Air ionization at rock surfaces and pre-earthquake signals

Friedemann T. Freund, Ipek G. Kulahci, Gary Cyr, Julia Ling, Matthew Winnick, Minoru M. Freund

1. Introduction

Seismologists often state that earthquakes cannot be predicted except within wide statistical margins, typically several years or decades (Geller, 1997; Kagan, 1997; Kellis-Borok, 2003). However, non-seismic signals that precede major earthquakes have been reported from essentially all tectonically active regions around the world. A partial list of these pre-earthquake signals includes: ionospheric perturbations (Hayakawa and Sazhin, 1992; Lipovsky et al., 2000; Pulinets and Boyarchuk, 2004; Shalimov and Gokhberg, 1998); thermal infrared anomalies (Saraf et al., 2008a, b; Tramutoli et al., 2005; Tronin, 2006); earthquake lights (Derr, 1973; St-Laurent, 2000); fog, haze and cloud formation (Aleksandrov et al., 2001; Guo and Wang, 2008); and unusual animal behavior (Tributsch, 1984).

1.1. Stress-activated electronic charge carriers

If these diverse and seemingly disjoint signals are truly precursory, the question arises how their generation may be linked to the earthquake preparation process.

In this report we present data from laboratory experiments that can help us gain insight into the generation of several of these pre-earthquake signals. Specifically we demonstrate that massive air ionization can take place at the surface of rocks, which are being stressed at one end. The build-up of stress within the Earth’s crust prior to major earthquakes may likewise lead to processes at the Earth’s surface, including massive air ionization, which can be expected to cause ionospheric perturbations and a host of other phenomena.

Pre-earthquake signals have been widely reported, including perturbations in the ionosphere. These precursory signals, though highly diverse, may be caused by just one underlying physical process: activation of highly mobile electronic charge carriers in rocks that are subjected to ever increasing levels of stress. The charge carriers are defect electrons associated with O2− in a matrix of O2. Known as positive holes or pholes h+, they flow out of the stressed rock into the unstressed rock volume, traveling meters in the laboratory, probably kilometers in the field. At the rock–air interface they cause: (i) positive surface potential, (ii) field-ionization of air molecules, (iii) corona discharges. The rate of formation of airborne ions can exceed 108 cm−2 s−1. Massive air ionization prior to major earthquakes increases the electrical conductivity in the air column and may cause ionospheric perturbations, earthquake lights, and unusual animal behavior as well as infrared emission.

© 2009 Elsevier Ltd. All rights reserved.

ARTICLE INFO

Article history:
Received 4 January 2009
Received in revised form 10 June 2009
Accepted 10 July 2009
Available online 11 August 2009

Keywords:
Pre-earthquake phenomena
Ionosphere
Air ionization
Corona discharges
Thermal infrared anomalies
Earthquake lights
Animal behavior

ABSTRACT

Pre-earthquake signals have been widely reported, including perturbations in the ionosphere. These precursory signals, though highly diverse, may be caused by just one underlying physical process: activation of highly mobile electronic charge carriers in rocks that are subjected to ever increasing levels of stress. The charge carriers are defect electrons associated with O2− in a matrix of O2. Known as positive holes or pholes h+, they flow out of the stressed rock into the unstressed rock volume, traveling meters in the laboratory, probably kilometers in the field. At the rock–air interface they cause: (i) positive surface potential, (ii) field-ionization of air molecules, (iii) corona discharges. The rate of formation of airborne ions can exceed 108 cm−2 s−1. Massive air ionization prior to major earthquakes increases the electrical conductivity in the air column and may cause ionospheric perturbations, earthquake lights, and unusual animal behavior as well as infrared emission.

© 2009 Elsevier Ltd. All rights reserved.
specifically peroxy defects, $O_3X–O–YO_3$, where $X \neq Si^{4+}, Al^{3+}$ etc. The peroxy defects in turn derive from small amounts of $O_3X–OH$ incorporated into nominally anhydrous minerals when they crystallize from $H_2O$-laden magmas or recrystallize in $H_2O$-laden high temperature metamorphic environments. The incorporation of "impurity" hydroxyl can be written as hydrolysis of an $O_3X–OH–YO_3$ bond:

$$O_3X–O–Y_2O+H_2O \rightleftharpoons O_3X–OH–YO_3$$  \hspace{1cm} (1)

During cooling, at temperatures below about 500 °C, hydroxyl pairs rearrange electronically in such a way as to change the valence of their oxygens from $O^2–$ to $O^–$, while reducing their protons from $H^+$ to $H$. The $H$ combine to molecular $H_2$, while the $O^–$ combine to form the peroxy bond. (Freund, 1985):

$$O_3X–OH–YO_3 \Rightarrow O_3X–OO–YO_3+H_2$$  \hspace{1cm} (2)

During mechanical deformation, dislocations are mobilized and/or generated, sweeping through the mineral grains. When they intersect peroxy defects, the $O^2––O^–$ bonds break (Freund, 2002). In the process an $O^2–$ from outside the peroxy bond donates an electron $e$, which is captured by the broken peroxy bond. The donor $O^2–$ thereby enters $O^–$; e.g. physically a defect electron in the $O^2–$ matrix, h', a mobile electronic charge carrier:

$$O_3X–OO–YO_3+\{XO_4\}^{2–} \rightleftharpoons O_3X–OwO–YO_3+\{XO_4\}^{2–}$$  \hspace{1cm} (3)

Here the donor $O^2–$ is represented by a structural unit $\{XO_4\}^{2–}$ changing to $\{XO_4\}^{4–}$.

As an electronic state, the h' "live" at the upper edge of the valence band, which consists primarily of $O 2sp$-symmetry energy levels. The h' can propagate along the valence band, presumably by phonon-assisted electron hopping. Since the valence bands of all grains in a rock form an energetic continuum, the h' can cross grain boundaries and propagate through sand and soils. Theoretically the phase velocity of an h' wave is given by the phonon frequency $\nu \sim 10^{12} \text{s}^{-1}$, times the hopping distance $d$ between oxygen sites, $\sim 2.8 \times 10^{-10} \text{m}$. Hence the h' phase velocity should be $v=d \sim 280 \text{m s}^{-1}$, consistent with the measured speed of propagation of h' waves through different igneous rocks, $300 \pm 100 \text{m s}^{-1}$ (Freund, 2002).

1.2. Rock battery

When stress is applied to a portion of a rock, the number density of electrons and holes inside the stressed rock volume increases. The h' charge carriers can flow out of the stressed rock and into an adjacent unstressed rock, while the electrons, e', stay behind. The reason for the electrons staying behind is that there are no energy levels in the unstressed rock, which they could access. Fig.1 shows a block of rock that is stressed at one end. With the outflow of h' a potential difference develops between the stressed and unstressed rock, equivalent to a battery voltage. The h' outflow continues until an equilibrium is reached between the electric field and the h' concentration gradient.

The situation is analogous to an electrochemical battery, which delivers two types of charge carriers, electrons and cations. The electrolyte allows cations to flow out. Thus, the electrolyte turns positive. For electrons to also flow out, a metal contact must be attached somewhere to the electrolyte. In the set-up depicted in Fig.1 the circuit is not closed.

1.3. Air ionization

An additional characteristic feature in Fig.1 is the build-up of a surface charge. It forms because, as the h' enter the unstressed rock, they repel each other electrostatically. They build-up a surface potential, which is a function of (i) the number density of h', and (ii) the dielectric constant (King and Freund, 1984).

The electric field reaches $400,000 \text{V cm}^{-1}$ on a flat surface of a dielectric with a dielectric constant $\varepsilon = 10$ containing $10^{16} \text{cm}^{-3}$ h' charge carriers, equivalent to 100 ppm. At edges and corners, where the radius of curvature is small, the electric field will be much higher, potentially exceeding the dielectric strength of air. This raises the question whether it might be possible to field-ionize air molecules at the rock surface.

2. Experimental

We used gabbro from Shanxi, China, a typical deep crustal, igneous rock, chemically identically to basalt, with ~40 modal% plagioclase, ~30 augitic clinopyroxene surrounded by alteration rims to amphibole and chlorite, plus ~25% opaques, a porosity of ~0.3%, and ~1% total water, mostly due to hydroxyl-bearing minerals such as amphiboles.

The experiments were conducted with $30 \times 15 \times 10 \text{cm}^3$ blocks with one polished surface and all other surfaces saw-cut. The blocks were placed inside an aluminum box ($50 \times 30 \times 30 \text{cm}^3$) acting as a Faraday cage and fitted with a steel bellow to apply the load. The pistons were in electrical contact with the rock but insulated from the Faraday cage and the hydraulic press by polyethylene sheets with $>10^{14} \text{Ohm}$ resistance as depicted in Fig.2a. Hardened stainless steel ball bearings, 6.3 mm diameter, were used to transmit load into the rock.

![Fig. 1. Stress applied to one end of a rock activates electrons and hole, e' and h'. The h' flow out of the stressed rock volume into the unstressed rock, creating a potential difference. The situation is similar to that of an open circuit electrochemical battery. The stressed volume is negative and the unstressed volume is positive. The h' charge carriers become trapped at the surface, leading to a positive surface charge.](image1)

![Fig. 2. (a) Configuration of the gabbro block inside the Faraday cage. (b) Detail view of the stainless steel pistons with rows of stainless steel ball bearings to act as stress concentrators.](image2)
glued onto the faces of the pistons as depicted in Fig. 2b. The ball bearings acted as stress concentrators causing the rocks to massively deform at the contact points. Because of the stress concentrators, we do not give the stresses in MPa but only report the load in pounds [lbs]. We loaded the rocks moderately fast, 200–300 lbs/s until failure, using a manually controlled hydraulic press.

During earlier impact experiments, we observed positive surface potentials accompanied by corona discharges along the rock edges (Freund, 2002). Similar positive surface potentials appeared in response to slow application of stress (Takeuchi et al., 2006). Values as high as +12 to +17 V were reported at high stress rates (Enomoto et al., 1993). In the present experiments, to detect surface potentials and measure airborne ions, we used an Al sheet (10 × 20 cm², 1 mm thick) as a capacitor plate or ion collector respectively. The Al sheet was placed above the rock on pieces of Styrofoam glued to the inside walls of the Faraday cage extending over the edges as depicted in Fig. 3. The air gap was ~5 mm. To collect positive and negative ions we biased the Al sheet at −90 and +90 V respectively, using ten 9 V batteries connected in series. Two photodiodes were aligned along one edge of the rock as depicted in Fig. 2a to capture flashes of light resulting from corona discharges.

Ion currents and surface potentials were recorded with a Keithley 487 picoammeter and a Keithley 617 electrometer respectively, using LabView 7.1. The photodiode output was recorded on a Tektronix TDS 224 oscilloscope at 100 MHz.

Fig. 3a shows the Faraday cage (with the front half of the cover removed) at the end of experiment #37 after the rock had slightly tilted. It shows the ion collector plate, which had initially been 5 mm above the rock and parallel to the rock surface. Fig. 3b shows details of the pistons with the ball bearings after a wedge-shaped section of the rock had cleaved off. Also seen are the polyethylene sheets used to electrically insulate the pistons from the press and one of the leads connecting the pistons to the picoammeter or electrometer.

### 3. Results

#### 3.1. Surface potentials

Fig. 4 shows an example of the surface potential using the Al plate as capacitor. The first sign of a positive potential appeared when the platen of the hydraulic press started to move upward (arrow 1). The stress caused by this modest acceleration was enough to activate some h’ charge carriers, which led to a positive surface potential of +300 mV. When the piston of the press made contact, the potential rose to +700 mV (arrow 2). When the press began loading the rock, about 200 s into the run (arrow 3), the surface potential reached +3.4 V, and soon dropped to negative values (~0.3 V).
surface potential rose rapidly reaching +3 V around 10,000 lbs. The surface potential remained high with further loading but fluctuated between +2.6 and +3.4 V. Discontinuities occurred in the loading curve, the first at about 280 s. They mark moments when the stainless steel ball bearings, e.g. the stress concentrators, are sinking into the rock, reducing temporarily the hydraulic pressure. Rock deformation episodes of this type were observed during all experiments described here.

Close to 300 s into the run, the surface potential around +3 V abruptly broke down as depicted in Fig. 4, turning negative to around –0.3 V with intermittent spikes. This general behavior, first a strong positive surface charge followed by abrupt transition to slightly negative values, was consistently observed during these experiments.

### 3.2. Positive air ions

To record positive ion currents across the air gap the Al plate was biased at +90 V. Fig. 5a/b shows a typical run with the inset depicting the electric circuit. Before loading, the background ion current was in the low pA range. It remained low at low loads. However, when the load approached 10,000 lbs, after the ball bearings had already sunk into the rock as indicated by the discontinuities in the load vs. time curve, a positive air ion current started to flow. Between 10,000 and 25,000 lbs several load vs. time discontinuities occurred. The positive air ion current increased, fluctuating in the 10–25 nA range. The discontinuities in the load vs. time curve are not correlated to spikes in the ion current. The area of the collector plate was 200 cm². An ion current of 20 nA corresponds to an average ion production rate on the rock surface on the order of $10^9$ cm⁻² s⁻¹.

About 2 s before failure, at about 30,000 lbs, a 55 nA spike occurred, indicating a burst of positive ions from the rock surface. The rock failed at the 75.5 min time mark. The rock slightly tilted but did not touch the collector plate. The collector plate continued to measure an ion current, which increased sharply to 450 nA and then decreased exponentially over the next 30 s.

Fig. 6 shows another run under negative bias. In this case the positive ion current started soon after loading, but the overall level remained moderate at around 10 nA (left inset). In spite of several episodes with the ball bearings sinking into the rock, the positive ion current did not increase. It even dropped to near-background level between 785 and 800 s. At 815 s a sharp, prominent current spike occurred, reaching 185 nA. The rock failed at 830 s, splitting across its width, causing the rock to drop from underneath the ion collector plate without touching it. The ion current increased moderately after failure and then decreased exponentially over the next 30 s (right inset).

The recurring high levels of positive ions immediately after fracture and their decay with a characteristic decay time around 30 s suggest that, during fracture, bursts of ions are generated, filling the Faraday box. The ions slowly drifted to the walls to be neutralized.

The short current spikes during loading are distinctly different. Observed often but without clear correlation to the episodes of rapid rock deformation, these spikes typically last for only a few seconds as depicted in Fig. 6. Since corona discharges had been observed during earlier impact experiments (Freund, 2002), we changed the bias to +90 V to collect electrons and negative ions. In addition, light emission was monitored with the pair of photodiodes situated as depicted in the inset in Fig. 2a.

### 3.3. Corona discharges

The results obtained under +90 V bias were characteristically different from those under -90 V bias. Positive ions typically started to form at relatively low loads as demonstrated in Figs. 5a/b. The results obtained under +90 V bias were characteristically different from those under -90 V bias. Positive ions typically started to form at relatively low loads as demonstrated in Figs. 5a/b.
and 6. By contrast negative ions and/or electrons recorded under +90 V bias appeared only at higher loads, typically ¼ to ½ the load needed to cause failure.

Fig. 7 shows an example. Up to 20,000 lbs, the current remained at background levels, in the low pA range. Abruptly, at the 2434 s mark, the current rose to 100–115 nA, followed by a continuously high level around 65–70 nA, accentuated by multiple short spikes. The spikes are accompanied by light pulses, each about 1.5 ms long, one of which is shown in the inset. Light pulses indicate corona discharges.

4. Discussion

The positive surface potentials confirm observations reported earlier (Freund, 2002) that, when an igneous rock is subjected to deviatoric stresses, electronic charge carriers are activated. These charge carriers are pholes h⁺ activated as described by Eq. (3) in Section 1.1. The activation is thought to involve dislocations that are mobilized in the stressed rock volume. The dislocations intersect pre-existing peroxy defects in the structure of rock-forming minerals and of the electronic charge carriers, which they engender when the rocks are stressed, holds the key to many reported pre-earthquake signals. Their discovery may be as fundamental for geophysics as was the discovery of semiconducting properties in amorphous materials for solid state and applied physics (Ovshinsky and Adler, 1976).

4.1. Surface potential, positive airborne ions and corona discharges

Stressing one end of a block of igneous rock such as gabbro leads to a series of processes at the unstressed end. First, positive surface potentials appear uniformly across the rock surface, increasing rapidly with increasing stress and reaching about +3 V. Second, massive amounts of positive airborne ions are collected above the unstressed end of the rock. Third, massive amounts of electrons and/or negative airborne ions are collected.

The positive surface potentials confirm observations reported earlier (Freund, 2002) that, when an igneous rock is subjected to deviatoric stresses, electronic charge carriers are activated. The charge carriers are pholes h⁺ activated as described by Eq. (3) in Section 1.1. The activation is thought to involve dislocations that are mobilized in the stressed rock volume. The dislocations intersect pre-existing peroxy defects in the matrix of minerals and cause them to break. Electrons e⁻ are activated alongside the pholes h⁺. However, while the pholes can flow out of the stressed rock volume, the electrons are unable to follow suit. Hence, as soon as h⁺ flow out, the stressed rock charges negatively relative to the unstressed rock. The unstressed rock becomes positively charged.

The situation is analogous to that in an electrochemical battery where cations spread into the electrolyte leaving behind a negative charge. The electrons, unable to spread into the electrolyte, can flow out via a metal contact. It is important to note that, even without closing the circuit, a potential difference develops, the “battery voltage”.

The difference between an electrochemical battery and the “rock battery” as presented here is that the positive charge carriers are not cations but positive holes, h⁺. Thus, when stressed, the rock turns into a type of semiconductor battery, which has not been previously described.

Positive charges on the surface of the rock appear even at low stress levels as demonstrated in Fig. 4. Under these conditions the number of h⁺ charge carriers available to flow out of the stressed rock volume may still be small but they suffice to build-up a surface potential (Freund et al., 2006). As the rock is loaded more, the surface potential increases, indicating a larger number of charge carriers activated in the stressed rock volume. At the moderately fast loading rates used here the surface potentials reach rapidly values around +3 V.

Fig. 8a/b shows the surface potential and associated electric fields calculated for a flat surface of a semi-infinite medium with the dielectric constant ε = 10 (King and Freund, 1984). Within the range of charge carrier concentration under consideration the surface potential is constant but the thickness (d) of the surface charge layer decreases with increasing h⁺ concentration. This means that the electric field, which builds up directly at the solid–air interface, increases as d decreases. At a relatively low charge carrier concentration of 10¹⁷ m⁻³ (~10 ppm), the calculated value for E is 120,000 V/cm. At a concentration of 10¹⁸ m⁻³, it increases to 400,000 V/cm. At still higher concentrations or at edges and corners, E is expected to soon exceed the dielectric break-down strength of air, ~2–3 x 10⁹ V/cm.

The electric fields due to the accumulation of charge carriers at and a few tens nanometer below the surface are not the same as the macroscopic vertical electric fields at the Earth surface that are normally given in units of V/m (Brown, 1985; Pulinets, 2009). The electric fields under consideration here are short-range and act only over short distances across the surface-to-air interface.

Fig. 8c shows a Monte Carlo simulation with 1000 mobile charges allowed to relax in a dielectric medium. The charge carrier concentration is very low inside, low on flat surfaces but high along edges and corners, predicting very high electric fields.

In a system as depicted in Fig. 1, where the battery circuit is not closed, the charge carriers flowing out of the stressed rock volume become stagnant in the unstressed rock. They should create a uniform and constant surface potential independent of the number of h⁺ charge carriers activated (King and Freund, 1984). However, as Fig. 4 shows, after the surface potential had increased to +3 V, it started to fluctuate between +2.6 and +3.4 V. Such fluctuations are consistent with a break-down of the open circuit approximation.

This suggests that, around +3 V, the electric field reaches high enough values to extract electrons from neutral gas molecules, e.g. to field-ionize them, in particular along edges and corners as depicted in Fig. 9. The most likely candidates for field-ionization are O₂ and H₂O, which have relatively low ionization potentials.

The measured ion current as shown in Fig. 5a is on the order of ~20 nA, equal to an ion generation rate ~10⁴ cm⁻² s⁻¹. For each air molecule that is field-ionized at the rock surface, an electron is deposited into the rock surface. This must lower the positive surface potential.

The charge carrier density needed to produce a 1 V surface potential is on the order of 10¹⁷ cm⁻² (Takeuchi and Nagahama,
Therefore, if $10^9 \text{cm}^2/\text{s}$ air molecules become field-ionized, they deposit $10^9 \text{cm}^2/\text{s}$ electrons into the rock surface. This is obviously sufficient to cause a significant reduction of the positive surface potential. At the same time, more $h^+$ charge carriers stream from the stressed portion of the rock to the unstressed portion and, hence, to the surface. These newly arriving $h^+$ rebuild the positive charge of the surface almost as fast as it is reduced. This causes the positive surface potential to fluctuate wildly as shown in Fig. 4.

Fig. 4 also shows that, at still higher load, the positive surface potential drops abruptly and turns negative. The emission of light shown in the inset in Fig. 7 indicates corona discharges. This reversal of the sign, together with the light blips, suggests the production of electrons by the corona discharges. These electrons “rain down” onto the rock surface and annihilate the positive surface charge more effectively than the field-ionization of air molecules was able to do. In addition, through attachment to neutral gas molecules, the electrons form negative airborne ions. Corona discharges are expected to manifest themselves in air ion current bursts, both under negative and positive bias. Concurrent electronic excitation and electron–ion recombination reactions will lead to the emission of visible light.

Ionization of air over the unstressed portion of the rock explains the break-down of the open circuit approximation of Fig. 1. As illustrated in Fig. 10, while electrons are delivered to the surface and recombine with $h^+$ charge carriers, positive ions will drift toward the negatively charged pistons, which are in contact with the stressed rock. This closes the battery circuit.

In summary, the laboratory experiments, designed to measure surface potentials and airborne ion currents, have provided...
insight into processes at the rock surface when a portion of the rock is subjected to deviatoric stresses. With increasing stress the following processes occur sequentially at or above the rock surface:

(i) trapping of h\(^+\) charge carriers and appearance of positive surface charges;
(ii) field-ionization of air molecules and generation of positive air ions; and
(iii) corona discharges with bursts of ion current and flashes of light.

4.2. Air ionization and pre-earthquake phenomena

We can use these laboratory observations to address phenomena, which have been observed in the field and have been linked to impending seismic activity.

We start from the simple, but very basic premise that earthquakes represent the final stage of a long drawn-out process deep below, in the future hypocentral volume, marked by increasing levels of deviatoric stresses. How the stresses evolve, how they increase, decrease or change directions will vary enormously from case to case. However, close to a catastrophic rupture, the stresses in the hypocentral volume can be expected to locally concentrate and to increase fast. Another generally valid statement is that the hypocentral volume, where the stresses evolve toward catastrophic rupture, will always be surrounded by a larger volume of less stressed or unstressed rocks.

Therefore, the experimental set-up used in the experiments described here can be used to draw parallels to pre-earthquake situations in the Earth's crust. This is particularly true for the electric analog of an open battery circuit with the outflow of h\(^+\) charge carriers from the most stressed rock volume being the dominant process.

Under such conditions h\(^+\) charge carriers activated deep below are expected to spread out of the prospective hypocentral volume into the surrounding less stressed or unstressed rocks, while the electrons cannot flow out. Some h\(^+\) will travel upward and reach the Earth's surface. The higher the stresses in the hypocentral volume, the larger the number of h\(^+\) charge carriers flowing out and the larger the number of h\(^+\) charge carriers reaching the Earth's surface. With more h\(^+\) charge carriers arriving, the electric fields at the surface will increase. Eventually those fields can be expected to reach values high enough to initiate air ionization at the Earth's surface (leading to positive ions) and corona discharges (leading to free electrons and negative ions alongside positive ions).

The measured ion currents suggest rates on the order of $10^4 \text{ to } 10^{10}$ ionization events per second and cm\(^2\). Extrapolated to km\(^2\), this is equivalent to an ion current on the order of 10–100 A rising off the Earth surface. Such currents, if they occur in nature, could constitute a significant part of the global electric circuit, which connects the solid Earth with the ionosphere (Roble and Tzur, 1986).

4.3. Perturbations in the ionosphere

Using a ground-based ionosonde to determine the total electron content (TEC) in the ionospheric foF2 layer, anomalies were recognized two to three days before most of the $M \geq 6.0$ earthquakes in the Taiwan area during 1994–1999 (Liu et al., 2000). Likewise, using GPS receivers, ionospheric anomalies were found to occur prior to 16 out of 20 $M \geq 6.0$ earthquakes in the Taiwan region between 1999 and 2002 (Liu et al., 2006). One day before the $M = 6.8$ Kythira earthquake in Greece on January 8, 2006, during a quiet period marked by the absence of magnetic storms, the TEC values above the epicentral region were found to have increased significantly, up to 50% relative to the surrounding region and the long-time average (Zakharenkova et al., 2007). In Japan statistically significant changes in the subionospheric propagation of low frequency radio waves prior to earthquakes have been recorded (Maekawa et al., 2006) as well as sporadic ionospheric E layer anomalies prior to the $M = 7.2$ Kobe earthquake of January 16, 1995 (Ondoh, 2003), and anomalies prior to the $M = 8.3$ Tokachi-oki of September 25, 2003 off the coast of Hokkaido (Hayakawa et al., 2005). Using satellite and ground-based GPS data anomalous fluctuations in the integrated TEC have also been described prior to the $M = 7.6$ Bhuj earthquake in Gujarat, India (Trigunait et al., 2004) and other events on the Indian subcontinent (Singh and Singh, 2007). Similar ionospheric variations were seen prior to the $M = 7.8$ Colima earthquake of January 21, 2003 in Mexico (Pulinets et al., 2005).

The ionospheric anomalies extend over several hundred to a few thousand kilometers. However, no consensus has been reached as to the cause or causes (Rishbeth, 2007). If a process exists that links the ground to the ionosphere, powerful enough to give rise to the reported perturbations, it must extend over large areas and it will most likely involve changes in thermospheric chemistry or in the electrodynamic drifts in the lower atmosphere.

Radon emission from the ground has been widely quoted as a possible cause for changes in the conductivity of the air linked to pre-earthquake ionospheric anomalies (Ondoh, 2003; Pulinets, 2007, 2009; Yasuoka et al., 2006). The basic concept is that, when stresses build-up across the “earthquake preparation zone”, radioactive radon is released from the ground. The “earthquake preparation zone” is assumed to be centered around the future epicenter with a radius $r$ given by the empirical relation $r = 10^{0.4 M} \text{ km}$, where $M$ is the magnitude of the earthquake (Dobrovolsky et al., 1979; Teisseyre, 1997). For magnitude 6, 7 and 8 events, the “earthquake preparation zones” would therefore cover areas as large as 500, 600 and 700 km across.

$^{222}\text{Rn}$ with its half-life of 3.82 days is a progeny of radium, $^{226}\text{Ra}$, which in turn derives from $^{238}\text{U}$, an element enriched in granitic rocks. $^{222}\text{Rn}$ decays by emitting 5.49 MeV alpha particles to $^{218}\text{Po}$ with a half-life of 3.11 min, which decays by emitting 6.00 MeV alpha particles to short-lived $^{214}\text{Pb}$ and other progeny. In air, each alpha particle of the $^{222}\text{Rn}$ decay generates 150,000 to 200,000 electron–ion pairs. The entire $^{222}\text{Rn}$ decay chain in air therefore produces at most $10^6$ ionization events.

While $10^6$ ionization events look like a large number, the rate at which radon is released from the ground, even in areas dominated by granitic rocks, is very small. On the average, air ion concentrations due to radon and its progeny range from about 25 cm\(^{-3}\) to 250 cm\(^{-3}\). Contributions from cosmic rays and secondary cosmic ray decay products at sea level are equally low, on the order of 2 ions cm\(^{-3}\) and 15 ions cm\(^{-3}\) respectively (Hoppel et al., 1986).

Measured over a 3-year period near Pune, India, the median radioactivity in air has reported to be 9.70 Bq m\(^{-3}\) due to radon itself and 2.84 Bq m\(^{-3}\) due to its progeny, producing a median air ion concentration of only 5.5 cm\(^{-3}\) s\(^{-1}\) (Nagaraja et al., 2003). Using a network of 20 radon measuring stations along the western part of the North Anatolian Fault, placed into 0.5–1 m deep trenches or holes, the count rates due to radon decay over a period of 1-year were found to vary from less than 0.3 to about 6 min\(^{-1}\) (Inan et al., 2008). At the same time, a total of 19 earthquakes of magnitudes 4–5 occurred in the region, some less than 20 km from a radon measuring station. At some stations $M > 4$
earthquakes correlate with a small increase in radon release from the ground but, more generally, the radon release patterns are found to be broad and not directly correlated to any specific seismic event.

In California radon transects were taken across creeping, locked, and freshly ruptured sections of the San Andreas Fault (King et al., 1993). Along the actively creeping section the radon release increased 6–11 times over the typically low background values. Increased radon release was also found over the fault itself, only tens of meters wide, indicating a very localized effect.

It thus appears that, while radon is definitely coming out of the ground, its release rate increases only by a factor of about 10 in seismically active regions, mostly in the immediate neighborhood of active faults. Hence, the contribution of radon to the air conductivity is of minor importance, in particular when averaged over areas as large as those predicted by the “earthquake preparation zone” concept and necessary to cause the reported pre-earthquake ionospheric perturbations.

By contrast, much higher rates of air ionization can be expected to occur when a large number of h− charge carriers, stress-activated at depth, arrive at the Earth surface. If our laboratory experiments are any indication of the potential magnitude of the effect, air ionization rates on the order of 105 cm−2 s−1 and higher appear possible. Even if a large fraction of these ions are lost to recombination events, air ion concentrations orders of magnitudes higher than those achievable by radon decay are to be expected.

High concentrations of airborne ions have indeed been reported. Prior to the October 30, 2007 Alum Rock, California, M5.4 earthquake ultra low frequency (0.01–12 Hz) pulsations were detected with a three axis induction magnetometer located 2 km from the epicenter. The 1–12 s wide pulsations were 10–50 times more intense than 2-year background noise levels, and occurred 10–30 times more frequently in the 2 weeks prior to the event than during the previous 22 months. An air conductivity sensor at the same site saturated for over 13 h during the night and much of the day before the earthquake. Compared to the previous year’s average, conductivity patterns at the site the conductivity was determined not to be caused by moisture contamination (Bleier et al., 2008, 2009). An increase in the air ionization was recorded at the same time at another station about 40 km from the epicenter, e.g. within the range of the “earthquake preparation zone”.

An extended network of dedicated air conductivity sensors in Japan, the PISCO network, has also produced a large amount of data pointing at episodes of dramatically increased air ionization, tentatively associated with seismic activity within a radius given by the “earthquake preparation zone” (Hattori et al., 2008; Wasa and Wadatsumi, 2003). One such record obtained by a station at Kanagawa near Yokohama on the Tokyo Bay is given in Fig. 11. The 28 day long record shows a series of positive ion maxima, mostly by small ions, occasionally accompanied by negative ion maxima, mostly by large ions. Both positive and negative ions reach concentrations on the order of > 50,000 cm−3. It is interesting to note that, while there are sometimes maxima of just positive ions, the negative ions seem to form only as a corollary to the positive ions. This is consistent with the sequence of first positive and then negative ion formation identified in our laboratory experiments.

If airborne ion concentrations at ground level increase significantly over such large areas, by many orders of magnitude over the normal background level, the ions can be expected to be transported upward, mostly due to thermal convection currents that will result from the release of the latent heat of condensation of water, when the ions act as condensation nuclei. Thus the electrical conductivity profile across the entire lower and middle atmosphere will be modified. This in turn will affect the electric field distribution between the lower edge of the ionosphere and the ground (Sorokin et al., 2006).

4.4. Thermal infrared anomalies

Night-time infrared (IR) satellite images have shown areas of enhanced IR emission from the epicentral regions prior to major earthquakes (Sarat et al., 2008a, b; Tramutoli et al., 2005; Tronin, 2000). Known as “thermal anomalies", they are also large areas, often 100–500 km across. They also reportedly appear a few days before major seismic events. They often fluctuate rapidly in areal extent and intensity. They reportedly disappear soon after the main shock and major aftershocks.

The rapidity with which the thermal anomalies appear and disappear rules out that they are caused by Joule heat rising from a source below and heating the rocks at the Earth’s surface. Several alternative processes have been invoked to account for the apparent temperature increase: rising fluids, due to regional stresses, leading to the emanation of warm gases (Gornyi et al., 1988), diffuse CO2 release from the ground, causing a local greenhouse effect (Quing et al., 1991; Tronin, 2000), and air ionization due to radon leading to the release of latent heat during condensation of water vapor (Pulinets et al., 2006). However, none of these explanations seem to adequately account for the characteristic features of the “thermal anomalies” as reported from satellite data.

A very different process has been proposed based on the activation of h− charge carriers and their diffusion to the surface (Freund et al., 2007). As outlined above in Eqs. (2) and (3), the dormant precursor of the h− charge carriers consists of peroxy links in which two O− are tightly bonded together. Breaking the peroxy link costs energy, estimated to be on the order of 2.4 eV. This energy is supplied by the mechanical work done in the stressed rock volume and leads to the activation of h−. When the h− arrives at the surface, they can recombine to restitute peroxy bonds. When this happens, energy on the order of 2.4 eV will be imparted to the two O− participating in the recombination process. The two O− will thus be “born” in a vibrationally highly excited state, equivalent to ~30,000 K. The best way for the two O− to get rid of this large excess vibrational energy is to radiatively de-excite by emitting photons corresponding to the transitions between the quantum levels of the vibrational manifold of the peroxy bond.

The energy for the transition from the 1st excited level to the 0th ground level is known (Ricci et al., 2001). It is equivalent to a narrow emission band at 930 cm−1 or 10.7 μm. Accordingly, transitions between higher excited levels, 2 ⇒ 1, 3 ⇒ 2, 4 ⇒ 3, etc., will lead to a series of narrow IR emissions on the long wavelength side of the 10.7 μm band as indeed observed. Because these narrow non-thermal IR bands occur within the spectral window of the broad thermal 300 K emission maximum and current satellite IR sensors cannot spectrally resolve them, they have been lumped together with the thermal emission. Hence the name “thermal anomalies”.

Rapid fluctuations, both in areal extent and intensity, are a characteristic feature of pre-earthquake “thermal anomalies” recorded by IR satellites. Similar fluctuations in the IR emission have been noted during the laboratory study (Freund et al., 2007). In both situations the fluctuations may be due to waves of h− charge carriers arriving at the surface, field-ionizing air molecules and thereby causing fluctuations of the positive surface potential as shown in Fig. 4. The surface potential fluctuations in turn lead to fluctuations in the number of h− that might be available for recombination and, hence, IR emission.
4.5. Earthquake lights

Fleeting, short-lived luminous phenomena arising from the ground and apparently related to earthquake activity have been reported since ancient times (Derr, 1973; Tributsch, 1984). They are often referred to as earthquake lights, EQLs. Based on reports from several events in Japan (Musya, 1931) stated: “the observations were so abundant and so carefully made that we can no longer feel much doubt as to the reality of the phenomena”. Nonetheless, doubts persisted in the scientific community at least until the late 1960s when EQLs were photographically documented during an earthquake swarm near Matsushiro, Japan (Yasui, 1973). EQLs have been reported from Mexico (Araiza-Quijano and Hernández-del-Valle, 1996), Canada (St-Laurent, 2000), and other seismically active regions of the world.

The most common explanations for EQLs invoke piezoelectricity as the physical process that can produce electric fields strong enough to cause a dielectric break-down of air and luminous discharges (Finkelstein et al., 1973; Johnston, 1991; Ouellet, 1990) or frictional heating of the fault during rupture (Lockner et al., 1983). However, on the basis of the experiments presented here a more likely explanation is that electric discharges at the Earth’s surface are caused by h+ charge carriers arriving at the ground-air interface, building up sufficiently strong electric fields to cause field-ionization of air molecules and corona discharges. Such ionization events are expected to occur over areas as large as h+ charge carriers are able to spread after they have been stress-activated at depth. Pervasive corona discharges may cause luminous phenomena and may also be the cause for increased noise in the radiofrequency range that reportedly disrupted the telemetric data transfer from a seismometer network in India prior to an earthquake swarm at a distance of about 150 km (Kolvankar, 2001).

Local outbursts of light from the ground, often called EQLs, may be due to a condition in the Earth’s crust that is theoretically predicted to occur when the number density of h+ in the rocks increases to a point where the electronic wavefunctions of the h+ charge carriers begin to overlap, creating a solid state plasma (St-Laurent et al., 2006). Such plasmas are expected to be inherently unstable, leading to a cloud of h+ charge carriers traveling outward at speeds around 200 m s⁻¹ (Freund, 2002). When the wave front breaks through the Earth’s surface, it will ionize the air and produce flashes of light. Depending on conditions yet to be fully understood, such process could also lead to “flames” coming out of the ground (Demetrescu and Petrescu, 1942; Galli, 1910; Mack, 1912) or to outbursts of light (St-Laurent, 2000).

4.6. Fog, haze and cloud formation

While there is a widespread tendency to associate fog, haze and cloud formations prior to earthquakes with air ionization due to radon emission (Araiza-Quijano and Hernández-del-Valle, 1996; Pulinets and Dunajec, 2006; Pulinets et al., 2006), air ionization at the rock-air interface due to stress-activated h+ charge carriers appears to provide a more effective mechanism for the generation of large numbers of airborne ions.

Since most airborne ions will be positive, at least during the early phase of surface potential build-up, they will be repelled by the surface and rise into the air. Acting as nuclei for the condensation of water, they will cause the release of latent heat, which will lead to thermal updrafts. Ionized and conventionally unstable columns of air may thus form “streamers”, which play a role in triggering cloud-to-ground lightning strikes (Aleksandrov et al., 2001).

4.7. Unusual animal behavior

High levels of positive air ions may also have biological consequences. Although effects of positive and negative airborne ions on humans and animals have been studied since the 1960s, the results have not been equivocal due to design faults (Krueger and Reed, 1976), or lack of proper controls (Hedge and Elfehteris, 1982). However, a majority of these studies have found that high levels of positive ions, ranging from 2 × 10⁶ to 1 × 10⁹ ions cm⁻³, have detrimental physiological and behavioral effects, including respiratory problems, possibly due to the swelling of the trachea, and increased levels of serotonin (5-hydroxytryptamine, 5-HT) levels in the blood (Krueger and Kotaka, 1969).

High serotonin levels can result in irritation, headaches, nervousness, and increased sensitivity to pain (Sulman et al., 1974). These effects are consistent with reported pre-earthquake changes in animal behavior. Laboratory data show changes in mouse circadian rhythms 24 h before earthquakes (Yokoi et al., 2003). Other causes quoted in literature for changes in animal behavior before earthquakes include low frequency vibrations, ground tilting, humidity changes, and magnetic field variations (Kirschvink, 2000).

Tributsch (Tributsch, 1982) suggested that changes in animal behavior before earthquakes may be caused by charged aerosol particles due to electrochemical glow. Hoenig (Hoenig, 1979) has shown that electrons and positive ions with current levels up to 10 pA are produced as rocks approach failure under laboratory settings. Although, in both cases the phenomena were attributed...
to piezoelectric effects, it is important to note that piezoelectricity always produces both positive and negative ions. The selective production of positive airborne ions as demonstrated in this paper may amplify the anomalous animal behavior before major earthquakes.

Acknowledgments

Supported by NASA grants "Earth Surface and Interior" #NNX08AG81G and "Exobiology" #NNX07AU04G, by the 2007 NASA Ames Academy, and a 2008 NSF–REU grant to the Department of Physics, SJSU. We thank the staff of the NASA Ames EEL (Engineering Evaluation Laboratory) Jerry Wang, Lynn Hofland, and Frank Pichay.

References


