Identification and discrimination of transient electrical earthquake precursors: fact, fiction and some possibilities

Andreas Tzanis a, *, Filippos Vallianatos b, Sylvie Gruszow c, 1

a Department of Geophysics — Geothermy, University of Athens, Zografou 157 84, Greece
b Technological Educational Institute of Crete, Chania Branch, Crete, Greece
c Laboratoire de Geomagnetisme et Paleomagnetisme, Institut de Physique du Globe de Paris, Paris Cedex, France

Received 7 May 1999; received in revised form 13 June 2000; accepted 16 June 2000

Abstract

The possibility of electrical earthquake precursors (EEPs) has long been appreciated, but to date there still exists neither a solid theory to describe their generation and expected characteristics, nor proven techniques to identify and discriminate true precursors from noise. Experimental studies have produced a prolific variety of signal shape, complexity and duration, but no explanation for the apparently indefinite diversity. Statistical analyses on the basis of such poorly constrained data were inconclusive, leading to scepticism and intense debate. The most objective means of EEP identification would be to construct generic models of their source(s) and compare the model predictions with field observations. We attempt to show the merits of this approach with two studies. The first study expands on the phenomena of spontaneous electric field generation during crack propagation (microfracturing), demonstrated by laboratory experiments. Large-scale microfracturing may appear at the terminal stage of earthquake preparation. We apply a generic, qualitative model, based on a kinetic theory of crack interaction and propagation. The model suggests that EEP signals from such a type of source may have a limited class of permissible waveforms, with characteristic bay- or bell-shaped curves of variable width and duration. We provide two examples consistent with this model: the VAN claims of precursors on 15/1/1983 and 18/1/1983. The magnetic field that may accompany an anomalous electric signal is the subject of the second study. This has been a grossly overlooked quantity, although valuable for identification and discrimination, because it is considerably less sensitive to distortion than the electric field, less sensitive to inhomogeneities along the propagation path, insensitive to the local geoelectric structure and sometimes, telltale of the source (for instance, external magnetic fields can only be generated by (sub)horizontal current configurations). We investigate the 18/4/1995 and 19/4/1995 electric and magnetic signals observed at Ioannina, Greece, used for the prediction of the 13/5/1995 M6.6 Kozani event by the VAN group. The electric and magnetic waveforms are inconsistent with the crack propagation model. By their observed characteristics, the magnetic signals preclude any (sub)vertical electrokinetic current. Using analytic formulations, we investigate whether they might have been generated by an electrokinetic source across a lateral interface, either at the focal area or locally, at Ioannina. We conclude that the magnetic field properties are also inconsistent with such a type of source. Conversely, we cannot rule out their local industrial origin. The examples presented herein indicate
that the successful identification and discrimination of EEP and noise may be possible by working out plausible theories of
the source. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Earthquake prediction; Earthquake precursors; Seismic electric signals

1. Introduction

The generation of transient electric potential during mechanical loading prior to rupture has been demonstrated in a number of laboratory experiments. Electrification by microfracturing, i.e. the appearance of spontaneous charge production and transient electric and electromagnetic (E-EM) emission associated with the opening and propagation of microcracks, has been discussed by several authors in connection with laboratory experiments. Some authors have also provided estimates of charge production rates and currents. For instance, Warwick et al. (1982) have measured current spikes from individual microcracks or the order $10^{-3}$ A, associated with crack opening times of the order of $10^{-6}$ s, thus providing a net charge density of $10^{-3}$ C/m$^2$. A similar value of $10^{-2}$ C/m$^2$ is reported by Ogawa et al. (1985), while Enomoto and Hashimoto (1990) measured a charge production of $10^{-9}$ C for cracks with surface of the order of $10^{-6}$ m$^2$, thus yielding a charge density of $10^{-3}$ C/m$^2$. More recent experiments (e.g. Fiffolt et al., 1993; Chen et al., 1994; Enomoto et al., 1994; Hadjicontis and Mavromatou, 1996; Yoshida et al., 1997) observe simultaneous acoustic and E-EM signals, confirming that electrification effects arise during microfracturing. Finally, Bella et al. (1994) observe simultaneous acoustic and E-EM under real world conditions in caves. Piezoelectricity has been shown to electrify quartz-bearing rocks (e.g. Nitsan, 1977; Warwick et al., 1982; Yoshida et al., 1997). Inasmuch as electrification has been observed in non-piezoelectric materials, additional mechanisms have also been considered. Contact or separation electrification is discussed by Ogawa et al. (1985). The motion of charged dislocations (MCDs) has been investigated both for the case of elastic rock deformation (e.g. Slifkin, 1993; Hadjicontis and Mavromatou, 1996), and for the case of non-elastic deformation, when dislocations move and pileup to form and propagate cracks (e.g. Ernst et al., 1993; Vallianatos and Tzanis, 1998, 1999a). Cress et al. (1987) also suggest that the ionisation of the void space within the crack and the acceleration of unbounded electrons may intensify charge production. Freund and Borucki (1999) demonstrate the existence of positive-hole dormant charge carriers in quartz-free or low-quartz rocks, that can be activated by low velocity impacts and suggest that similar activation may take place by the acoustic waves or direct impulse during crack propagation.

The electrokinetic effect (EKE), i.e. electrification due to the flow of water driven through permeable rock by crustal strain or gravity, has long ago been demonstrated by laboratory experiments (e.g. Morgan et al., 1989 and references therein; Jouniaux and Pozzi, 1995, 1997, etc.). The EKE is consistent with the wet models of the earthquake preparation process (for instance, the dilatancy-diffusion model (Scholz, 1990)). Consequently, it is a frequently quoted mechanism of precursory electric and magnetic fields (e.g. Mizutani et al., 1976; Fitterman, 1979; Dobrovolsky et al., 1989; Bernard, 1992; Fenoglio et al., 1995; Molchanov, 1999, and many others).

Other mechanisms rigorously promoted as an explanation of electrical earthquake precursor (EEP) phenomena have actually never been verified by laboratory experiments. The piezostimulated depolarisation current (Varotsos and Alexopoulos, 1986) requires the polarisation of point defects of the form ‘anion+cation vacancy’, by some external electric field. The polarised defects can change their orientation through jumps of the neighbouring cations to the vacancy with a relaxation time $\tau=\tau_0 \exp(g_m/kT)$, where $g_m$ is the Gibbs’ energy for the migration process, $k$ the Boltzmann’s constant and where $T$ is the temperature. A massive change in point defect orientation is expected to stimulate a macroscopic, short-lived current. The relaxation time decreases exponentially with temperature, but in order for it to decrease with pressure $P$ at a constant temperature, one requires that the migration volume $v_m=\partial g_m/\partial P<0$, a property hitherto never observed in lithospheric materials.
Moreover, the origin of the external electric field with duration considerably longer than \( \rho \) required to polarise the point defects is rather obscure, at least in the case of non-piezoelectric materials. Varotsos et al. (1999a) presented a theoretical refinement, but the concept is still unverified. Teisseyre (1997) proposed that the MCDs provide the electric field by which to polarise the point defects and the rapid stress drop during crack opening allows for the generation of a depolarisation current that enhances or reduces the electric field of the MCD. This is the \textit{piezo-stimulated dilatancy current}, but there is also no indication that this mechanism is realisable.

Although electrification is clearly observed in controlled experiments, the scaling up from laboratory size specimens to the enormous and heterogeneous rock volumes involved in the preparation of earthquakes is an altogether different and as yet unsolved problem. In addition, the earthquake source may host a number of different electrification phenomena, whose spatial and temporal sequence is not clear and which may be synergistic or competitive in a complex manner that is inadequately understood. Accordingly, the nature and properties of possible EEPs are still poorly understood, to the point that their very existence can be debated. Hitherto theoretical attempts to address this problem were usually generically associated with a particular electrification mechanism and different source geometries and propagation/decay laws (e.g. Dobrovolsky et al., 1989; Bernard, 1992; Slifkin, 1993; Varotsos et al., 1993; Molchanov and Hayakawa, 1998; V allianatos and Tzanis, 1998, 1999a). The difficulty in understanding what produces the EEP and how it reaches the observer raises a more important question: how can we tell what is an EEP from what is not?

The only group to have attempted a systematic resolution of this question was the VAN team in Greece, who have devised a set of ad hoc empirical rules to distinguish between local and remote sources of signals (Varotsos and Lazaridou, 1991; Varotsos et al., 1993). The VAN criteria have never been tested for their performance and limitations. As Nagao et al. (1996) state, “the physical meaning of these rules is straightforward,” and as such, they have been accepted by many researchers of electrical precursory phenomena (Park, 1996; Uyeshima et al., 1998, and others). A critical appraisal by Tzanis and Gruszow (1998) shows that the criteria may be misleading due to the strong dependence of the electric field on the geoelectric structure, both local and along the propagation path: noise may be identified as distant signal and vice versa. When local sources are concerned, the criteria can only recognise known emitters; new or shifting ones can deceive them as above. It is apparent, however, that the VAN and any similar criteria, irrespective of their performance, do not provide any means of recognising a genuine EEP save for the \textit{abstract} assumption that if the signal is not local, then it should have been generated by some distant earthquake source.

The validation (or refutation) of anomalous electric signals as EEP has also been tried with statistics, both by the VAN group and a significant number of other researchers. Specifically, an attempt was made to establish a ‘beyond chance’ association between presumed ‘seismic electric signal’ (SES) and earthquake activity. However, in the absence of well-constrained experimental data and an articulate and testable theory of the observable quantities, the use (or misuse) of statistics may lead to such an intense debate and controversy, as that appearing in the GRL Special Issue on VAN (Volume 23, No. 11, May 1996), but will not provide definite answers.

Experimental studies on EEP have produced a prolific variety of signal forms, as any survey of the international literature will show. There have been reported ELF-ULF pulses of variable duration with shapes like spikes, delta functions or boxcars, designated as ‘single’ SES (e.g. Varotsos and Lazaridou, 1991; Kawase et al., 1993; Maron et al., 1993; Varotsos et al., 1993; V allianatos and Tzanis, 1998, 1999a); ULF transient, bay-like variations with durations of a few minutes to a few hours, also classified as single SES (e.g. Varotsos and Alexopoulos, 1984a,b; Varotsos and Lazaridou, 1991; Maron et al., 1993; Varotsos et al., 1993; V allianatos and Tzanis, 1996a); ULF multiple pulses of variable duration and shapes as above, appearing either discretely in time, or in a cascade succession, designated as \textit{SES activities} (e.g. Varotsos et al., 1996a); long-lasting variations (e.g. Sobolev, 1975; Sobolev et al., 1986), including the gradual variations of the electric field (e.g. Meyer and Pirjola, 1988; Varotsos et al., 1993). A variety of additional forms have also been reported, but will not be reviewed for the sake of brevity. Let alone the
difficulty in explaining the origin of the signal, the seemingly limitless diversity of forms and shapes is confusing.

The absence of a magnetic field was considered one of the most salient features of EEP and the SES in particular. For instance, Varotsos and Alexopoulos (1984a) stated that “no significant variation is produced by the signal,” although they did not possess the appropriate instrumentation to detect it, while Park (1996) asserts that “…the mechanism generating the SES does not result in observable magnetic fields regardless of location,” albeit the ‘mechanism’ was unknown and the statement was made on the basis of limited and indirect evidence. Yet, the possibility of magnetic fields had already been established (e.g. from lateral EKEs, Fitterman, 1979; Fenoglio et al., 1995), while there was compelling experimental evidence for ULF magnetic fields of lithospheric origin (e.g. Fraser-Smith et al., 1990; Kopytenko et al., 1993; Kawate et al., 1998; Dea and Boerner, 1999; Hayakawa et al., 1999), which should have electric counterparts, either primary or induced. In principle, the lithospheric EM data could be distinguished from fields of different origin (e.g. Pilipenko et al., 1999).

Unfortunately, in most of the cases cited above, simultaneous electric and magnetic field measurements were not made. This ‘(mal)practice’ used to be quite common, as also noted by Johnston (1997). Although experimental facilities improve, there is still limited experimental information about the possible magnetic companions of EEP. This is regrettable, given the progress made towards defining the received characteristics and relationship between seismogenic ULF electric and magnetic fields (e.g. Molchanov et al., 1995) and the fact that the presence and properties of quasi-static magnetic fields may provide valuable information about the nature of the source, in some cases more significant than the electric field (also see Section 3.1).

It is apparent that a great deal more is required before one can decide on the nature of some anomalous electric or magnetic field variation. To this effect, the most solid approach would be to expand our understanding of the processes that we are trying to detect, i.e. build appropriate physical models for the generation and propagation of EEP and simulate their received characteristics. The comparison and possible agreement of theory with observations may provide a basis for the recognition of some classes of EEP. The shape of the EEP waveform enters these considerations naturally: it may be the principal means by which to authenticate a signal (for instance, like seismologists can tell an earthquake from a quarry blast). Herein, we explore a few ideas about what should be the shape and duration of EEP signals generated during the nucleation and propagation of cracks. We conclude that the very nature of microfracturing allows for a limited class of signal waveforms with bay-like shapes. Our model successfully describes a number of observed signals (the claims of precursors on 15/1/1983 and 18/1/1983, by Varotsos and Alexopoulos (1984a)), and if indeed it is a valid description of natural processes, it may facilitate the identification of some types of EEP signals.

EEP signals may also be generated by electrokinetic mechanisms. Apparently, if the pressure difference driving the EKE results from the stress drop accompanying crack propagation, then the shape of the signal should be determined by the evolution of the microfracturing process. It is also reasonable to assume that, within a complex system stressed to the limit of failure, there may appear pressure differentials of origin other than microfracturing, with arbitrary time functions. It has also been proposed that EKE may be triggered remotely, by transfer of stress from the earthquake focus (Dobrovolsky et al., 1989; Bernard, 1992). Such signals may have arbitrary shapes and their identification is an altogether different problem.

In view of the fact that EK fields are dc or quasi-static, some problems of evaluating electric signals with arbitrary shape can be made by considering the properties of their companion magnetic signals (if any). The final part of this paper is an attempt to appraise the usefulness of magnetic fields in the analysis and identification of anomalous transient variations, by studying the electric and magnetic signals received on 18 and 19 April 1995, prior to the 15 May 1995, M6.6 Kozani–Grevena earthquake. These signals were the basis of a formal prediction statement (see Varotsos et al., 1996a), and in addition to demonstrating the merits of simultaneous electric and magnetic field measurements, our analysis comprises an a posteriori appraisal of the prediction and shows the degree of scrutiny required in evaluating the observations.
2. Transient signals associated with formation and propagation of cracks

2.1. Feasibility of long-distance electric and magnetic fields due to crack propagation

Crack propagation is inherent to brittle failure, while crack dynamics and interactions comprise the basis of all theories attempting to describe the processes leading to rupture. In earthquake seismology, the precipitous increase of crack production shortly prior to rupture is predicted by volume dilatancy models (e.g. Myachkin et al., 1975; Scholz, 1990), damage mechanics (e.g. Voight, 1989) and the critical point earthquake rupture model (e.g. Sornette and Sornette, 1990; Sornette and Sammis, 1995). The initiation and duration of dynamic crack propagation in large heterogeneous rock volumes depends on the mechanical and thermal history and the present state of the stressed materials, and may vary between places even within the same seismogenic volume. Nevertheless, all the theories and models of precursory phenomena indicate that stress and strain accumulation should become non-linear near the end of the loading cycle, producing greatly accelerated effects in the last 1 to several days prior to rupture (e.g. Stuart, 1988; Varnes, 1989; Voight, 1989; Scholz, 1990; Sornette and Sornette, 1990; Sornette and Sammis, 1995). By all accounts, if any transient electric precursor is generated during crack propagation, it should appear a few hours to several days prior to the nucleation of the earthquake.

The electrification of individual cracks is very short-lived. For common petrogenetic mineral and rock resistivities ($\rho$) and dielectric permittivities ($\varepsilon_d$), any charge and electromagnetic fluctuations with source dimension $|\approx|10^{-4}-10^{-1}$ m will disappear after a time $\varepsilon_d \rho \approx 10^{-5}-10^{-7}$ s (if no external sources are applied). This is comparable to the duration of crack opening ($10^{-4}-10^{-7}$ s). Charge production inside the crack is quickly destroyed by redistribution of the displacement currents and the electric polarisation emerges only while the crack is opening. If any long-lasting EEP is to be observed, it will have to be generated by the superposition of the signals from a large number of consonant, simultaneously propagating cracks and evolve in time just like the crack propagation process. However, when cracks begin to propagate, they are expected to do so in unison responding to the same average stress field, and therefore, with the same average geometry. All proposed electrification mechanisms indicate that the field generated during crack opening is dipole in nature and almost all concepts of EEP generation involving the propagation of cracks expand on this premise (e.g. Slifkin, 1993; Yoshida et al., 1997; Molchanov and Hayakawa, 1998; Vallianatos and Tzanis, 1998, 1999a). It follows that the electric field resulting from the superposition of a large number of consonant cracks is expected to have the same dipole nature and average geometry of the fields of individual cracks.

At a point $r$ and time $t_j$, the measured electric field may be qualitatively expressed as

$$E_c(r, t_j) \approx \sum_{i=1}^{n(t_j)} E_i(r, r_i) \cdot \left[ u(t_j) - u\left( t_j - \frac{l_i}{v_i} \right) \right] \quad (1)$$

where $n(t_j)$ is the number of cracks propagating at time $t_j$, $E_i(r, r_i)$ the dipole field due to a crack of length $l_i$ opening at point $r_i$ with velocity $v_i$ and where $u(t_j)$ is the Heaviside step function, so that the right-hand factor in the sum allows the $i$th crack to contribute only as long as it is opening.

The collective effect of many small, distributed, simultaneous cracks can be a very efficient EEP source. For an estimate of the actual current through an individual crack, we rely on published laboratory results and assume that $I \approx 10^{-3}$ A, as measured by Warwick et al. (1982). The typical lengths of microcracks are of the order of $10^{-4}-10^{-1}$ m, so let us assume a mean length $l=10^{-3}$ m. Then, a representative crack dipole moment would be $I=10^{-6}$ A m. Let this be horizontal and buried in a 100 $\Omega$ m half-space, at the depth of 5000 m. Then, at a period of 1000 s and distance 100 km perpendicularly to the dipole axis, the total horizontal electric and magnetic fields are, respectively, $E_h=2.5 \times 10^{-20}$ V/m and $B_h=3.2 \times 10^{-23}$ T, using the full analytic solution of King et al. (1992, pp. 155–159). Now, consider that the maximum number of cracks containable in a unit volume is controlled by their size. Gershenzon et al. (1989) provide the relationship $N_{MAX}=(3-t_c \cdot v)^{-3}$, where $t_c$ is the crack opening time. Assuming $v=10^2$ m/s (constant and comparable to that of Rayleigh waves), we find $t_c=10^{-6}$ s and $N_{MAX}=3.7 \times 10^7$ m$^{-3}$. If $N_{MAX}$ millimetric-size cracks are simultaneously excited in a
volume \( V \sim 10^7 \) m\(^3\), then at a distance of 100 km from the source, the received horizontal electric and magnetic fields will be \( E_c \sim N_{\text{MAX}} V E_h \sim 9.1 \times 10^{-6} \) V/m and \( B_c \sim N_{\text{MAX}} V B_h \sim 1.2 \times 10^{-8} \) T. Such amplitudes are observable with standard equipment and require the maximum coherent excitation of an effective volume with dimensions 215 m \( \times \) 215 m \( \times \) 215 m. The estimate of \( N \) is a determinative factor. For instance, if \( N \sim 10^8 \) and all other parameters are kept constant, the dimensions of the effective volume \( V \) will rise to 1 km\(^3\). We do not know what a realistic \( N \) may be and the order of magnitude variations should be expected within the earthquake preparation zone. Accordingly, the effective volume will be larger, or the signal weaker. It follows that, in order to observe long-distance EEP, we require the excitation of distributed cells with kilometric-order dimensions. Let \( l_0 \) (in kilometres) be a characteristic dimension for the strongly deforming volume \( V_0 \), which is related to the moment magnitude as \( \log(l_0^3) = M_w - 4 \) for a 3 MPa stress drop (Kanamori and Anderson, 1976). For \( M_w = 6.5 \), we find \( l_0 \approx 18 \) km and the minimum \( V_0 \) is approximately \( 18^3 = 5800 \) km\(^3\), i.e. \( 10^3 \sim 10^6 \) times larger than the elementary volume sufficient to provide a 9 mV/km/12 nT ULF variation, even if only a part of it is activated. Note, however, that in cases of low deformation rates (insufficient microfracturing), small excited volumes, or small earthquakes, signals may not be observed even at close ranges, because they may be very weak to detect.

2.2. Dynamics of crack propagation processes

It is now accepted that brittle failure (fracturing, fragmentation and rupture) is self-similar with respect to its geometry and critical point with respect to its dynamics (e.g. Sornette and Sornette, 1990; Turcotte, 1997 and references therein). It begins at the microscopic scale and cascades to the macroscopic by co-operative crack growth and coalescence in such a way, that fracturing at one scale (or level of the crack hierarchy) is part of the damage accumulation at a larger scale. Once microfracturing begins, the number of propagating cracks (and the electric field) is first expected to rise sharply, but as the sustainable crack density is approached or stress/strain levels drop below a threshold value, it will decelerate and decline to 0 when no more cracks can be produced. The duration of this process is unknown, but conceivably, it may require any time up to a few hours, depending on the size, mechanical and thermal state of the deforming volume. In the following, we will attempt to construct an expression for the time function of the EEP source, which will be consistent with the phenomenology of brittle fracture, while accommodating a wide spectrum of allowed signal durations.

A simple approach towards modelling this sequence of events is to consider a point process, i.e. the evolution of cracks in terms of their number \( n(t) \) and irrespective of their size, which can be expressed as

\[
\dot{n} = \frac{dn(t)}{dt} = f(t, n, p_i), \quad t > 0, i = 1, 2, 3, \ldots \tag{2}
\]

where the parameters \( p_i \) represent the factors controlling the nucleation and propagation of cracks (material properties, stresses, temperature, interaction probabilities etc.). Note that \( \dot{n} \) is identical to \( n(t) \) in (1). The function \( f \) characterises a very complex system, but for the moment, we assume that it depends only on \( n(t) \) through a feedback mechanism, and \( n(t) \) in turn is controlled by all other parameters of the system. Expanding \( f \) in a Taylor series of \( n \) and neglecting higher-order terms, we can write (2) as

\[
\dot{n} = n(t)(a_0 - b_0 \cdot n(t)), \quad t > 0 \tag{3}
\]

where \( a_0 \) and \( b_0 \) are positive constants, respectively representing the gains (production rate) and losses (reduction rate) of cracks. It can be shown analytically that \( n(t) \) and \( \dot{n} \) behave according to the initial value \( n(t=0) = N_0 \). If \( N_0 < a_0 / b_0 \) (gains higher than losses), \( n(t) \) initially increases with time and then levels off, approaching a constant value when no more cracks can be produced (Fig. 1a). Respectively, \( \dot{n} \), which controls the amplitude of \( E_c(t) \), increases, maximises and decreases to 0 when no more cracks are produced (Fig. 1b). If \( N_0 > a_0 / b_0 \) (losses higher than gains), \( n(t) \) decays towards a constant value; new cracks are not formed, but instead, their numbers are reduced. This is possible (e.g. by healing), but does not apply to the case of pre-seismic microfracturing we consider herein, at least to this extent. Eq. (3) and Fig. 1b indicate the general form of the expected time functions \( \dot{n}(t) \), which should be variants of a bell-shaped curve. In the following, we shall attempt to address this problem in more detail.
Very few attempts to study the dynamics of crack propagation from first principles have appeared in the international literature. A kinetic approach to the solution of this difficult problem has been attempted by only a handful of researchers. Petrov et al. (1970) proposed a theory which for its complexity could only be solved under drastic assumptions leading to a restrictive model: microcracks nucleate but cannot propagate, unless they join with freshly nucleated cracks in their region of mutual instability. In another example, Newman and Knopof (1983) have only considered a predator–prey form of coupled non-linear differential equations representing the two end members of the crack hierarchy, which is inadequate for our purposes. Molchanov and Hayakawa (1994) develop a simpler model requiring an increase in crack nucleation until, after a critical density, the production of new cracks declines while existing cracks grow by extension of their lengths. Molchanov and Hayakawa (1998) present an improved version, in which cracks are normally distributed in their size-space and undergo multistage redistribution with size development according to a sub-critical stress corrosion process. This model reasonably describes the source spectrum of purported ULF precursors, but does not allow for the interactive evolution of hierarchical crack populations, which is intrinsic to microfracturing processes.

Czechowski (1991, 1995) has developed a more complete and comprehensive theory of crack dynamics, which expands on assumptions similar to those of Boltzman’s and amounts to the kinetic equation

\[
\frac{\partial f(x, l, t)}{\partial t} + \frac{\partial [vpf(x, l, t)]}{\partial l} = \frac{1}{2} \int_{0}^{l} f(x, l_1, t)[f(x, l - l_1, t) - f(x, l, t)] \, dl_1 - f(x, l, t) \int_{0}^{\infty} f(x, l_1, t) \, dV \, dl_1 + N(l) \quad (4)
\]

where \(f(x, l, t)\) is a size distribution function of cracks defined so, that \(f(x, l, t) \Delta x \Delta l\) is the number of cracks at a time \(t\) within a volume element \(\Delta x\) around a point \(x\) and has sizes within \(\Delta l\) around size \(l\), \(p\) is the probability and \(v\) is the velocity of propagation. The
The left-hand side of Eq. (4) expresses the changes in the number and size of cracks, as resulting from the interactions described by the right-hand side. Specifically, the first term on the right-hand side of (4) is the total number of ‘gains’, i.e. the number of binary interactions whereby cracks of (smaller) sizes \( l_1 < l \) collide and merge with cracks \( l \rightarrow l_1 \) to produce cracks with sizes \( l \), where \( \sigma = (l(l_1, \sigma) \) is the cross-section of collisions, \( \sigma \) an average stress field and where the factor \((1/2)\) prevents from counting an interaction twice. The second term on the right-hand side is the number of ‘losses’, i.e. the number of binary interactions whereby cracks of any size \( l_1 \) forming a beam with flux density \( \lim_{l \rightarrow l_1} \) collide with crack \( l \) and consume it. \( N(i) \) is the nucleation term. The kinetic equation describes how cracks propagate and join each other with probability depending on the total cross-section of collisions between cracks. The quantities \( s, \nu, \) and \( \rho \) are functions of crack size, stress field and rock properties. We are particularly interested in an analysis that discretizes (4) in the size-space of cracks, according to

\[
n_i(t) = \int_{L_{i-1}}^{L_i} f(l, t) \, dl
\]

so that the total number of cracks is divided into \( m \) populations \( n_i, i = 1, 2, \ldots, m \) with respect to their size. The case \( m = 10 \) has been studied in Czechowski (1995), subject to the constraints \( 0 = L_0 < L_1 < \ldots < L_9 < L_{10} = \infty \) and \( L_i - L_{i-1} = 1 \) for \( i = 1, 2, \ldots, 9 \). Successive integrations of (4) over the intervals \( (0, L_1], (0, L_2], \ldots, (L_9, \infty) \), produces a set of ordinary differential equations:

\[
\begin{align*}
  n_1 & = 0.5(1-k_1)n_1^2 - n_1N + n_1 \\
n_2 & = 0.5k_1n_1^2 - n_1n_2 - n_2N + n_2 \\
n_3 & = k_2n_1n_2 - (1-k_2)n_1n_2 - n_2N + n_2 \\
n_4 & = k_3n_1n_3 + 0.5k_{1n}n_3^2 + (1-k_3)n_1n_3 + (1-k_3)n_1n_2 - n_3N + n_4
\end{align*}
\]

Equations (5) describe the balance of gains and losses of any given group of cracks by merging \( (n_i n_j \) denotes the fusion of cracks \( n_i \) with cracks \( n_j \)) and by propagation \( n_i \) of cracks with velocity \( v \) and probability \( p \), where \( n_i = (i; v_i p_i)^{-1} \frac{dN}{dt} \), and \( N = \sum_{i=1}^{10} n_i \) is the total number of cracks:

\[
n_i = \frac{P_i}{p_i} \int [f(L_{i-1}, t) - f(L_i, t)]
\]

is the propagation term, and the factors \( k_j, j = 1, 2, \ldots, 25 \) enter into the estimations of integrals of the type

\[
I = \int_{L_i}^{L_{i+1}} f(l) \int_{L_{k+1}}^{L_{k+1}} f(l_1) \, dl_1 \, dl
\]

and determine the span of interactions between any two crack populations, while the factors \( (1-k_j) \) represent the extent of losses due to healing. For a decreasing \( f(l) \), \( 0 < k_j < 1/2 \), with \( k_j = 1/2 \) for \( f(l) \) constant.

Assuming a constant production rate for the smallest crack population, in Fig. 2, we essentially reproduce the result of Czechowski (1995, Figure 11.7.2a), but we also include the total number of cracks present and propagating at any time. Observe that the successive crack populations appear with a power law delay such that the total number of cracks exhibits a sharp rise followed by asymptotic convergence towards a constant value, as the crack density approaches saturation. The complete time function may be approximated with a step function of the form

\[
n(t) = N_0(1 - e^{-\alpha t})
\]
Since only the active (propagating) cracks are electric field sources, their time function should be
\[ \dot{n}(t) = N_0 \gamma \frac{t^\gamma - 1}{\gamma} e^{- \gamma(t)^\gamma} \]

(7)

where \( \alpha \) is a characteristic relaxation time and the exponent \( \gamma \) determines the shape. Note that (6) is in reality a Weibull cumulative distribution function and (7) the corresponding probability density function. Alternatively, an empirical description can be adopted, using a half-step function such as is the error function (for \( t>0 \)) for the rise time of the source, and assuming an exponential decay:
\[ \dot{n}(t) = \text{erf}((At)^\beta) e^{-\gamma(At)^\gamma} u(t) \]

(8)

where \( u(t) \) is the Heaviside step function with \( u(t)=1 \) for \( t>0 \) and \( u(t)=0 \) for \( t<0 \), assuring causality. The constant \( \beta \) determines the slope of the rise time and \( A \) is a characteristic time of the accelerating crack production; they both should depend on material properties. Examples of (8) for different parameters \( A, \alpha \) and \( \beta \) are shown in Fig. 3; these are characteristic shapes expected from both of the generically related functions (7) and (8). Variations of crack counts with a bell-shaped envelope have often been seen prior to rupture, in recent experiments involving large rock samples (e.g. Ponomarev et al., 1997; Feng and Seto, 1998, 1999; Baddari et al., 1999). Although much work is still needed in order to define the details, it appears that (7) and (8) may comprise a representative phenomenological description of crack propagation processes over a wide spectrum of time scales.

2.2.1. The received electric signal

The electric signal generated by microfracturing electrification processes will result from the convolution of the source time function \( \dot{n} \) and the resultant \( E_c(t) \) of the ensemble of the simultaneously electrified cracks. From (1), the time constant of \( E_c \) is comparable to the crack opening times \( t_c \) (\( \sim 10^{-7}-10^{-4} \) s). This implies that \( E_c(t) \) may be approximated by a Dirac-\( \delta \) type of function, with a flat spectrum up to a corner frequency in the VLF-HF bands. It is therefore expected that, when the source time function is much slower, for instance in the ULF band, its waveform
will predominate and determine the waveform of the resulting EEP. In consequence, only the long periods of the electric field are allowed to propagate. This ‘natural selection’ process has an advantage in that the long periods propagate farther with less attenuation, and they have sufficiently long wavelengths, so as to experience less distortion, either by mutual interaction, or by small-scale geological structures.

2.3. Two examples from the January 1983, M7 Kefallinia earthquake sequence (Ionian Sea)

One of the largest events to have occurred in the Ionian Sea region this century, this earthquake occurred offshore to the SW of Kefallinia island, Greece, at 12.41 GMT on 17 January 1983, at co-ordinates 38.09°N, 20.19°E and a focal depth of 9 km (see Baker et al. (1997) for a review). Varotsos and Alexopoulos (1984a) claim to have recorded an electrical precursor to this earthquake at their PIR station, approximately 130 km SE of the epicentre, which they illustrate in Fig. 7 of their paper (see Figs. 4 and 8). We have reproduced a digital version of the longer periods of the signal by scanning their Fig. 7, enhancing the image and digitising it on a high-resolution monitor. The digitised raw signal comprises a transient beginning at approximately 14.00 h on 15 January 1983 and lasting for 1.5–2 h, superimposed on a non-linear variation of the background (Fig. 5). On removing the background, we obtain a very strong E–W component (25 mV over 50 m), but a very weak N–S (Fig. 5 bottom). The E–W waveform has an asymmetric bell shape, with faster rise time and a slower, exponential-type decay; for most of its duration, it stands clearly above noise, the peak amplitude of which is approximately 20% of the peak signal amplitude. The later times of the signal, however, are obscured and there is no real way of telling the exact duration of the decay phase. The long period E–W components can be easily fitted with functions of the forms (7) and (8). Recall that both functions are phenomenological descriptions of the signal shape only, since we cannot as yet estimate the
Fig. 4. Epicentres and Harvard CMT focal mechanisms of the 17/1/1983 M7 and the 19/1/1983 M5.6 Kefallinia events. PIR is the location where the ‘precursory’ electrical signals were detected.

Fig. 5. The upper panel shows the digitised signal recorded at 14.00 GMT on 15 January 1983 at Pyrgos, Greece, and reported by Varotsos and Alexopoulos (1984a) as a precursor to the 17 January 1983 Kefallinia earthquake (Δ=130 km). The lower panel shows the transient signal after removing the background. Hour 0 in the time axis corresponds to 13.00 GMT.
amplitude. Therefore, we may only attempt to fit and
the signal and the model normalised with respect to
their maximum values. In Fig. 6, we present a model
based on Eq. (8), with \( \gamma = 1 \) (fixed), \( A \approx 5.3 \times 10^{-4} \, \text{s}^{-1} \),
\( \beta \approx 2.1 \) and \( \alpha \approx 9.9 \times 10^{-4} \, \text{s}^{-1} \) (\( 2\pi/\alpha \) is approx-
imately the duration of the model and \( 1/\alpha \) is a char-
acteristic relaxation time).

A large \( M = 5.6 \) aftershock of this event occurred
at 00.02 GMT on 19 January 1983, at 38.11°N and
20.25°E (Fig. 4). Varotsos and Alexopoulos (1984a)
again claim to have recorded a precursor at PIR, which
they illustrate in Fig. 8 of their paper. This signal was
also reproduced digitally. The E–W component is
shown in Fig. 7a and b (broken line, after removing a
linear trend). Again, it comprises an asymmetric-bell
shaped variation with very fast rise time and a slower
exponential decay, beginning at approximately 14.30
GMT on 18 January 1983 and lasting for almost
50 min. The solid line in Fig. 7b is a model based
on Eq. (8), with \( \gamma = 1 \), \( A \approx \alpha \approx 3.15 \times 10^{-3} \, \text{s}^{-1} \), and
\( \beta \approx 0.74 \); here as well, \( 2\pi/\alpha \approx 1990 \, \text{s} \) (55 min)
is approximately the duration of the signal and
model.

It is important to note that both these signals be-
long to the small ensemble of transients used by
Varotsos and Alexopoulos (1984a) to construct their
amplitude–magnitude empirical scaling law, of the
form \( \log(\Delta V) = cM + d \), with \( c = 0.3–0.4 \) a uni-
versal value. A number of authors have independently
argued, or shown that this law derives from the fund-
amental fractal scaling of the electric field sources
(Sornette and Sornette, 1990; Molchanov, 1999; Val-
lianatos and Tzanis, 1999b). Such properties are
not likely to have been generated by anthropogenic
noise and indicate that both signals may be a real,
long-range EEP.

2.4. A brief discussion

If the electrification by microfracturing observed in
laboratory experiments can scale up, it may comprise
a very efficient source of EEP. We have attempted to
model such a process, using an approach consistent with the phenomenology of brittle failure. We have concluded that (micro)fracturing is likely to evolve with a bay- or bell-shaped time function, which may be asymmetric and skewed towards the early time of the process. Due to the very short time constant of individual crack electrification, any resulting macroscopic electric field should have a time function almost identical with the microfracturing process. The model successfully applies to a certain class of short-term transient signals, but addresses only a part of the complex processes leading to earthquake nucleation. Moreover, microfracturing electrification may be inhibited by the factors outlined below.

A determinative parameter is the magnitude of $n(t)$, which must be large enough to form a macroscopic field. Consider, however, that, in fault systems controlled mainly by friction, brittle behaviour is limited and may be insufficient to build up an observable signal. Conversely, strong signals may be expected during large-scale brittle deformation, which is more likely to occur in compressive or transpressive stress regimes (as for instance in the area of Kefallinia). This hypothesis may need to be investigated in more detail. The dependence on rock resistivity is also very important. The discharge constant $\varepsilon_d \rho^{-1} \approx 10^{-5} - 10^{-7}$ s quoted in Section 2.1 is calculated for resistivities of the order of $10^6 - 10^4 \Omega \text{m}$. If the resistivity in the neighbourhood of the crack decreases, charge redistribution will occur much faster than crack opening and a macroscopic field will not build up, unless the number of cracks increases by a forbiddingly large factor. This is consistent with the majority of laboratory experiments observing electric signals in dry (i.e. resistive) rock samples. Thus, it must be emphasised that this type of EEP should not always be expected prior to an earthquake. EKEs may step into action in the case of wet rocks (e.g. Fenoglio et al., 1995; Yoshida et al., 1998), but as will be discussed shortly, this process should be detectable only at short to intermediate ranges. A
A lot of work is required to clarify such problems. It is clear, however, that some observed signals can indeed be described with generic theories of the source.

3. Appraising the usefulness of the magnetic field

The properties of magnetic fields from dc sources and their propagation in a conducting earth medium are well known from the theory of the Magneto-Metric Resistivity exploration method (see Edwards and Nabighian, 1991 and references therein). Only a summary is provided here for the benefit of our discussion.

External to the earth magnetic fields can only be generated by subsurface dc current configurations with a significant horizontal component. Vertically directed dc current distributions do not generate external magnetic fields. Point dc sources and horizontal current sheets such as might flow in thin layers above a resistive basement may only produce horizontal external fields. It follows that the presence of an external magnetic field can be diagnostic of its source. Approximately similar conditions apply in the quasi-static case, where the lateral wave produced at the surface generates non-zero external magnetic fields at non-zero frequencies, but these are extremely weak and undetectable at intermediate to long distances from the source (for a comprehensive account see King et al. (1992)). Furthermore, the surface magnetic field from any type of buried quasi-static source is independent of the geoelectric structure in a homogeneous or layered half-space (e.g. Stefanescu, 1929), and therefore, unaffected by overlying formations like the water table and clay beds (that may strongly attenuate the electric field). In a laterally inhomogeneous (or anisotropic) structure, an anomalous magnetic field is generated, proportional to the reflection coefficient (or to the anisotropy factor) across an interface. The distortion may be a few to several tens percent, but it certainly cannot be as large as the corresponding electric field variations across high resistivity contrasts.

It appears that the magnetic field can be a reliable indicator of the characteristics of the subsurface current distribution, and in some cases, provide more significant information than the electric field, both with respect to the nature and the distance of the source. In the following, we will attempt to demonstrate the potential usefulness of magnetic field measurements in identifying the origin of anomalous electric signals.

3.1. ULF electric and magnetic field observations at Ioannina, Greece

The area of Ioannina is located in a NW–SE trending basin in NW Greece, embedded in a Neogene neritic limestone environment that has been extensively karstified (Fig. 8). A telluric station (JAN E) has been installed near the village of Likotrikion, 4.5 km from the VAN station IOA and in a similar geological context. The electric field sensors comprised two sets of grounded lines with non-polarisable Pb/PbCl₂ electrodes, in a NS–EW configuration. The sampling rate was 20 s. Magnetic field measurements were carried out with a high sensitivity (0.25 nT), observatory-type fluxgate variometer, installed at the Ioannina airport (JAN M, in Fig. 8). Details are given in Gruszow et al. (1995).

On 18 and 19 April 1995, anomalous transient electric and magnetic field variations were simultaneously recorded at JAN E and JAN M. These signals were also recorded at IOA and comprised the basis of a formal prediction statement by VAN (for details, see Varotsos et al., 1996a). Two possible epicentral areas were proposed, one near Kyllini — Vartholomio (West-Peloponnesus, see Figs. 4 and 8) and the other at a few tens of kilometres NW of Ioannina. Following the Kozani–Grevena earthquake in Macedonia, N. Greece (08.47 GMT, 13 May 1995, Mₛ=6.6, M_w=6.5, φ=40.16°N, λ=21.67°E, Δ~90 km NE of Ioannina, H≈14 km), the VAN group declared that their prediction was to be related to this event (e.g. Varotsos et al., 1996a).

Fig. 9a shows five components (two electric, three magnetic) of the anomalous signal recorded on 18 April 1995 at JAN E and JAN, while Fig. 9b presents a 24 min-long portion of the recording, after removing all spectral components with a period longer than 5 min. The anomalous electric variation stands clearly above noise, but the magnetic signal is rather weak, with a mean amplitude of 0.8 nT in Z, 0.4–0.5 nT in D (E–W) and practically absent from H (the N–S component). The electric and magnetic fields are coherent, while the vertical and E–W magnetic components vary exactly out-phase and the amplitude ratio |Z|:|D|≈2:1.
These signals and their relationship with the prediction were scrutinised by Gruszow et al. (1996), who argued that their shapes, durations, geographical distribution, polarisations and magnetic field amplitudes favour a local industrial origin. However, they could not locate the source. This interpretation was rigorously contested by Varotsos et al. (1996b, 1999b), mainly on the basis of the magnetic companion. Pham et al. (1999) attributed the signals to a transmitter located somewhere due north of JAN E, but they could also not pinpoint the source. Herein, we attempt an in-depth analysis of the magnetic companion, in order to assess whether the signals may reasonably be attributed to some natural electrification mechanism, or they are more likely an industrial artefact.

### 3.2. Inquiring the natural origin of the magnetic field observed at JAN M

To begin with, all electrification mechanisms involving crack propagation in the Kozani–Grevena seismogenic zone may be ruled out on the basis of the signal waveform(s), according to the model developed in Section 2. Piezostimulated depolarisation currents have never been produced in lithospheric materials, and for this reason we are inclined to reject interpretations involving this mechanism. Likewise, we consider unlikely the piezostimulated dilatancy current mechanism (Teisseyre, 1997), because it involves crack propagation in which case the waveform is also constrained according to the arguments developed in Section 2. The only remaining possibility is an intermittent electrokinetic mechanism. The most apparent source of EK fields is the seismogenic volume in itself. However, under a wide range of conditions, the amplitude of EK fields is rather weak and may not be detectable at intermediate to long distances from the source (Fitterman, 1979; Dobrovolsky et al., 1989). In remedy of this problem, it was proposed that an electrokinetic field might be generated near the observer by stress/strain propagation from the earthquake zone, either directly (Dobrovolsky et al., 1989), or indirectly by triggering pressure instability in small-scale metastable or
Fig. 9. (a) Five components (two horizontal electric, three magnetic) of the anomalous signal of 18 April 1995, as recorded in JAN E and JAN M. (b) A 24 min-long portion of the signal in Fig. 9a, after removing the spectral components with a period longer than 5 min.
unstable fluid reservoirs (Bernard, 1992). Another possibility is the appearance of local electrokinetic phenomena unrelated to tectonic processes. In the following, we investigate these possibilities.

Electrokinetic phenomena do not generate a magnetic field observable at the surface of a homogeneous or layered half-space, irrespective of the spatial or temporal distribution of excess pore fluid pressure (Fitterman, 1979; Dobrovolsky et al., 1989). In consequence, we can exclude from further consideration all phenomena involving (sub)vertical water flow. Non-zero external magnetic fields can arise only in the case of pressure differences across a lateral inhomogeneity of the electrokinetic coupling coefficient (EKCC) and are also influenced by coincident changes in rock conductivity. Under steady-state conditions, the problem is mathematically equivalent to the electrostatic problem of dipolar surface charge at the boundary between two regions. Analytic solutions for the electric and magnetic fields from this type of source were produced by Fitterman (1979, 1981). This problem admits an equivalent solution by replacing the surface charge with a distribution of parallel physical electric dipoles, subject to the constraint $\Delta V = \Delta C = \Delta P = II_L \rho / A$, where $\Delta V$ is the potential difference, $\Delta C$ the EKCC contrast, $\Delta P$ the stress differential, $A$ the total area of the source region (the interface), $I$ the total current across it and $L$ lengths. In this case, the observed electric and magnetic fields results from the superposition of the fields of the distributed dipoles. This approach was also adopted by Fitterman (1981) and has certain advantages, as will be seen below.

Consider the right-handed co-ordinate system ($x$ — top, $y$ — right, $z$ — down) and let a horizontal grounded dipole source of current moment $IL$ be buried in a conductive half-space at the point $(0, 0, -z')$, looking at the positive-$y$ direction. As Edwards and Nabighian (1991) indicate, at a point $(x, y, z)$, the observed magnetic field comprises the superposition of the magnetic field $B^d$ due to the current flowing in the dipole:

$$B^d_x = \frac{\mu IL}{4\pi} \left[ \frac{z - z'}{r^3} \right], \quad B^d_y = 0,$$

$$B^d_z = -\frac{\mu IL}{4\pi} \cdot \frac{x}{r^3}$$  \hspace{1cm} (9a)

and the magnetic field $B^d$ due to current flow in the conductive half-space:

$$B^s_x = \frac{\mu IL}{4\pi} \left\{ \left( \frac{x^2 - y^2}{t^4} \right) + \left( \frac{y^2(z + z')}{t^2s^3} \right) \right\},$$

$$B^s_y = \frac{\mu IL}{4\pi} \left\{ \left( \frac{1 - z + z'}{s^2} \right) \left( \frac{2xy}{t^4} \right) - \left( \frac{xy(z + z')}{t^2s^3} \right) \right\},$$

$$B^s_z = 0$$  \hspace{1cm} (9b)

where $r^2 = x^2 + y^2 + (z - z')^2$, $t^2 = x^2 + y^2$ and $s^2 = t^2 + (z + z')^2$. The surface magnetic fields are as given by (9) even if the dipole is embedded in a layered medium; this point is clearly illustrated by Howland-Rose et al. (1980), with an analogue model experiment. Equations (9a) show that only $B_x$ and $B_z$ components are generated directly by the source, while current flow in the ground will produce $B_y$ but no $B_z$. Fitterman (1979, 1981) produced solutions equivalent to (9a) for the surface charge problem, which did not include the components due to current flow in the ground. Equations (9b) suggest that in the general case, $B_x$ and $B_z$ may convey small influences from lateral geoelectric inhomogeneities affecting the current flow. The vertical magnetic field, however, is a direct effect of the source, and therefore, more suitable for numerical simulations and comparisons. Since $\Delta V = II_L \rho / A$, the numerical solution for any distribution of dipoles on the interface can be made equal to Fitterman’s analytic solution by adjusting the value of the current moment $IL$.

The Kozani–Grevena earthquake occurred on a pure normal fault with strike N240°–250°, dip 31–41° to NW, rake –85° approximately and length $\approx 27$ km (e.g. Hatzfeld et al., 1997). Since only the horizontal component of any stress difference (hence electrokinetic current) across the fault can generate an external magnetic field, the results will not change if instead we consider a vertical plane of equivalent area and apply Fitterman’s solution for the vertical field due to a vertical contact. Our calculations assume an outcropping vertical fault with length 30 km, depth extent 15 km and azimuth N240°, which is the direction of positive-$x$ axis of the co-ordinate system assumed for the calculations (see Fig. 10). We let the conductivity of the host rocks be 0.01 S/m and consider a favourable scenario in which the source region comprises the entire area of the fault, the EKCC contrast is 100 mV/MPa and the pore pressure difference is 10 MPa. Fig. 10 shows a map view of the vertical magnetic field ($B_z$). At JAN M, we obtain $B_z \approx -0.05$ nT,
Fig. 10. Map view of the vertical magnetic field due to an electrokinetic source comprising the entire Kozani–Grevena fault (details in the text). The focal sphere is 30 km in diameter (approximately the length of the fault) and located at the epicentre. The fault coincides with the north dipping plane in the direction 240° N, taken to be the positive-\(x\) axis of the right-handed coordinate system (positive-\(z\) is downward). A Lambert conformal conic (isogonic) projection is used.

so that \(|B_z|\) is approximately 20 times lower than observed. In order to obtain 1 nT at JAN M, we need to increase the conductivity of the host rocks by a factor of 20 which is impossible (the focal area would be a 5 \(\Omega\) m conductor). Else, we require very high pressure differences (up to 200 MPa), or EKCC differences (up to 2000 mV/MPa). We consider these scenarios very unlikely, because they demand the existence of such extreme conditions over the entire area of the fault (about 450 km\(^2\)).

EKCC differences of the order of 100 mV/MPa may be reasonable, but they may also comprise an upper limit, especially at depths comparable to the nucleation depths of large earthquakes (e.g. see Morgan et al., 1989). Furthermore, it is now understood that conductive pore fluids inhibit the EKCC and its increase as a function of permeability (Morgan et al., 1989; Jouniaux and Pozzi, 1995). This is particularly important in view of the fact that pore fluids are very saline in granitic and metamorphic rocks (i.e. conductive, usually <3 \(\Omega\) m — such brines are very important fluid phases and were actually found in the deep upper crust by the Russian Kola and German KTB ultra-deep drilling projects). It follows that any lateral changes in permeability by dilatancy or other means may be compensated for by high fluid conductivity in the deep(er) rocks. Shallow(er) sedimentary rocks could be expected to exhibit larger lateral EKCC differences due to their higher inhomogeneity, but these may also be reduced by surface conductivity effects and high conductivity formations such as strata rich in clay minerals. Consider also that the average steady-state pressure differential across the fault is not expected to exceed the level of the static stress drop (i.e. the excess stress that produced the earthquake), which averages to about 3 MPa, although values as low as 0.03 MPa and as high as 30 MPa are rarely observed (e.g. Scholz, 1990, pp. 180–189). The static stress drop due to the Kozani–Grevena event is expected not to be high, given also the normal nature of the rupture. For a
very rough estimate, we use the well-known relationship $M_0 = (16/7)\pi \Delta \sigma \chi^3$, with $\chi$ being the radius of the rupture (assumed to be circular). $M_0$ is the scalar moment, estimated between 6.2$\times$10$^{18}$ Nm (Hatzfeld et al., 1997) and 7.6$\times$10$^{18}$ Nm (Harvard CMT solution). Letting $\chi = 15$ km (one-half the length of the fault), we obtain $\Delta \sigma \approx 0.8$–1 MPa, which is lower but comparable to the average stress drop. Thus, the difference of 10 MPa used in calculating the results of Fig. 10 is rather high. It is possible that very high horizontal stress differentials exist in the vicinity of active faults (e.g. 100 MPa), but these should be expected to comprise patches distributed on the fault zone (e.g. in asperities), rather than large contiguous fault segments. Moreover, it is not apparent why lateral pressure differences considerably higher than a few MPa should appear in near-surface rocks.

In consequence of the above reasoning, we consider a patch embedded at 6 km, with dimensions 2 km$\times$2 km. This would give $|B_\perp|=0.009$ nT at JAN M, if the potential difference across it was 20 V. This means that at least 110 such patches should be simultaneously active in order to yield 1 nT at JAN M, with a total area of 440 km$^2$, which is comparable to the area of the fault. The amplitude of $B_\perp$ is linear with patch area if the other parameters remain unchanged, and thus, we run into an impasse: in order to observe $|B_\perp|=0.8$ nT, we require a potential difference of 20 V across a total fault area greater that 300 km$^2$. We posit that such conditions are physically very unlikely to occur and argue against an electrokinetic source in the fault zone of the Kozani–Grevena earthquake.

It is straightforward to rule out the possibility of a local EKE by direct strain propagation from the seismogenic zone. For a rough demonstration, let $l_0 \approx 18$ km be the characteristic dimension for the strongly deforming volume $V_0$, as calculated after Kanamori and Anderson (1976) — see Section 2.1. Assuming that the crust behaves elastically through the relatively short duration of a precursor event, the strain varies with distance $R$ according to the law $\varepsilon = \varepsilon_0 l_0^3/R^3$, where $\varepsilon_0$ is the deformation inside $V_0$ (e.g. Bernard, 1992). The local pore pressure due to $\varepsilon$ is $P \approx Y \varepsilon$, where $Y$ is the bulk modulus. Thus, the pressure gradient is $dP/dr \approx -3Y\varepsilon_0 l_0^3/R^4$. When $R=80$ km (approximately the distance between the west end of the fault and JAN M), $Y = 9.4 \times 10^{10}$ Pa for Ioannina limestones, and assuming that $\varepsilon_0 = 10^{-5}$ (10% the average co-seismic strain), we obtain $dP/dr \approx 0.4$ Pa/m (400 Pa/km), which is obviously insufficient to drive any appreciable electrokinetic current.

We are left with two alternatives. Either the pressure gradient from the earthquake source triggered horizontal flow of fluid between metastable reservoirs (as per Bernard, 1992), or the EK current was a purely local phenomenon, unrelated to the earthquake. Both these hypotheses must account for their driving mechanism because they require the prior existence of sufficiently high pore pressures, so as to generate differentials of 1 to a few MPa over kilometric-scale distances. Some information is available from the logs of one deep well drilled at the site of the city’s water purification plant and kindly provided by the operator, Project Studies and Mining Development Corporation S.A. (W1 in Fig. 8, approximately 2 km west of JAN M). The thickness of the Neogene sedimentary sequence at W1 is 340 m. The bedrock to the bottom of the well (1523 m) comprises a sequence of karstified limestones and shales. Fresh water was found in several layers, the more significant located just beneath the Neogene sediments and at 1000–1100 m. The latter is better seen in the well resistivity log, with short-normal and long-normal apparent resistivities of 100 and 400 $\Omega$m, respectively. Quite higher resistivities were found in the other segments of the well. This may indicate a low degree of interconnection between subsurface void space, while the low salinity and temperature of the water also contribute in maintaining relatively lower conductivity contrasts. According to the drilling operations manager (Dr. D. Papademetriou, personal communication), the deep water tables exhibited artesian pressures of up to 5 atm (~0.5 MPa) at the head of the drilling bit. Given also the geological structure of the broader area, it is likely that the magnitude of the artesian pressure at depths of a few hundreds of metres does not change significantly over kilometric-scale distances. It is not at all clear how and why there may be horizontal pressure differences higher than the vertical (artesian), or near-surface hyper-pressureised fluid reservoirs in the area of Ioannina, given also the absence of any significant tectonic activity.

We again consider a favourable scenario in which the pressure difference is 0.5 MPa and the EKCC difference is as high as 100 mV/MPa. Let the conductivity of the host rocks be 0.005 S/m, consistent
with the resistivity log and let the source be a vertical contact directly beneath JAN M and W1, with length 3 km, depth of burial 0.2 km and depth extent 2 km (comparable to the scale of Ioannina basin). It turns out that the maximum amplitude $B_z \approx -0.06$ nT at the edge of the contact, so that $|B_z|$ is almost 16 times too low. The (absolute) observed values will be achieved by decreasing the resistivities of the host rocks by a factor of 10, which may be rejected on the basis of the resistivity logs at W1, or, increasing the EKCC contrast by a factor of 10, which is much higher than reasonable values quoted from international literature for limestones, or finally, by raising the pressure difference by a factor of 10, which is not justifiable due to the absence of significant tectonic activity in the area of Ioannina basin, at least during the late Quaternary.

The arguments and calculations presented above indicate that the 18 and 19 April 1995 signals were generated by neither a remote nor a local electrokinetic source. Both signals are inconsistent with the expected characteristics of signals generated by all the electrification processes discussed herein, a fact that strongly indicates against their source being natural. In the following, we will investigate whether it may be anthropogenic.

3.3. Inquiring the anthropogenic origin of the magnetic field at JAN M

Electric and magnetic signals with characteristics such as observed at Ioannina are commonly generated by industrial activity. For instance, it has long been recognised that high-tension dc electric railways generate very similar stray currents (e.g. Jones and Kelly, 1966; Fournier and Rossignol, 1974), that are a dreaded source of noise in MT sounding. There are no electric railways in NW Greece, but similar effects may arise from switching of grounded machines, or variable loads on an unbalanced power distribution grid, providing time-variable line sources and wide-band contamination; such electric and magnetic variations can be produced by a number of industrial activities (e.g. Kishinovye, 1951). On the basis of our observational data, we can outright exclude point sources, for they may only generate an azimuthal horizontal primary field (e.g. Edwards and Nabighian, 1991). This may produce a vertical secondary field on high contrast interfaces, which cannot exceed the primary in amplitude. In order to make some order of magnitude estimates, we can only consider grounded wire sources.

Let us first consider a grounded horizontal dipole source in a conductive half-space, whose magnetic field components are given by Equations (9). We do not know the distance $r$ from JAN E and IOA to the source, but according to the arguments developed in Gruszow et al. (1996), it is not very far. Due to the distribution of the intensities and polarisations of the signal, it seems safe to suppose that it is of the order of the distance between IOA and JAN E, say 4 km. When $z=0$ and $z'=0$ and $r=4$ km, then in order to observe $B_z \approx B_h \approx 1$ nT, the current moment $I_z$ would be $1.6 \times 10^5$ A m. For $r=2$ km, the corresponding current moment would be $4 \times 10^4$ A m. If the source is indeed short, then it is powerful (~100 kW), but not at all improbable in an industrial area. However, this is an upper limit. It is not at all necessary for the source be a physical dipole. It may as well be a long wire configuration. Let such a source be a $y$-directed, $2L$ km-long wire, carrying an electric current $I$, grounded at both ends A and B, with a current source at $(0, L, z_1)$ and an equal but opposite current sink at $(0, -L, z_2)$, where $z_1$ and $z_2$ are arbitrary (Fig. 11). The magnetic field components on the surface can be derived from...
the law of Biot–Savart:

\[
B_x = \frac{\mu_0 I}{4\pi} \left( \frac{\cos \theta_1 - \cos \theta_2}{R_1} \right)
\]
\[
= \frac{\mu_0 I}{4\pi} \left( \frac{y + L}{x^2 + (y + L)^2} - \frac{y - L}{x^2 + (y - L)^2} \right) 
\]
\[\text{(10a)}\]

\[
B_y = \frac{\mu_0 I}{4\pi} \left( \frac{\sin \theta_1 - \sin \theta_2}{R_1} \right)
\]
\[
= \frac{\mu_0 I}{4\pi} \left( \frac{x}{x^2 + (y + L)^2} - \frac{x}{x^2 + (y - L)^2} \right) 
\]
\[\text{(10b)}\]

\[
B_z = \frac{\mu_0 I}{4\pi x} \left( \sin \varphi_2 - \sin \varphi_1 \right)
\]
\[
= \frac{\mu_0 I}{4\pi x} \left( \frac{y - L}{x + 2 \sqrt{(y - L)^2}} - \frac{y + L}{x + 2 \sqrt{(y + L)^2}} \right) 
\]
\[\text{(10c)}\]

For a given source–receiver separation, the longer the source, the smaller the current required to generate the same magnetic field amplitude. In the case of a current source at the near end and a current sink at the far end of a very long wire, less than 7 A should leak into the ground, in order to observe \(B_z \approx B_h \approx 0.7 \text{ nT}\) at a distance of 2 km.

The relationship between the observed \(B_h\) and \(B_z\) may also assist in understanding several aspects of the source. We have observed that \(|B_z| > B_h\). This immediately excludes the possibility of the source being a grounded dipole on the surface of the earth \((z=0, z' = 0)\). It is straightforward to determine that in this case, we should observe \(|B_z| < B_h\) except for when \(y=0\), in which case they should be equal. Things, however, change dramatically when the dipole is buried. For instance, when \(z=0\) and \(z'=2\) m, \(|B_z| > B_h\) in the entire region shown in Fig. 12, which at a distance of 4 km from the source develops into a 400 m-wide zone. For a long wire source, it is clear from (10) that \(|B_z| > B_h\) for the entire length of the wire. These results (obtained for a half-space), are perturbed in the presence of heterogeneous structures, as is the region of Ioannina, but in a fashion likely to increase the \(|B_z|/B_h\)

Fig. 12. The shaded area represents the domain where \(|B_z| > B_h\), in the case of a physical horizontal dipole buried at the depth of 2 m. The arrow indicates the position and orientation of the dipole.
ratio (see Edwards and Nabighian, 1991). Such details have apparently not been accounted for by Varotsos et al. (1996b, 1999b), who attempt to refute the arguments of Gruszow et al. (1996) by considering the case of an ideal dipole on the surface of the Earth.

(Powerful) industrial noise sources are abundant in and around Ioannina. It should be emphasised that Ioannina is a city of more than 100,000 and the industrial and cultural centre of NW Greece. IOA is located only 6–7 km ENE from the city centre, in the immediate vicinity of Perama town. Industrial-scale electrical installations within a radius of 2 km from IOA and JAN M include the village of Krya, a heavy duty water pumping station supplying the entire urban complex of Ioannina (Krya), the tourist attractions of Perama cave and the large military camp in which IOA is located. There is no heavy industry, but within a radius of 5 km, small and medium size factories are numerous, including two foundries, a waste water purification plant with heavy electrical machinery, dozens of manufacturing and agricultural industries, several quarries, the airport and an unspecified number of broadcasting and communications transmitters. The bulk of the other industrial installations is located along the 8 km-highway between Ioannina and the town of Eleousa, 1.5 km W of JAN M, where there is a large power distribution substation with high-to-medium tension transformers. These are grounded via a resistance and occasionally discharged through the ground, as a part of the network maintenance procedures. The road system of the area forms a dense rectangular grid longitudinal and transverse to the basin. Live power lines of various capacities run parallel to every single road, supplying the large number of consumers everywhere. Within this context, it is easy to see how short and long grounded wire, and other types of sources may appear. In fact, line, single-electrode sources and grounded loops are more likely than short dipoles. Line sources can be due to failing connections or current leakage through buried lifelines. Single-electrode sources may arise from failing groundings; these are actually line sources, with the second electrode a long distance away. Finally, there is a large variety of cable configurations which may affect grounded loops. In all the above cases, the current required to produce the observed effect may actually be quite small and difficult to locate.

3.4. Brief discussion

The 18 and 19 April 1995 electric and magnetic signals recorded at Ioannina (Greece), independently by Gruszow et al. (1996) and Varotsos et al. (1996a,b), were pronounced by the latter to be EEP and used as the observational basis for a formal earthquake prediction statement. For this reason, they have also been subject to debate (see Section 3.1).

We made an extensive effort to verify whether these signals are natural or not. We have based a substantial part of our analysis on the magnetic field, because it is considerably less sensitive to distortion than the electric field, less sensitive to inhomogeneities along the propagation path, insensitive to the local geoelectric structure and in several cases, telltale of the source. All the tests we performed could not verify the natural origin of the signals, unless some rather bizarre physical conditions were assumed. The evidence in favour of their anthropogenic origin is compelling; however, we cannot locate the source, and therefore, we did not prove this hypothesis in a formal sense. What is certain, however, is that if such rigorous analysis was done at the outset (April 1995), it would have appeared unwarranted if not unreasonable, to issue a formal prediction statement on the basis of such a poorly constrained ‘precursory’ data set. We also believe that we have shown the usefulness of the magnetic field in identifying the nature of a candidate precursor, which in itself is an important result.

4. Discussion and conclusions

We believe that the difficult task of identification and discrimination of EEPs (or any precursors indeed), requires a good understanding of the physical processes leading to their generation. With the possible exception of the piezoelectric and EKEs, most of the proposed EEP mechanisms have been hypothetical or qualitative, leading to a diversity of models, the majority of which could not stand up to scrutiny. In consequence, the establishment of a relationship between precursors and earthquakes remained mostly empirical, and at least in the case of VAN, relied heavily on parametric statistics. However, in the absence of definitive and rigorous physical constraints (the rules of the game), parametric statistical analysis may
lead to very different conclusions depending on the point of departure. The inconclusive ' Debate on VAN' (GRL Special Issue, Volume 23, No. 11, 1996), has shown that it might be possible to associate or dissociate any two sparse random sequences of 'signals' and earthquake occurrences, by formulating different null hypotheses and testing them with different statistical models. In the particular case of the EEP, there has been additional uncertainty and equivocality, because the signals reported in the literature came in such a great variety of shapes and duration that defied intuition and so lacking in interpretation, as to generate scepticism and negative feedback. For example, Park et al. (1996) assert that “the ability to distinguish SES from other electric field variations using objective criteria appears to be established.” However, in their reply to the unduly negative thesis of Geller et al. (1997), Aceves and Park (1997) remark that “… The key to assuring that these experiments are valuable is to design them to objectively define anomalies, differentiate between natural signals and noise, elucidate physical mechanisms, and provide a data set amenable to statistical analysis …. ” The latter statement is clearly more representative of the present state of affairs.

The question of identification and discrimination of true EEP signals from noise will be open, while there is no convincing explanation(s) of their origin and we note that there are no objective criteria to identify an EEP based on measurements of the electric field only (e.g. Tzanis and Gruszow, 1998). In this paper, without any recourse to statistics, we attempt to show that the successful discrimination of EEP and noise may be possible by working out plausible theories of the source. To this effect, we present two examples of profoundly different signals reported as earthquake precursors by the same research group, who in both cases were content with a vague rationalisation under the context of their piezostimulated depolarisation current theory and did not attempt to justify why the same process should be manifested in two such profoundly different modes. In the first example (15 and 18 January 1983), there may be a plausible explanation of their being true EEP signals, based on a model of electrification during crack propagation at the terminal stage prior to rupture. In the second example (18 and 19 April 1995), however, the signal can be rejected on the basis of the crack propagation model. It can also be rejected on the basis of EKEs by its inconsistent magnetic field properties. It turns out that the presence and properties of a companion magnetic field may greatly assist in identifying the nature of the signal and its measurement should be standard in all future monitoring for ULF precursors.

Earthquake processes are very complex and occur in a manner completely indifferent to the requirements of human science and society, sparingly providing information and insights into their secrets. For this reason, experimental methods alone, however elaborate, cannot provide viable answers without hard physical constraints and the research should now focus on deciphering the physics of the earthquake preparation process. Although progress is being made, the existing knowledge and theories of the EEP source are still very early and imperfect, our work own work included. There are no definite answers to the basic questions and research must be given time to progress, albeit with vigilance and adherence to the rules of the scientific game. We cannot assess when, or if it will be possible to predict earthquakes and we also note that earthquake prediction is not a mature discipline. It is science in the making and this should not be forgotten by both its critics and defenders.

References


Jouniaux, L., Pozzi, J.-P., 1997. Laboratory measurements of anomalous 0.1–0.5 Hz streaming potential under geochemical changes: implications for electrotelluric precursors to earthquakes. J. Geophys. Res. 102, 15335–15343.


Kawase, T., Uyeda, S., Uyeshima, M., Kinoshita, M., 1993. Possible correlation between geoelectric potential change in Izu–Oshima Island and the earthquake swarm off the east Izu Peninsula, Japan. Tectonophysics 224, 83–93.


