality of the EQs with Ms≥4.0 used by MG.

and that the probability did not change with time. Therefore the Poisson distribution can be used only when main shocks (but not aftershocks) are considered. However, MG used in the calculation of μ both main shocks and aftershocks, e.g., they used 42 or so aftershocks with Ms≥4.0 that occurred within 3 weeks after the 5.8 main shock of May 18, 1988 and 54 or so aftershocks (M≥4.0) which occurred within 2 weeks after the 6.0 main shock of October 16, 1988, etc. As the restriction of not using aftershocks is widely known in the literature we shall not discuss this point further here.

5. Backwards correlation when aftershocks are considered

MG claim that they "find the existence of a backward time association significantly beyond chance for EQs with Ms≥5.8 and Δt=11 days in the non-homogeneous region taking ΔM=±0.7 for predictions together with Δr≈30 km (s.l.≈0.03) and ΔM=±0.7 from EQs together with Δr≈120 km (s.l.=0.04)."

It is clear that when we have main shocks followed by a number of aftershocks, and we have issued successful predictions both for the main shocks and their aftershocks, a "backwards association" is also good. This however, cannot be misinterpreted as indicating that successful predictions are "post-seismic effects." We now show that MG's claims are invalid for both of the aforementioned cases.

First case: Table 4 of MG indicates that 2 out of the 19 EQs with (actual) M$_{EQs}$≥5.1 were followed by predictions with M_{pred}≥5.8. An inspection of the data shows that the MG s.l.-value of ≈0.03 (Δt≤11 days, Δr≤30 km) is based only on the following two "successful backwards associations" (see also the data included in Table 3 of MG): (a) the prediction of October 21, 1988 with the 6.0 EQ of October 16, 1988 and (b) the prediction of August 24, 1989 with the 5.9 EQ of August 20, 1989. Although both of these "backwards associations" are questionable, it is enough to discuss here only one, e.g. association (a), because when excluding one of them increases MG's s.l.-value of 0.03 obtained for n=2 from Eq.(1) to s.l.=0.2 (>>0.05) for n=1, invalidating the MG-claim.

The text of the prediction of Oct. 21, 1988 presented by MG (see Appendix C of their paper) reads "predicted epicenter 240 km west of Athens with Ms = 6.3 - 6.5 (or 400 km NW of Athens with Ms=5.5)." However the summary of this prediction, given in Varotsos and Lazaridou [1991] reads

"predicted epicenter, several tens of km away from W 240 with Ms = 6.3 - 6.5." (The full text of this prediction can be found in p. 201 of Dologlou [1993] which explains that "a displacement of the epicenter of the event of October 16, 1988 is expected by some tens of kilometers..." due to the fact that "the ratio of the two [SES] components is different"). Thus any association of this prediction with an EQ within Δr≤30 km from the location 240 km west of Athens should be excluded. In spite of this categorical exclusion, MG proceeded to the "successful backwards association" (a) mentioned above. (On November 8, 1988, a 5.3 EQ actually occurred with an epicenter 170 km SW of Athens. This epicenter was indeed several tens of km away from the previous one; Varotsos and Lazaridou [1991] however considered this prediction as unsuccessful due to the large value of ΔM.)

Second case: Table 5 of MG indicates that 3 out of the 18 predictions with M_{pred}≥5.1 correlate "backwards" (within Δr≤120 km, Δt≤11 days) with EQs M$_{EQs}$≥5.8. Their s.l.-value of 0.04 is based on the following three "successful backwards associations" (see also their Table 3): (i) the prediction of October 21, 1988 with the 6.0 EQ of October 16, 1988, (ii) the prediction of May 21, 1988 with the 5.8 EQ of May 18, 1988, and (iii) the prediction of August 24, 1989 with the 5.9 EQ of August 20, 1989.

Although two of these three "backwards associations" can be questioned, it is enough to exclude only one of them because the s.l.-value of 0.04 obtained by MG, for n=3, changes to s.l.≈ 0.2 (>>0.05) for n=2, invalidating MG's-conclusion. The "association" (ii) must be categorically excluded from the calculation of the s.l.-value (by means of Poisson distribution) because the time series in this case involves non-independent events. Note that the main 5.8 EQ of May 18, 1988 was successfully preceded by the 5.3 prediction of May 15 and the 5.5 aftershock of May 22 was also successfully preceded by the 5.3 prediction of May 21; association (ii), in fact, refers to a backward "successful" correlation of the prediction of an aftershock with the main EQ (which is expected if a method, e.g., an IPEPM, issues successful predictions both for the main shocks and their aftershocks). We clarify once again that neither the time series of the EQs nor the time series of the predictions follow Poisson distribution when aftershocks (and their predictions respectively) are involved in the calculation.

Table 2. Evaluation of IPEPM by MG-procedure

EQ. range	Total pred.	Correct pred.	No. of events	μ	s.l.
All	29	29	547	317.3	1.00
M≥5.0	29*	15	31**	18.0	0.79
M≥5.3	29*	11	11**	6.4	0.06
M≥5.5	29*	10	10**	5.8	0.07
M≥5.8	24*	5	5**	2.4	0.10

Rules of the game: Magnitude of earthquakes ±0.7, Δt≤22 days, Δr≈0 km.

* All predictions issued with M_{pred}≥M-0.7, where M denotes the value of the first column. (Recall that it was agreed that predictions are issued only when the expected Ms≥5.0.)

** All EQs with Ms≥M. Note that all EQs with Ms≥5.3 have been successfully predicted by the IPEPM.

The number 547 refers to all EQs with Ms≥4.0 used by MG.

Appendix

Evaluation of an ideally perfect earthquake prediction method by adopting different magnitude thresholds for EQs and predictions

We follow MG's-procedure exactly in order to evaluate the fictitious Ideally Perfect Earthquake Prediction Method (IPEPM) which by definition achieves the prediction of all earthquakes above a certain magnitude (Ms≥5.3) with a reasonable accuracy in time ($\Delta t \leq 22$ days), epicenter ($\Delta r \approx 0$) and magnitude ($|\Delta M| \leq 0.7$). An IPEPM does not issue any false alarms.

In order to evaluate such an IPEPM we construct Tables 1 and 2 exactly in the same way (and under the same headings) as Tables 1 and 2 of MG. Here we feel that some explanation may be helpful for readers on the way that MG's Tables 1 and 2 were constructed (their rules of the game). In Table 1, M in the first column is the predicted magnitude and the number of events in the 4th column is for all EQs with Ms≥M-0.7. In contrast, in Table 2, M in the first column is the actual Ms and "Total pred." in the 2nd column is the number of all predictions issued with $M_{pred} \geq M-0.7$. In order to avoid any misunderstanding we consider exactly the same number of EQs for various magnitude thresholds as in MG and also the same period of almost 3 years, i.e., $T \approx 1100$ days.

According to its definition we assume that the IPEPM correctly predicted all (i.e., 11) EQs with Ms≥5.3; furthermore, we assume that it was agreed that this IPEPM issues predictions only when the expected magnitude is Ms≥5.0. The total number of the (successful) predictions is 29 and naturally a number of them (e.g., 15) correlate with EQs having (actual) Ms≥5.0 and the rest with smaller EQs (i.e., 4.3≤Ms<5.0). Note that the value in the last row of the 2nd column is changed from 5 to 3 in MG's Table 1. This change was made for the same reason as above, i.e., although 5 (successfully predicted) EQs with $M_{EQs} \geq 5.8$ occurred, three of them had an expected magnitude larger than or equal to 5.8 and the other two had an expected magnitude smaller than 5.8 (but, of course, $|\Delta M| \leq 0.7$). The number 24 in Table 2 refers to the number of predictions issued by the IPEPM with $M_{pred} \geq 5.1$.

An inspection of Tables 1 and 2 shows that in all cases s.l.>0.05 and hence the successful predictions of the IPEPM can be ascribed to chance. This result is, of course, not physically acceptable (especially if we assume that, at least, the 11 EQs with Ms≥5.3 were totally independent, i.e., isolated in time and space).

Acknowlegments

References

Dologlou, E., A three year continuous sample of officially documented predictions issued in Greece using the VAN method: 1987-1989, *Tectonophysics*, 224, 189-202, 1993.

Gershenzon, N., and M. Gokhberg, On the origin of electrotelluric disturbances prior to an earthquake in Kalamata, Greece, *Tectonophysics*, 224, 169-174, 1993.

Hamada, K., Statistical evaluation of the SES predictions issued in Greece: alarm and success rates, *Tectonophysics*, 224, 203-210, 1993.

Lazarus, D., Note on a possible origin for seismic electric signals, *Tectonophysics*, 224, 265-267, 1993.

Mulargia, F., and P. Gasperini, Evaluating the statistical validity beyond chance of VAN earthquake precursors, *Geophys. J. Int*, 111, 32-44, 1992.

Nishizawa, O., X. Lei, and T. Nagao, Point Process analysis of Seismic Electric Signals, in "*Electromagnetic Phenomena Related to Earthquake Prediction*", edited by M. Hayakawa and Y. Fujinawa, pp 459-474, TerraPub, Tokyo, 1994.

Park, S., M. Johnston, T. Madden, D. Morgan, and F. Morrison, Electromagnetic precursors to earthquakes in the VLF hand: A review of observations and mechanisms, *Rev. Geophys*, 31, 117-132, 1993.

Shnirman, M., S. Schreider, and O. Dmitrieva, Statistical evaluation of the VAN-predictions issued during the period 1987-1989, *Tectonophysics*, 224, 211-221, 1993.

Slifkin, L., Seismic electric signals displacement of charged dislocations, *Tectonophysics*, 224, 149-152, 1993.

Takayama, H., Comment on "Evaluating the statistical validity beyond chance of VAN earthquake precursors" by Mulargia F. and Gasperini P., *Geophys. J. Int*, 115, 1197-1198, 1993.

Uyeda, S., Greece; A country where earthquake prediction is in practice (in Japanese), *Zisin*, 44, 391-405, 1991.

Varotsos, P., and K. Alexopoulos, Physical properties of the variations of the electric field of the earth preceding earthquakes, I, *Tectonophysics*, 110, 73-98, 1984a

Varotsos, P., and K. Alexopoulos, Physical properties of the variations of the electric field of the earth preceding earthquakes, Determination of epicenter and magnitude, II, *Tectonophysics* 110, 99-125, 1984b.

Varotsos, P., and K. Alexopoulos, Physical properties of the variations of the electric field of the earth preceding earthquake, *Practica Athens Academy 59*, 51-116, 1984c.

Varotsos, P., and K. Alexopoulos, Stimulated current emission in the earth and related geophysical aspects, in *Thermodynamics of Point Defects and their Relation with Bulk Properties*, edited by S. Amelinckx, R. Gevers and J. Nihoul, pp. 136-142, 403-406, 410-412, 417-420, North Holland, Amsterdam, 1986.

Varotsos P., and M. Lazaridou, Latest aspects of earthquake prediction in Greece based on Seismic Electric Signals, *Tectonophysics*, 188, 321-347, 1991.

Varotsos, P., K. Alexopoulos, and M. Lazaridou, Latest aspects of earthquake prediction in Greece based on seismic electric signals, II, *Tectonophysics*, 224, 1-37, 1993.

Varotsos P., K. Eftaxias, M. Lazaridou, G. Antonopoulos, and J. Makris, Summary of the five Principles suggested by Varotsos et al. [1996] and the additional questions raised in this debate, *Geophys. Res. Lett*, this issue, 1996.

P. Varotsos, K. Eftaxias and M. Lazaridou, Solid State Section, Department of Physics, University of Athens, Knossou Str. 36, Ano Glyfada 16561, Athens-Greece (e-mail: pvaro@leon.nrcps.ariadne-t.gr)

F. Vallianatos, Technical University of Creta, Chania, Greece.