# The dilatancy-diffusion hypothesis and earthquake predictability

**Citation for published version:**
Main, IG, Bell, AF, Meredith, PG, Geiger, S & Touati, S 2012, 'The dilatancy-diffusion hypothesis and earthquake predictability' Geological Society of London Special Publications, vol 367, pp. 215-230. DOI: 10.1144/SP367.15

**Digital Object Identifier (DOI):**
10.1144/SP367.15

**Link:**
Link to publication record in Edinburgh Research Explorer

**Document Version:**
Peer reviewed version

**Published In:**
Geological Society of London Special Publications

DOI: 10.1144/SP367.15



# The dilatancy-diffusion hypothesis and earthquake predictability

Ian G. Main[1,*], Andrew F. Bell[1], Philip G. Meredith[2], Sebastian Geiger[3] and Sarah Touati[1]

*(1) University of Edinburgh, School of GeoSciences, West Mains Road, Edinburgh.*

*(2) University College London, Department of Earth Sciences, Gower Street, London.*

*(3) Heriot-Watt University, Institute of Petroleum Engineering, Riccarton, Edinburgh.*

**Email: ian.main@ed.ac.uk*

**Abstract.** The dilatancy-diffusion hypothesis was one of the first attempts to predict the form of potential geophysical signals that may precede earthquakes, and hence provide a possible physical basis for earthquake prediction. The basic hypothesis has stood up well in the laboratory, where catastrophic failure of intact rocks has been observed to be associated with geophysical signals associated both with dilatancy and pore pressure changes. In contrast the precursors invoked to determine the predicted

earthquake time and event magnitude have not stood up to independent scrutiny, such that an expert IASPEI panel could find no compelling field-based evidence for earthquake precursors. There are several reasons for the lack of simple scaling between the laboratory and the field scales, but key differences are those of scale in time and space, and in material boundary conditions, coupled with the sheer complexity and non-linearity of the processes involved. 'Upscaling' is recognised as a difficult task in multi-scale complex systems generally, and specifically in oil and gas reservoir engineering, and may provide a clue as to why simple local laws for dilatancy and diffusion do not scale simply to bulk properties at a greater scale, even when the fracture system that might dominate the mechanical and hydraulic properties of the reservoir rock is itself scale-invariant.

## Introduction

Ernest Masson Anderson developed his theory for the structure of faults and fractures primarily from matching observation in nature made by pioneers such as Hutton and Lyell to hypotheses developed by Navier, Coulomb, and Mohr, in the 19$^{th}$ century, citing some early controlled experiments on analogue materials such as layered clay or mastic (Anderson, 1905). Experiments on actual rocks in compression were available in the mining engineering literature by the time of the publication of his book (Anderson, 1942), but largely corroborated the inferences already made. Anderson extrapolated these results more or less linearly to the crustal scale (Fig 1). Some features scale remarkably well, notably the typical orientation of the angle of deformation in shear (controlled by the internal frictional properties) and in tension (opening against the least resistance or minimum principal stress). More recently a much broader range of structural properties of populations of faults and fractures have been shown to scale remarkably well from laboratory failure to crustal scales, as observed directly in field outcrop (Bonnet et al., 2001) or inferred from the scaling of earthquake stress drops and frequency-magnitude scaling (Main, 1996). Given this *structural* scaling, it might at first glance seem natural to assume that other aspects of the *physics* of catastrophic failure will scale linearly from the lab to natural earthquakes, but is this appealing notion really how nature works?

Early papers on the role of dilatancy in the earthquake cycle were based on geological observation (Mead, 1925; Frank, 1965). The dilatancy-diffusion hypothesis itself was developed from the observation of changes in geophysical properties associated with dilatant strain in laboratory tests (Nur, 1972). It was one of the first to be put forward as a physical basis for purported earthquake precursors, also assuming (implicitly) linear scaling of the physics involved from lab to field (Scholz et al., 1973). This paper contains what now seem like wildly over-optimistic statements on the existence of earthquake precursors ('occur before many, and perhaps all earthquakes') and the prospects for earthquake predictability ('the mechanism of premonitory changes appears to lead to prediction which is deterministic rather than probabilistic'), given subsequent experience. However, in science we often learn more from hypothesis failure than confirmation: in a classical example, Einstein's special theory of relativity followed the failure of the hypothesis of the 'ether' as a fixed reference frame for the propagation of light in the Michaelson-Morley experiment. In this paper we re-examine the failure of the dilatancy-diffusion hypothesis with the benefit of hindsight, and suggest new areas to explore in constraining the physics of earthquakes and the prospects for predictability.

The dilatancy-diffusion hypothesis has not yet been validated at a crustal scale, primarily due to the general absence of the predicted dilatancy-related precursors (including seismic velocity, seismicity, electrical conductivity, or radon release, e.g. Jordan et al., 2011). Scholz (1997) has argued instead that the search for dilatancy-related precursors has become biased, with the mainstream community too ready to dismiss or not look for evidence of precursors. However, since then data recorded in real time, even at well-monitored borehole sites near the 2004 Parkfield earthquake rupture (Bakun et al., 2005), as well as other significant events in California (e.g. Loma Prieta, Northridge, Landers, Superstition Hills) have all failed to show any direct evidence for detectable precursory behaviour. Amongst this predominantly negative evidence, Niu et al. (2008) observed two large excursions in the travel-time data that are coincident with two earthquakes (magnitudes 3 and 1 respectively) that are among those predicted to produce the largest coseismic stress changes at the SAFOD drilling site. The two excursions started approximately 10 and 2 hours before the events, respectively. Niu et al. (2008) suggesting that they may be related to pre-rupture stress induced changes in crack properties, as observed in early laboratory studies. More recently satellite interferometry has confirmed more directly the absence of any significant precursory strain recorded at the Earth's surface in the case of the 2009 L'Aquila earthquake in Abruzzo, Italy [Amoruso & Crescentini (2010)].

While the search for precursors continues despite this, we take a different tack here and ask instead how such rigorous 'negative' observations may nevertheless be used instead in a positive way as a significant physical constraint on the actual physics of the process involved. In particular a careful 'upscaling' exercise remains to be done for the dilatancy-diffusion hypothesis, i.e. to take account of differences in loading and sample boundary conditions, spatial and temporal scale, and the material, structural, mechanical and hydraulic complexities involved. For example Nur (1975) pointed out that various forms of dilatancy (microcrack, existing fractures, granular) could be expected in the Earth's seismogenic crust, and they would be expected to have different stress sensitivities. For example the hydrofracture dilatancy reported in the vicinity of some fault zones (Sibson 1981) requires pore pressure in excess of the minumim principal stress ($p>\sigma_1$). This can only be achieved under low levels of differential stress ($\sigma_1 - \sigma_3$) < 4T), where T is the tensile failure stress. This is in direct contrast with the high levels of differential stress required for microcrack dilatancy in the laboratory, and possibly also the high pore pressure (Sibson, 2009) required for microfracture in nature.

Laboratory tests typically utilise intact uniform samples of rock, in order to produce as uniform a stress field as possible in a controlled test. This introduces a kind of ‘sample bias’ or epistemic error not accounted for in linear scaling arguments, because it is not representative even of the small-scale heterogeneity in the Earth. In contrast it is clear that the majority of moderate-to-large crustal earthquakes involve repeated reactivation of existing faults (e.g. Holdsworth et al., 1997) which may have very different properties to those of an intact rock sample.

Another potential source of epistemic error in the application of laboratory-scale experiments to the Earth is the laboratory testing protocol itself, which typically involves increasing the axial stress $\sigma_1$ on a right-cylindrical specimen an constant strain rate under hydraulic compression ($\sigma_2=\sigma_3$). The mean stress and fault frictional strength are therefore increasing with time, corresponding in nature to load-strengthening behaviour during the loading of a reverse fault with $\sigma_3$ vertical. However, during crustal extension the loading of a normal fault to failure involves progressive reduction of $\sigma_3$ while the vertical stress $\sigma_1$ stays fixed. In this case the mean stress and fault strength are decreasing while shear stress on the fault and differential stress are both increasing (load-weakening behaviour). This may help to account for the observation that foreshock activity is more commonly associated with normal faults than with reverse (Abercrombie & Mori, 1996). In the case of strike-slip faults, loading to failure may be either load-strengthening or load-weakening. In the Earth loading typically also involves a concomitant stress relaxation in the minimum stress direction (decreasing $\sigma_3$). Recognising the importance of E.M. Anderson's inferences, some laboratories have examined the effect of 'true' triaxial stresses or rock strength (Haimson & Chang, 2000) and its effect on geophysical properties such as shear-wave birefringence due to aligned microcracks (Crawford et al., 1995).

In addition to these mechanical and spatial scaling arguments, recent laboratory results have demonstrated a systematic decrease in bulk sample dilatancy as strain rates are lowered towards more realistic values for crustal-scale deformation (Heap et al., 2009, fig. 7). The results are consistent with the absence of strong dilatancy-related precursors associated with large earthquakes. A further suite of even slower deformation experiments is planned to test the extrapolation, to fill in an important gap in our understanding of the temporal scaling of brittle-field rheology.

**The Dilatancy-Diffusion hypothesis**

The hypothesis was based on solid and repeatable evidence of primarily mechanical and geophysical precursors to failure in the laboratory, associated with measured changes in sample volume after the yield point in crystalline rocks. Typically this occurs at around or above half of the ultimate strength of the rock sample: dilatancy is a 'high-stress' phenomenon. Such bulk dilatancy, due to microcracking of the type shown also in sedimentary rocks (Figure 1), was associated in the laboratory with changes in seismic velocity, electrical resistivity, and acoustic emission event rate and the scaling of event size, expressed by the exponent *b* (the '*b*-value') in the Gutenberg-Richter relation for the frequency $F$ of events of magnitude $m$, of the form $F(m)=a-bm$. We might then expect such sample dilatancy to affect pore fluid volume and/or pressure, depending on the permeability of the medium, the local strain rate, and the experimental boundary conditions.

At sufficiently high volumetric strain rates dilatancy in a relatively impermeable crystalline rock in the Earth's subsurface would be expected initially to produce a local decrease in pore pressure, and a concomitant increase in the effective normal stress, resulting in material hardening and delaying failure (Paterson & Wong, 2005). Implicit in this scenario is that the rate of dilatancy (volumetric strain rate) must remain higher than that which will allow pore water to diffuse into the new cracks to restore the pore pressure and induce concominant material softening. In practice the low strain rates at the onset of dilatancy means there will be a finite lag time between the onset of dilatancy and local pore pressure reduction. Assuming a supply of fluid from outside the dilatant zone, and a deceleration in the rate of dilatancy associated with the hardening effect, the drop in pore pressure would be followed by a slow recovery by fluid flow from the surrounding undilated region. This recovery would ultimately trigger dynamic failure.

Scholz et al (1973) presented no direct measurements of fluid pressure variations from the laboratory, but instead inferred such a decrease then recovery, solving a simple diffusion law for transient pressure recovery in a spatially uniform medium with a constant diffusivity to estimate the duration of the recovery time. The hypothesis predicted systematic *qualitative* changes in geophysical signals associated with stages of elastic loading, dilatant yield, pore pressure recovery, dynamic failure, and post-seismic relaxation as illustrated in Figure 4.

The basic coupled process has been replicated to some extent (though not exactly) in the laboratory, including contemporary measurement of actual pore pressure change and its impact on seismic (acoustic emission) precursors under constant strain rate loading. Figure 2 (from Sammonds et al., 1992) shows an example of two tests, one nominally 'dry' and one completely saturated in a constant volume of water, held at pressure in the sample under 'undrained' conditions, with sample boundaries sealed to fluid flow in either direction. The dry sample shows an acceleration in event rate (related to *a*) and a decrease in the seismic *b*-value associated with an increase in stress. Similar behaviour is seen in a 'drained' test held at constant boundary pore pressure, allowing fluid flow at the sample boundary (fig. 2c in Sammonds et al., 1992). In the undrained test of Figure 2b the pore pressure, measured at the sample boundary, first increases due to crack and pore closure associated with an increase in mean stress, and then decreases up to the failure time due to shear-enhanced dilatancy. The inferred dilatancy hardening with zero-permeability boundary conditions does indeed significantly extend the post-peak stress deformation phase and delay the failure time. No pore pressure recovery is seen because of the sealed boundary, but the event rate flattens off and the *b*-value recovers in an extended strain-softening phase, before dropping to a minimum at the final stage near dynamic failure. We might imagine an experimentally-challenging test with intermediate boundary conditions (contemporaneous change in both pore fluid volume and pressure) that would show intermediate behaviour between the drained and undrained extremes, but this has (to the authors' knowledge) yet to be done: the 'diffusion' or final pore pressure recovery phase has yet to be demonstrated in such an open system in the laboratory.

The dilatancy-fiffusion hypothesis is based on the assumption of a finite-sized 'preparation zone' within which microcrack damage is occurring, and that the size of the preparation zone is related to the eventual size of the mainshock. A preparation zone is well-defined in the laboratory by the sample boundaries, but remains elusive in field data: behaviour identified after the earthquake as anomalous is often available only at one or at most a handful of selected sites, even for spatially very extensive mainshocks. In some cases this has been argued to be a consequence of low instrument density. To get round this problem, Scholz et al. (1973) estimated the size of the upcoming event (based on the extent of the aftershock zone) from its correlation with and the duration of the reported precursor, based on the literature then available. Interestingly, this correlation could be explained by the duration of the inferred diffusive pore pressure recovery phase in a uniform medium, albeit with an inferred diffusivity higher than a typical laboratory test for a low-porosity crystalline rock under similar pressure conditions. The hypothesis remains unproven because the predicted precursors failed to materialise convincingly in a consistent and reliable way in field evidence (Wyss and Booth, 1997; Bakun et al., 2005).

At this point it is useful to note that the notion of dilatancy-diffusion does have an important bearing on dynamic failure processes. Rudnicki & Chen (1998) developed a coupled model to explain how rapid frictional slip may be stabilised on an otherwise weakening fault by dilatancy hardening. Under constant flow rate boundary conditions the same coupled model predicts a dynamic 'suction pump' effect, where fluids are actively channelled into the zone of rapid pore pressure drop in dilating fault zone. The results of a numerical model for this dynamic effect compare favourably with those of a laboratory experiment at similar conditions (Grueschow et al., 2003). The suction generated by dynamic dilatancy in the fault zone is manifest by a drop in the inlet pressure required to push fluid in at a constant rate at the sample boundary (Fig. 3). Such seismic pumping, repeated over many cycles, is consistent with the observation of mineral deposits formed by episodic channelling of hydrothermal fluids along faults and fractures in meso-thermal conditions (Sibson et al., 1975), but this interpretation is not unique (e.g. Sibson, 1981; 2001). In any case such *dynamic* effects appear to scale better between the laboratory and the brittle Earth than the *quasi-static* loading phase where precursors might be expected. However, this dynamic coupling is consistent with short-lived dilatancy concentrated very near the fault zone, rather than the longer-term quasi-static regional dilatancy invoked by Scholz et al. (1973), or related theories based on seismic anisotropy and extensive fracture dilatancy outside the nominal mainshock 'preparation zone' (Crampin et al., 1984).

**The flawed search for earthquake precursors**

There has been much discussion of this issue in the literature, and only a brief summary can be given here. An excellent and accessible summary of the repeated conflict between an otherwise reasonable hypotheses and data, along with an interesting and very relevant discussion of the social, human and even political dimensions that are very much part of the story, is given by Hough (2009). Following the most comprehensive study to date by an expert panel convened by the International Association for Seismology and Physics of the Earth's Interior (IASPEI), Wyss and Booth (1997) concluded that there were no candidate precursors that satisfied all of the criteria set by the panel for a physically and statistically reasonable precursory signal (for example any anomaly must be seen at more than one site to be acceptable). This means the 'precursor' durations used to determine the magnitude correlation by Scholz et al. (1973), and hence the inferred quantitative value of the fluid pressure diffusion constant, were based on questionable published data.

The correlation between a reported fluctuation in a geophysical parameter and a subsequent earthquake itself may have other more mundane causes, for example retrospective selection bias in a noisy signal (Mulargia, 2001) as illustrated in the example shown as a tutorial in Fig 5. Fig 5(a) reproduces a figure from Scholz et al. (1973) cited as evidence for changes in event rate *a* and in the scaling exponent *b* prior to a magnitude 3 earthquake. At first glance this selected data seems to show a convincing minimum and recovery in event rate prior to the magnitude 3 mainshock time identified on the diagram, consistent with the predictions of the dilatancy-diffusion model of Fig 4. For reference, Fig 5 (c) shows fluctuations in event rate for a random (Poisson) process with a similar average number (50) of events per day, sampled at 12-hour intervals as in Fig 5(a). Fig 5(c) illustrates the large relative fluctuations expected from simple counting errors of the number of events in a random process with this average, and the tendency to cluster rather than produce the flat graph expected for an infinitely-sampled process. Fig 5(b) is a blow-up of one of the minima in Fig 5(c), illutrating how minima such as Fig 5(a) could occur simply by finite sampling of a random process. In another example Main et al (2008) showed that the non-linear statistics of seismicity (exemplified by the Gutenberg-Richter law and exacerbated by earthquake triggering not considered in Fig 5) can lead to very large samples (several thousand) being required even to get a stable value of average total event rate and its standard deviation. Therefore the simplest interpretation consistent with the data of Fig 5(a) is a finite (small) sample of a random process, with one of several candidate magnitude 3 earthquakes selected retrospectively within a time window that effectively introduces two additional free parameters (start time and end time) to the search.

In fact most candidate precursors fail as potential predictors because of poor hypothesis testing protocols, notably examining and selecting data in retrospect and not accounting for the resulting sample bias in assessing the significance of any correlation identified. Such data selection is perfectly valid in developing a hypothesis, but not in testing or validation. As a consequence clinicians developed the prospective ‘double-blind’ test as the standard, and only acceptable, method of testing in medical sciences (Modell & Houde, 1958). Ultimately any hypothesis must be put at risk in a situation where the outcome is not known a priori. In our case this means actual forecasting in real time is needed to evaluate fully the significance of any precursor and its quantitative impact on earthquake predictability. This aspect has now been fully embraced by the global seismological community, with a range of regional testing centres now set up by the Collaboratory for the Study of Earthquake Predictability (http://www.cseptesting.org/).

**Using negative evidence as a constraint**

The dilatancy-diffusion hypothesis was based on direct evidence of dilatant strain in laboratory samples. In the laboratory dilatant strain can be measured by strain gauges placed directly on the sample, or more recently through changes in the volume of the pore fluid or the fluid confining medium. Modern satellite interferometry data (the synthetic aperture radar technique) now provide extremely sensitive measurements of strain at the Earth’s surface. For example during the two years before the $M_W$ 6.3 earthquake struck the city of L'Aquila, Italy on April 6, 2009, no anomalous precursory strain larger than a few tens of nano-strain units is visible, limiting the volume of the possible earthquake ‘preparation zone’ to less than 100 $km^3$ (Amoruso & Crescentini, 2010), or a linear dimension of 4.6 km. This is much smaller than the 10-20 km or so rupture length for a magnitude 6.3 mainshock, calling into question the generality of the notion of a ‘preparation zone’ similar to the sample dimensions of a laboratory test.

Even seconds before the 2009 L'Aquila earthquake, "strain is stable at the $10^{-12}$ level and pre-rupture nucleation slip in the hypocentral region is constrained to have a moment less than $2 \times 10^{12}$ Nm, i.e. 0.00005% of the main shock seismic moment". Assuming a scale invariant strain change with a typical stress drop of 30 bar for a continental earthquake or $10^{-4}$ units of strain for the earthquake itself, the nucleation zone is restricted to a scale length of at most 100m, and this likely at typical earthquake nucleation depths of 10 km or so. At this localised scale, likely related to re-fracture of a healed, locked asperity, we might expect to see the same physics as we observe in a laboratory test, but this is going to be hard to detect. Clearly the nucleation zone is much, much smaller than the eventual rupture, and the two need not be directly related.

While precursory dilatant strain has not yet been observed directly and systematically for continental earthquakes, there is evidence that post-seismic strain relaxation is clearly visible, for example following the 26 December 2003 Bam earthquake in Iran (Fielding et al., 2009). Using satellite-based InSAR observations, and after accounting for poro-elastic effects, they identify a localised zone of dilatant strain recovery near (within 200m or so) the centre of the mapped fault trace where the co-seismic slip was greatest. Such dilatancy is therefore much more likely to be due to co-seismic dilatancy of the type modelled by Rudnicki & Chen (1988) and observed in the laboratory by Grueschow et al. (2003), rather than any residual memory of precursory dilatant strain. This confirms the inference that actual precursory dilatant strain is quantitatively much less, and/or much more highly localised than the bulk behaviour of a laboratory test.

While more difficult to measure and subject to much debate, the inferred shear stresses involved in crustal loading prior to earthquake rupture could also provide a constraint on the type and amount of dilatancy that might be expected in the seismogenic crust (e.g. Brune & Thatcher, 2003). One view holds that the ambient effective differential stresses are low, on the order of 100 bars or less, implying almost total stress relaxation during rupture. Another holds that shear stresses are high, on the order of 1 kbar or more, comparable to those where dilatancy is seen in crystalline rocks in the laboratory. The absence of a clear dilatancy signal from microcracking around seismogenic faults is consistent with relatively low-stress (shear stress < 100 bar) rupturing on existing, relatively weak, structures.

In summary direct observation of dilatant strain implies that the dilatancy-diffusion process does apply well, and on a large scale, to the co-seismic and post-seismic phases, and may apply to earthquake nucleation on a very small scale up to a few hundred m. This geodetic constraint is supported by recent seismic evidence (Bouchon et al., 2011) of an accelerating signal concentrated on a very localised zone, identified by cross-correlation techniques prior to the 1999 $M_w$ 7.6 Izmit (Turkey) earthquake. The signal consisted of a succession of small foreshocks in the form of repetitive seismic bursts, accelerating with time in the 2 minutes preceding the event, and increased low-frequency seismic noise in the 44 minutes preceding the event. Any one of these foreshocks is located within 20 m or less from the majority of the other events, comparable to the size of the largest events (25m). These results confirm a very short duration, very localised, but nevertheless detectable nucleation phase for this event. Modern techniques of data assimilation applied to continuously-recorded broadband seismic data will be required to confirm the generality or otherwise of this intriguing observation.

## Up-scaling of a complex system in space and time

Ultimately the reason for the failure of the dilatancy-diffusion hypothesis to scale simply to crustal processes is due to the complexity and non-linearity of the processes, and large differences in space and time between the laboratory and the field case. We address these issues separately below.

### *(a) Complexity and predictability*

In a laboratory test such as illustrated in Fig 2 the sample is initially chosen for its uniformity, and loaded first by increasing the isotropic stress (axial stress and confining pressure) to a given level, and then by increasing the axial stress alone at a constant strain rate to change the differential stress from zero. In this sense the sample is loaded from a very sub-critical state (zero differential stress) to a more critical (high-stress) state near the dynamic failure time. However, in the Earth the spatial structure is highly heterogeneous, and the tectonic stress maintains the system perpetually in a state much nearer its critical value than the starting conditions of such laboratory tests, making the system much more sensitive to small stress perturbations. Amongst other drivers, such complexity has led to a completely alternative view on earthquake mechanics proposed by Bak et al. (1987), who postulated that earthquakes occurred in a state of self-organised criticality. This hypothesis neatly explained much of the phenomenology of earthquakes, including the Gutenberg-Richter law, the scale-invariant distribution of faults, the relatively low and constant stress drop, the ease with which small natural and man-made stress perturbations can induce earthquakes, and the long-term stationarity inferred for seismic hazard calculation (Main, 1995, 1996). Unfortunately this came at the expense of degraded predictability – the size of an event in a near-critical system is determined by small details of the avalanche-like response, so that event size would be primarily determined during, and not before, the event (Main, 1995). This is consistent with the small size of the nucleation patch inferred by Amoruso and Crescentini (2010) and Bouchon et al. (2011) described above. These inferences and observations are all consistent with the relatively low correlation between magnitudes estimated from the early part of the seismogram and the eventual magnitude of the earthquake used in earthquake 'early-warning' systems, including the recent Mw 9.0 tsunamogenic earthquake in northeastern Japan (Cyranoski, 2011). The notion of self-organised criticality, or near-criticality, implies that any hope for deterministic prediction of earthquakes is remote (Main, 1997). Nevertheless the finite (albeit small) stress drop of earthquakes implies a slightly sub-critical system, where a small but finite degree of forecasting power might be expected, albeit of a probabilistic nature (see Nature website debate at http://www.nature.com/nature/debates/earthquake/equake_frameset.html and a discussion on the role of dissipation on maintaining a near-but-subcritical state in Main & Naylor, 2008).

Such self-organised 'near but strictly *sub'*-criticality is also consistent with recent data from earthquake repeat times in palaeoseismic data from the San Andreas fault (Scharer et al., 2010), which show to first order the temporally random recurrence of a purely critical system, but a second-order quasi-periodic component to the stress renewal process expected from a system with finite stress drop. Any probability gain due to this effect is therefore extremely subtle. In retrospective mode quasi-periodic renewal models have been suggested to provide a factor 2-5 probability gain over a temporally random process (Imoto, 2004), but this is likely to be an upper bound to a true prospective forecasting scenario. This marginal probability gain has led to its effectiveness as an operational tool being questioned both in California (Chui, 2009) and in Japan (Geller, 2011).

Much higher probability gains (over background) are possible with space-time clustering associated with earthquake triggering in a system operating near its critical point. For example the long-term background seismic risk (of more than 100 fatalities) in the L'Aquila area is on the order of $10^{-6}$ per day (van Stiphout et al., 2010). From the clustering properties of swarm activity preceding the 2009 *M* 6.3 L'Aquila earthquake, they estimated this probability was increased by a factor 30 or so prior to the mainshock. To put these numbers into perspective, the typical estimated probability of dying in an earthquake for an individual person in the next 24 hour was temporarily elevated to $10^{-9}$, whereas the average probability of dying in a car accident in Italy in any 24 hours period is $2.7*10^{-9}$. Taking this a step further van Stiphout et al. (2010) developed a quantitative cost-benefit analysis to demonstrate that the negative consequences of an evacuation (integrated over all such swarms, including the many false alarms, and funds diverted from other potential life-saving activities) outweighed the positive benefits. Dealing with such high probability gain, but still low-probability forecasts, remains a generic subject of research at the interface between the natural and social sciences (http://www.protezionecivile.it/cms/attach/ex_sum_finale_eng1.pdf)

**(b) Scaling in space**

Like resistivity, hydraulic diffusivity (or the related permeability) spans a huge number of scale ranges. For example Fig 6(a) shows calculations of the flow velocity, in a real fracture network mapped in detail at the surface where the fractures are 1 mm wide (Geiger & Emmanuel, 2010). These range from $10^{-16}$ m s$^{-1}$ to $10^{-4}$ m s$^{-1}$. A reservoir engineer must then estimate a representative single permeability from a block of this size - shown in Fig 6(b) - typically comprising a horizontal resolution of approximately 100 m and a vertical one of approximately 10m.

Because it is computationally not feasible to perform reservoir simulations at the resolution of individual fractures and/or sedimentary layers, major research efforts have been dedicated to develop best practices for computing effective coarse-scale permeabilities that preserve the average fine-scale flow behaviour through individual sedimentary beds (c.f. Christie 1996, 2001; Renard and de Marsily, 1997). Common analytical methods in reservoir engineering include arithmetic permeability averaging (for flow parallel to layering), harmonic permeability averaging (for flow perpendicular to layering), and geometric permeability averaging (for flow in randomly correlated permeability fields). Flow-based permeability averaging can be used to compute the effective permeability of more complex geological structures. Here, a steady state pressure field is computed for a sub-section of the reservoir model using the fine-scale permeability field and known boundary conditions to obtain the total volumetric flux through the model. Using the total flux through the model and the known boundary conditions, the average permeability can be computed straightforwardly from Darcy's law although the final value will be sensitive to the applied boundary condition (e.g., no-flow vs. leaky boundaries parallel to the main flow direction). It has become increasingly common to estimate the error introduced by upscaling a priori by computing a measure for heterogeneity (usually containing permeability, porosity, and flow rate) in all sedimentary layers comprising the geological model and comparing it on a layer-by-layer basis (King et al., 2006). This allows the reservoir engineer to generate optimised simulation grids that non-uniformly group different geological layers of similar heterogeneity while preserving others that have a major impact on flow; coarsening the detailed geological model uniformly, for example by grouping every ten vertical layers of the geological model into one single layer for the flow simulation model, provides little control on the upscaling error. Other methods to validate the quality of upscaling include streamline comparisons between the original fine-scale geological model and coarse-scale flow simulation model (Samier et al., 2002). Hence, the common view among reservoir engineers is that sedimentary heterogeneities in clastic rocks can be upscaled reliably and classical benchmark studies such as the $10^{th}$ SPE Comparative Solution Project (Christie and Blunt, 2001) appear to confirm this view: here it has been demonstrated that a fine-scale geological model of a fluvial North Sea reservoir containing over $10^6$ cells can be simulated equally well with a wide range of upscaled flow models containing between $7x10^4$ and $10^3$ cells. It needs to be pointed out that the upscaled flow models used vastly different upscaling methods and different simulators, which allowed to fine-tune an upscaled flow model for a certain simulator; if on the other hand different upscaling methods and simulators are applied to a flow model with fixed number of grid cells, then results varied significantly.

However, upscaling remains a fundamental problem if fractures are present in the reservoir and pose the philosophical question: can something as inherently discrete as the fracture network of Fig 6(a), which is embedded in a permeable rock matrix, be described by a 'representative elemental volume' that can be modelled by a continuum theory with averaged parameters such as fluid pressure diffusion such that all relevant time- and length scales, spanning several orders of magnitude, are retained? Formally the answer is no – the correlation length of the fractures is much larger than the block size, and may even approach crustal scales, based on evidence from borehole logs (Dolan et al., 1998; Berkowitz 2002). Still, for practical purposes, fracture networks are upscaled in reservoir simulations using the so-called Oda's method, which attempts to compute an effective permeability tensor of the fracture network based on the aperture and connectivity of the individual fractures (Dershowitz et al., 2000). The effective fracture permeability is then employed in a dual-porosity simulation, which assumes that the fractures comprise the flowing domain of the porous media while the rock matrix is stagnant and provides the fluid storage (c.f. Warren and Root, 1963). Fluid exchange between fracture and matrix is modelled using a transfer function, which attempts to account for the physics of the fracture-matrix fluid transfer and the geometry of the fracture network. The dual-porosity approach can be extended to simulate geomechanical effects where changes in fracture permeability and matrix porosity are computed individually (Bagheri and Settari, 2008). Yet, defining the appropriate scale, i.e. grid cell size in the flow simulation model, to compute the effective fracture permeability tensor remains a major challenge because it must preserve the connectivity and permeability of the original fracture network, which evolves if the fractured porous media is deforming; hence Dershowitz et al. (2002) concluded that “if there is no grid cell scale that can reproduce the connectivity of the [discrete fracture network], [dual porosity] continuum simulation results, will need to be treated with caution”.

The ‘up-scaling’ problem is exacerbated by the fact that the detailed information from the total geological exposure of a fracture system observed at the earth’s surface and modelled in Fig 6(a) is not available, and an initial estimate of bulk permeability must be made from very limited and quasi-1D data available from core samples and logs in well-bores. As a result the up-scaled permeability estimate for a reservoir block in Fig 6(b) is often obtained empirically, by combining the qualitative geological and structural interpretation of the reservoir with quantitative geo-statistical simulation of the flow field (i.e., permeability and porosity fields), conditioned to the sparse field data (e.g., Strebelle, 2002), and finally calibrated to the observed flow rates by complex history matching algorithms (e.g., Oliver & Chen, 2011). Practically such models are used by engineers to manage the hydrocarbon field by changing fluid injection or production rates, and shutting in or drilling new wells. Interestingly reservoir engineers are also beginning to realise that history matching is insufficient – and moving towards predictive tests (e.g., Christie et al., 2006; Heffer et al., 2010).

### (c) Scaling in time

Earthquakes occur at strain rates that are several orders of magnitude lower than those achievable in the laboratory. A typical laboratory test at constant strain rate loading is on the order of $10^{-5}s^{-1}$, and a very slow 'creep' test (loaded at constant stress) may take a few weeks or months, slowing the strain rate down to $10^{-8}s^{-1}$ or so. In contrast earthquake strain rates in continental zones occur under regional strain rates on the order of $10^{-15}s^{-1}$ (Jackson & McKenzie, 1988) though locally these can be higher ($10^{-12}$ $s^{-1}$; Sibson, 1982). It is therefore quite possible that different physical, and physico-chemical, processes may actually be involved across these enormous scale ranges.

To illustrate this Figure 7 shows recent results (Ojala et al., 2004) from a suite of experiments at different strain rates aimed at determining the process of acceleration to failure due to the mechanism of stress corrosion associated with dissolution of silica in a sandstone sample (Ojala et al., 2003). In the quasi-static phase the acoustic emission event rate shows a systematic acceleration to failure of the near-asymptotic form $a{=}a_0(t_m{+}\mathrm{c}{-}t/t_m)^{-p'}$, where $t_m$ is the mainshock (dynamic failure) time, $p'$ is a positive exponent and $c$ is a characteristic time that keeps the event rate finite at the main event time. This form is consistent with the predictions from accelerated stress corrosion cracking (Main, 1999; 2000), where $p' = 1{-}2/(n\text{-}2)$ for event rate (assuming event rate is proportional to crack growth rate) and the stress corrosion index $n$ is defined by the empirical observation that the velocity of subcritical crack growth scales as the $n$'th power of the stress intensity (Charles' law: Meredith & Atkinson, 1983). Figure 7 shows that the absolute time for a warning that an acceleration has started (in real time) decreases systematically as the strain rate decreases, from a few minutes at a strain rate of $10^{-5}$ s$^{-1}$ to a few tens of seconds at $10^{-8}$ s$^{-1}$. This may be due to variations in the rate-limiting step for the process (e.g. the diminishing effect of the slow transport rate of reactive fluid to the fresh crack tip, Atkinson, 1987). The net effect is that the system becomes more non-linear, with a shorter detectable precursor duration, and overall less predictable in real time, as the strain rate decreases. This non-linearity in the normalised event rate occurs because more acoustic events are concentrated later in the loading history – the absolute number of events remains relatively insensitive to strain rate (Ojala et al., 2004).

Fujii et al. (1998) provide a more direct clue as to the diminishing role of dilatant strain at slower strain rates. By carrying out constant strain rate loading experiments between $10^{-8}$ s$^{-1}$ and $10^{-3}$ s$^{-1}$ on Kimachi sandstone (their fig. 13), Inada granite and Noboribetsu tuff (their fig. 15), they showed a systematic decrease in the critical tensile strain (most strongly associated with dilatancy) with lower strain rate. For example the tensile strain decreases in Inada granite from 0.075% to 0.060% between $10^{-4}$ s$^{-1}$ and $10^{-8}$ s$^{-1}$ respectively, albeit with a large uncertainty due to sample variability of 0.5%. This is consistent with the much lower dilatant volumes that can be inferred from Amoruso and Crescentini's (2010) results in that single example.

In summary the signals from dilatant strain associated with with microcracking diminish systematically as the strain rate diminishes, even under laboratory conditions. Likewise the inferred much lower ratio of dilatancy (volumetric strain) rate to volumetric fluid flow rate at the very low strain rates applicable in the Earth is also likely to reduce the dilatancy hardening effect invoked to explain the duration of earthquake precursors. Both processes may contribute significantly to the lack of scaling of the dilatancy-diffusion process from the laboratory to the field case, with associated degradation of predictability.

## Conclusion

The dilatancy-diffusion hypothesis as originally proposed failed for several reasons, largely associated with the validity of the assumptions on which it was based. Primarily the anecdotal data used as evidence of geophysical precursors similar to those observed in the laboratory has not stood up to subsequent more rigorous testing: systematic precursors remain elusive, and the community has now moved on to true prospective testing of earthquake forecasting, specifically in order to avoid the retrospective selection bias inherent in previous literature. The hypothesis assumed a linear scaling (after renormalisation of the parameters) of the physics from tests of small uniform lab tests with well-defined boundaries, loaded from zero to critical shear stress intensities, whereas the Earth is much more complex, has no such clear boundaries, and is maintained by plate tectonics in a near-critical effective stress state with relatively small stress fluctuations between events and a strong sensitivity to even smaller stress perturbations. Nevertheless quasi-periodic stress 'renewal' models can provide a statistically-significant, though small in absolute terms, probability gain over a purely random process.

The concept of a large-scale 'preparation zone', indicating the likely magnitude of a future event, remains as ethereal as the ether that went undetected in the Michaelson-Morley experiment. There appears to be little correlation even of aspects of the early part of rupture with the eventual magnitude of an event, consistent with the complexity and the critical or domino-like cascade or the rupture process. In contrast, recent geodetic and seismic data reveal in some cases the existence of a small but finite nucleation zone of around a few hundred m at earthquake nucleation depths. Perhaps the basic features of precursory dilatancy and diffusion do occur locally on this scale, but are essentially impractical to detect reliably.

A significant upscaling exercise in space and time is needed to account for the multi-scale physics involved in extrapolating from laboratory tests to crustal scales, involving large scale computational simulation to handle the several orders of magnitude differences in spatial and temporal scales. New laboratory tests at very slow strain rates are needed to bridge the gap to natural ones, and to explore of the effect of the ratio of volumetric strain rate to volumetric fluid flow rate on the coupled behaviour in the precursory phase.

Finally the hypothesis is not a complete failure: coupled dilatancy-diffusion processes remain a prime candidate for coseismic and post-seismic processes localised on or near the fault rupture plane and validated by geodetic and geological observation, and perhaps for localised nucleation processes inferred or constrained by geodetic or seismic data. These are interesting topics for study, irrespective of their implications for earthquake forecasting.

**Acknowledgements**

During the writing of this paper Ian Main was funded by a Royal Society of Edinburgh Research Fellowship, Andrew Bell by NERC grant NE/H02297X/1, and Sarah Touati by an EPSRC studentship. We are grateful to two anonymous reviewers for their constructive and insightful comments on a previous manuscript.

## References

Abercrombie R.E. & J. Mori 1996. Occurrence patterns of foreshocks to large earthquakes in the western United States, Nature 381, 303-307.

Amoruso A. & Crescentini L., 2010. Limits on earthquake nucleation and other pre-seismic phenomena from continuous strain in the near field of the 2009 L'Aquila earthquake, Geophysical Research Letters, 37, L10307.

Anderson, E.M. 1905. The dynamics of faulting. Trans. Edin. Geol. Soc. 8, 387-402.

Anderson, E.M. 1942. The dynamics of faulting, Oliver & Boyd, Edinburgh.

Atkinson, B.K., 1987. Fracture mechanics of rock, B.K. Atkinson (ed.), Academic press, London, 1-23.

Bagheri, M. & A. Settari, 2008. Modeling of geomechanics in naturally fractured reservoirs. SPE Reservoir Evaluation and Engineering, 11(1), 108-118.

Bak, P. & C. Tang, 1989. Earthquakes as a self-organized critical phenomenon, J. Geophys. Res., 94, 15,635-15,637.

Bakun, W.H. et al., 2005. Implications for prediction and hazard assessment from the 2004 Parkfield earthquake, Nature 437, 969-974.

Berkowitz, B., 2002. Characterizing flow and transport in fractured geological media: A review. Advances in Water Resources, 25, 861-884.

Bonnet, E., O. Bour, N.E. Odling, P. Davy, I. Main, P. Cowie & B. Berkowitz, 2001. Scaling of fracture systems in geological media, Rev. Geophys. 39, 347-383.

Bouchon, M., H. Karabulut, M. Aktar, S. Özalaybey, J. Schmittbuhl & M.-P. Bouin, 2011. Extended Nucleation of the 1999 Mw 7.6 Izmit Earthquake, Science, 331, 877-880.

Brune, J.N. & W. Thatcher, 2003. Strength and energetics of active fault zones, International Handbook of Earthquake and Engineering Seismology, Academic Press, v. 81A, pp. 569-588).

Christie, M.A., 1996. Upscaling for reservoir simulation. Journal of Petroleum Technology. November, 1004-1010.

Christie, M.A., 2001. Flow in porous media – scale up of multiphase flow. Current Opinion in Colloid & Interface Science, 6, 236-241.

Christie, M.A. & M.J. Blunt, 2001. Tenth SPE comparative solution project: A comparison of upscaling techniques. SPE Reservoir Evaluation and Engineering, 4(4), 308-317.

Christie, M.A., V. Demyanov & D. Erbas, 2006. Uncertainty quantification for porous media flows. Journal of Computational Physics, 217, 143-158.

Chui, G., 2009. Shaking up earthquake theory, Nature 461, 870-872.

Crampin, S., Evans, R. & B.K. Atkinson, 1984. Earthquake prediction: a new physical basis, Geophys. J. R. Astron. Soc. 76, 147-156.

Crane R. & D. Sornette, 2008. Robust dynamic classes revealed by measuring the response function of a social system, Proceedings of the National Academy of sciences of the USA, 105, 41, 15649-15653.

Crawford, B.R., B.G.D. Smart, I.G. Main & F. Liakopoulou-Morris (1995). Strength characteristics and shear acoustic anisotropy of rock core subjected to true triaxial compression, *Int. J. Rock Mech. Min. Sci*. 32, 189-200.

Cyranoski, D., 2011. Japan faces up to failure of its earthquake preparations, Nature 471, 556-557.

Dershowitz, B., P. LaPointe, T. Eiben & L. Wei, 2000. Integration of discrete fracture network methods with conventional simulator approaches. SPE Reservoir Evaluation and Engineering, 3(2), 165-170.

Dolan, S.S., C.J. Bean & B. Riollet, 1998. The broad-band fractal nature of heterogeneity in the upper crust from petrophysical logs, Geophys. J. Int. 132 489-507.

Fielding, E.J., P. R. Lundgren, R. Bürgmann, & G. J. Funning, 2009. Shallow fault-zone dilatancy recovery after the 2003 Bam earthquake in Iran, Nature 458, 64-68.

Frank, D.C.,1965 : On dilatancy in relation to seismic sources. Reviews of Geophysics and Space Physics 3: 485-503.

Geiger, S. & S. Emmanuel, 2010. Non-fourier thermal transport in fractured geological media. Water Resources Research, 46, W07504.

Geller, R.J., 2011. Shake-up time for Japanese seismology, Nature, doi:10.1038/nature10105 Comment.

Grueschow, B., O. Kwon, I.G. Main & J.W. Rudnicki, 2003. Observation and modelling of the suction pump effect during rapid dilatant slip, Geophys. Res. Lett., 30(5), 1226.

Haimson B. & C. Chang, 2000. A new true triaxial cell for testing mechanical properties of rock, and its use to determine rock strength and deformability of Westerly granite, Int. J. Rock Mech. Min. Sci. 37, 285-296.

Heap, M. J., P. Baud, P.G. Meredith, A. F. Bell, and I. G. Main, 2009. Time-dependent brittle creep in Darley Dale sandstone, J. Geophys. Res., 114, B07203.

Heffer, K.J., Greehough, J. & I.G. Main, 2010. The accuracy of production forecasts: conventional models versus a statistical model for short-term predictions allied to reservoir monitoring and characterization, in proc., Petex

2010, London, 23-25 November (extended abstract).

Holdsworth, R.E., C.A. Butler, A.M. Roberts, 1997. The recognition of reactivation during continental deformation, J. Geol. Soc. 154, 73-78.

Hough, S., 2009. Predicting the Unpredictable: The Tumultuous Science of Earthquake Prediction, Princeton University Press, 272pp.

Imoto, M., 2004. Probability gains expected for renewal process models, Earth Planets Space, 56, 563–571.

Jackson, J. & D. Mckenzie, 1988. The relationship between plate motions and seismic moment tensors, and the rates of active deformation in the Mediterranean and middle-east, Geophys. J. 93 45-73.

Jordan, T., Y.-T. Chen, P. Gasparini, R. Madariaga, I. Main, W. Marzocchi, G. Papadopoulos, G. Sobolev, K. Yamaoka & J. Zschau, 2011. Operational earthquake Forecasting: State of Knowledge and Guidelines for Utilization, Annals of Geophysics 54, 4, 316-391.

King, M.J., K.S. Burn, P. Wang, V. Muralidharan, F. Alvarado, X. Ma & A. Datta-Gupta, 2006. Optimal coarsening of 3D reservoir models for flow simulation. SPE Reservoir Evaluation and Engineering, 9(4), 317-334.

Main, I.G., 1995. Earthquakes as critical phenomena: Implications for probabilistic seismic hazard analysis, Bull. Seismol. Soc. Am. 85, 1299-1308.

Main, I.G., 1996. Statistical physics, seismogenesis, and seismic hazard, Rev. Geophys. 34, 433-462.

Main, I.G., 1997. Earthquakes - long odds on prediction, Nature 385, 19-20

Main, I.G., 1999. Applicability of time-to-failure analysis to accelerated strain before earthquakes and volcanic eruptions, Geophys. J. Int. 139, F1-F6.

Main, I.G., 2000. A damage mechanics model for power-law creep and earthquake aftershock and foreshock sequences, Geophys. J. Int. 142, 151-161.

Main, I.G & M. Naylor, 2008. Maximum entropy production and earthquake dynamics, Geophys. Res.Lett. 35, L19311.

Main, I.G., L. Li, J. McCloskey & M. Naylor, 2008. Effect of the Sumatran mega-earthquake on the global magnitude cut off and event rate, Nature Geoscience 1, p142.

Mair, K., I.G. Main & S.C. Elphick, 2000. Sequential development of deformation bands in the laboratory, J. Struct. Geol. 22, 25-42.

Mead, W.J., , 1925. The geologic role of dilatancy, J. Geol., 33, 685-698.

Meredith, P.G. & B.K. Atkinson, 1983. Stress corrosion and acoustic emission during tensile crack propagation in Whin sill dolerite and other basic rocks, Geophys. J. R. Astr. Soc. 75, 1-21.

Modell W. & R.W. Houde, 1958. Factors influencing clinical evaluation of drugs - with special reference to the double-blind technique, *J. American Med. Assoc.*, 167, 2190-2199.

Mulargia, F., 2001. Retrospective selection bias (or the benefit of hindsight), Geophysical Journal International 146, 489–496.

Niu, F., P. G. Silver, T. M. Daley, X. Cheng & E. L. Majer, 2008. Preseismic velocity changes observed from active source monitoring at the Parkfield SAFOD drill site, Nature 454, 204-208.

Nur, A., 1972. Dilatancy, Pore Fluids, and Premonitory Variations of ts/tp Travel Times, Bull. Seism. Soc. Am. 62, 1217-1222.

Nur, A., 1975. A note on the constitutive law for dilatancy. Pure App Geophys. 113, 197-206).

Ojala, I., B.T. Ngwenya, I.G. Main & S.C. Elphick, 2003. Correlation of microseismic and chemical properties of brittle deformation in Locharbriggs sandstone, J. Geophys. Res. 108 (B5), 2268.

Ojala I O., I. G. Main & B. T. Ngwenya, 2004. Strain rate and temperature dependence of Omori law scaling constants of AE data: implications for earthquake foreshock-aftershock sequences. Geophys. Res. Lett., 31, L24617.

Oliver, D. S., & Y. Chen, 2011. Recent progress on reservoir history matching: A revew. Computational Geosciences, 15(1), 185-221.

Paterson, M. S. & T.-F. Wong, 2005. Experimental rock deformation – the brittle field (2nd ed.), Springer, Berlin, Heidelberg.

Renard, P. & G. de Marsily, 1997. Calculating equivalent permeability: a review. Advances in Water Resources, 20(5-6), 253-278.

Rudnicki, J. W., and C.-H. Chen, 1988. Stabilization of rapid frictional slip on a weakening fault by dilatant hardening, J. Geophys. Res., 93, 4745– 4757.

Samier, P., L. Quettier & M. Thiele, 2002. Applications of streamline simulations to reservoir studies. SPE Reservoir Evaluation and Engineering, 5(4), 324-332.

Sammonds, P.R., P.G. Meredith & I.G. Main, 1992. Role of pore fluids in the generation of seismic precursors to shear fracture, Nature 359, 228-230.

Scharer, K.M., G. P. Biasi, R. Weldon & T.E. Fumal , 2010. Quasi-periodic recurrence of large earthquakes on the southern San Andreas Fault, Geology, 38, 555-558.

Scholz, C.H., Sykes, L.R. and Aggarwal, Y.P., 1973. Earthquake prediction – a physical basis, Science 181, 803-810.

Scholz, C.H., 2002. The mechanics of earthquakes and faulting, Cambridge University Press.

Scholz, C.H., 1997. Whatever happened to earthquake prediction? Geotimes, 42, 16-19.

Sibson, R.H., 1982. Fault zone models, heat flow, and the depth distribution of earthquakes in the continental crust of the Unites states, Bull. Seism. Soc. Am. 72, 151-163.

Sibson, R.H., 2009. Rupturing in overpressured crust during compressional inversion-the case from NE Honshu, Japan, Tectonophysics 473, 404–416

Sibson, R.H., 2001. Seismogenic framework for hydrothermal transport and ore deposition, Reviews in Economic Geology 14, 25-50.

Sibson, R.H., J. Mc. M. Moore & A.H. Rankin, 1975. Seismic pumping—a hydrothermal fluid transport mechanism, Journal of the Geological Society, 131, 653-659.

Stephen, K. D., G. E. Pickup & K. S. Sorbie, 2001. The local analysis of changing force balance in immiscible incompressible two-phase flow. Transport in Porous Media, 45, 63-88.

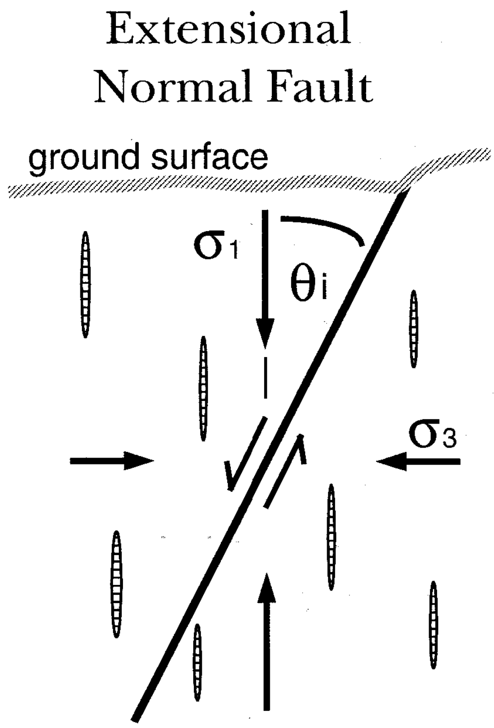
Strebelle, S., 2002. Conditional simulation of complex geological structures using multi-point statistics. Mathematical Geology, 34(1), 1-21.

Van Stiphout, T., S. Wiemer & W. Marzocchi, 2010. Are short-term evacuations warranted? Case of the 2009 L'Aquila earthquake. Geophys. Res. Lett. 37, L06306.

Warren, J.E. & P.J. Root, 1963. The behaviour of naturally fractured reservoirs. Society of Petroleum Engineers Journal, 3, 245-255.

Wyss, M. & Booth, D.C., 1997. The IASPEI procedure for the evaluation of earthquake precursors, Geophys. J. Int., 131, 423-424.

**Figures**



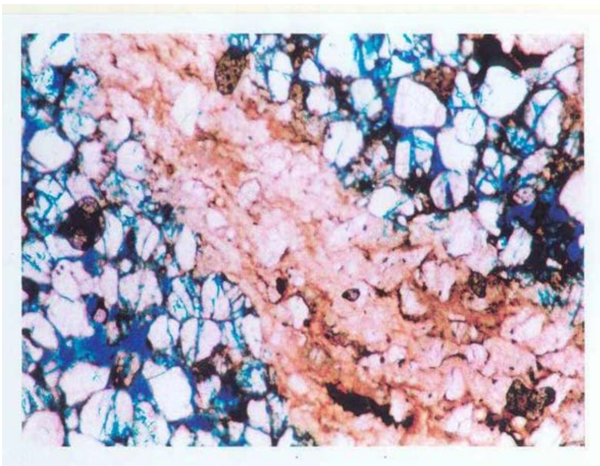



**Figure 1:** (Upper diagram) Shear fault and extensional fracture orientations predicted by E.M. Anderson's model for failure in the brittle crust, in the case of a vertical maximum principal stress (after Sibson, 2001, image provided by Richard Sibson). (Lower diagram) Orientations of a shear band and local tensile microcracks on the grain scale (around 300 microns) in a porous sandstone, also in the case of a vertical maximum principal stress (after Mair et al., 2000).

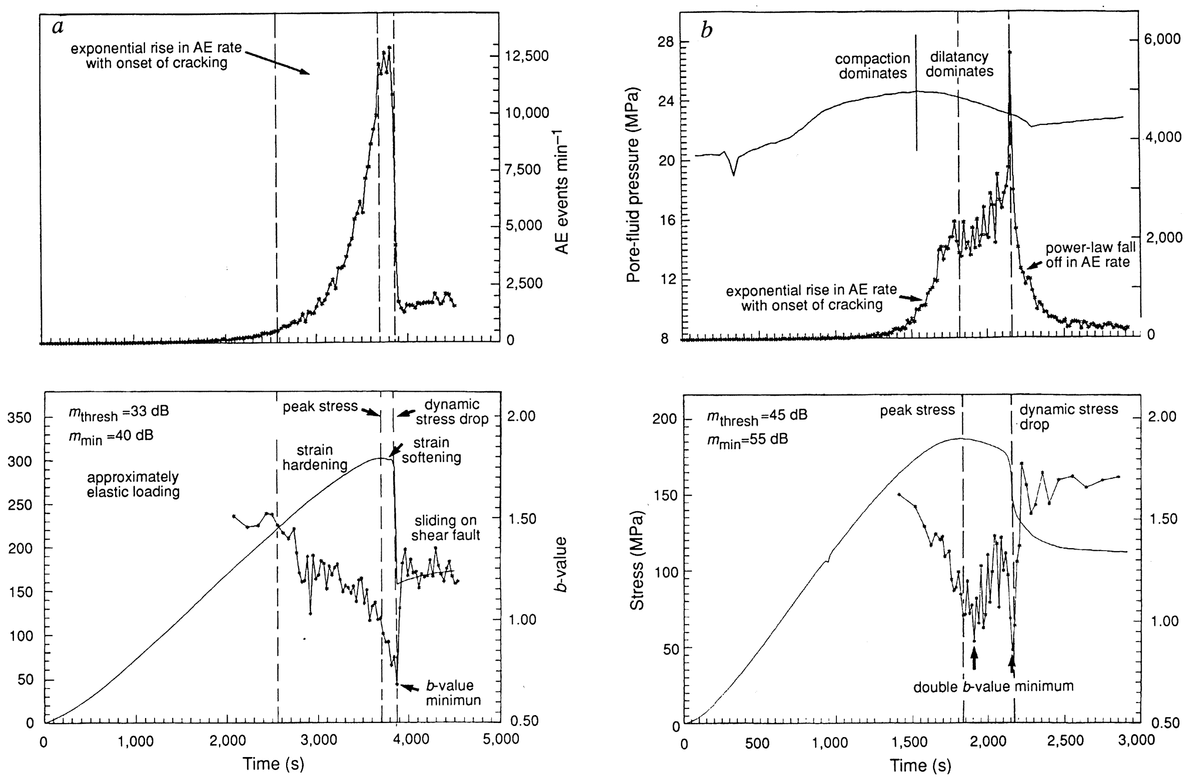



**Figure 2:** Acoustic emission event rate (AE events min$^{-1}$) and pore-fluid pressure (in MPa) (upper diagrams) and variations in the differential stress (in MPa) and the Gutenberg-Richter *b*-value (lower diagrams) for (a) nominally dry and (b) water-saturated samples of Darley-Dale sandstone (after Sammonds et al., 1992).

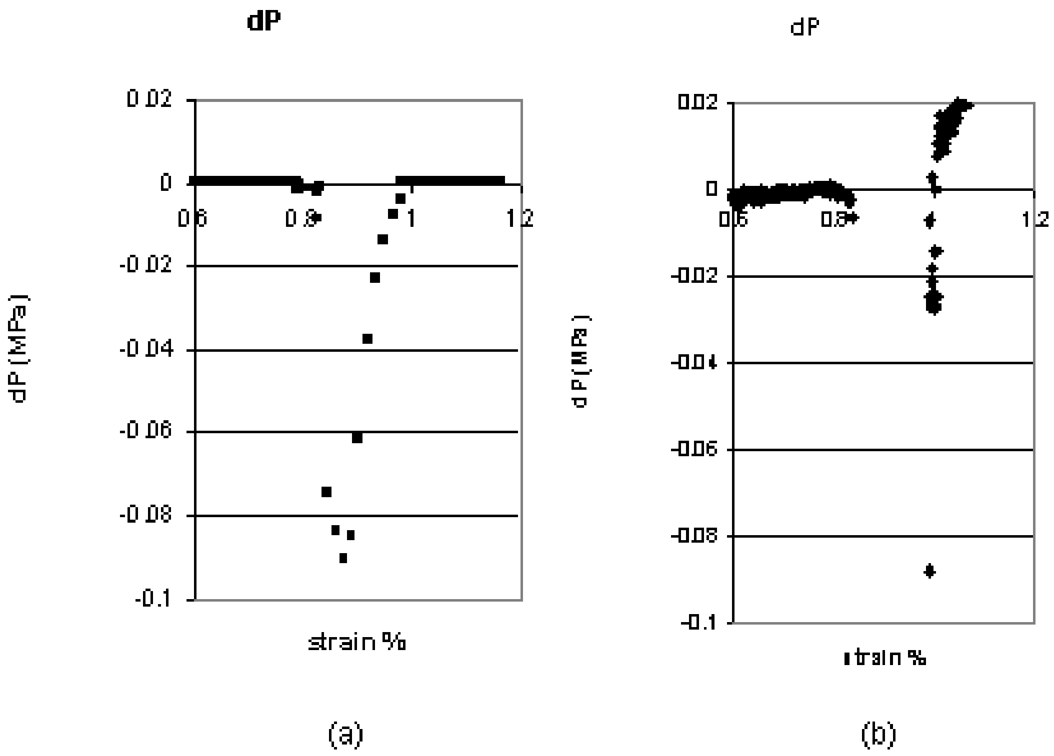



**Figure 3:** The dilatant ‘suction pump’ in action: (a) a dynamic model and (b) observation for pore pressure drop during dynamic failure of a sample of Clashach sandstone under constant input fluid flow rate, using unreactive oil as a permeant (after Grueschow et al., 2003).

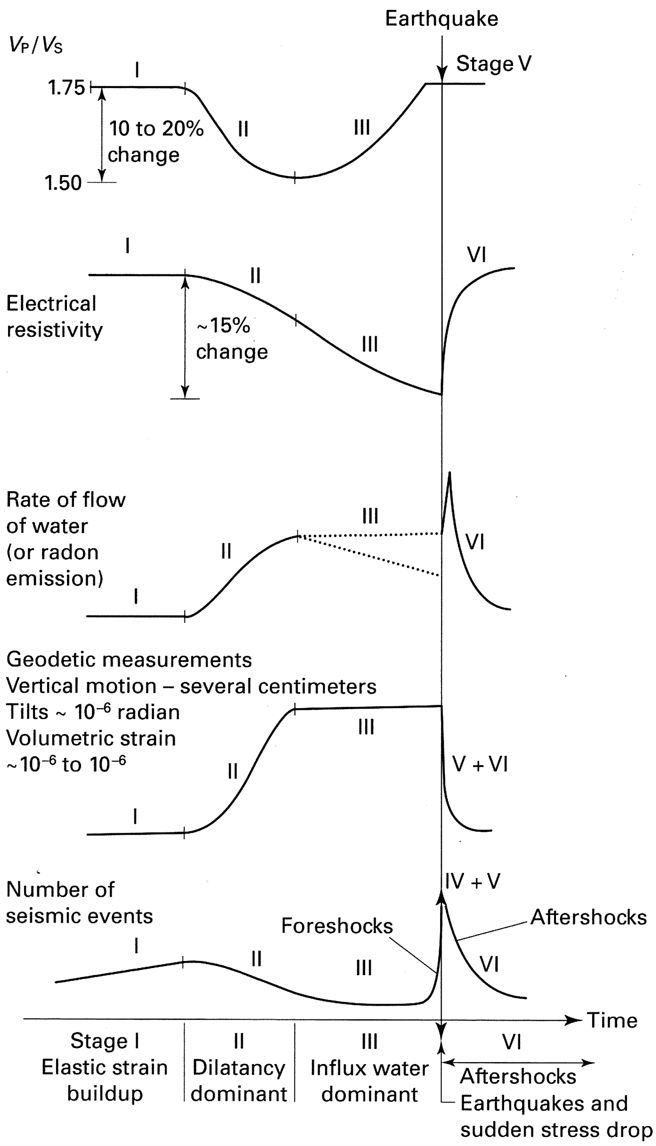



**Figure 4:** Predictions of anomalies in geophysical signals associated with elastic loading, dilatancy, diffusion, earthquake and post-seismic periods (after Scholz et al, 1973, 2002).

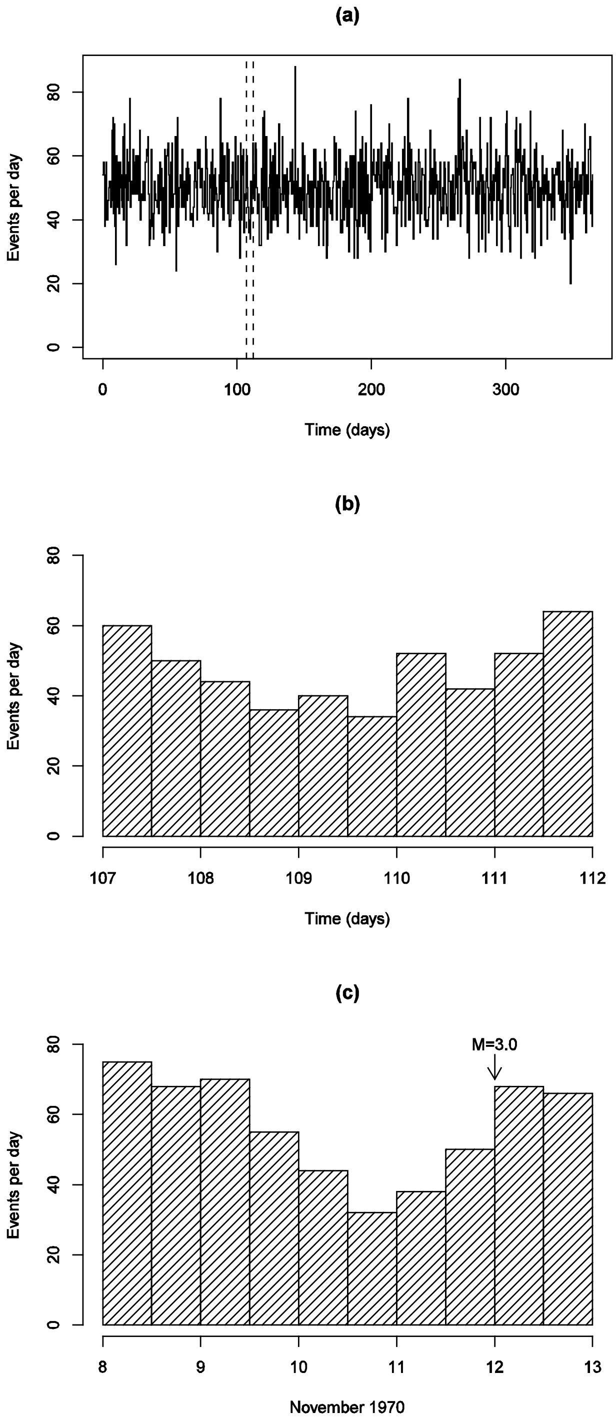



**Figure 5:** (a) Fluctuations in a random (Poisson) process with an average daily event rate of 50, sampled at 12 hour intervals. (b) Blow-up of one of the minima, in between the two vertical dashed lines, in (a). (c).Variation in seismic event rate per day over a similar timescale as (b), also sampled in 12 hr intervals, with a similar average event rate as in (a), after Scholz et al, 1973.

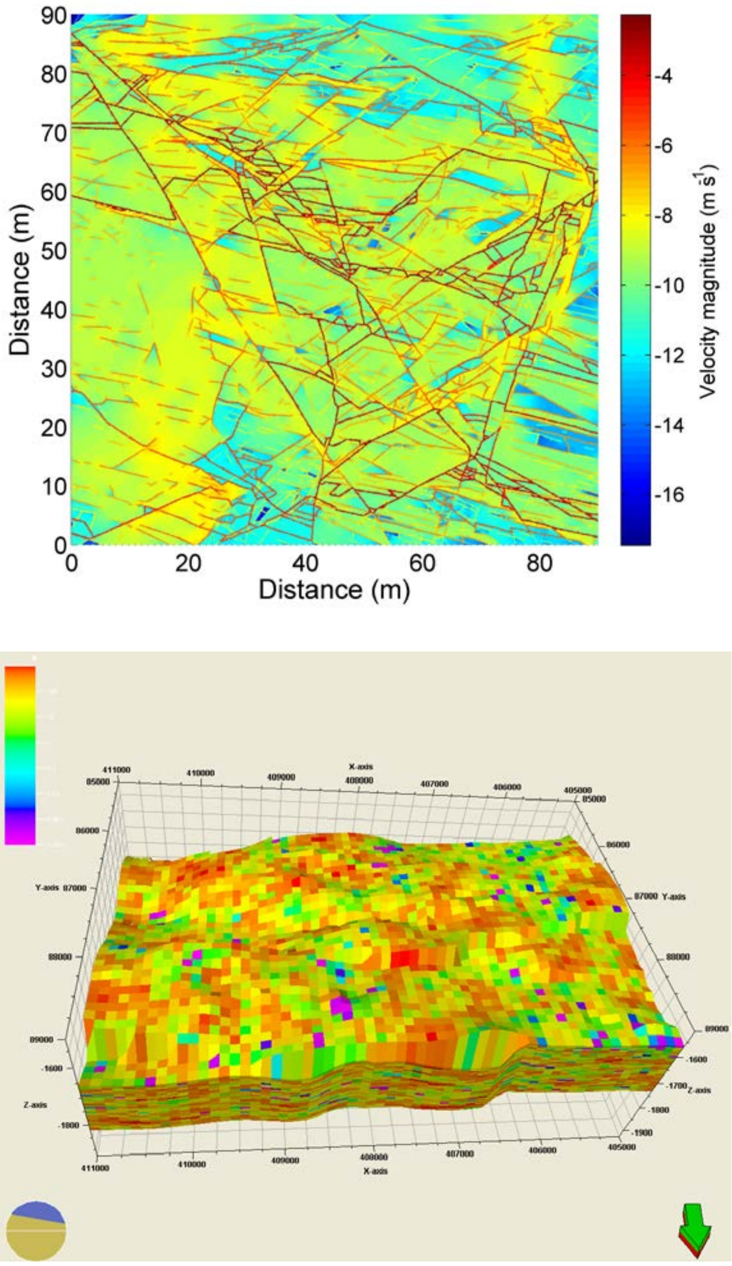



**Figure 6:** Illustration of spatial complexity in material properties on different scales. (upper diagram) Spatial variations in computed flow rates, expressed as the magnitude of the Darcy velocity (note the Log10 scale), on a heavily-fractured mapped outcrop (after Geiger and Emmanuel, 2010). (lower diagram) A typical reservoir model, where the spatial extent of a single voxel is also on the order of 100m and the colour coding denotes the variation in the reservoir property (courtesy of Viswa Chandra, Heriot-Watt University).

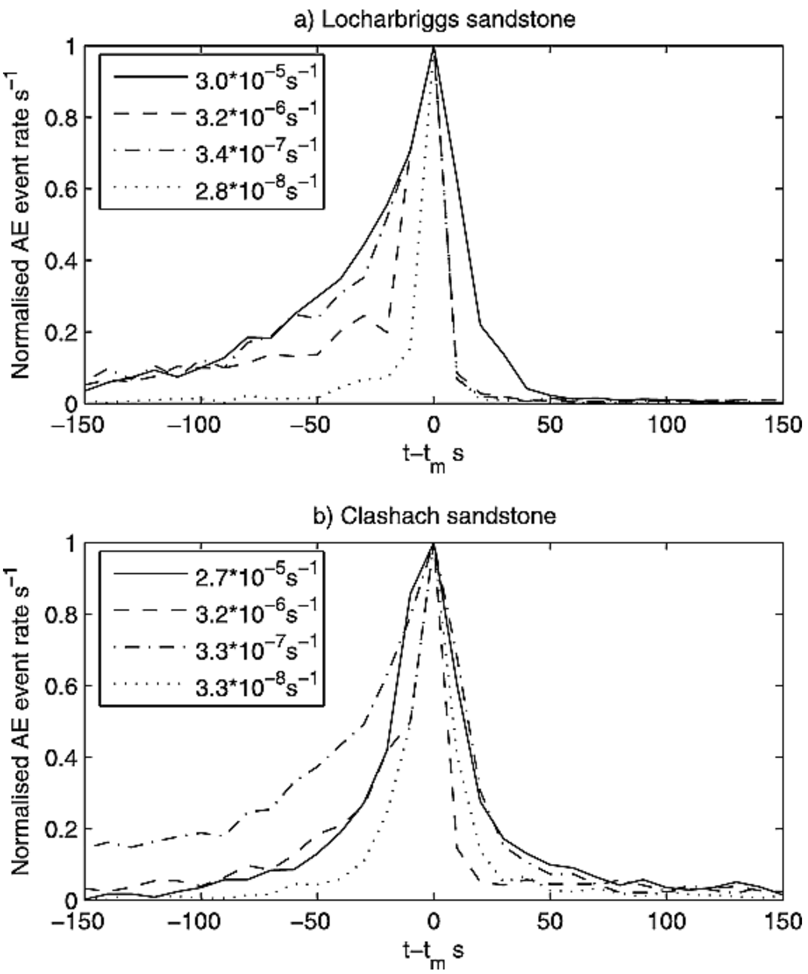



**Figure 7:** Evolution of the normalised acoustic emission event rate *a* for the four tests at strain rates between $10^{-5}$ to $10^{-8}$ $s^{-1}$ (a) Locharbriggs and (b) Clashach sandstones, all run at a temperature of 80°C.