Extreme events and predictability of catastrophic failure in composite materials and in the Earth

Citation for published version:
Main, I & Naylor, M 2012, 'Extreme events and predictability of catastrophic failure in composite materials and in the Earth', *European Physical Journal - Special Topics*, vol. 205, no. 1, pp. 183-197.
https://doi.org/10.1140/epjst/e2012-01570-x

Digital Object Identifier (DOI):
10.1140/epjst/e2012-01570-x

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
European Physical Journal - Special Topics

Extreme events and predictability of catastrophic failure in composite materials and in the Earth

Ian G. Main and Mark Naylor

School of GeoSciences
The King's Buildings
University of Edinburgh
West Mains Road
Edinburgh, UK

This is the author's final draft as submitted for publication. The final version was published in European Physical Journal: Special Topics by Società Italiana di Fisica and Springer (2012)

Cite As: Main, I & Naylor, M 2012, 'Extreme events and predictability of catastrophic failure in composite materials and in the Earth' *European Physical Journal - Special Topics*, vol 205, no. 1, pp. 183-197.

DOI: 10.1140/epjst/e2012-01570-x

Extreme events and predictability of catastrophic failure in composite materials and in the Earth

Ian Main and Mark Naylor

University of Edinburgh

Abstract

Despite all attempts to isolate and predict extreme earthquakes, these nearly always occur without obvious warning in real time: fully deterministic earthquake prediction is very much a 'black swan'. On the other hand engineering-scale samples of rocks and other composite materials often show clear precursors to dynamic failure under controlled conditions in the laboratory, and successful evacuations have occurred before several volcanic eruptions. This may be because extreme earthquakes are not statistically special, being an emergent property of the process of dynamic rupture. Nevertheless, probabilistic forecasting of event rate above a given size, based on the tendency of earthquakes to cluster in space and time, can have significant skill compared to say random failure, even in real-time mode. We address several questions in this debate, using examples from the Earth (earthquakes, volcanoes) and the laboratory, including the following. How can we identify 'characteristic' events, i.e. beyond the power law, in model selection (do dragon kings exist)? How do we discriminate quantitatively between stationary and non-stationary hazard models (is a dragon likely to come soon)? Does the system size (the size of the dragon's domain) matter? Are there localising signals of imminent catastrophic failure we may not be able to access (is the dragon effectively invisible on approach)? We focus on the effect of sampling effects and statistical uncertainty in the identification of extreme events and their predictability, and highlight the strong influence of scaling in space and time as an outstanding issue to be addressed by quantitative studies, experimentation and models.

Introduction

Catastrophic rupture of a variety of composite natural or man-made ceramic materials under slow loading conditions is usually preceded by chemically-activated, physically-enhanced microscopic processes that both weaken the material around its initial flaws and result in greater stress concentration there (Lawn, 1993). This leads to a strong positive feedback effect in promoting accelerating crack growth, for example as in the process of sub-critical crack growth by stress corrosion cracking (e.g. Atkinson, 1987). For a single crack in an otherwise uniform medium (a single crystal) reaction rate theory on a crystal lattice containing a single initial flaw predicts an exponential relationship between crack growth velocity V and the stress intensity factor $K=Y\sigma\sqrt{x}$ (a measure of stress concentration at the crack tip that depends on the geometric loading factor Y, the stress σ and the crack length x), of the form $V\sim\exp(K/K_0)$ (Lawn, 1993), where K_0 is a characteristic stress intensity factor.

In composite or polycrystalline materials the behaviour is instead a power-law known as Charles' law: $V\sim K^n$, most likely due to the additional stochastic component from the spatial distribution of the local stress and strength field in a non-uniform medium. The power-law exponent n is known as the stress corrosion index. Under constant stress loading, above the threshold stress for subcritical crack growth, Charles' law predicts an accelerating crack growth of the singular form

$$x = x_0(1-t/t_f)^{-\nu} \qquad [1]$$

where x_0 is the initial crack length, t_f is the failure time and the exponent $\nu=2/(n-2)$. In practice catastrophic failure occurs before t_f, i.e. while the rupture velocity is finite near the inertial limit. The same equation can be derived assuming a random strength distribution in the site percolation model near the percolation threshold or critical point, and by empirically-based constitutive laws based on observation of power-law creep experiments in the laboratory (Main, 1999). More complex models (e.g. with log-periodic oscillation superposed, e.g. Sornette & Sammis, 1995) have been proposed for this acceleration, but laboratory rock deformation tests rarely show such behaviour in variables such as acoustic emission event rate, instead having a monotonic increase (e.g. Sammonds et al., 1992, Fig 2a) except where pore pressure fluctuations can lead to a single temporary hiatus in acceleration and associated precursors to failure (e.g. Sammonds et al., 1992, Fig 2b).

Catastrophic failure in the laboratory is often associated with prior localisation of acoustic emission hypocentres along the eventual rupture plane. In a recent set of experiments with acoustic monitoring to determine the full seismic moment tensor of local strain, Graham et al. (2010) showed localisation of failure in a compression test was also associated with the transition from predominantly tensile to predominantly shear-type moment tensors, marking a transition from dilatant (positive volumetric strain) to double-couple (little or no volumetric strain) focal mechanisms. Tensile crack growth is associated with bulk sample dilatancy, once though to be the key to earthquake prediction (Scholz et al., 1973). Dilatant microcracking in a compressional stress field results in a negative feedback inhibiting further crack growth, representing a temporary strain hardening process. In this case the stress intensity is negatively correlated to crack length and hence

$$x = x_0 (t + T)^m \quad : \qquad 0<m<1 \qquad , \qquad [2]$$

where T is a characteristic transient time (Main, 2000). This is eventually overtaken by an unstable positive feedback due to localising shear deformation of the form [1]. Main (2000) also showed that the transition such negative to positive feedback in crack growth dynamics subject to constant loading (creep) resulted in an emergent constant strain rate due to the superposition of these two processes. The constant-strain rate phase has a power-law rheology between stress and strain rate, as observed for many composite materials even in the microscopically brittle field [Ngwenya et al., 2001; Heap et al., 2009]. This discussion illustrates a degree of universality in the constitutive equations between the phenomenology of failure in a variety of natural and man-made composite materials, some aspects of natural geophysical systems, and potential physical explanations in the transition from sub-critical to critical crack growth.

A key aspect of near-critical point dynamics such as equation [1] is the emergence of power law frequency-size distributions, for example the incremental frequency of rupture or crack length x

$$F(x) \sim x^{-d} \quad . \qquad [3]$$

The exponent d can be measured after a laboratory test by measuring the crack length distribution in section, or by inference from inversion of the radiated acoustic waves in three dimensions [Hatton et al., 1993]. Equation [3] is consistent with the Gutenberg-Richter law for earthquake recurrence $F(m)=a-bm$, where F here is an incremental frequency, the magnitude m is the logarithm of the maximum radiated amplitude, corrected for attenuation, and a and b are model parameters. The exponent b depends both on the level of stress (Scholz, 1968), which affects the stress intensity, and on the material heterogeneity (Mogi, 1962), which affects the fracture toughness or critical stress intensity K_c. In the laboratory the parameter b is negatively correlated to the normalised stress intensity K/K_c (Meredith & Atkinson, 1983) thereby providing a quantitative unification of the effect of applied stress and material heterogeneity on the exponent b.

A major difference between the laboratory and natural processes is the order of magnitude differences in scale in space and time, allowing the possibility of a range of physical and chemical processes not observable on a laboratory time scale (e.g. fault and fracture healing, stress interaction between faults, strong and systematic material heterogeneity due to previous geological history on all scales). In additions the boundary conditions are very different, especially the absence of a large stiffness contrast between the rock sample and the loading medium in a lab test: in the Earth rock is loaded by more rock, not steel or hydraulic fluid. In the lab the ultimate maximum magnitude of a rupture event is pre-determined by the sample size. In the laboratory scaling exponents, such as the b-value, vary systematically between rock type and with stress as described above, whereas in natural seismicity the b-value is remarkably constant in space and time, within its uncertainty, at a value near 1 for shallow earthquakes [Frohlich & Davis, 1993; Kagan, 1999].

The Earth is permeated by pre-existing faults and fractures on all scales, and the presence of fault healing and reloading allows deformation at a remarkably steady rate, with a long-term balance in the energy flux, in a state permanently near the critical point. After an appropriate 'run-in' period, this near-critical, steady-state dynamics occurs spontaneously, without the need for external 'tuning' within the spectrum of possibilities that can be more fully explored in the transition from sub-critical to critical crack growth in the laboratory. This state of 'self-organised criticality', proposed by Bak & Tang [1988] for avalanches in sandpiles loaded at a constant mass flux rate, and soon after by Bak et al [1989] and Sornette & Sornette [1989] for earthquakes, neatly explains the presence of power law statistics such as equations [1] and [3], the observed low stress drop in natural seismicity compared to the ambient state, and the ease with which earthquakes can be triggered by man-made stress perturbations [Main, 1996]. The cellular automaton type models used to model the process initially used only local interactions, analogous to the Ising model for magnetism. In the Earth however long-range stress interactions are also important. In this case power-law statistics emerge not just in the avalanches (or earthquakes), but in the permanent structures left behind, i.e. the faults [Cowie et al, 1993; Sornette et al., 1994].

Here we address the problem of prediction of individual extreme events associated with localised, catastrophic rupture in composite materials and in the Earth, using a number of examples primarily from our recent published work, and some new analysis represented in Figures 1-3, to illustrate particular concepts or questions.

Do Dragon Kings exist?

Dynamic failure in a laboratory test fractures the whole sample, such that the final crack or fault is much bigger than would be expected from extrapolation of the previous crack population using equation [3]. In this paper this is what we mean by an extreme event being a 'Dragon King': the largest rupture is not just bigger than anything seen before (an extreme event in this time interval) but is also a genuine outlier from what would be expected from extrapolation of the overall population of smaller fractures, with and an enhanced probability of occurrence of the largest events above the power-law trend. One model with this property is the 'characteristic earthquake' model, with power-law scaling at small magnitudes and a dragon king (characteristic earthquake of a given repeatable size) that occurs more often than expected from extrapolation of this smaller population, often inferred from a big gap in the frequency-magnitude plot between the magnitude of the largest event and the next largest [Schwartz & Coppersmith, 1984]. A second aspect of this model is the notion of quasi-periodic recurrence, related to the concept of elastic rebound and stress renewal [Reid, 1910; Shimazaki & Nakata, 1980] as well as the related concept of a 'seismic gap' [McCann et al., 1979], i.e. a segment of a fault or plate boundary with previous history of a large earthquake but not in the recent past. The temporal behaviour of renewal models and their applicability to Earthquakes is discussed in the next section. Here we address the question of whether we can clearly identify 'characteristic' events from frequency-magnitude data; given a finite, and perhaps not representative, data sample from the short timescales we have access to in instrumental, historical and palaeoseismic data.

One example of a gap between the largest and next largest event in a population is Bath's law for aftershocks, where the largest aftershock is (on average) around 1.2 magnitude units less than that of the mainshock, irrespective of the mainshock magnitude. Aftershocks are conventionally modelled statistically with an Epidemic-Type Aftershock Sequence (ETAS, Ogata, 1988) model. This model is 'seeded' with random events and then adds triggered events whose event rate decays in time according to the first derivative of equation [2], a form known as the modified Omori law Physically such empirically-observed power-law decay is therefore consistent with the time-dependent response of a complex medium to the stress transient caused by the parent or triggering event (Main 2000). In the ETAS model both the temporally-random seeded events and their associated triggered events are assigned magnitudes at random from the Gutenberg-Richter law [equivalent to equation 3]. In this model triggered events can (occasionally) be bigger than the triggering event due to this, and so are not labelled as conventional 'aftershocks'.

Helmstetter & Sornette (2003) examined this problem, and showed under some conditions that Bath's law is an emergent property of finite sampling of the ETAS model, i.e. the gap in magnitude in a single parent-daughter sequence is a statistical average that would be expected from a parent process with a pure power-law frequency-size distribution of the form [3]. In this example visually quite convincing magnitude gaps are very much consistent with a pure power law parent distribution. As a consequence great care must be taken in assigning significance to potentially illusory 'dragon kings' identified by a magnitude gap or by comparing the observed frequencies with the trend extrapolated from the distribution at smaller magnitude . Sample bias is not confined solely to the largest events in power law systems. On timescales longer than a single aftershock sequence simple metrics such as the mean event rate (of all sizes) show very slow, non-Gaussian convergence to the mean rate. For example Main et al. (2008) demonstrated that the mean event rate per month in real data from the Centroid Moment Tensor catalogue tends to increase systematically as more data is collected. The average event rate is now relatively stable at a global level, but the standard deviation is increasing, in contradiction to the central limit theorem. This slow, non-Gaussian, convergence is also an emergent property of the ETAS model (Naylor et al. 2009a), and implies that even simple properties such as mean event rate or inter-event time cannot be determined without sampling bias even with relatively large (thousands) numbers of events in a catalogue.

The identification of 'characteristic earthquakes' from frequency-magnitude data is also fraught with problems of counting errors arising from 'statistics of small numbers'. However, it is possible to account for this in model discrimination using a

Poisson distribution for the uncertainty in the number of events counted. To identify a dragon king, we must first reject the null hypothesis that the candidate dragon is not part of the power-law population of smaller dragons that we have seen before, while accounting for the counting errors in both. Leonard et al. (1999) introduced a method for doing this, applying a Bayesian Information Criterion for model selection using a generalised linear regression model that introduces an appropriate statistical penalty for models with greater numbers of free parameters, while simultaneously allowing for Poisson error statistics in the frequency at a given magnitude. This assumption was tested a posteriori and the resulting 95% confidence limits were consistent at that level with the scatter in data. In the limit of small bin widths in the regression, this combination is consistent with a similar combination of the alternate (Akaike) Information criterion and Poisson errors assumed in the method of inversion conventionally used for estimating the best fit (maximum likelihood) parameters in the ETAS model (Ogata, 1988).

The identification of 'dragon kings' requires a simpler null hypothesis (i.e. the unbounded Gutenberg-Richter distribution) to be rejected, and to show that the largest event really is a statistical outlier at a given level of confidence (Kagan, 1993). As we go to ever higher magnitudes, the probability of randomly sampling an event of that magnitude decreases, but is never zero - there is always a finite chance of sampling a large event from the unbounded parent distribution. To illustrate this Figure 1a plots the superposed frequency-magnitude distribution for 5000 random samples of 50,000 events, each from a parent Gutenberg-Richter distribution with a *b*-value of 1. The spread of these Monte-Carlo outcomes visually illustrates the non-Gaussian nature of the statistics, especially at high magnitudes, in particular the 'blowing up' of the statistical scatter, and the discrete or 'raster' pattern due to sampling whole numbers of points on the histogram associated with the statistics of small numbers. The red line on Fig 1a shows the known parent distribution and the blue lines the 95% confidence intervals assuming a Poisson distribution in the counting errors. The largest event in all of the random samples plots a clear 3 orders of magnitude above the parent trend, but is not a real outlier given the spread of outcomes, despite a relatively large but finite sample of total number of 50,000 events. If we plot the Monte-Carlo simulation containing this extreme event (data points circled in red in Fig 1a) the magnitude of the extreme event is above the low-to medium magnitude trend by some three orders of magnitude. Thus, when we analyse a real catalogue containing a known potential outlier event, identified in retrospect, we cannot avoid introducing a similar magnitude of sample bias in any attempt to reject the null hypothesis of the Gutenberg-Richter distribution this event emerged from.

The problem is exacerbated when we consider the impact of the event. Fig 1b shows the frequency scaled by the seismic moment. On a linear scale (Fig 1c) the extreme event dominates the total moment release. The sample containing the extreme event appears very much like a 'dragon king' (Fig 1d), with over 4 orders of magnitude more total moment release in the random sample of 1000 events more than the mean, and around 2.5 orders of magnitude outside the 97.5% percentile. In conclusion the strongly non-gaussian statistics of sampling a pure power can produce very realistic-looking 'dragon kings' in terms of impact.

To examine this more closely we examine the probability distribution of the magnitude gap predicted by the characteristic earthquake hypothesis. The distribution has the form of an exponential, with the incremental (Fig 2a) and cumulative (Fig 2b) distributions having the same form. For the population as a whole a magnitude gap of greater than 1.7 would lie outside the 95% confidence limits for a power law parent distribution, and hence a characteristic distribution may be inferred as more likely (the magnitude gap is real) at that level of confidence. This represents the statistical uncertainty. Next we consider how sample bias may affect the search for characteristic earthquakes. In looking for dragon kings it would be otherwise quite logical to begin the search with the largest events. However, in Fig 2c the largest events have the biggest magnitude gap, so the statistics of the whole population are not representative for these events. The 'search' itself introduces a hidden bias, and any statistical test of the significance of a large magnitude gap must take this into account.

The proportion of seismic moment that is released by the largest event in a given sample of 5000 events is shown in Fig 3. The figure shows a bimodal distribution, with a broad peak skewed to the right at 40%, and a narrow one near 100%. The latter is due to the samples with apparent 'dragon kings', and the former to the tendency to a more pure power law appearance with the bulk of the samples. This random synthetic result for finite sampling of a pure power law can be compared to natural data. For example, in a study of the Alpine-Mediterranean belt using some eighty years of instrumentally-recorded earthquakes, Jackson and McKenzie (1988) estimate the proportion of seismic moment released by the largest event in nine different tectonic zones. The average and standard deviation of these individual data points is 35% and 18% respectively, compared to 55% and 22% respectively for the synthetics of Fig 1. Quite large significances can be assigned to such observations , whereas Fig. 3 shows that their observed variability in natural seismicity can be reproduced very well by sample bias effects.

Naylor et al. (2009) examined the problem of finite sampling formally for earthquakes on the Sumatra-Andaman Islands mega-thrust and other subduction zones. First they showed that Poisson confidence intervals do indeed correctly account for the counting errors in the histogram data. Using these confidence intervals they then demonstrated that it is currently not possible to reject the null hypothesis of the Gutenberg-Richter law in favour of the characteristic earthquake model, even for a regional catalogue containing a moment magnitude 9.3 event (the 2004 Boxing day earthquake resulting in the great Indian Ocean tsunami). Furthermore, if the parent distribution is an unbounded power law, the Poisson errors for the largest events remain constant as the sample size increases, while the event rates for the smaller events do converge. Of course there must be some upper bound to earthquake magnitude to maintain a finite energy flux (Main, 1996). In this case the statistics would eventually converge for the largest events, but this may require very long (thousands of years) timescales to acquire sufficient data to sample the rare mega-earthquakes sufficiently. This does not mean dragon kings do not exist in earthquake frequency-magnitude data, just that we cannot yet tell, and simpler models with a single population following the power-law trend, for the time being, formally provide more Information (see also Kagan, 1993).

We can also consider the time evolution of the magnitude catalogue rather than a snapshot. Starting from a small sample size, the confidence intervals at low counts trivially shift towards higher magnitudes as the catalogue grows. In a catalogue which grows following a Gutenberg-Richter distribution, we must expect periods when the frequency-magnitude plot looks characteristic and periods when it looks like it has rolled off at high magnitude purely as part of the inherent sampling variability. In contrast to Gaussian convergence, such sampling fluctuations are the norm when sampling from an unbounded power law.

So far we have considered only the implications of statistical models of earthquake populations where earthquake magnitude is assigned randomly. This is consistent with the notion of self-organised criticality, where event size is determined by fine details of the processes governing the avalanche dynamics. In this case there is nothing about the nucleation process (a single grain landing in the sandpile and starting grains around it rolling) that can be used to predict the eventual size of the event: the avalanche does not ‘know’ how big it is going to be at the start (Bak, 1996). However, in real earthquake data there is a small but statistically-significant correlation between the predominant frequency of the early part of a seismogram and the eventual size of an earthquake, albeit with a large scatter due to the variability of the source time function (Olsen & Allen, 2005; Abercrombie, 2005). This scatter means in practice it is hard to predict the eventual size of a given event accurately from its nucleation phase. In fact the 95% confidence limits on the best fit magnitude for a given predominant nucleation phase frequency is on the order of ±2 magnitude units (Olsen & Allen, 2005, fig 3), with significant implications for the accuracy of earthquake early warning systems based on the properties of the early part of the seismogram. This an individual nucleation event does to some extent ‘know’ how big an event it is going to start is, but only within 4 orders of magnitude in the predicted rupture area. The large data scatter is consistent to first order with the notion of self-organised criticality, but the finite positive (albeit weak or second order) correlation between nucleation and eventual rupture size is not.

Finally the existence of a characteristic earthquake event size may also be determined by long-term geological boundaries or segments on the fault that play the role of the finite sample size in a laboratory test. This is analogous the behaviour of real sandpiles on real (finite-size) tables, where the frequency-size distribution is elevated above the power-law trend at the largest avalanche sizes, often associated with quasi-periodic recurrence of large avalanches (Held et al., 1990). In fact one of the original proponents of the Characteristic earthquake model points out that the concept was developed exactly for finite fault segments, whereas much of the statistical analysis on earthquake catalogues is done on regions where a pure power-law is much more likely to emerge (D. Schwartz, 2011, pers. comm.). In some areas there is quite strong evidence for extremely similar slip repeating at a given site several times. For example Klinger et al. (2011) used stream channel offsets at the source of 5 historically recorded strike-slip earthquakes on the Fuyun fault, Xinjiang, China to determine a rather constant slip in each of 6.3 ±1.2 m, consistent with the concept of a characteristic earthquake for that region. Thus at least some fault or fault segment boundaries may have the same effect as the finite table size in a sandpile experiment.

The problem with an approach based on identifying fault segments with a hard horizontal constraint on earthquake size is the ‘black swan’: there is always a finite probability that a fault segment may be breached in a given earthquake to produce an even bigger event than that recorded to date. For example during the 28 June 1992 Landers, California earthquake the rupture propagated through a strong bend in the fault break mapped at the surface (Aydin & Du, 1995). Similarly in the recent Tohoku, Japan earthquake previously identified ‘characteristic’ earthquake source zones all failed in a single event (Geller, 2011). In accounting for both possibilities Parsons and Geist (2009) analyse instrumental and palaeoseismic data together for California, and find no strong preference for either the Gutenberg-Richter or the Characteristic earthquake model. They suggest a ‘middle way’ in which both possibilities are allowed for.

In the vertical direction the maximum dimension is more predictable in the sense that the largest events are confined predominantly to a seismogenic zone strongly associated with a thermal cut-off where the material behaves in a more ductile fashion. For example a clear characteristic peak was seen in the frequency magnitude statistics of the earthquakes preceding the eruption of Mount St Helens in 1980 (Main, 1987), most likely due to the strong restriction on event size caused by an elevated geothermal above a magma chamber at a time of rapid inflation. The effect of system size (the size of the dragon’s domain) on the dynamics is covered in more detail in a separate section below.

In summary the evidence for characteristic earthquakes is far from compelling, though their existence cannot be ruled out either. As a consequence its utility as an operational tool has been called into question, particularly when undue faith is placed on the model in seismic hazard assessment, for example prior to the occurrence of the great Tohoku earthquake in Japan (Geller, 2011).

Is a Dragon likely to come soon?

There is (by definition) no predictability in the random or Poisson component of the ETAS model. However, its parameters may be uncertain because of the slow non-Gaussian convergence of event rate and Poisson counting errors discussed above, and the total event rate is likely to be biased to low values. If the background rate is sufficiently high to produce a large proportion of overlapping aftershock sequences, there is a significant bias in the best-fit ETAS solution, and known solutions cannot be determined by inversion without labelling (assigning marks to) 'parent' and 'daughter' events (Touati et al., 2009, 2011). These systematic and random uncertainties are important because the spatially-localised and random in time model is the basis for seismic hazard calculations, and also for the calculation of 'probability gain' over a random process that is a key metric of the skill of a given probabilistic forecasting strategy. It is relatively easy to show that the ETAS model outperforms the purely Poisson process despite the additional model parameters, commonly achieving transient probability gains of two or three orders of magnitude over the background rate. This can be used during aftershock sequences to estimate a conditional probability for future aftershocks at a given magnitude (Marzocchi and Lombardi, 2009). The ETAS model therefore provides an important benchmark against which to test alternate forecasting strategies, i.e. an appropriate null hypothesis.

Can we do better than forecasting aftershocks or more generally triggered events? Sykes & Menke (2006) use a Bayesian technique to examine earthquake recurrence on major plate boundaries using the coefficient of variation as a metric for quasi-periodic behaviour based on the renewal model. Using relatively small numbers of instrumentally and historically recorded events, they obtain repeat times with a variance less than mean, citing this as evidence for the seismic 'gap' hypothesis, itself based on the notion of fault segments at plate boundaries that fail in characteristic earthquakes in a quasi-periodic way. This retrospective analysis is not borne out however by retrospective tests of prospective forecasts made using seismic gap hypothesis (Rong et al. 2003). Using a much longer-term data set Scharer et al [2010] examined a 3000-year combined palaeoseismic record of 29 ground-breaking ruptures from the Southern San Andreas Fault in California at Wrightwood, and used statistical (Kolmogorov-Smirnov) tests to show that the temporally-random Poisson model could be rejected in favour of a weakly quasi-periodic recurrence model, with the caveat that the slow sedimentary deposition rate restricts the minimum frequency 'bin' size to 10 years (Scharer, pers. comm.). Despite its intuitive appeal, the probability gain in terms of predicted ground motion from quasi-periodic models in California using the historical and palaeoseismic record is rather local and relatively modest (±30% or so) over a Poisson process (Petersen et al., 2007). This degradation of periodicity, which might be associated with notions of an isolated fault with constant material properties (breaking and sliding stresses), implies that such material properties may not be reproducible between events, and/or that stress interactions (short and long-range) between the population of faults add a further degree of stochastic complexity that almost eradicates any periodic component. If dragon Kings do exist in the frequency-size distribution, they don't appear very regularly in the best record we have for earthquake repeat times, i.e. on the San Andreas fault (Scharer, 2010).

Few, if any, scientists believe the accurate, deterministic prediction of individual earthquakes is a reasonable scientific goal for the foreseeable future http://www.nature.com/nature/debates/earthquake/equake_frameset.html. Largely this is a consequence of the reality check provided by the absence of clear and reliable precursors when subjected to a set of reasonable 'rules of the game' (Wyss, 1991; Hough, 2009), strong negative evidence (hypothesis falsification) from highly-instrumented sites such as the Parkfield earthquake experiment (Bakun et al., 2005; Jackson & Kagan, 2006), and by the general acceptance of the complexity and non-linearity of the process, and its degrading effect on deterministic prediction (Main, 1992; Geller et al., 1997). However, even in a chaotic state there is still some probabilistic short-term predictability. Self-organised criticality has itself been defined sufficiently loosely to allow some systematic variability in stress state due to a low but finite stress drop, providing then potential of a degree of predictability in a probabilistic sense, possibly associated with systematic precursors associated with much smaller stress fluctuations (relative to the breaking stress) than seen in a typical laboratory test (Main & Naylor, 2006, 2010). The main reason such hypotheses of 'intermittent' criticality remain unproven is the lack of evidence for precursory signals that may be associated with such stress fluctuations, in particular the poor quality of the statistical evidence for quantitative evidence of critically-accelerating strain of the form of equation (1) (Hardebeck et al., 2008; Greenhough et al., 2009). The jury remains out on the possibility of probabilistic earthquake forecasting based on precursors, but for now the current frontier in the global Collaboratory for the Study of Earthquake Predictability (CSEP) Experiment is in determining the best model for earthquake triggering, perhaps combined with a small degree of periodicity in the stress renewal model (http://www.cseptesting.org/results). Once this is established, it will then be possible to use that as a null hypothesis for predictability above that level.

In contrast volcanic eruptions have often provided much more systematic warning prior to eruption. The 1980 explosive eruption of Mount St Helens, Washington, occurred while the National park was closed due to a series of phreatic (steam) eruptions, concentrated seismicity where none had been seen before, and where the flank of the mountain was inflating by up to 1 m per day in the run-up to the event (Lipman et al., 1981). The location was known, but the magnitude, eruption directionality and the precise time could not have been predicted accurately in advance, since the eruption was triggered by an unanticipated landslide event on the volcano flank that preceded a strong lateral blast. More effusive basaltic eruptions are also often preceded by clear precursors associated with accelerating seismic event rate, with stacked data following equation [1] closely (Chastin & Main, 2003). However, similar sequences are also seen prior to intrusive events. This is not surprising because we might expect the physics of fracturing and faulting ahead of an intruding magma body is similar for an event that breaches the percolation threshold than one which is just below (Gudmundsson et al., 2006). The estimated ratio of erupted to intruded volumes of magma in the Etna Volcano in Italy in the period 1980–1995 is 0.13, whereas the same ratio for the Krafla Volcano in Iceland in the period 1975–1984 is 0.30 (Gudmundsson et al., 2006). The ‘false alarm’ rate may therefore significant (70-87% for these volcanoes), but not so high as to preclude clear operational decisions to be made during a crisis, up to and including evacuation. In contrast the vast majority of potentially precursory earthquake swarms in Italy, such as the one preceding the 2009 L’Aquila earthquake, do not end in a significant earthquake, and evacuation as a mitigation strategy cannot be justified by operational decision-making tools such as a formal cost-benefit analysis (van Stiphout et al., 2010).

Volcanic eruptions may be easier to predict (in a probabilistic sense) for a variety of reasons. The location is known, removing two degree of freedom from the problem. The source is shallower (a few km), and the deformation accelerated by magma injection, so that strain rates are higher, intermediate between laboratory strain rates and those of earthquake loading, and hence easier to detect. Clear precursors have been seen, though the magnitude and time of the eruption can usually only be predicted approximately, and the false alarm rate is high. Nevertheless this is sufficient to be of some practical use. The even higher false alarm ratio for earthquake forecasting based on earthquake clustering has significant implications for operational decision-making (van Stiphout, et al, 2010). However even low (in an absolute sense) probability forecasts may have practical uses in issuing low level alerts or advisories summarised in the findings and recommendations of the International Commission on earthquake forecasting for civil defence http://www.protezionecivile.it/cms/attach/ex_sum_finale_eng1.pdf and discussed more fully by Jordan and Jones (2010) and justified in more detail by Jordan et al. (2011).

Does the size of the dragon's domain matter, and is the dragon effectively invisible on approach?

The scale-invariant nature of the geometry and some aspects of the physics of brittle field rupture has been used by many (largely implicitly) to argue for a linear scaling of physical processes between the laboratory and the field case. However, the failure of the Earth is different in many respects to that of a laboratory test. No significant bulk volumetric strain (dilatancy) is in fact observed prior to or during real earthquakes, as commonly observed in laboratory tests on brittle materials. This may be because it is not possible yet to detect strain associated with the nucleation process of earthquakes, i.e. dilatancy is local, rather than distributed throughout the finite sized sample as in the laboratory. The absence of significant co-seismic dilation argues for a much lower role of volumetric strain at tectonic scales than in the laboratory. In other words the system size (the dragon's domain) matters a great deal, and this is the determining factor on the practical visibility of the dragon on approach.

In the two years before the M_W 6.3 earthquake struck the city of L'Aquila, Italy on April 6, 2009, no anomalous strain signal larger than a few tens of nanostrain units was visible, limiting the volume of the possible earthquake 'preparation zone' to less than 100 km^3 (Amoruso & Crescentini, 2010). 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 2×10^{12} Nm, i. e. 0.00005% of the main shock seismic moment". The dragon may be coming imminently, but is effectively invisible on approach. In contrast surface deformation due to dilatant inflation of the carapace due to magma emplacement is clearly seen prior to some volcanic eruptions (e.g. Jordan & Keiffer, 1981) almost always associated with changes in seismicity (Qamar, et al., 1983; Main, 1987). Such differences in scaling in space (size) and time (strain rate) have a big effect on the detectability and reliability of precursors, and this has not yet been fully addressed by multi-scale, multi-physics models, in such a way that preserves the observed structural scale-invariance.

It is hard to imagine that the processes observed in laboratory tests do not apply at least at small scales in the Earth, e.g. in the fracturing of healed asperities or contact points on the fault. However, these may be very local, as highlighted by the discussion above on the strain observations prior to the 2009 L'Aquila earthquake. The upper limit of 2×10^{12} Nm on the length scale of any precursory slip patch inferred by Amoruso and Crescentini (2010) implies an object no bigger than 100m, operating at a nucleation depth of typically 10 km or so in a continental setting, with the nucleation point possibly changing from even to event (as it has done for example in the last two Parkfield earthquakes, Bakun et al., 2005). 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 Mw 7.6 Izmit (Turkey) earthquake. The signal consisted of a succession 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 event (25m). These results confirm a very short duration, very localised, detectable nucleation phase for this event, much smaller than the eventual rupture size. This is a singular observation at present: it is not yet known how common such events are, or how often they occur without triggering a larger event.

In summary physical models require a nucleation process, but nucleation may be very local, and the process highly complex and highly non-linear, involving the interactions of several elements. The eventual size of the rupture is only very weakly dependent on the nucleation process, and highly uncertain. This is quite different from a laboratory test where a clear 'preparation zone' (equivalent to the eventual rupture size) is determined a priori by the sample boundaries. Identifying such a small nucleation patch reliably in advance prior to natural earthquakes might be like looking for a ‘needle in a haystack’. In summary reliable earthquake precursors remain a ‘black swan’ – not yet seen clearly but not possible to rule out, at least at some scale. The question is not the theoretical possibility, but the practical issue of reliable identification, although recent constraints on the size of the nucleation patch go a long way in explaining why reliable precursors have not been observed. So far the goal of identifying dragon kings in earthquake populations, and seeing them coming in advance, with sufficient reliability for a strong operational response, has proven elusive. For now the model to beat (where significant probability gains over a random process can be demonstrated in real time) remains the earthquake triggering model.

Conclusion

There is no compelling evidence for dragon kings in the sense of outliers from the power-law size distribution of earthquake event size. In contrast, some volcanic sequences show clear characteristic behaviour, and nearly all laboratory tests show an extreme event at the sample size, well outside the population of acoustic emissions that largely indicate grain scale processes until very late in the cycle. Clearly differences in scale in space and time, and sample boundary conditions, play a strong role. Identifying real outliers from power law distributions is inherently difficult: the largest event in an unbounded system is always equally uncertain, and the statistics for smaller events is to first order Poisson, converging only slowly to the Gaussian central limit as more data are collected. As a consequence finite-size sampling bias can lead to significantly appealing ‘characteristic’-type plots, even for relatively large sample sizes taken at random from a pure power law. In terms of the prediction of failure time, it is a mistake to use scale-free geometry as an argument for scale-free physics. While physical models argue for a degree of probabilistic predictability associated with identifying the dragon’s approach, recent work on direct strain and cross-correlation seismic analysis has shown that any nucleation signal may be highly localised in comparison with the size of the ensuing event. The dragon king is usually highly visible on approach in the laboratory, and before some volcanic eruptions, but for now our best forecasting model for earthquake populations is stochastic: based on fault interaction and triggering phenomena rather than any true precursory processes.

Acknowledgements

We are grateful to Andrew Michael and an anonymous reviewer for their thoughtful reviews of an earlier manuscript, and David Shwartz for a discussion of the evidence base for characteristic earthquakes. Both authors were funded by fellowships awarded by the Royal Society of Edinburgh during the preparation for, and writing of, the paper.

References

Abercrombie, R. (2005). Seismology: The start of something big? Nature 438, 171-173.

Atkinson, B.K. (1987). Fracture mechanics of rock, Academic Press, London.

Amoruso, A. & L. Crescentini (2010). Limits on earthquake nucleation and other pre-seismic phenomena from continuous strain in the near field of the 2009 L'Aquila earthquake, Geophys. Res. Lett., 37, L10307.

Aydin, A. & Y. Du (1995). Surface rupture at a fault bend: the 28 June 1992 Landers, California, earthquake, Bull. Seism. Soc. Am., 85: 111 - 128.

Bak, P. (1996). How Nature Works: The Science of Self-Organised Criticality Copernicus Press, New York.

Bak, P. & C. Tang (1989). Earthquakes as a self-organized critical phenomenon, J. Geophys. Res, 94, 15635-15637

Bak, P, C. Tang & K. Wiesenfied (1988). Self-organized criticality, Phys. Rev. A 38, 364-374.

Bakun, W.H., B. Aagaard, B. Dost, W. L. Ellsworth, J. L. Hardebeck, R. A. Harris, C. Ji, M. J. S. Johnston, J. Langbein, J. J. Lienkaemper, et al., Implications for prediction and hazard assessment from the 2004 Parkfield earthquake, Nature, 437, 969-974.

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.

Chastin, S.F.M. & I.G. Main (2003). Statistical analysis of daily seismic event rate as a precursor to volcanic eruptions, Geophys. Res. Lett. 30(13), 1671.

Cowie, P.A., C. Vanneste & D. Sornette (1993). Statistical physics model for the spatio-temporal evolution of faults, J. Geophys. Res., 98, 21,809-21,821.

Frohlich C. & S.D. Davis, J. (1993), Teleseismic *b*-values, or much ado about *b*=1, Geophys. Res., 98, 631-644.

Geller, R.J., D.D. Jackson, Y.Y. Kagan & F. Mulargia (1997). Earthquakes Cannot Be Predicted, Science, 275, p. 1616.

Geller, R.J. (2011). Shake-up time for Japanese seismology, Nature, 472, 407–409.

Graham, C.C., S. Stanchits, I. G. Main & G. Dresen (2010). Source analysis of acoustic emission data: a comparison of polarity and moment tensor inversion methods, Int. J. Rock Mech. Min. Sci. 47, 161–169.

Greenhough, J., A. Bell & I.G. Main (2009). Comment on “Relationship between accelerating seismicity and quiescence, two precursors to large earthquakes” by Arnaud Mignan and Rita Di Giovambattista, Geophys. Res. Lett. 36, L17303,

Gudmundsson A. & Philip S.L. (2006). How local stress fields prevent volcanic eruptions, J. Volcanol. Geotherm. Res., 158, (2006), 257-268.

Hardebeck, J. L., K. R. Felzer, and A. J. Michael (2008), Improved tests reveal that the accelerating moment release hypothesis is statistically insignificant, J. Geophys. Res., 113, B08310.

Hatton, C.G., I.G. Main & P.G. Meredith (1993). A comparison of seismic and structural measurements of fractal dimension during tensile subcritical crack growth, J. Struct. Geol. 15, 1485-1495.

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.

Held, G.A., D. H. Solina, H. Solina, D. T. Keane, W. J. Haag, P. M. Horn, and G. Grinstein (1990). Experimental study of critical-mass fluctuations in an evolving sandpile, Phys. Rev. Lett. 65, 1120–1123.

Helmstetter, A. & D. Sornette, (2003). Bath's law Derived from the Gutenberg-Richter law and from Aftershock Properties, Geophys. Res. Lett., 30, 2069.

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

Jackson, D.D. & Y.Y. Kagan (2006). The 2004 Parkfield earthquake, the 1985 prediction, and characteristic earthquakes: Lessons for the future, Bull. Seism. Soc. Am., 96, S397-S409 DOI: 10.1785/0120050821.

Jackson, J. & D. McKenzie, 1984. Active tectonics of the Alpine-Himalayan belt between western Turkey and Pakistan, Geophys. J. R. Astr. Soc., 77, 185-264.

Jordan, R. & H.H. Keiffer (1981). Topographic changes at Mount St Helens: Large-scale photogrammetry and digital terrain models, in The 1980 eruption of Mount St. Helens, Washington, U.S.G.S. professional paper 1250, 135-141.

Jordan, T.H. & L.M. Jones (2010). Operational Earthquake Forecasting: Some Thoughts on Why and How, Seismological Res. Lett., 81, 571-574.

Jordan, T., Y. 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). doi:10.4401/ag-5350.

Kagan, Y.Y. (1993). Statistics of characteristic earthquakes, Bull. Seism. Soc. Am., 83, 7-24.

Kagan Y.Y. (1999). Universality of the seismic moment-frequency relation, Pure App. Geophys.,155: 2-4, 537-573.

Klinger, Y., M. Etchebes, P. Tapponier & C. Narteau (2011). Characteristic slip for five great earthquakes along the Fuyun fault in China, Nature Geoscience 4, 389-392.

Lawn, B. (1993). Fracture of brittle solids, 2nd ed., Cambridge Solid State Science Series, Cambridge.

Lipman, Peter W., and Mullineaux, Donal R. (eds.) (1981). The 1980 Eruptions of Mount St. Helens, Washington: U.S. Geological Survey Professional Paper 1250, 844p.

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

Main, I.G. (1987). A characteristic earthquake model of the seismicity preceding the eruption of Mount St. Helens, Phys. Earth Planet. Int. 49, 283-293.

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

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 & M. Naylor (2010). Entropy production and self-organized (sub) criticality in earthquake dynamics, Phil. Trans. Roy. Soc. Lond. A 2010 368, 131-144.

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.

Main, I.G., T. Leonard, O. Papasouliotis, C.G. Hatton & P.G. Meredith (1999). One slope or two? Detecting statistically-significant breaks of slope in geophysical data, with application to fracture scaling relationships, Geophys. Res. Lett. 26, 2801-2804.

Marzocchi, W. & A.M. Lombardi (2009). Real-time forecasting following a damaging earthquake, Geophys. Res. Lett., 36, L21302.

McCann, W.R., S.P. Nishenko, L.R. Sykes, et al. (1979). Seismic gaps and plate tectonics – seismic potential for major plate boundaries, Pure App. Geophys., 117,1082-1147.

Mogi, K. (1962), Bull. Earthq. Res. Inst. Tokyo University, 40, 831.

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.

Naylor, M., I.G. Main & S.Touati (2009a). Quantifying uncertainty on mean earthquake inter-event times for a finite sample, J. Geophys. Res., 114, B01316.

Naylor, M., J. Greenhough, J. McCloskey, A. F. Bell & I.G. Main (2009b). Statistical evaluation of Characteristic Earthquakes in the frequency-magnitude distributions of Sumatra and other subduction zone regions, Geophys. Res. Lett., 36, L20303.

Ngwenya, B.T., I.G. Main, S.C. Elphick, B.R. Crawford & B.G.D. Smart (2001). A constitutive law for low-temperature creep of water- saturated sandstones, J. Geophys. Res. 106, 21811-21826.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes, J. Amer. Statist. Assoc. 83, 9-27.

Olson, E.L. & R. M. Allen, The deterministic nature of earthquake rupture, Nature 438, 212-215.

Parsons, T. & E. Geist (1999). Is there a basis for preferring characteristic earthquakes over a Gutenberg-Richter distribution in probabilistic earthquake forecasting?, Bull. Seism. Soc. Am. 99, 2012-2019.

Petersen, M.D., T. Cao, K.W Campbell & A.D. Frankel (2007). Sesimic hazard assessment for the state of California: uniform California earthquake rupture forecast model 1.0, Seism. Res. Lett., 78, 99-109.

Reid, H.F. (1910). The mechanism of the earthquake, in The California Earthquake of April 18 1906, Report of the State Investigation Commission, Vol. 2, Washington D.C.: Carnegie Institution 1-192.

Rong, Y., D.D. Jackson & Y.Y. Kagan (2003). Seismic gaps and earthquakes, J. Geophys. Res. 108, 2471.

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., (1968). Frequency-magnitude relation of microfracturing in rock and its relation to earthquakes, Bull. Seism. Soc. Am., 58, 399-415.

Scholz, C.H., L.R. Sykes and P. Aggarwal (1973). Earthquake Prediction: A Physical Basis, Science, 181 no. 4102 pp. 803-810.

Schwartz, D. & K. Coppersmith (1984). Fault Behavior and Characteristic Earthquakes: Examples From the Wasatch and San Andreas Fault Zones, J. Geophys. Res., 89, B7, 5681-5698.

Shimazaki K. & T. Nakata (1980). Time-predictable recurrence model for large earthquakes, Geophys. Res. Lett., 7, 279-282.

Sornette, D. & A. Sornette (1989). Self-organized criticality and earthquakes, Europhys. Lett., 9, 197-202.

Sornette, D., and C.G. Sammis (1995). Complex critical exponents from renormalisation group theory of earthquakes: implications for earthquake predictions, J. Phys. I, 607-619.

Sornette, D., P. Miltenberger and C. Vanneste (1994). Statistical physisc of fault patterns self-organized by repeated earthquakes, Pure App. Geophys., 142, 492-527.

Sornette, D. P. Davy & A. Sornette (1990). Structuration of the lithosphere in plate tectonics as a self-organized critical phenomenon, J. Geophys. Res. 95, 15353-17361.

Touati, S., M. Naylor & I. G. Main (2009). The origin and non-universality of the earthquake inter-event time distribution, Phys. Rev. Lett. 102, 168501.

Touati, S., M. Naylor, I.G. Main & M. Christie (2011). Masking of earthquake triggering behaviour by a high background rate and implications for ETAS inversions, J Geophys. Res. (in press).

Qamar, A., W. St. Lawrence, J.N. Moore & G. Kendrick (1983). Seismic signals preceding the explosive eruption of Mount St. Helens, Washington on 18 May 1980, Bull. Seism. Soc. Am., 73, 1797-1813.

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

Wyss, M., Second round of evaluation of proposed earthquake precursors, Pure Appl. Geophys. 149, 3-16 (1991).

Figures

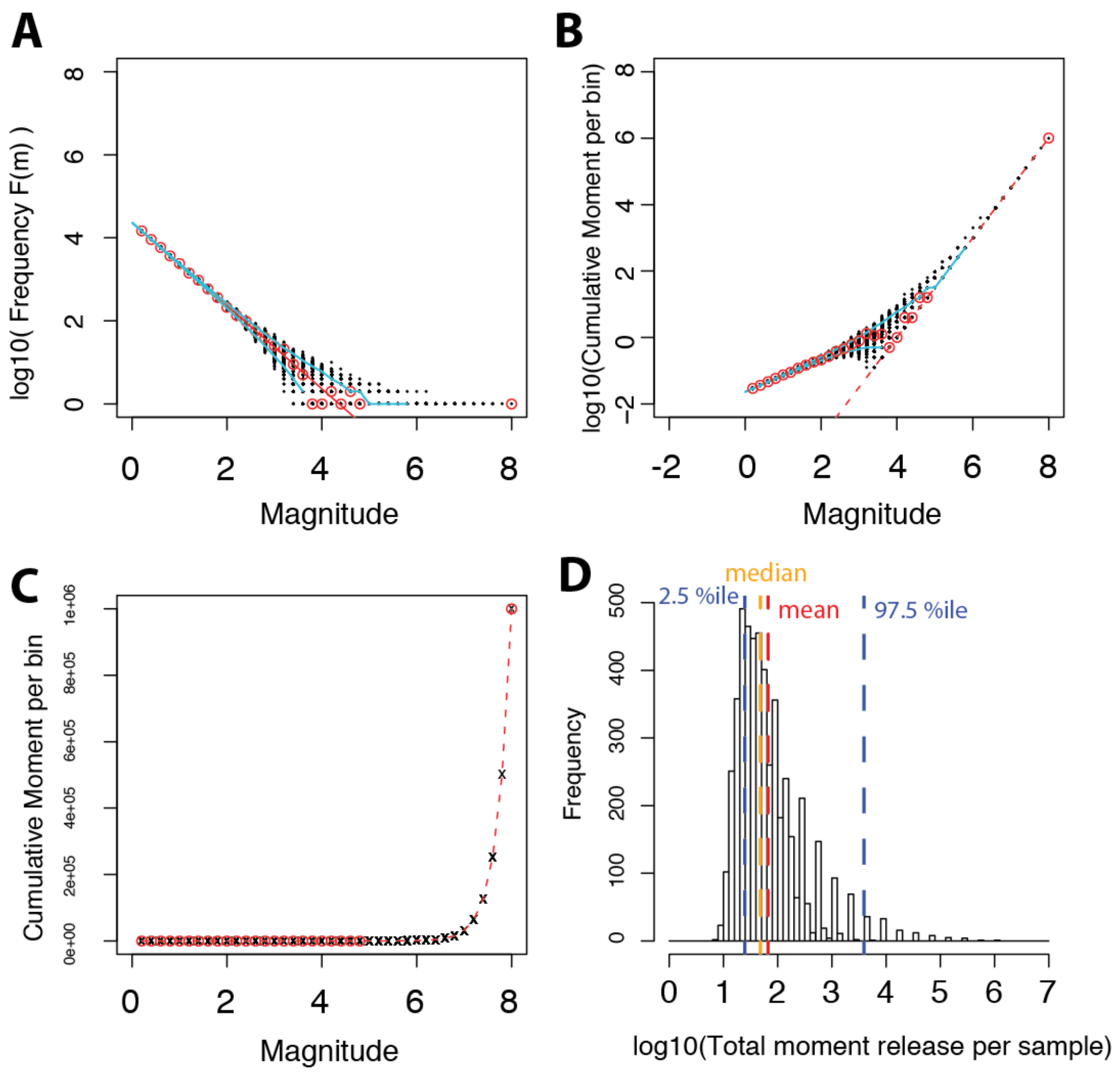

Figure 1: (a) Simulated incremental frequency-magnitude data from 5000 random samples of 50,000 events, drawn from a parent Gutenberg-Richter distribution with a b-value of 1, and plotted on top of each other, with magnitude increments or bin widths of 0.2. The red line indicates the known parent distribution and the blue lines the 95% confidence intervals calculated using Poisson count statistics. Data points circled in red represent the sample containing the largest event which is some 3 orders of magnitude greater than the next biggest event in this one sample. (b) Frequencies weighted by seismic moment plotted as in (a). The dashed line indicates the cut-off for whole numbers of events (n=1). (c) As in (b) but with a linear scale on the y-axis. (d) Histogram of the total moment released in each of the 5000 random samples. The red line is the mean, the orange line is the median, and blue lines are the 2.5 and 97.5 percentiles.

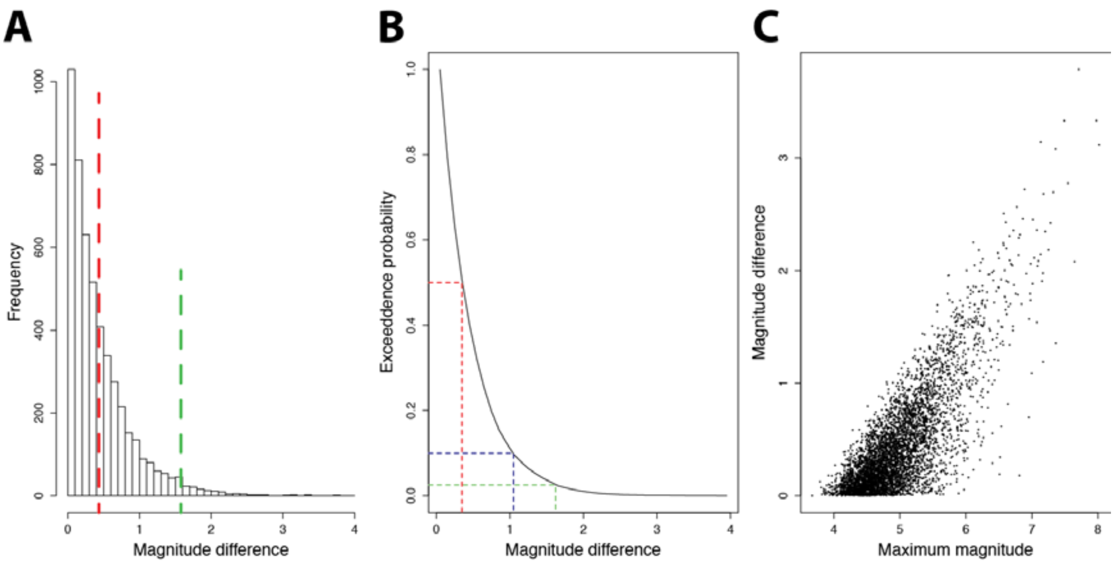

Figure 2: (a) Frequency of the magnitude difference between the largest and the second largest event from a simulation of a pure power-law distribution of seismic moment, obtained by the same simulations used to plot Fig 1. (b) Cumulative frequency expressed as a probability. (c) Scaling of the magnitude difference with respect to the maximum magnitude of the sample catalogue.

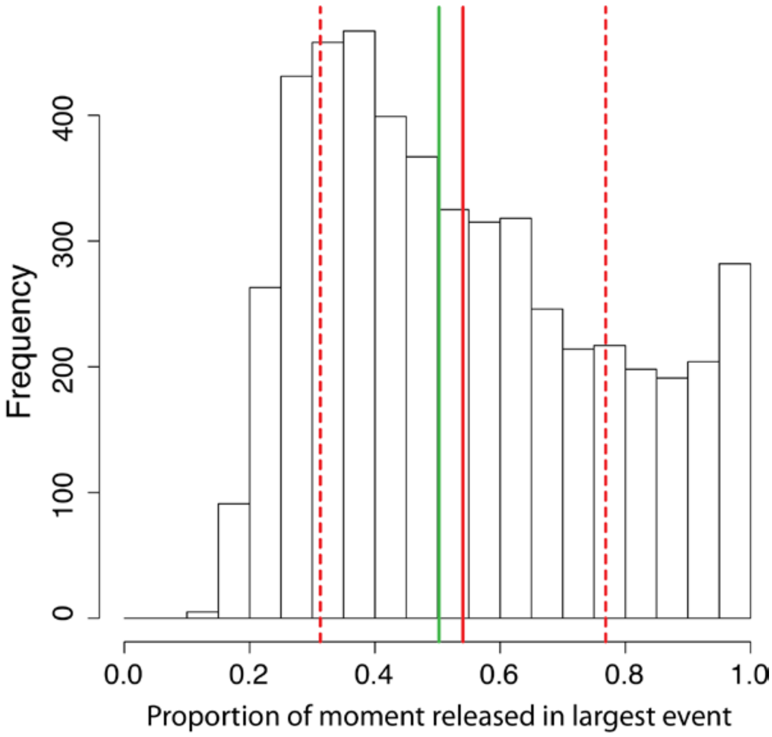

Figure 3: Histogram of the proportion of seismic moment released in the single largest event for all of the individual simulations shown in Fig 1, also indicating the mean (red line), median (green line) and one standard deviation range (red dashed lines).