

Rescaled earthquake recurrence time statistics: application to microrepeaters

Christian Goltz,^{1,2} Donald L. Turcotte,² Sergey G. Abaimov,² Robert M. Nadeau,³ Naoki Uchida⁴ and Toru Matsuzawa⁴

¹Kiel University, Germany. E-mail: sgabaimov@gmail.com

²Department of Geology, UC Davis, Davis, CA, USA

³Berkeley Seismology Laboratory, UC Berkeley, Berkeley, CA, USA

⁴Research Center for Prediction of Earthquakes and Volcanic Eruptions, Graduate School of Science, Tohoku University, Sendai, Japan

Accepted 2008 October 3. Received 2008 October 2; in original form 2007 December 20

SUMMARY

Slip on major faults primarily occurs during ‘characteristic’ earthquakes. The recurrence statistics of characteristic earthquakes play an important role in seismic hazard assessment. A major problem in determining applicable statistics is the short sequences of characteristic earthquakes that are available worldwide. In this paper, we introduce a rescaling technique in which sequences can be superimposed to establish larger numbers of data points. We consider the Weibull and log-normal distributions, in both cases we rescale the data using means and standard deviations. We test our approach utilizing sequences of microrepeaters, micro-earthquakes which recur in the same location on a fault. It seems plausible to regard these earthquakes as a miniature version of the classic characteristic earthquakes. Microrepeaters are much more frequent than major earthquakes, leading to longer sequences for analysis. In this paper, we present results for the analysis of recurrence times for several microrepeater sequences from Parkfield, CA as well as NE Japan. We find that, once the respective sequence can be considered to be of sufficient stationarity, the statistics can be well fitted by either a Weibull or a log-normal distribution. We clearly demonstrate this fact by our technique of rescaled combination. We conclude that the recurrence statistics of the microrepeater sequences we consider are similar to the recurrence statistics of characteristic earthquakes on major faults.

Key words: Probability distributions; Earthquake dynamics; Statistical seismology.

INTRODUCTION

Because of their complexity, earthquakes satisfy several scaling laws with considerable robustness. The most famous of these is the Gutenberg–Richter frequency–magnitude scaling. It can be shown that this is equivalent to a fractal (power law) scaling between the number of earthquakes and the linear dimension of rupture. The cumulative number N_c with rupture areas greater than A_r scales as $N_c \propto A_r^{-D/2}$, where the fractal dimension $D \approx 2$.

Another property of seismicity that has been shown to satisfy a universal scaling law is the statistical distribution of interoccurrence times. All earthquakes in a specified region and specified time window with magnitudes greater than m are considered to be point events. Based on studies of properties of California seismicity, a unified scaling law for the temporal distribution of recurrence times between successive earthquakes was proposed by Bak *et al.* (2002). Two distinct scaling regimes were found. For short times, aftershocks dominate the scaling properties of the distribution, decaying according to the modified Omori’s law. For long times, an

exponential scaling regime was found that can be associated with the occurrence of main shocks. To take into account the spatial heterogeneity of seismic activity, it has been argued that the second regime is not an exponential but another power law (Corral 2003).

Shcherbakov *et al.* (2005) showed that the observed distribution of interoccurrence times can be well approximated by a distribution derived from assuming that aftershocks follow non-homogeneous Poisson statistics with the rate given by Omori’s law.

The main focus of this paper is the scaling of recurrence times. It is important to distinguish between the statistical distribution of recurrence times between earthquakes on a specific fault or fault section, and the interoccurrence times between earthquakes on all of the faults in a region. This difference is clearly illustrated by the behaviour of the San Andreas fault in California. Quasi-periodic sequences of great or moderate earthquakes include great earthquakes on the northern section of the San Andreas fault (the 1906 San Francisco earthquake), the southern section of the San Andreas fault (the great 1857 earthquake) and moderate earthquakes ($m \approx 6$) on the Parkfield section of the San Andreas fault in

central California. Offsets on these faults are dominated by the largest earthquakes. These earthquakes are generally known as characteristic earthquakes. Detailed discussions of the relationship between recurrence times and characteristic earthquakes have been given by Turcotte *et al.* (2007) and Abaimov *et al.* (2007a,b, 2008).

Characteristic earthquakes are associated with quasi-periodicity, but can also have considerable variability. A measure of the variability of recurrence times on a fault or fault segment is the coefficient of variation C_V (the ratio of the standard deviation σ to the mean μ). For strictly periodic earthquakes on a fault or fault segment, we would have $\sigma = C_V = 0$. For a random (i.e. exponential with no memory) distribution of recurrence times, we would have $C_V = 1$, (i.e. $\sigma = \mu$). Ellsworth *et al.* (1999) analysed 37 series of recurrent earthquakes and suggested a provisional generic value of the coefficient of variation $C_V \approx 0.5$. A number of alternative distributions have been proposed for the purpose of specifying the statistics of recurrence times. These include the exponential (Poisson), the log-normal, Brownian passage-time (inverse Gaussian) and Weibull (stretched exponential) distributions (Davis *et al.* 1989; Sornette & Knopoff 1997; Ogata 1999; Matthews *et al.* 2002).

The statistical distribution of recurrence times of characteristic earthquakes is important in defining the seismic hazard (Working Group on California Earthquake Probabilities 2003). For example, the last characteristic earthquake on the northern section of the San Andreas fault in 1906 was responsible for the destruction of much of San Francisco. It has been almost 100 yr since that characteristic earthquake and the mean recurrence interval of these characteristic earthquakes is estimated to be about 200 yr. If these characteristic earthquakes are periodic, then the next earthquake can be expected in about 100 yr. If these characteristic earthquakes are Poissonian, that is, have an exponential distribution, then the future probability distribution would be exponential with a mean of about 200 yr independent of when the last earthquake occurred.

Ideally, observed sequences of earthquakes on a fault should be used to establish the applicable statistical distribution for recurrence times. However, the number of events in observed earthquake recurrence sequences is not sufficient to establish definitively the validity of a particular distribution (Savage 1994). In this paper, we introduce a rescaling technique in which sequences of interval times can be superimposed. The result is that larger numbers of data points can be used to test alternative statistical distributions. Our approach can be used for the Weibull and log-normal distributions but not for the Brownian passage-time distribution.

To test our approach we consider the recurrence time statistics of microrepeaters, small earthquakes ($m \approx 1-2$) which recur at the same locations on a fault. We consider sequences that are not associated with aftershocks and find that the recurrence time statistics are stationary and have a similar behaviour to the recurrence time statistics of characteristic earthquakes on major faults. Schaff *et al.* (1998), Peng *et al.* (2005) and Peng & Ben-Zion (2006) have carried out such studies during aftershock sequences and find that the interval statistics are consistent with Omori's law.

APPLICABLE DISTRIBUTIONS

One objective of this paper is to determine whether there is a preferred distribution for recurrence time intervals of microrepeaters. Two widely used statistical distributions for recurrence time intervals are the log-normal and Weibull. We will compare each of them with our data.

A primary objective of this paper is to superimpose recurrence times for large numbers of microrepeaters to test alternative statistical distributions. We will illustrate how this will be done for the Weibull and log-normal distributions.

The probability distribution function (pdf) for a Weibull distribution is given by (Patel *et al.* 1976)

$$p(t) = \frac{\beta}{\tau} \left(\frac{t}{\tau}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\tau}\right)^\beta\right], \quad (1)$$

where β and τ are fitting parameters. The mean μ and the aperiodicity (coefficient of variation) C_V of the Weibull distribution are given by

$$\mu = \tau \Gamma\left(1 + \frac{1}{\beta}\right), \quad (2)$$

$$C_V = \left\{ \frac{\Gamma\left(1 + \frac{2}{\beta}\right)}{\left[\Gamma\left(1 + \frac{1}{\beta}\right)\right]^2} - 1 \right\}^{1/2}, \quad (3)$$

where $\Gamma(x)$ is the gamma function of x . The corresponding cumulative distribution function (cdf) for the Weibull distribution is given by

$$P(t) = 1 - \exp\left[-\left(\frac{t}{\tau}\right)^\beta\right]. \quad (4)$$

If $\beta = 1$ the Weibull distribution becomes the exponential (Poisson) distribution with $\sigma = \mu$ and $C_V = 1$. In the limit $\beta \rightarrow +\infty$, the Weibull distribution becomes a δ -function with $\sigma = C_V = 0$. In the range $0 < \beta < 1$, the Weibull distribution is often referred to as the stretched exponential distribution.

An important property of the Weibull distribution is the power-law behaviour of the hazard function

$$h(t_0) = \frac{\text{pdf}}{1 - \text{cdf}} = \frac{\beta}{\tau} \left(\frac{t_0}{\tau}\right)^{\beta-1}. \quad (5)$$

The hazard function $h(t_0)$ is the probability that an event will occur during an interval δt_0 at a time t_0 after the occurrence of the last event. For the Poisson distribution, $\beta = 1$, the hazard function is constant $h(t_0) = \tau^{-1}$ and there is no memory of the last event. For $\beta > 1$, the hazard rate increases as a power of the time t_0 since the last event. We argue that the hazard rate must increase after a characteristic earthquake. A characteristic earthquake reduces the regional stress. Tectonic loading slowly increases the regional stress until the next characteristic earthquake occurs. Complexity introduces fluctuations in stress but the risk of the next earthquake must increase with time. As a specific example consider 1906 San Francisco earthquake. The plate motion is loading the San Andreas fault resulting in an increased probability of the next great San Francisco earthquake.

Our rescaled combination approach for the Weibull correlations is carried out as follows. We consider m sets of recurrence times for m sets of microrepeater earthquakes. For the j th set, we have n_j recurrence times t_{ij} with $i = 1, 2, \dots, n_j$. We order these times from the shortest to the longest and construct the cumulative distribution of the t_{ij} . We then obtain the best fit of the cumulative Weibull distribution from eq. (4) to these values using a maximum likelihood test and obtain τ_j and β_j . Using these values we have a renormalized set of recurrence times given by $(t_{ij}/\tau_j)^{\beta_j}$. This is repeated for the m sets of microrepeaters.

The values of the rescaled times are next ordered from smallest to largest for all sets of microrepeaters. The cumulative distribution of these values is then obtained. If the distributions are well approximated by the Weibull then there will be a good fit to the exponential relation given in eq. (4).

The pdf for a log-normal distribution of recurrence times t is given by (Patel *et al.* 1976)

$$p(t) = \frac{1}{(2\pi)^{1/2}\sigma_y t} \exp\left(-\frac{(\ln t - \mu_y)^2}{2\sigma_y^2}\right). \quad (6)$$

The log-normal distribution can be transformed into a normal distribution by making the substitution $y = \ln t$; μ_y and σ_y are the mean and standard deviation of this equivalent normal distribution. The mean μ , standard deviation σ and aperiodicity (coefficient of variation) C_V for the log-normal distribution are given by

$$\mu = \exp\left(\mu_y + \frac{\sigma_y^2}{2}\right), \quad \sigma = \mu_y \sqrt{e^{\sigma_y^2} - 1} \quad \text{and} \\ C_V = \frac{\sigma}{\mu} = \sqrt{e^{\sigma_y^2} - 1}. \quad (7)$$

The corresponding cdf $P(t)$ (fraction of recurrence times that are shorter than t) for the log-normal distribution is

$$P(t) = \frac{1}{2} \left\{ 1 + \operatorname{erf} \left[\frac{\ln t - \mu_y}{2^{1/2}\sigma_y} \right] \right\}, \quad (8)$$

where $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy$ is the error function.

Our rescaled combination approach for the log-normal correlations is very similar to the approach used for the Weibull correlations. Again we consider the m sets of recurrence times for m sets microrepeater earthquakes. The recurrence times t_{ij} are ordered for each j set and the cumulative distribution is plotted. We then obtain the best fit of the cumulative log-normal distribution from eq. (8) to the distribution by using a maximum likelihood test and we find μ_{yj} and σ_{yj} . Using these values we have a rescaled set of recurrence times given by $(\ln t - \mu_{yj})/2^{1/2}\sigma_{yj}$. This is repeated for each of the m sets of microrepeater earthquakes.

The values of the renormalized times are next ordered from smallest to largest for all sets of microrepeaters. The cumulative distribution of these values is then obtained. If the distribution is well represented by the log-normal distribution then there will be a good fit to the error function relation given in eq. (8).

Although the Weibull and log-normal distributions are the most widely used distributions for recurrence time statistics, the Brownian passage-time distribution has recently been used (Matthews *et al.* 2002). We do not consider this distribution in this paper because it is not possible to rescale it as we have done for the Weibull and log-normal distributions. The reason for this is that the mean μ enters the Brownian passage-time distribution both as an additive and multiplicative factor. In the Weibull and log-normal distributions, the fitting parameters are either multiplicative or additive, they are not both. Thus it is not possible to rescale the Brownian passage-time distribution as we have done for the Weibull and log-normal distributions.

CALIFORNIA MICROREPEATER EARTHQUAKES

Repeating sequences of small earthquakes in central California have been widely studied. Some 300 clusters of these repeating earthquakes on the San Andreas fault have been recognized (Nadeau

et al. 1995; Schaff *et al.* 1998; Nadeau & McEvilly 1999, 2004). Some clusters contain as many as 20 similar regularly occurring ‘characteristic’ micro-earthquakes. Each sequence has a common hypocentre and very similar waveforms. They appear to be the result of periodic ruptures of an asperity on a creeping section of a fault. Time intervals between characteristic micro-earthquakes range from months to years with the intervals scaling with earthquake magnitude (Nadeau & Johnson 1998; Chen *et al.* 2007). For similar repeaters on the Calaveras fault, Vidale *et al.* (1994) and Marone *et al.* (1995) correlated time intervals with rupture duration.

In some cases, the microrepeaters occur during aftershock sequences and rates of occurrence scale with Omori’s law (Schaff *et al.* 1998; Peng *et al.* 2005; Peng & Ben-Zion 2006). In this paper, we will concentrate our attention on sequences that exhibit stationarity. We eliminate sequences that exhibit systematic trends in the recurrence times. We will also only consider sequences that include at least 10 intervals to obtain significant statistics for the distribution of recurrence times.

We will consider sequences on the creeping section of the San Andreas fault in central California. These sequences have been previously studied by Nadeau *et al.* (1995), Nadeau & Johnson (1998) and Nadeau & McEvilly (2004). We have studied in detail 28 sequences. The locations of these sequences are given in Fig. 1. Each sequence contained at least 11 events and covered about 20 yr, the earliest event was in 1984 and the most recent in 2005.

The majority of locations are on the creeping section north of the Parkfield segment with only three locations on the locked Parkfield segment. The numbering scheme seen in Fig. 1 will be used to discuss individual sequences below.

Some sequences, for example, 21 and 24, show a shortening of waiting times towards the end, that is, acceleration in the failure process. There is also evidence of deceleration and missing data in sequence 12.

We estimate the number of missing characteristic events to be on the order of 5 per cent in general. The reason is that both catalogue entries and waveforms from the Northern California Seismic Network (NCSN) are required to identify members of microrepeater sequences (characteristic sequences, CSs), however, the procedure is not yet optimized for the smaller microrepeating events. Catalogue entries are generally available before their corresponding waveforms. The waveforms are available only after an analyst reviews the data and finalizes the event, and in some cases, waveforms are not made available even after review due to the small magnitudes of the events considered. However, as with any seismological network, when large event aftershock sequences occur, small events are missed due to network saturation and in addition the review and finalization process falls behind. For the data sets considered here, we therefore expect about 15–20 per cent missing events during the first month after the 2003 December 22 San Simeon $M = 6.5$ earthquake for example.

Only six sequences could be considered to be stationary without any data manipulation. These are marked in red in Fig. 1. No satisfactory fits of any distribution can be obtained for recurrence time data of non-stationary sequences. In the following, we will concern ourselves with stationary data only.

Properties of the six sequences we consider are given in Table 1. Included in this table are the set number, the number of events n , mean magnitude \bar{m} , mean recurrence time μ , coefficient of variation C_V and the Weibull parameters τ and β for each sequence. The number of events range from $n = 12$ –14, the magnitudes range

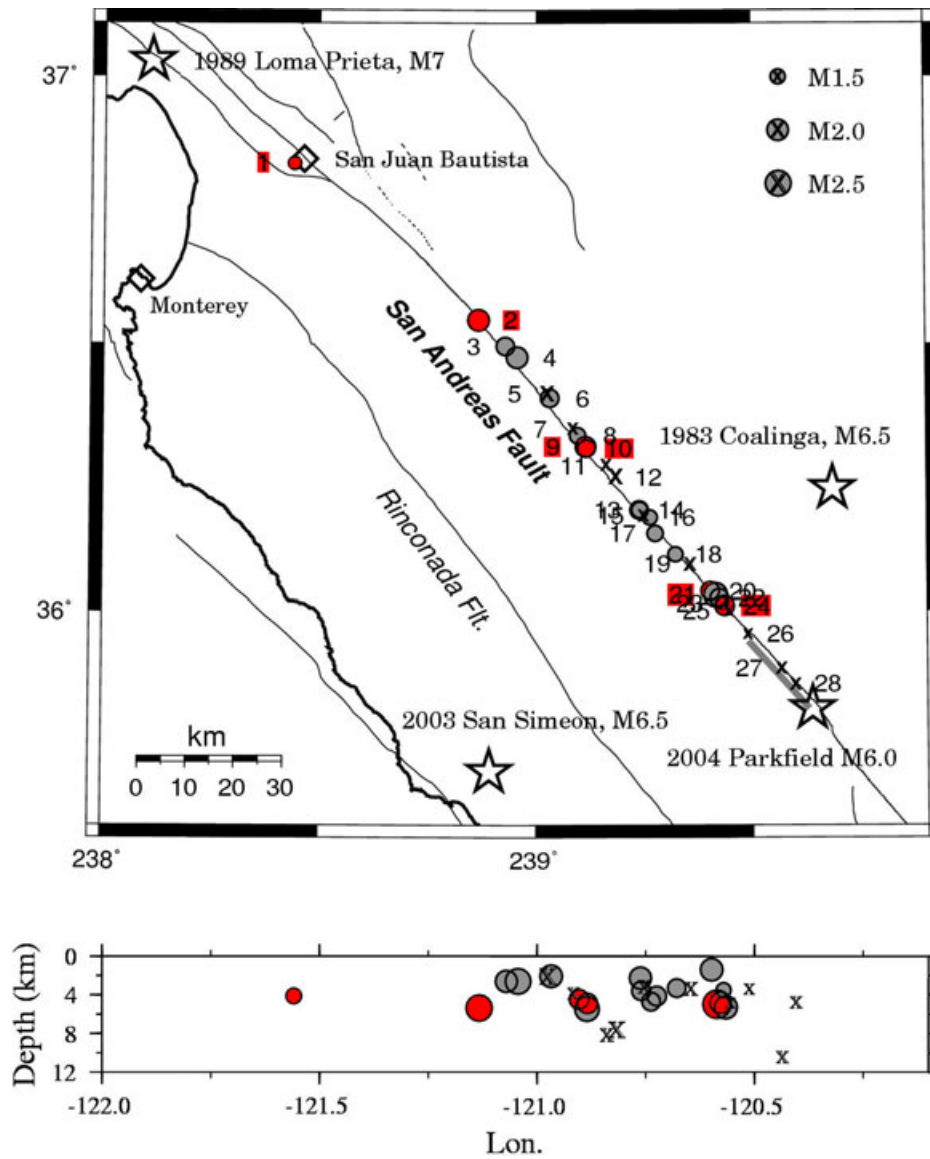


Figure 1. Locations of microrepeater sequences that occurred during the last ≈ 20 yr on the Parkfield segment and the neighbouring creeping section of the San Andreas fault. Each numbered location corresponds to a sequence of at least 10 repeating events. Data of sufficient stationarity are marked in red. Note small average magnitudes.

Table 1. Properties of the six sets of repeating micro-earthquakes from the Parkfield region that we consider.

Set	n	\bar{m}	μ (days)	C_V	τ (days)	β
2	14	2.07	559	0.25	611	4.78
8	13	1.62	501	0.24	598	4.96
9	13	1.36	509	0.38	572	2.91
10	14	1.84	561	0.27	618	4.36
14	14	1.95	500	0.20	541	5.42
23	12	2.22	654	0.33	728	3.35

Note: included are the set number, number of events n , mean earthquake magnitude \bar{m} , mean recurrence time μ , coefficient of variation C_V and the Weibull parameters τ and β .

from $\bar{m} = 1.36 - 2.22$, the mean recurrence times range from $\mu = 500 - 655$ days, and the coefficients of variation range from $C_V = 0.20 - 0.38$. The mean of the coefficients of variation is $C_V = 0.278$.

As a typical example we illustrate set 9 in Fig. 2. The cdf $P(t)$ of the 13 recurrence times t is given. Also included is the best-fit Weibull distribution from eq. (4). The fitting parameters are $\tau = 572$ days and $\beta = 2.91$. The values of $(t/\tau)^\beta$ are then obtained for the 13 recurrence times. For example, the shortest recurrence time is $t = 252$ days so that $(t/\tau)^\beta = 0.092$. When this process is carried out for the six sets we have 80 values of $(t/\tau)^\beta$. The cdf $P[(t/\tau)^\beta]$ of these values of $(t/\tau)^\beta$ is given in Fig. 3. Based on the applicability of the eq. (4) the fit of these data to the exponential distribution is also given in Fig. 3. The log likelihood of this fit is -74.0 .

We next consider the fit to a log-normal distribution. To illustrate our approach we again consider set 9 illustrated in Fig. 2. From Table 1 we have $\mu = 509.92$ days and $\sigma = 195.31$ d. From eq. (7) we find that $\mu_y = 6.166$ and $\sigma_y = 0.370$ with μ and σ in days. The values of $[(\ln t - \mu_y)/2^{1/2}\sigma_y]$ are then obtained for the recurrence times. For example, the shortest recurrence time is $t = 252$ days so that $[(\ln t - \mu_y)/2^{1/2}\sigma_y] = -0.741$. When this process is carried

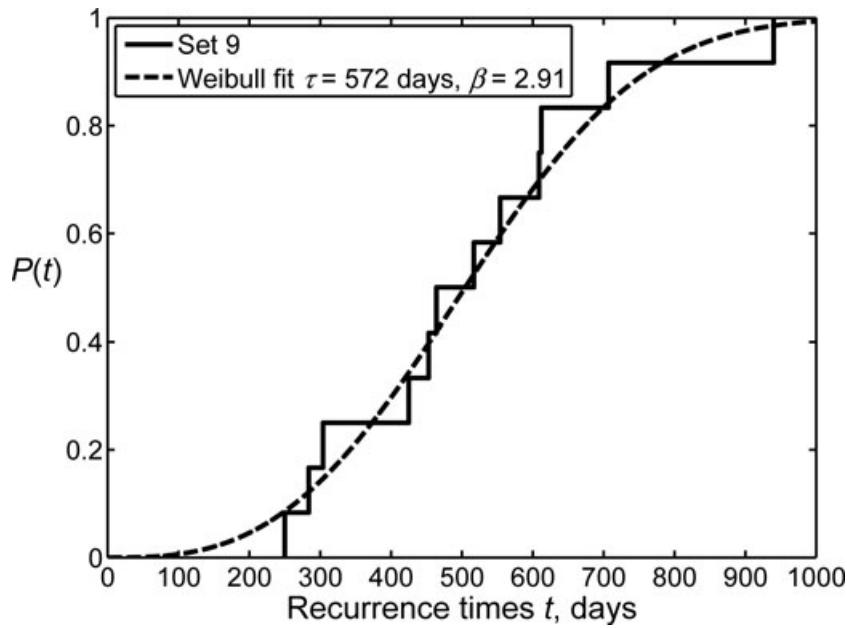


Figure 2. The cdf $P(t)$ of the 13 recurrence times t for data set 9 of Parkfield repeating micro-earthquakes. Also included is the best fit of the cumulative Weibull distribution from eq. (4) taking $\tau = 572$ days and $\beta = 2.91$.

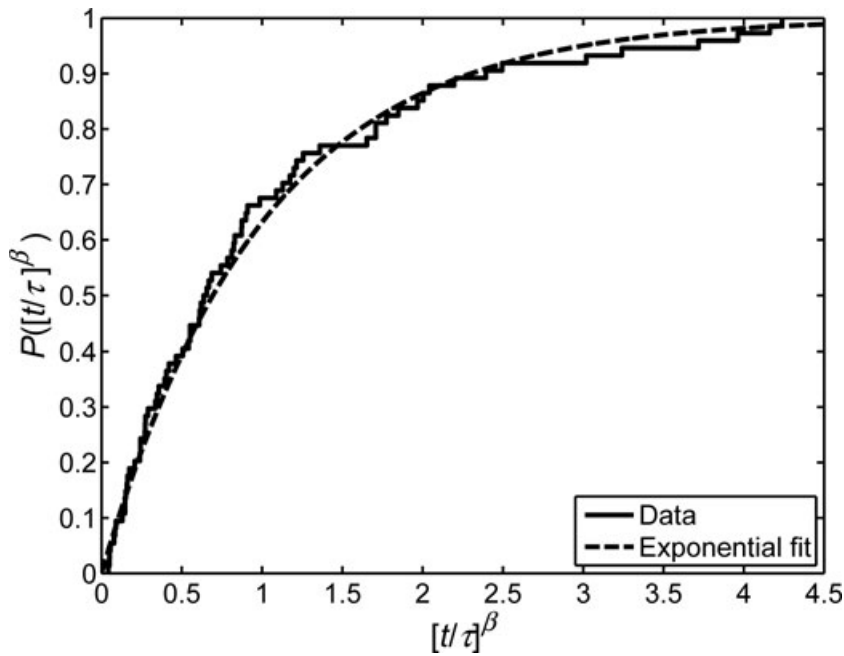


Figure 3. The cdf $P([t/\tau]^\beta)$ of the 80 values of $[t/\tau]^\beta$ for the six Parkfield data sets that we have considered. Also included is the exponential dependence given in eq. (4) for a Weibull distribution.

out for the six sets we have 80 values of $[(\ln t - \mu_y)/2^{1/2}\sigma_y]$. The cdf $P([(\ln t - \mu_y)/2^{1/2}\sigma_y])$ of these values is given in Fig. 4. Based on the applicability of eq. (8) the best fit of the data to the error function distribution is also given in Fig. 4. The log likelihood of this fit is -76.2 . This value is very close to the value -74.0 for the Weibull distribution so that it is not possible to clearly prefer one distribution over the other.

MICROREPEATER EARTHQUAKES IN THE NORTHEASTERN JAPAN SUBDUCTION ZONE

Repeating sequences of small earthquakes in the northwestern subduction zone have also been widely studied (Matsuzawa *et al.* 2002; Igarashi *et al.* 2003; Uchida *et al.* 2003, 2006). More than 300

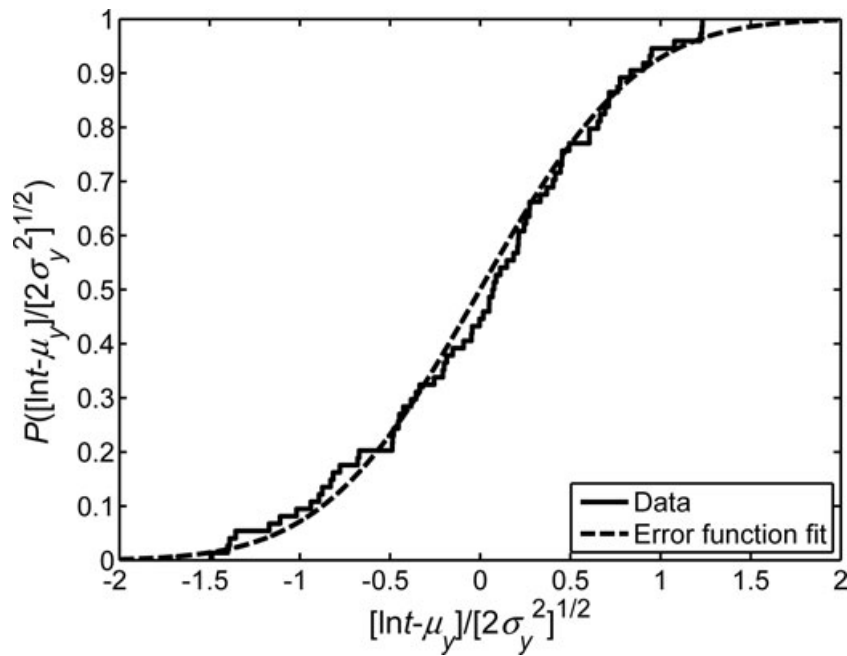


Figure 4. The cdf $P([\ln(t-\mu_y)/2\sigma_y^2]^{1/2})$ of the 80 values of $(\ln t-\mu_y)/2^{1/2}\sigma_y$ for the six Parkfield data sets that we have considered. Also included is the expected fit to the error function dependence given in eq. (8) for a log-normal distribution.

clusters of these repeating earthquakes have been recognized. It has been hypothesized that these earthquakes are caused by slips on small asperities surrounded by aseismic slip on the plate boundary. We have applied the same approach described for California to select 27 data sets that exhibit steady-state behaviour.

Properties of 27 data sets we consider are given in Table 2. Included in this table are the set number, the number of events n , the mean magnitude \bar{m} , mean recurrence time μ and the coefficient of variation for each sequence. The number of events ranges from $n = 10-28$, the mean magnitudes range from $\bar{m} = 2.30-3.54$, the mean recurrence times range from $\mu = 276$ days to 825 days and the coefficients of variation range from $C_V = 0.362-1.17$. The mean of the coefficients of variation is $C_V = 0.564$. It is interesting to note that this is more than twice as large as the value for the California data. The variability of the recurrence times on the strike-slip fault is much smaller than the subduction zone fault. The difference may be attributed to the greater complexity of the subduction zone microrepeater earthquakes relative to the near planar strike-slip behaviour of the California microrepeater earthquakes.

As a typical example we illustrate set 294 in Fig. 5. The cdf $P(t)$ of the 23 recurrence time t is given. Also included is the best-fit Weibull distribution from eq. (4). The fitting parameters are $\tau = 370$ days and $\beta = 3.89$. The values of $(t/\tau)^\beta$ are then obtained for the 23 recurrence times. When this process is carried out for the 27 data sets we have 405 values of $(t/\tau)^\beta$. The cdf $P([t/\tau]^\beta)$ of these values of $(t/\tau)^\beta$ is given in Fig. 6. Based on the applicability of the eq. (4) the fit of these data to the exponential distribution is also given in Fig. 6. The log likelihood of this fit is -377 .

We next consider the fit to a log-normal distribution. For set 294 illustrated in Fig. 5, we have $\mu = 334$ days and $\sigma = 102$ days from Table 2. From eq. (7) we find that $\mu_y = 5.77$ and $\sigma_y = 0.297$ with μ and σ in days. The values of $[(\ln t-\mu_y)/2^{1/2}\sigma_y]$ are then obtained for the 23 recurrence times in this set. When this process is carried out for the 27 data sets we have 405 values of $[(\ln t-\mu_y)/2^{1/2}\sigma_y]$. The cdf $P([\ln t-\mu_y)/2^{1/2}\sigma_y]$ of these values is given in Fig. 7. Based on the applicability of eq. (8) the best fit of the data to the error function

Table 2. Properties of the 27 sets of repeating micro-earthquakes from the Tohoku region, Japan, that we consider.

Set	n	\bar{m}	μ (d)	C_V
139	10	2.83	608	0.365
174	17	2.78	290	0.361
177	10	2.59	347	0.414
243	13	2.46	562	0.656
250	14	2.95	520	0.616
252	20	2.77	325	0.621
260	14	3.37	521	0.473
274	28	2.54	275	0.393
276	12	3.14	484	0.643
282	11	3.48	558	1.166
293	10	3.57	782	0.708
294	23	3.31	334	0.303
397	20	2.74	394	0.570
399	13	3.25	588	0.813
400	15	2.30	411	0.896
428	10	2.83	803	0.662
432	22	3.05	338	0.526
434	26	2.62	279	0.436
523	14	3.22	464	0.390
545	14	2.96	477	0.553
550	11	2.95	725	0.582
574	10	2.45	825	0.432
580	15	3.19	516	0.415
673	11	3.05	650	0.459
676	14	2.80	554	0.567
781	15	3.17	477	0.639
800	13	3.43	565	0.567

Note: included are the set number, number of events n , mean earthquake magnitude \bar{m} , mean recurrence time μ and coefficient of variation C_V .

distribution is also given in Fig. 7. The log likelihood of this fit is -390 . Again this value for the fit of the log-normal distribution is close to the value -377 for the Weibull distribution so again it is not possible to definitely prefer one distribution over the other.

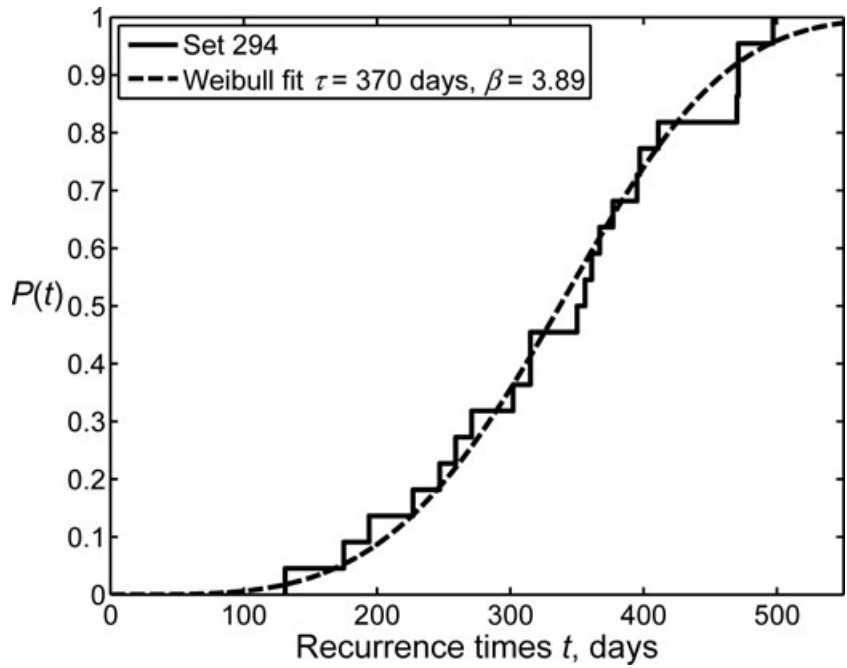


Figure 5. The cdf $P(t)$ of the 23 recurrence times t for data set 294 of the Tohoku repeating micro-earthquakes. Also included is the best fit of the cumulative Weibull distribution from eq. (4) taking $\tau = 370$ days and $\beta = 3.89$.

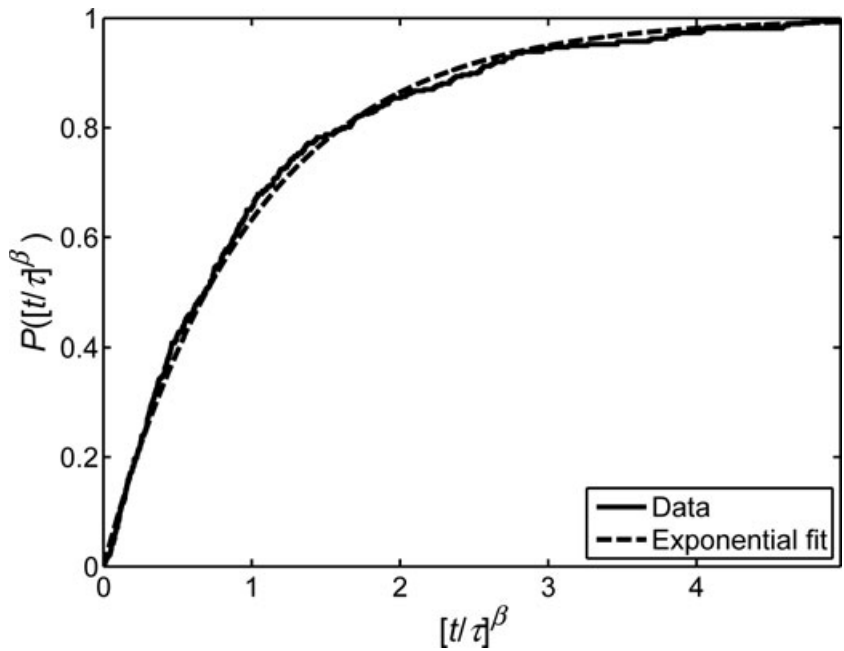


Figure 6. The cdf $P([t/\tau]^\beta)$ of the 405 values of $[t/\tau]^\beta$ for the 27 Tohoku data sets that we have considered. Also included is the exponential dependence given in eq. (4) for a Weibull distribution.

DISCUSSION

The purpose of this paper has been to consider two aspects of recurrence time statistics of characteristic earthquakes on a fault or fault segment. These are (1) the application of a rescaling technique and (2) the behaviour of the recurrence time statistics of microrepeater earthquakes. Recurrence time statistics of characteristic earthquakes play an important role in seismic hazard assessment. But the number of well-established sequences of characteristic earthquakes is small and the number of earthquakes in each

sequence is also small. Thus it has been difficult to establish the validity of alternative statistical distributions. In this paper, we introduce a rescaling technique in which sequences of interval times can be superimposed. The result is that larger numbers of data points can be used to test alternative statistical distributions. Our approach can be used for the Weibull and log-normal distributions but not for the Brownian passage-time distribution. The reason for this is that the two fitting parameters in the Weibull and log-normal distributions are either additive or multiplicative but not both. For the Weibull distribution we rescale utilizing the rescaled times $t_{jrs} = (t_j/\tau_j)^\beta$

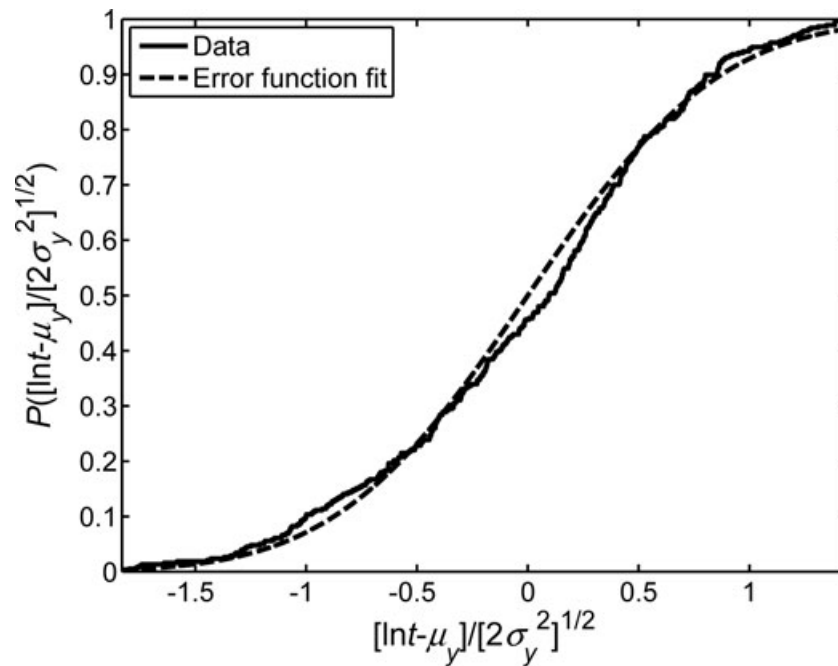


Figure 7. The cdf $P[(\ln t - \mu_y)/2^{1/2}\sigma_y]$ of the 405 values of $(\ln t - \mu_y)/2^{1/2}\sigma_y$ for the 27 Tohoku data sets that we have considered. Also included is the expected fit to the error function dependence given in eq. (8) for a log-normal distribution.

where t_{ij} are the i interval times of data set j and τ_j and β_j are the Weibull fitting parameters for set j . For the log-normal distribution we rescale utilizing the rescaled times $t_{ijrs} = (\ln t_{ij} - \mu_{yj})/2^{1/2}\sigma_{yj}$, where μ_{yj} and σ_{yj} are the log-normal fitting parameters for data set j . For the Brownian passage-time distribution the fitting parameter μ additive and multiplicative (see eq. 12 in Matthews *et al.* 2002) so that rescaling is not possible.

As a test of our rescaling approach we consider six stationary sequences of microrepeater earthquakes on the San Andreas fault in central California and 27 stationary sequences of microrepeater earthquakes in northeastern Japan.

We consider the two sets separately and introduce our rescaled technique for both the Weibull and the log-normal distribution. For each sequence we obtain the best-fit fitting parameters for the two distributions. This constitutes a rescaling and the six (27) sequences are then combined to form a single statistical set each. For the Weibull distribution the cdf of the values of $(t/\tau)^\beta$ is expected to fit an exponential distribution. These fits are given in Fig. 3 for the California data and in Fig. 6 for the Japanese data. For the log-normal distribution the cdf of the values of $[(\ln t - \mu_y)/2^{1/2}\sigma_y]$ is expected to fit an error function distribution. These fits are given in Fig. 4 for the California data and in Fig. 7 for the Japanese data. The log-likelihood fits for the four data sets are: California Weibull -74.0 and log-normal -76.2 , Japan Weibull -377 and log-normal -390 . Although the fits are quite good it is not possible to clearly establish a preference for one of the two distributions based on these data sets. This is despite analysis of data sets of unprecedented lengths obtained from our novel method of rescaled combination. The main reason is the absence of extremely long waiting times, that is, the lack of samples from the tails of the distributions. The tails of the two distributions differ while they are very similar otherwise. It would be interesting to apply rescaled combination to other time-series such as volcanic eruption sequences.

It is also of interest to compare the statistical behaviour of the recurrence times of the microrepeater earthquakes with related phenomena. We have shown that the distributions of recurrence times for our two sets of data can be well approximated by either the Weibull or the log-normal distribution. This is also the case for characteristic earthquakes under a variety of circumstances (Abaimov *et al.* 2007a,b, 2008; Turcotte *et al.* 2007). A measure of the variability of recurrence times is the coefficient of variation. For the six sequences we considered for the San Andreas fault we found that the mean value of the coefficient of variation is $C_V = 0.278$. The coefficient of variation for the seven characteristic earthquakes that have occurred on the Parkfield section of the San Andreas fault is $C_V = 0.378$ (Abaimov *et al.* 2008). These values tend to be somewhat lower than the usual value near $C_V = 0.5$ (Ellsworth *et al.* 1999). Abaimov *et al.* (2007a) studied the recurrence time statistics of slip events on the creeping section of the San Andreas fault in central California. The slip events typically had magnitudes of a few millimetres and were aseismic. The distributions were shown to fit Weibull statistics to a good approximation. The 67 events recorded at 10min intervals at the Cienega Winery creepmeter had a coefficient of variation $C_V = 0.554$. Clearly the values for the creepmeter slip events are higher than the values for the microrepeater earthquake sequences. For the 27 sequences of microrepeater earthquakes we studied from northwestern Japan the mean value of the coefficient of variation is $C_V = 0.564$. This is a value that is typical of sequences of characteristic earthquakes.

ACKNOWLEDGMENTS

CG is funded through a EU Marie Curie Outgoing International Fellowship and currently hosted by U. C. Davis. RMN was funded through National Science Foundation Grants EAR-0337308, EAR-0544730 and EAR-0510108.

REFERENCES

- Abaimov, S.G., Turcotte, D.L. & Rundle, J.B., 2007a. Recurrence-time and frequency-slip statistics of slip events on the creeping section of the San Andreas fault in central California, *Geophys. J. Int.*, **170**, 1289–1299.
- Abaimov, S.G., Turcotte, D.L., Shcherbakov, R. & Rundle, J.B., 2007b. Recurrence and interoccurrence behavior of self-organized complex phenomena, *Nonlinear Proc. Geophys.*, **14**, 455–464.
- Abaimov, S.G., Turcotte, D.L., Shcherbakov, R., Rundle, J.B., Yakovlev, G., Goltz, C. & Newman, W.I., 2008. Earthquakes: recurrence and interoccurrence times, *Pure appl. Geophys.*, **165**, 777–795.
- Bak, P., Christensen, K., Danon, L. & Scanlon, T., 2002. Unified scaling law for earthquakes, *Phys. Rev. Lett.*, **88**, 178501.
- Chen, K.H., Nadeau, R.M. & Rau, R.J., 2007. Towards a universal rule on the recurrence interval scaling of repeating earthquakes? *Geophys. Res. Lett.*, **34**, L16308.
- Corral, A., 2003. Local distributions and rate fluctuations in a unified scaling law for earthquakes, *Phys. Rev. E*, **68**, 035102(R).
- Davis, P.M., Jackson, D.D. & Kagan, Y.Y., 1989. The longer it has been since the last earthquake, the longer the expected time till the next, *Bull. seism. Soc. Am.*, **79**, 1439–1456.
- Ellsworth, W.L., Matthews, M.V., Nadeau, R.M., Nishenko, S.P., Reasenberg, P.A. & Simpson, R.W., 1999. A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities, USGS, Open-File Report 99–522.
- Igarashi, T., Matsuzawa, T. & Hasegawa, A., 2003. Repeating earthquakes and interplate aseismic slip in the northeastern Japan subduction zone, *J. geophys. Res.*, **108**, 2249.
- Marone, C., Vidale, J.E. & Ellsworth, W.L., 1995. Fault healing inferred from time-dependent variations in-source properties of repeating earthquakes, *Geophys. Res. Lett.*, **22**, 3095–3098.
- Matsuzawa, T., Igarashi, T. & Hasegawa, A., 2002. Characteristic small-earthquake sequence off Sanriku, northeastern Honshu, Japan, *Geophys. Res. Lett.*, **29**, 1543.
- Matthews, M.V., Ellsworth, W.L. & Reasenberg, P.A., 2002. A Brownian model for recurrent earthquakes, *Bull. seism. Soc. Am.*, **92**, 2233–2250.
- Nadeau, R.M. & Johnson, L.R., 1998. Seismological studies at Parkfield VI: moment release rates and estimates of source parameters for small repeating earthquakes, *Bull. seism. Soc. Am.*, **88**, 790–814.
- Nadeau, R.M. & McEvilly, T.V., 1999. Fault slip rates at depth from recurrence intervals of repeating microearthquakes, *Science*, **285**, 718–721.
- Nadeau, R.M. & McEvilly, T.V., 2004. Periodic pulsing of characteristic microearthquakes on the San Andreas Fault, *Science*, **303**, 220–222.
- Nadeau, R.M., Foxall, W. & McEvilly, T.V., 1995. Clustering and periodic recurrence of microearthquakes on the San-Andreas fault at Parkfield, California, *Science*, **267**, 503–507.
- Ogata, Y., 1999. Estimating the hazard of rupture using uncertain occurrence times of paleoearthquakes, *J. geophys. Res.*, **104**, 17 995–18 014.
- Patel, J.K., Kapadia, C.H. & Owen, D.B., 1976. *Handbook of Statistical Distributions*, Marcel Dekker, New York.
- Peng, Z.G. & Ben-Zion, Y., 2006. Temporal changes of shallow seismic velocity around the Karadere-Duzce branch of the North Anatolian fault and strong ground motion, *Pure appl. Geophys.*, **163**, 567–600.
- Peng, Z.G., Vidale, J.E., Marone, C. & Rubin, A., 2005. Systematic variations in recurrence interval and moment of repeating aftershocks, *Geophys. Res. Lett.*, **32**, L15301.
- Savage, J.C., 1994. Empirical earthquake probabilities from observed recurrence intervals, *Bull. seism. Soc. Am.*, **84**, 219–221.
- Schaff, D.P., Beroza, G.C. & Shaw, B.E., 1998. Postseismic response of repeating aftershocks, *Geophys. Res. Lett.*, **25**, 4549–4552.
- Shcherbakov, R., Yakovlev, G., Turcotte, D.L. & Rundle, J.B., 2005. Model for the distribution of aftershock interoccurrence times, *Phys. Rev. Lett.*, **95**, 218501.
- Sornette, D. & Knopoff, L., 1997. The paradox of the expected time until the next earthquake, *Bull. seism. Soc. Am.*, **87**, 789–798.
- Turcotte, D.L., Abaimov, S.G., Shcherbakov, R. & Rundle, J.B., 2007. Nonlinear dynamics of natural hazards, in *Nonlinear Dynamics in Geosciences*, pp. 557–580, eds Tsonis, A.A. & Elsner, J.B., Springer, New York.
- Uchida, N., Matsuzawa, T., Hasegawa, A. & Igarashi, T., 2003. Interplate quasi-static slip off Sanriku, NE Japan, estimated from repeating earthquakes, *Geophys. Res. Lett.*, **30**, 1801.
- Uchida, N., Matsuzawa, T., Hirahara, S. & Hasegawa, A., 2006. Small repeating earthquakes and interplate creep around the 2005 Miyagi-oki earthquake ($M = 7.2$), *Earth Planets Space*, **58**, 1577–1580.
- Vidale, J.E., Ellsworth, W.L., Cole, A. & Marone, C., 1994. Variations in rupture process with recurrence interval in a repeated small earthquake, *Nature*, **368**, 624–626.
- Working Group on California Earthquake Probabilities, 2003. Earthquake probabilities in the San Francisco Bay Region, 2002–2031, U.S. Geological Survey, Open-File Report 2003-214.