

# Record-breaking earthquake intervals in a global catalogue and an aftershock sequence

M. R. Yoder<sup>1</sup>, D. L. Turcotte<sup>2</sup>, and J. B. Rundle<sup>1,2,3</sup>

<sup>1</sup>Department of Physics, University of California, Davis, California, 95616, USA

<sup>2</sup>Department of Geology, University of California, Davis, California, 95616, USA

<sup>3</sup>Santa Fe Institute, Santa Fe, New Mexico 87501, USA

Received: 29 October 2009 – Revised: 14 January 2010 – Accepted: 18 January 2010 – Published: 31 March 2010

**Abstract.** For the purposes of this study, an interval is the elapsed time between two earthquakes in a designated region; the minimum magnitude for the earthquakes is prescribed. A record-breaking interval is one that is longer (or shorter) than preceding intervals; a starting time must be specified. We consider global earthquakes with magnitudes greater than 5.5 and show that the record-breaking intervals are well estimated by a Poissonian (random) theory. We also consider the aftershocks of the 2004 Parkfield earthquake and show that the record-breaking intervals are approximated by very different statistics. In both cases, we calculate the number of record-breaking intervals ( $n_{rb}$ ) and the record-breaking interval durations  $\Delta t_{rb}$  as a function of “natural time”, the number of elapsed events. We also calculate the ratio of record-breaking long intervals to record-breaking short intervals as a function of time,  $r(t)$ , which is suggested to be sensitive to trends in noisy time series data. Our data indicate a possible precursory signal to large earthquakes that is consistent with accelerated moment release (AMR) theory.

Tata (1969) introduced a basic theory of record-breaking statistics for events that occur randomly. Tata’s paper addressed record-breaking statistics for a sequence of variables drawn from a continuous, independent identically distributed (iid) process. Glick (1978) published several applications of Tata’s method, including a brief study of daily temperatures. Benestad (2004, 2008) and Redner and Petersen (2006) further developed meteorological applications in the context of global warming; Vogel et al. (2001) applied the method to flooding in the United States, and Van Aalsburg et al. (2010) applied the method to global earthquake magnitudes.

Time series, such as maximum or minimum temperatures on a specified day of the year, are not truly random (iid) sequences. Important deviations include temporal correlations and temporal trends. Temporal correlations, in many naturally occurring time series, exhibit long-range correlations and self affinity (Turcotte, 1997). A standard measure of these correlations is the power-law dependence of the power spectral density  $S$  on frequency  $f$

$$S \sim f^{-\beta} \quad (1)$$

If  $\beta = 0$ , the time series is a white noise, comprised of a random (iid) sequence of values. In the range  $0 < \beta < 1$  the correlations are weak and the time series is weakly stationary. For the daily time series of temperatures, we typically observe  $\beta \approx 0.5$ , a Hurst exponent,  $Hu \equiv \frac{(\beta+1)}{2} = 0.75$  (Pelletier and Turcotte, 1999). Simulations show that for this value of  $\beta$  the iid theory of record-breaking statistics is a good approximation for the weakly correlated time series.

A second deviation of a time series from a random (iid) sequence involves a trend in the expected values. Simulations show that record-breaking statistics are sensitive to such trends. A specific example is the association of record-breaking temperature statistics with global warming. Benestad (2004, 2008) studied monthly maximum temperatures on a global basis. The number of record-breaking temperatures were determined both with time running forward and with

## 1 Introduction

A record-breaking event is defined to be one that is larger (or smaller) than all previous events. A typical example is the sequence of record-breaking temperatures (either highest or lowest) on a specified day of the year at a specified monitoring station. The rate at which records are broken is an important characteristic of the sequence; studies involve both the number of record-breaking temperatures and their values. The ratio of the number of record-breaking high temperatures to record-breaking low temperatures has been interpreted as a measure of global warming (Meehl et al., 2009).



Correspondence to: M. R. Yoder  
(yoder@physics.ucdavis.edu)

time running backwards. Significantly more forward record-breaking temperatures were found than backward record-breaking temperatures. This ratio can be quantitatively related to global warming. Redner and Petersen (2006) conducted a similar study, calculating the numbers of record breaking maximum and minimum temperatures in Philadelphia for each day of the year over a 120-year period. They present a framework for record-breaking climatological analysis.

The concept of record-breaking events can also be applied to earthquakes. A catalogue, study area, starting time and minimum magnitude must be specified. Van Aalsburg et al. (2010) considered record-breaking magnitudes of earthquakes in the global Centroid Moment Tensor (CMT) catalogue with moment magnitude  $M_W \geq 5.5$  over fifteen sequential, non-overlapping two-year periods between 1977 and 2006. For their study, a record-breaking earthquake magnitude is greater (or smaller) than the magnitude of any previous earthquake in the study region since the chosen starting time. Van Aalsburg et al. (2010) showed that the mean numbers of record-breaking earthquake magnitudes ( $n_{rb}$ ) and mean record-breaking magnitudes ( $M_{rb}$ ), determined from the CMT catalogue, agree closely with the iid theory.

The primary purpose of this paper is to determine whether record-breaking statistics can distinguish background sequences of main-shocks from correlated aftershocks. Because the frequency-magnitude distributions of main shocks and aftershocks are very similar, if not identical, record-breaking magnitude statistics cannot be used to separate the two classes of earthquakes. Instead, we will utilize the record-breaking interval statistics.

We construct catalogues by selecting all earthquakes within a region, with magnitudes greater than a specified minimum value. We consider the sequence of interval times between successive earthquakes in our catalogues. Record-breaking long intervals are the sequence of interval times longer than any previous interval times. Record-breaking short intervals are the sequence of interval times shorter than any previous interval times. The interval between the first and second earthquake is, by definition, the first record-breaking long interval. The next interval, longer than this interval, is the second record-breaking long interval, and so on. Similarly, the first interval, between the first and second earthquakes, is by definition also the first record-breaking short interval. The next interval shorter than this interval is the second record-breaking short interval, and so on. Intervals can be taken either forward or backward in time. We first consider global earthquakes with magnitudes greater than 5.5 and show that the record-breaking intervals are well estimated by a Poissonian (random) theory. In this case, the number of record-breaking long intervals ( $n_{rb-long}$ ) are statistically identical to the number of record-breaking short intervals ( $n_{rb-short}$ ).

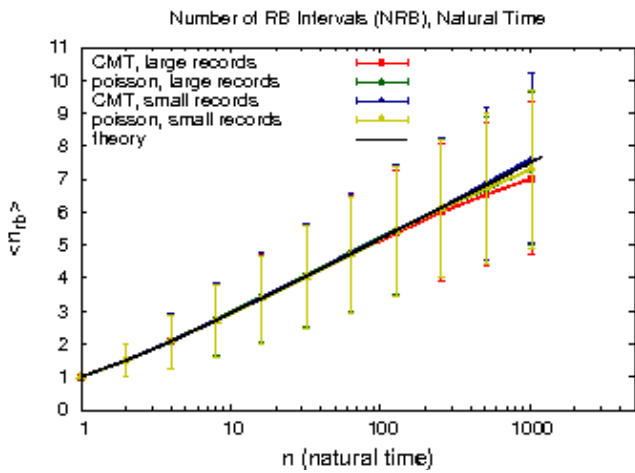
We also consider the aftershocks of the 2004 Parkfield earthquake and show that the record-breaking intervals are characterised by very different statistics. Because of the applicability of Omori's law, the interval times in an aftershock sequence become systematically longer. Thus, after the main shock, there are many more record-breaking long intervals than record-breaking short intervals ( $n_{rb-long} > n_{rb-short}$ ).

For both the sequence of global earthquakes and the sequence of aftershocks, we first determine the number of record-breaking intervals  $n_{rb}$ , both long and short ( $n_{rb-long}$  and  $n_{rb-short}$ , respectively) as a function of "natural time"  $n$ , the number of elapsed intervals. Second, we determine the record-breaking interval durations,  $\Delta t_{rb}$ , both long and short ( $\Delta t_{rb-long}$  and  $\Delta t_{rb-short}$ ) as a function of "natural time"  $n$ . Third, we calculate the ratio of the number of record-breaking long intervals to the number of record-breaking short intervals as a function of time  $r(t)$ , which is suggested to be sensitive to trends in noisy time series data (Benestad, 2004, 2008; Redner and Petersen, 2006). We will also show that our data indicate a possible precursory signal to large earthquakes that is consistent with the accelerated moment release (AMR) theory.

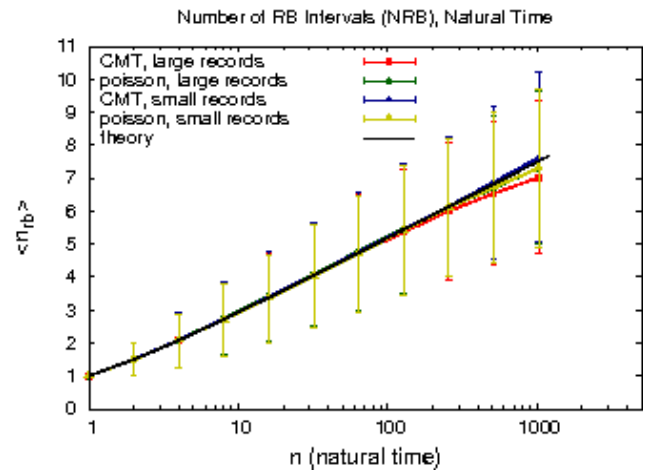
## 2 Record-breaking intervals in the CMT global catalog

We first calculate the record-breaking statistics of earthquake intervals for global earthquakes during the period January 1977–December 2006. We consider earthquakes with  $M_W \geq 5.5$  from the CMT catalogue over windows of  $n=1024$  intervals. Initially, we consider the 1024 intervals between the first 1025 earthquakes. Within this subsequence, we calculate the number of record-breaking long and record-breaking short intervals ( $n_{rb-long}$  and  $n_{rb-short}$ ) separately for each  $n_i = 2^i$  ( $i=0, 1, 2, \dots, 10$ ) elapsed intervals. Similarly, we calculate the record-breaking long and record-breaking short interval durations ( $\Delta t_{rb-long}$  and  $\Delta t_{rb-short}$ ) as a function of  $n_i$ . We advance the window one event at a time and repeat the above procedure to obtain 10 592 values for  $n_{rbi}$  and  $\Delta t_{rbi}$  (both long and short) for each  $n_i$ . We then determine the means and standard deviations of  $n_{rbi}$  and  $\Delta t_{rbi}$ , for both the longest and shortest intervals in the sequences. The mean values of  $n_{rb}$  as a function of the number of elapsed events (natural time)  $n$  are given in Fig. 1. Results are given for both longest record-breaking and shortest record-breaking interval times. The two results are almost identical and  $\langle n_{rb} \rangle \sim \ln(n)$  appears to be a good approximation. The mean lengths of both longest and shortest record-breaking intervals  $\langle \Delta t_{rb} \rangle$  and their standard deviations are given as a function  $n$  in Fig. 2. Again,  $\langle \Delta t_{rb} \rangle \sim \ln(n)$  appears to be a good approximation.

Next, we generate a synthetic catalogue of random event intervals. To do this, we utilize the cumulative distribution of the interval times in a homogeneous Poisson process (HPP)



**Fig. 1.** Mean numbers of record-breaking intervals  $\langle n_{rb} \rangle$  and their standard deviations are given as a function of event number  $n$  (natural time). Results are given for the longest and shortest intervals from the CMT catalogue and the synthetic random catalogue. Also included is the theoretical prediction for an iid process from Eq. (4).



**Fig. 2.** Mean lengths of record-breaking intervals  $\langle \Delta t_{rb} \rangle$  and their standard deviations are given as a function of event number  $n$  (natural time). Results are given for the longest and shortest intervals from the CMT catalogue and the synthetic random catalogue.

given by

$$F(\Delta t) = 1 - e^{-\frac{\Delta t}{\langle \Delta t \rangle}} \quad (2)$$

We consider a sequence of intervals with the same number of events as the CMT catalogue that we used. For each interval, we determine  $F(\Delta t)$  as a random number in the range 0 to 1 and solve Eq. (2) for the corresponding random interval  $\Delta t$ . The mean record-breaking numbers,  $\langle n_{rb} \rangle$ , and time intervals,  $\langle \Delta t_{rb} \rangle$ , are calculated from the synthetic catalogue as described above and included in Figs. 1 and 2. The CMT catalogue results are in good agreement with the random simulations.

### 3 Record-breaking events of an independent, identically distributed (iid) process

We will now show that the numbers of record-breaking intervals  $\langle n_{rb}(n) \rangle$ , as a function of the number of intervals  $n$ , as given in Fig. 1, is well approximated by the statistical analysis of a random process. The basic theory for random record-breaking events was developed by Tata (1969) and was clearly explained by Glick (1978). Their results are valid for any random process that has a continuous distribution of values; the results are independent of the particular distribution of values. We will apply this iid analysis both to the distribution of maximum intervals and to the distribution of minimum intervals.

We consider a sequence of random values,  $x_i$  ( $i = 1, 2, \dots, n$ ), selected from a continuous distribution. The first

element is always a record-breaking event. The second variable is larger or smaller than the first, with equal probability:

$$p_{rb}(2) = \frac{1}{2} \quad (3)$$

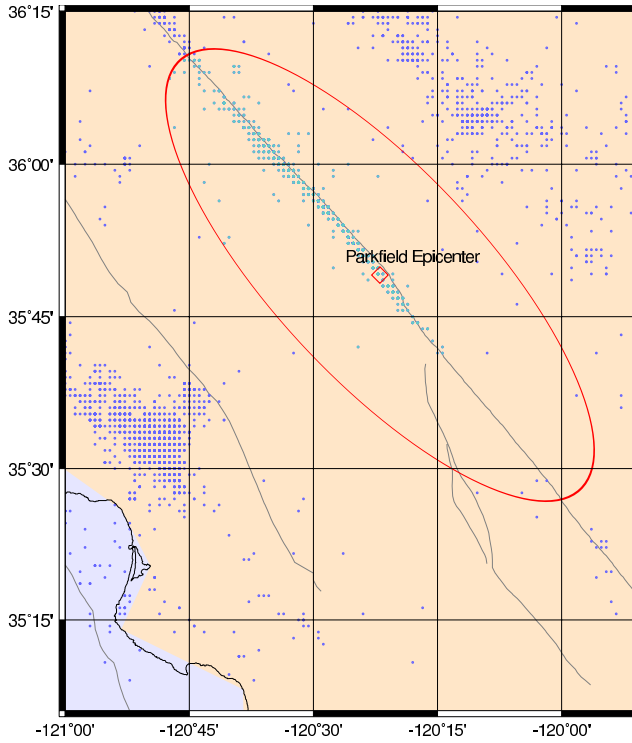
Because the sequence is random, the probability that any one element  $x_i$  occupies the  $j$ -th ( $j \in i$ ) position is  $1/n$ . For  $n = 2$ , the probability that the larger element terminates the sequence is  $1/2$ . For  $n = 3$ , the probability that the final element is the largest (or smallest) value is  $1/3$ . Accordingly, the expected number of record-breaking events  $\langle n_{rb} \rangle$  in an iid sequence is:

$$\langle n_{rb} \rangle_n = 1 + 1/2 + 1/3 + \dots + 1/n \quad (4)$$

For large  $n$ , we have approximately

$$\langle n_{rb} \rangle_n \approx \gamma + \ln(n) \quad (5)$$

where  $\gamma = 0.577215$  is the Euler-Mascheroni constant. For example, with  $n = 100$  we have  $\langle n_{rb}(100) \rangle = 5.187$  from Eq. (4) or  $\langle n_{rb}(100) \rangle = 5.183$  from Eq. (5). Even for  $n = 4$ , we have  $\langle n_{rb}(4) \rangle = 2.08$  from Eq. (4) and  $\langle n_{rb}(4) \rangle = 1.96$  from Eq. (5). It must be emphasized that the values given in Eqs. (4) and (5) represent the expected mean values for many realizations; Vogel et al. (2001) provide a more thorough analysis of the statistical moments of record-breaking sequences. The values from Eq. (4) are also included in Fig. 1 and are seen to be in good agreement with the global earthquake values and the random simulations.



**Fig. 3.** Earthquakes included in the Parkfield aftershock catalogue are enclosed by the red ellipse and events occurred after the 28 September 2004 mainshock ( $M_W = 5.96$ ); epicenter  $35.818^\circ$  N,  $-120.366^\circ$  W, inclined  $50^\circ$  south of east (Shcherbakov et al., 2006).

#### 4 Record-breaking intervals in the 2004 Parkfield aftershock sequence

Next, we consider the statistics of record-breaking maximum intervals in an aftershock sequence. As a specific example, we consider the 2004 Parkfield, CA earthquake. We expect record-breaking behaviour in an aftershock sequence to deviate substantially from the stationary Poisson process. Specifically, we expect intervals to increase in time, when time is measured forward, according to the modified form of Omori's law (Shcherbakov et al., 2006)

$$\frac{dn}{dt} = \frac{1}{\tau(m_c)} \cdot \frac{1}{[1+t/c(m_c)]^p} \quad (6)$$

where  $n$  is the number of aftershocks as a function of time  $t$  after the main shock and  $\tau(m_c)$  and  $c(m_c)$  are characteristic times obtained empirically from the data and depend on the minimum magnitude,  $m_c$ , considered.

We use the same definition of Parkfield aftershocks given by Shcherbakov et al. (2006). These are illustrated in Fig. 3; for this study, we obtained interval data from the Advanced National Seismic System (ANSS) catalogue. We start counting events from 0.01 day after the main-shock, to mitigate the effects of early aftershocks being masked by the coda of the main-shock, and measure time forward through 6 April

2009. We consider the record-breaking maximum intervals from a single pass through the entire aftershock sequence for minimum aftershock magnitude thresholds of  $m_c = 1.5, 2.0$ , and  $2.5$ . The numbers of record-breaking maximum intervals  $n_{rb}$  and the lengths of the record-breaking maximum intervals  $\Delta t_{rb}$  are given as functions of the number of aftershocks  $n$  (natural time) in Figs. 4 and 5, respectively. To a first approximation,  $n_{rb} \sim n$  and  $\Delta t_{rb} \sim \exp(\alpha n)$  (where  $\alpha$  is a fitting constant), in strong contrast to  $n_{rb} \sim \ln(n)$  and  $\Delta t_{rb} \sim \ln(n)$  shown for global earthquakes in Figs. 1 and 2.

#### 5 Non-Homogeneous Poisson Processes (NHPP)

The probability of some interval duration  $\Delta t = t_{i+1} - t_i$  is equal to the probability that zero events occurred between times  $t_{i+1}$  and  $t_i$ . For a Poisson process, the cumulative probability distribution function (CDF) can be expressed as (Ross, 2003; Shcherbakov et al., 2005; Yakovlev et al.):

$$F(\Delta t_i, t_i) = 1 - e^{-\int_0^{\Delta t_i} \lambda(t_i+v)dv} \quad (7)$$

where  $\lambda(v)$  is a rate. For the special case where  $\lambda(v) = \lambda_0$ , a constant, we recover the homogeneous Poisson process (HPP), Eq. (4), where  $\lambda_0 = 1/(\langle \Delta t \rangle)$ . When  $\lambda$  is not constant, we say the Poisson process is non-homogeneous.

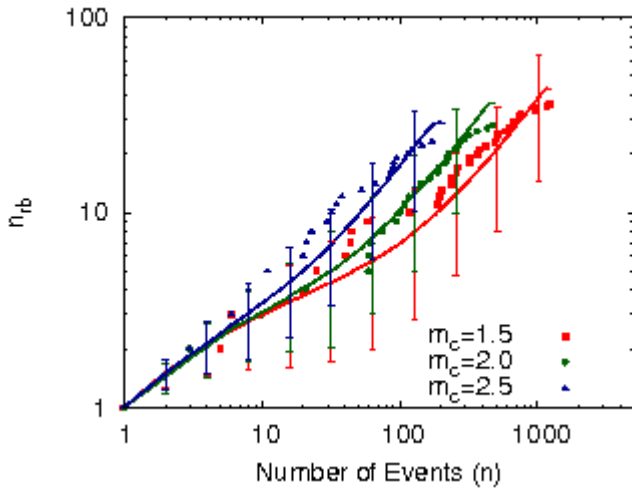
Substituting Omori's Law, Eq. (6), for  $\lambda(v)$  into Eq. (7) and assuming for simplicity  $p = 1$ , we integrate with the result

$$F(\Delta t, t) = \left( \frac{c(m_c) + t + \Delta t}{c(m_c) + t} \right)^{-\frac{c(m_c)}{\tau(m_c)}} \quad (8)$$

Solving for  $\Delta t$  and replacing  $F(\Delta t, t)$  with  $u$ , a random number in the range 0 to 1, we generate a time series from the relation:

$$\Delta t(t) = \left( u^{-\tau(m_c)/c_m} - 1 \right) \cdot (c(m_c) + t) \quad (9)$$

Taking the values  $\tau(m_c)$  and  $c(m_c)$  for  $m_c = 1.5, 2.0$ , and  $2.5$  from Shcherbakov et al. (2006), we produce a NHPP series of interval times for the period considered above (Shcherbakov et al., 2005, 2006). For each  $m_c$ , we find the mean and standard deviation of  $n_{rb}(n)$  and  $\Delta t_{rb}(n)$  over 1000 simulations and compare the results with the Parkfield data in Figs. 4 and 5. In Fig. 4, the general dependence of  $n_{rb}$  on  $n$  is approximately the power law with  $n_{rb}(n) \sim n^{0.5 \pm 0.1}$  for the number of longest record-breaking intervals as a function of the number of intervals. The deviation between the observed data and the NHPP simulations can be attributed, in part, to aftershocks of aftershocks which we have not included in the application of Omori's law, Eq. (6). Clearly, a strong, late aftershock can introduce a sequence of short intervals which will delay the occurrence of the next record-breaking long interval. In Fig. 5 the general dependence of  $\Delta t_{rb}$  on  $n$  is approximately exponential,  $\Delta t_{rb}(n) \sim \exp(\alpha n)$ , where  $\alpha$  is a fitting constant for the length of longest record-breaking intervals as a function of the number of elapsed intervals.



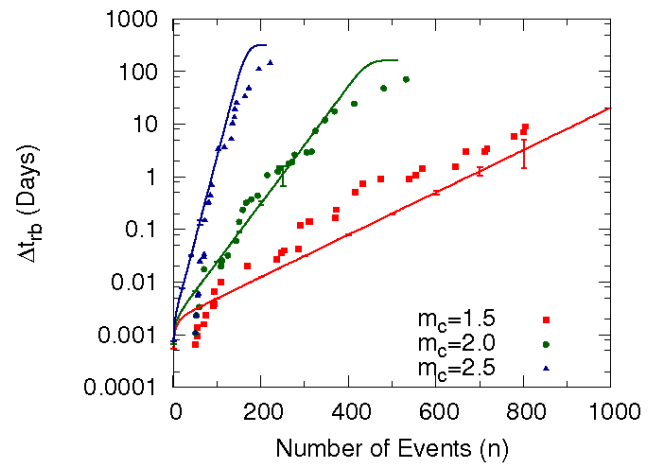
**Fig. 4.** Numbers of record-breaking longest intervals,  $n_{rb}$ , are given as a function of event number  $n$  (natural time). Results are given for the Parkfield aftershock sequence, starting 0.01 days after the main-shock, compared to a simulated non-homogeneous Poisson process (NHPP). The blue, green and red points represent data from the Parkfield aftershock sequence for  $m_c=1.5, 2.0, 2.5$ , respectively. The solid lines and error bars of the same colour represent the mean and standard deviation, over 1000 simulations, from the corresponding NHPP.

## 6 Record-breaking temperatures at the Mauna Loa Observatory (MLO), Hawaii

As an example of relating record-breaking events to trends, we will consider the statistics of the record-breaking maximum high and minimum low temperatures observed at the NOAA MLO, Big Island, Hawaii for the period 1977–2006. This observatory, remotely located and situated at an altitude of 3397 m a.s.l., provides a wide range of atmospheric data relatively unperturbed by local tropospheric, biospheric and anthropogenic activities (NOAA, 2008). Keeling et al. (1976) showed, from observations at MLO, that atmospheric levels of  $\text{CO}_2$  are systematically increasing.

In this context, a new record-breaking high temperature occurs when the maximum temperature for a given day is higher than all subsequent maximum temperatures on that calendar day. Similarly, a new record-breaking low temperature occurs when a day’s minimum temperature is lower than all preceding minimum temperatures. For each calendar day of the year, excluding leap years, we calculated the number of record-breaking high and record-breaking low temperatures since the same calendar day in 1976, our chosen starting date. We then averaged over the 365 days of each year to produce a mean number of record-breaking events as a function of time in one year increments:

$$\langle n_{rb}(\text{year}) \rangle = \frac{1}{365} \sum_{\text{day}=1}^{365} n_{rb,\text{day}}(\text{year}) \quad (10)$$



**Fig. 5.** Lengths of record-breaking longest intervals,  $\Delta t_{rb}$ , are given as a function of event number  $n$  (natural time). Results are given for the Parkfield aftershock sequence, starting 0.01 days after the main-shock, compared to a simulated non-homogeneous Poisson process (NHPP). The blue, green and red points represent data from the Parkfield aftershock sequence for  $m_c = 1.5, 2.0, 2.5$ , respectively; these data points are shifted 50 events to the right to further mitigate seismographic anomalies in the coda. The solid lines and error bars of the same colour represent the mean and standard deviation, over 1000 simulations, from the corresponding NHPP.

In Fig. 6, we show the mean numbers of record-breaking maximum high temperatures  $\langle n_{rb,\text{max}}(t) \rangle$  and record-breaking minimum low temperatures  $\langle n_{rb,\text{min}}(t) \rangle$  as a function of time measured forward from 1977 to 2006. Also included in Fig. 6 are the predicted values for an iid process from Eq. (4). We see that  $\langle n_{rb,\text{max}}(t) \rangle$  is systematically greater than  $\langle n_{rb,\text{min}}(t) \rangle$ . Following Redner and Petersen (2006), we introduce the ratio  $r$  of the values

$$r(t) = \frac{\langle n_{rb,\text{max}}(t) \rangle}{\langle n_{rb,\text{min}}(t) \rangle} \quad (11)$$

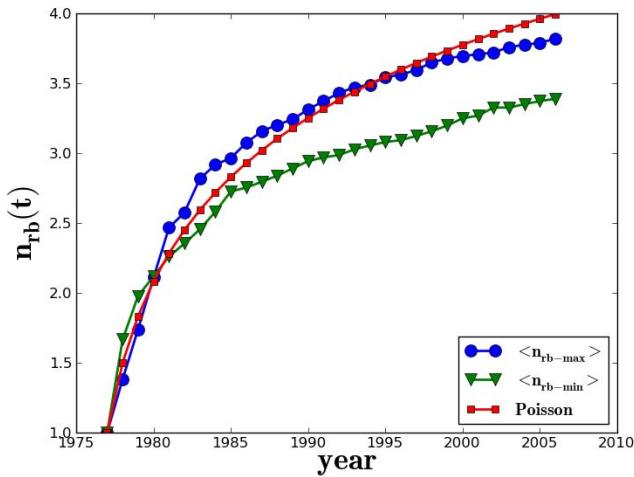
Values of this ratio as a function of time measured forward are given for the MLO data in Fig. 7. We find a near constant value in the range  $1.13 < r < 1.15$  for the period 1977–2006, indicating systematic global warming.

## 7 Ratios of record-breaking earthquake intervals

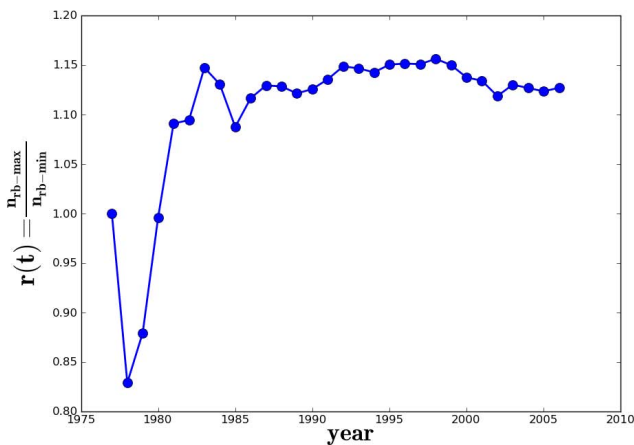
We now extend the concept of a ratio of record-breaking temperatures introduced in Eq. (11) to sequences of earthquake interval times. We introduce the ratio  $r(n, t)$  of record-breaking longer intervals to record-breaking shorter intervals

$$r(n, t) \equiv \frac{n_{rb\text{-long}}(n, t)}{n_{rb\text{-short}}(n, t)} \quad (12)$$

where  $n$  is the number of events in the sequence,  $t$  is the time at which the sequence terminates and, unless otherwise



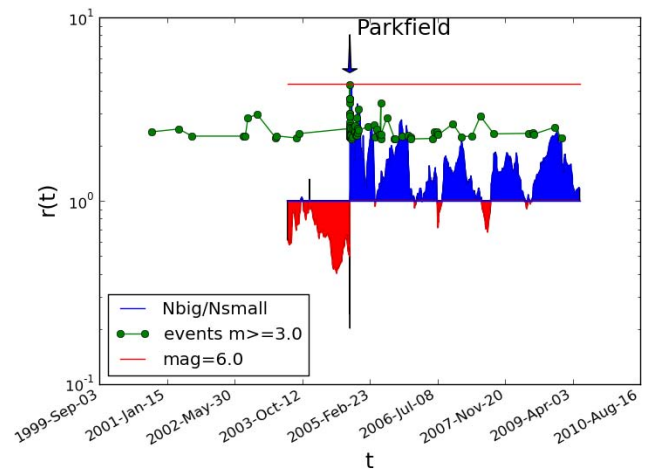
**Fig. 6.** Mean number of record-breaking maximum temperatures  $\langle n_{rb-max} \rangle$  and record-breaking minimum temperatures  $\langle n_{rb-min} \rangle$  from MLO are given as a function of time measured forward from 1977 to 2006. Also included is the theoretical prediction for an iid process from Eq. (4).



**Fig. 7.** Values of the ratio  $r(t)$ , defined in Eq. (11), from MLO are given as a function of time measured forward from 1977 to 2006.

specified, we assume that time is measured forward. For our global study, illustrated in Fig. 1, we see that  $\langle n_{rb-long}(n, t) \rangle$  is statistically identical to  $\langle n_{rb-short}(n, t) \rangle$  for background seismicity, so  $r(n, t)$  is expected to fluctuate around 1. For our Parkfield study, illustrated in Fig. 4, we see that  $n_{rb-long}(n, t)$  is systematically greater than  $n_{rb-short}(n, t)$ , so  $r(n, t)$  is expected to be predominantly greater than 1. It may be possible to use this difference to separate aftershocks from background seismicity.

To explore this possibility, we consider earthquakes with minimum magnitude  $m_c \geq 1.5$  and use a moving window of 256 intervals. We pick an event, an earthquake that occurred at  $t = t_0$ , and consider the 256 intervals between the preceding 256 earthquakes. Starting with the first of these earthquakes, we obtain the number of record-breaking intervals

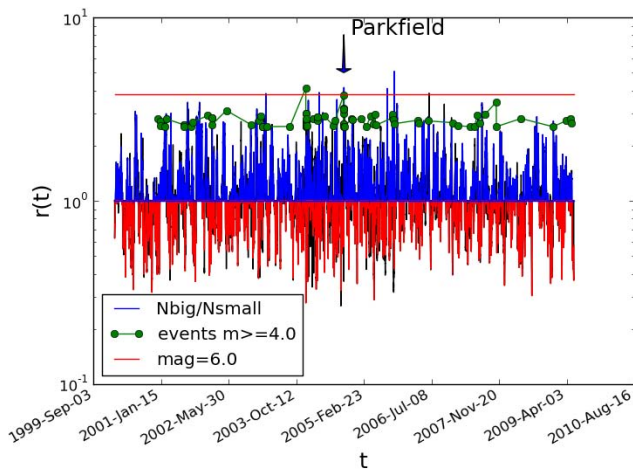


**Fig. 8.** The ratios of record-breaking longest intervals to record-breaking shorter intervals,  $r(t)$ , in the Parkfield aftershock sequence (illustrated in Fig. 3) with minimum magnitude  $m_c \geq 1.5$  are shown as a function of date. Large and small record breaking intervals are counted forward starting from the 256-th interval preceding the interval ending at the time  $t$ . Each value,  $r(n = 256, t_i)$  is smoothed by averaging over the preceding 16 values. Blue regions, where  $r(n, t) > 1$ , imply aftershocks, or at least some form of decreasing mean seismicity; red regions, where  $r(n, t) < 1$ , suggest an increasing rate of seismicity.

$n_{rb-long}(256, t_0)$  and  $n_{rb-short}(256, t_0)$ . We obtain the value of  $r(256, t_0)$ , from Eq. 12, and assign it to time  $t_0$ . We repeat this process with subsequent earthquakes,  $t > t_0$ , to obtain a time series,  $r(256, t_0)$ . We first consider the area of Parkfield aftershocks defined in Fig. 3. The values of  $r$  as a function of time are given in Fig. 8. We see that, after the Parkfield earthquake, the values of  $r$  are predominantly greater than unity, as expected.

We next consider earthquakes for the same period of time and the same minimum magnitude in a  $4^\circ \times 4^\circ$  region centered on the Parkfield epicenter. The values of  $r$  as a function of time are given in Fig. 9; a random-like behaviour is clearly illustrated. Using the broader catalogue, the Parkfield aftershock is obscured, presumably by aftershocks from uncorrelated earthquakes. This is consistent with our expectations from Figs. 1 and 4, which suggest that aftershock sequences produce more large than small interval records compared to a broader, background catalogue.

The variability and short periods of  $r < 1$  after the mainshock, in Fig. 6, presumably indicate aftershocks of events in the primary aftershock sequence (aftershocks of aftershocks). Of particular interest is the period preceding the Parkfield main-shock where consistently  $r < 1$ . Early results indicate that  $r < 1$  indicates some sort of precursory seismic acceleration; further study will consider whether this is a systematic phenomenon that can be quantified. It may be possible to tune this method to resolve events of different



**Fig. 9.** Same as Fig. 8 but for a  $4^\circ \times 4^\circ$  region centered on the Parkfield epicenter.

magnitudes by varying the number of events in the record-breaking window or the spatial geometry of the area being considered.

## 8 Conclusions

We have studied the record-breaking statistics of two very different earthquake catalogues – a global dataset taken from the CMT catalogue and an isolated aftershock sequence, taken from the ANSS catalogue for the 2004 Parkfield earthquake. In each case, we find that the number of record-breaking intervals is consistent with an established theory and with simulations. Specifically, we find that the global catalogue produces record-breaking behaviour that is well estimated by a homogeneous Poisson process (HPP), where  $n_{rb}(n) \sim \ln(n)$ . For an isolated aftershock sequence, in which intervals become systematically longer with time, we find that  $n_{rb}(n) \sim n^{0.5 \pm 0.1}$ , which is in reasonable agreement with a non-homogeneous Poisson process (NHPP). Our results are consistent with the sensitivity of record-breaking statistics to trends, as discussed in the introduction. For the examples given above, we have shown a strong distinction between the record-breaking statistics of a large global earthquake catalogue and a well-defined aftershock sequence.

This method shows promise as a simple, computationally efficient test for trends in time series data. In particular, we suggest that we can characterise a given catalogue of earthquakes as being dominated by background, aftershock, or possibly AMR seismicity. To separate correlated aftershock sequences from a broader catalogue, one requires a method to systematically select likely sub-catalogues, which can then be tested by the methods described in this paper. In the case of our Parkfield example, knowledge of the location

and geometry of the rupture was important for our analysis. To first order, aftershock sequences of many past events can be isolated visually, by simply selecting spatially clustered earthquakes in the vicinity of the epicentre and excluding seismicity clustered around neighbouring large events; elliptical regions appear to be a reasonable first estimate. Preliminary record-breaking interval studies of the 1999 ( $M=7.1$ ) Hector Mine event suggest that this simple approach is sufficient to produce convincing signal to noise in Eq. (12). Numerical two point correlation methods might also be employed, and of course more complex geometries, for example, ruptures on multiple faults, present greater challenges. It may be possible to develop a method by which one starts with a simple geometry and systematically adds and removes events in order to optimize a correlation metric and converge upon a complete aftershock catalogue.

Forecasting applications require more sophisticated, efficient methods to locate possible rupture epicentres and geometries. Clearly, our successful retrospective forecast of the 2004 Parkfield earthquake, as one might interpret Fig. 8 to suggest, benefits from advance knowledge of the rupture epicenter and geometry. Anticipating this geometry, however, is non-trivial and has so far proven an unresolved obstacle to AMR based forecasts. Calculating all rupture geometries at all locations of a large map is computationally impractical, and fault models are incomplete and unreliable indicators of the locations of future epicentres. It might be possible to use seismic rate-based hazard maps, for example, Relative Intensity (RI) or Pattern Informatics (PI) (Holliday et al., 2006), to reduce the problem to one that is computationally feasible. Epicentres can be too constrained to “hot-spot” regions likely to experience seismicity; the size of the rupture region, where aftershocks and AMR are clustered, can be estimated as a function of mainshock magnitude. Presumably, the record-breaking sequence  $n$  is also related to the magnitude of the event being forecast. Likely rupture geometries can be inferred by convolving along clusters or contours of the hazard map. Again, by adjusting the geometry of the region being analysed to optimize some metric, for example, Eq. (12), it may be possible to converge upon the precise parameters of likely rupture area geometries in advance of mainshocks. In both the forecasting and retrospective cases, the simplicity and computational efficiency of record-breaking methods may contribute to the computational feasibility of comprehensive solutions.

*Acknowledgements.* We wish to acknowledge valuable discussions with B. Malamud, G. Yakovlev, and W. Newman. We also wish to acknowledge the suggestions of two anonymous reviewers who greatly improved the presentation of this paper. This work has been supported by NASA Grant NNX08AF69G.

Edited by: J. Davidsen

Reviewed by: two anonymous referees

## References

- Benestad, R. E.: Record-values, nonstationarity tests and extreme value distributions, *Global Planet. Change*, 44, 11–26, doi:10.1016/j.gloplacha.2004.06.002, 2004.
- Benestad, R. E.: A simple test for changes in statistical distributions, *EOS T. Am. Geophys. Un.*, 89, 389–390, doi:10.1029/2008EO410002, 2008.
- Glick, N.: Breaking records and breaking boards, *Am. Math. Mon.*, 85, 2–26, 1978.
- Holliday, J. R., Rundle, J. B., Tiampo, K. F., and Turcotte, D. L.: Using earthquake intensities to forecast earthquake occurrence times, *Nonlin. Processes Geophys.*, 13, 585–593, 2006, <http://www.nonlin-processes-geophys.net/13/585/2006/>.
- Keeling, C. D., Bacastow, R. B., Bainbridge, A. E., Ekdahl, C. A., Guenther, P. R., Waterman, L. S., and Chin, J. F. S.: Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii, *Tellus*, 28, 538–551, 1976.
- Meehl, G. A., Tebaldi, C., Walton, G., Easterling, D., and McDaneil, L.: The relative increase of record high maximum temperatures compared to record low minimum temperatures in the U.S., *Geophys. Res. Lett.*, 36, L23701, doi:10.1029/2009GL040736, 2009.
- NOAA: Oceanic and Atmospheric Administration, Earth System Research laboratory (ESRL), [Online], mauna Loa hourly temperature data for 1 January 1977 to 7 February 2007, available at: <ftp://ftp.cmdl.noaa.gov/met/> (last access: December 2008), 2008.
- Pelletier, J. D. and Turcotte, D. L.: Self-affine time series: II. Applications and models, *Adv. Geophys.*, 40, 91–166, 1999.
- Redner, S. and Petersen, M. R.: Role of global warming on the statistics of record-breaking temperatures, *Phys. Rev. E*, 74(6), 061114, doi:10.1103/PhysRevE.74.061114, 2006.
- Ross, S. M.: *Probability Models*, Academic Press, 8th edn., 2003.
- Shcherbakov, R., Yakovlev, G., Turcotte, D. L., and Rundle, J. B.: Model for the distribution of aftershock interoccurrence times, *Phys. Rev. Lett.*, 95, 218501, doi:10.1103/PhysRevLett.95.218501, 2005.
- Shcherbakov, R., Turcotte, D. L., and Rundle, J. B.: Scaling properties of the Parkfield aftershock sequence, *Bull. Seis. Soc. Am.*, 96, S376–S384, doi:10.1785/0120050815, 2006.
- Tata, M.: On outstanding values in a sequence of random variables, *Z. Wahrsch. verw. Geb.*, 12, 9–20, doi:10.1007/BF00538520, 1969.
- Turcotte, D. L.: *Fractals and Chaos in Geology and Geophysics*, Cambridge University Press, 1997.
- Van Aalsburg, J., Turcotte, D. L., Newman, W. I., and Rundle, J. B.: Record-breaking earthquakes, *B. Seimol. Soc. Am.*, in press, 2010.
- Vogel, R. M., Zafirakou-Koulouris, A., and Matalas, N. C.: Frequency of record-breaking floods in the United States, *Water Resour. Res.*, 37, 1723–1731, doi:2001WR900019.0043-1397/01/2001WR900019, 2001.
- Yakovlev, G., Newman, W. I., Turcotte, D. L., and Gabrielov, A.: An inverse cascade model for self-organized complexity and natural hazards, *Geophys. J. Int.*, 163, 433–442.