



Seismicity rate immediately before and after main shock rupture from high-frequency waveforms in Japan

Zhigang Peng,^{1,2} John E. Vidale,¹ Miaki Ishii,³ and Agnes Helmstetter⁴

Received 13 March 2006; revised 17 September 2006; accepted 27 October 2006; published 17 March 2007.

[1] We analyze seismicity rate immediately before and after 82 main shocks with the magnitudes ranging from 3 to 5 using waveforms recorded by the Hi-net borehole array in Japan. By scrutinizing high-frequency signals, we detect ~ 5 times as many aftershocks in the first 200 s as in the Japan Meteorological Agency catalogue. After correcting for the changing completeness level immediately after the main shock, the aftershock rate shows a crossover from a slower decay with an Omori's law exponent $p = 0.58 \pm 0.08$ between 20 and 900 s after the main shock to a faster decay with $p = 0.92 \pm 0.04$ after 900 s. The foreshock seismicity rate follows an inverse Omori's law with $p = 0.73 \pm 0.08$ from several tens of days up to several hundred seconds before the main shock. The seismicity rate in the 200 s immediately before the main shock appears steady with $p = 0.35 \pm 0.50$. These observations can be explained by the epidemic-type aftershock sequence (ETAS) model, and the rate-and-state model for a heterogeneous stress field on the main shock rupture plane. Alternatively, nonseismic stress changes near the source region, such as episodic aseismic slip, or pore fluid pressure fluctuations, may be invoked to explain the observation of small p values immediately before and after the main shock.

Citation: Peng, Z., J. E. Vidale, M. Ishii, and A. Helmstetter (2007), Seismicity rate immediately before and after main shock rupture from high-frequency waveforms in Japan, *J. Geophys. Res.*, 112, B03306, doi:10.1029/2006JB004386.

1. Introduction

[2] Large shallow earthquakes are typically followed by increased seismic activity, known as "aftershocks", which diminish in rate approximately as the inverse of the elapsed time since the main shock [Omori, 1894]. The aftershock decay rate $R(t)$ is well described by the modified Omori's law [Utsu et al., 1995]

$$R(t) = \frac{K}{(t+c)^p} \quad (1)$$

where t is the time elapsed since the main shock and K , p , c are empirical parameters. In addition, earthquakes are sometimes preceded by statistically accelerating seismic activity, known as "foreshocks" [Jones and Molnar, 1979; Abercrombie and Mori, 1996]. Recent studies have shown that the increased rate of foreshocks can be described by an inverse power law with the same functional form as the modified Omori's law for aftershocks, but the values of the

parameters are different [Jones and Molnar, 1979; Maeda, 1999; Helmstetter and Sornette, 2003].

[3] Among the three parameters in the modified Omori's law, the K value describes aftershock productivity, which scales with main shock magnitude m [Felzer et al., 2004; Helmstetter et al., 2005], and depends on the cutoff magnitude of the events considered as aftershocks. Observed value for the exponent p is, in general, around 1.0 [e.g., Reasenbergh and Jones, 1989, 1994; Utsu et al., 1995], although it varies for different aftershock sequences [Wiemer and Katsumata, 1999]. The variation in the p value has been related to crustal temperature [Mogi, 1962], heat flow [Kisslinger and Jones, 1991], degree of heterogeneity in the fault zone [Mikumo and Miyatake, 1979], and fractal dimension of the preexisting fault system [Nanjo et al., 1998]. However, it is still not clear which factors play the major roles in controlling the p value. The parameter c is an apparent time offset, commonly a fraction of an hour or a day. It eliminates the singularity in the aftershock rate at zero time.

[4] The c value is a controversial quantity [e.g., Utsu et al., 1995; Kisslinger, 1996]. Although it is claimed to scale with the main shock magnitude and the lower magnitude cutoff for different aftershock sequences [Shcherbakov et al., 2004], and the recurrence time of the main shock [Narteau et al., 2002], the commonly accepted view is that the c value is an artifact due to incompleteness of early aftershocks in a catalogue [Kagan, 2004; Kagan and Houston, 2005; Lolli and Gasperini, 2006]. Immediately after a large earthquake, many aftershocks are missing in the catalogue due to overlapping with coda from the main

¹Department of Earth and Space Sciences, University of California, Los Angeles, California, USA.

²Now at School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA.

³Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA.

⁴Observatoire des Sciences de l'Université de Grenoble, Laboratoire de Géophysique Interne et Tectonophysique, Grenoble, France.

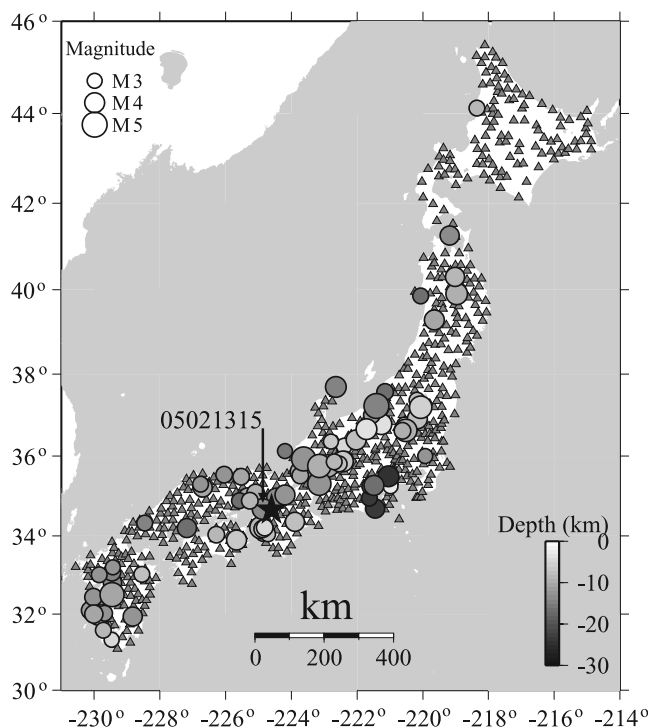


Figure 1. Hypocentral locations of 82 earthquakes (circles) with $3 \leq m \leq 5$ used in this study. The size of each circle scales with the magnitude listed in the JMA catalogue, and the shade corresponds to its hypocentral depth. The 692 Hi-net array stations are plotted as gray triangles. The waveforms recorded by station KNHH (star) generated from an $m = 4.1$ event 05021315 (pointed by an arrow) are shown in Figure 3.

shock, and overload of processing facilities [Kagan, 2004]. Thus the characteristics of the aftershock decay rate immediately after a main shock remains uncertain. Yet this period holds valuable information on the transition from main shock to aftershocks, and the underlying earthquake physics that controls the time-dependent behavior of aftershocks. The short-term property of aftershock decay rate is also very important for evaluating and forecasting short-term earthquake probability [Kagan, 2004; Gerstenberger et al., 2005; Helmstetter et al., 2006].

[5] In comparison, foreshocks are less affected by catalogue incompleteness. However, only a few studies have focused on the foreshock behavior on the timescale of seconds to minutes before the main shock [e.g., Maeda, 1999; McGuire, 2003; McGuire et al., 2005]. This is because the total number of foreshocks observed is smaller than the number of aftershocks. Often, no foreshock is observed before a main shock. Thus it is difficult to examine the behavior of foreshock activity based on just a few earthquake sequences. To overcome such difficulty, a stacking method is generally applied to obtain a statistical space-time distribution of foreshocks, with the assumption that every main shock has a similar behavior of foreshock activity [Jones and Molnar, 1979; Maeda, 1999; Reasenberg, 1999; Helmstetter and Sornette, 2003]. A

clearer understanding of the foreshock behavior would provide keys to a better understanding of the physical mechanism of foreshock occurrence, and may also allow us to distinguish foreshock activity from fluctuations of background seismicity, which may be useful in predicting large earthquakes [Geller et al., 1997; Chen et al., 1999].

[6] More complete catalogues are needed for a better constraint on the seismicity rate immediately before and after main shocks. An effective approach is to go beyond conventional catalogues and analyze continuous waveforms recorded by high-quality seismic stations close to the main shock [Vidale et al., 2003; Peng et al., 2006; Enescu et al., 2007]. For example, Vidale et al. [2003] have found several times more aftershocks in the first few minutes than are reported in existing catalogues from high-pass filtered seismograms for several moderate to large main shocks in California and Japan. Unfortunately, unclipped broadband recordings from stations close to large earthquakes (e.g., $m \geq 6$) are rare. The seismicity rate immediately after a large earthquake is likely to be very high, which may cause overlapping phase arrivals, making identification of individual aftershocks difficult. Furthermore, the analysis is complicated by a relatively long main shock rupture, and a broad spatial distribution of aftershocks. Therefore it is difficult to assign magnitudes to aftershocks identified by a single or a few nearby stations, and assess the completeness of the catalogue in terms of event time and minimum magnitude. On the other hand, small earthquakes (e.g., $m \leq 3$) typically have small number of aftershocks. This implies that a large number of main shocks are required to accumulate enough aftershocks for statistically significant conclusions. Avoiding these two cases, we use waveforms from many moderate-size earthquakes (e.g., $3 \leq m \leq 5$) recorded by high-quality seismometers to investigate seismic activities immediately before and after main shocks.

[7] In this study, we analyze waveforms of 82 earthquakes with $3 \leq m \leq 5$ that are recorded by the High Sensitivity Seismograph Network (Hi-net) in Japan [Okada et al., 2004]. The Hi-net array consists of ~ 700 stations (Figure 1). Most of them are installed in boreholes at a depth of 100 to 300 m. Each station consists of a three-component velocity seismometer with a natural frequency of 1 Hz and a sampling rate of 100/s. In sections 2 and 3, we first describe the data set and the analysis procedure. We summarize the results in section 4. The interpretation is given in section 5.

2. Data

[8] We systematically searched the Japan Meteorological Agency (JMA) catalogue during an 18-month period (1 December 2003 to 20 June 2005), for shallow earthquakes in the magnitude range $3 \leq m \leq 5$, and within a rectangular area (longitude between 129°E and 146°E, and latitude between 30°N and 46°N). We then examined the waveform data that are recorded by the closest 10 stations in the Hi-net array starting ~ 200 s before until ~ 900 s after each main shock.

[9] To control the data quality, the following criteria are applied to select a subset of earthquake sequences for analysis. First, we require that the best recording station in the Hi-net array be within 30 km of the main shock

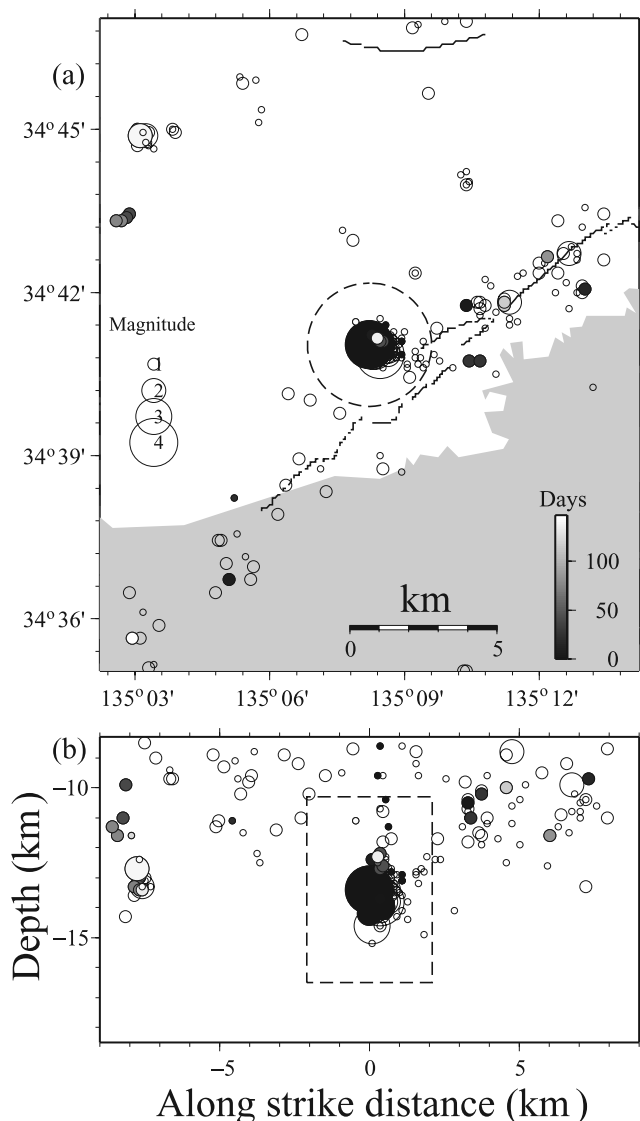


Figure 2. (a) Map view of the seismicity (circle) from the JMA catalogue around the $m = 4.1$ event 05021315. Earthquakes that occurred before the main shock are shown by open circles, and those after the main shock are shown with shade denoting the time since the main shock. The size of each circle scales with the magnitude listed in the JMA catalogue. The large dashed circle marks radius that is used to select foreshocks and aftershocks. (b) Cross-sectional view of the seismicity. The dashed box marks the spatial window used in this study.

epicenter. The best station, which is typically the nearest station, is defined as the one with the combination of a low preevent noise level, a high signal-to-noise ratio, and a fast decaying main shock coda. The distance criterion is applied to minimize the duration of the main shock coda, which falls off with time more rapidly at closer distance. Only crustal events with hypocentral depths less than 30 km are analyzed in this study because they are closer to the nearest station and tend to have more aftershocks than deeper events [Kagan, 1991; Mori and Abercrombie, 1997]. On the other hand, events with hypocentral depth less than 2 km

are mostly associated with volcanic activities, and their behavior has been asserted to differ from that of tectonic earthquakes. So only events with their hypocentral depths within 2–30 km are included in the analysis. We also require that at least one event with $m \leq 1.5$ be present in the JMA catalogue within the time period of the retrieved waveform for magnitude calibration. The detailed calibration procedure is given in section 3.

[10] Finally, we select as main shocks those earthquakes that were not preceded by large events with $m \geq 3$ by at least R km and 100 days. The purpose of this selection criterion is to ensure that our main shocks are not influenced by, or direct consequences of, previous large events. The spatial influence length R is defined as 2 times the rupture length $L(m) = 0.01 \times 10^{0.5m}$ (km) [Working Group on California Earthquake Probabilities, 2003] of an earthquake with magnitude m . The minimum value of R is set to be 3 km. In addition, we eliminated one sequence that showed strong swarm-like behavior [Vidale and Shearer, 2006].

[11] We then select all events within 100 days of a main shock as aftershock sequences (Figure 2). The 100-day time window represents a compromise between a need for a sufficiently long window to accurately estimate the long-term aftershock rate, and a need to minimize the contamination from background seismicity. The spatial window for aftershocks is defined as a circle around the main shock epicenter, with the radius equal to one-rupture length $L(m) + 1$ km horizontal location error. We also require the depth of aftershocks to be within one-rupture length $L(m) + 2$ km vertical location error (Figure 2). Foreshocks are defined as all events within 545 days (as limited by the availability of the JMA catalogue) before the main shock within the same area.

[12] As described before, we only require that a main shock not be preceded by another event that is potentially a main shock itself. We do not require that aftershocks be smaller than the main shock. We believe that our selection process does not impose a priori assumption on the temporal variation of seismicity rate around a main shock, but help to identify those sequences with high data quality. After the selection process, we obtain 82 events. Their hypocentral locations are widely distributed across the Japanese Islands (Figure 1).

3. Analysis Procedure

[13] The three-component seismograms recorded by the best station are high-pass-filtered using a two-pass Butterworth filter with a corner frequency of 30 Hz. Next, we compute the envelopes of the high-pass-filtered seismograms, stack the three-component envelopes, take the logarithm, and smooth the resulting envelope by a median operator with a half width of 0.1 s. All procedures are done using subroutines in the Seismic Analysis Code [Goldstein *et al.*, 2003]. The results are similar for 10 to 30 Hz high-pass filters, but the data filtered with higher frequency have sharper onsets and coda that decay more rapidly. P wave arrivals are clearer on the vertical component, but we include the two horizontal components to produce a flat background noise level, and also to minimize disturbances that sometimes appear only on one channel.

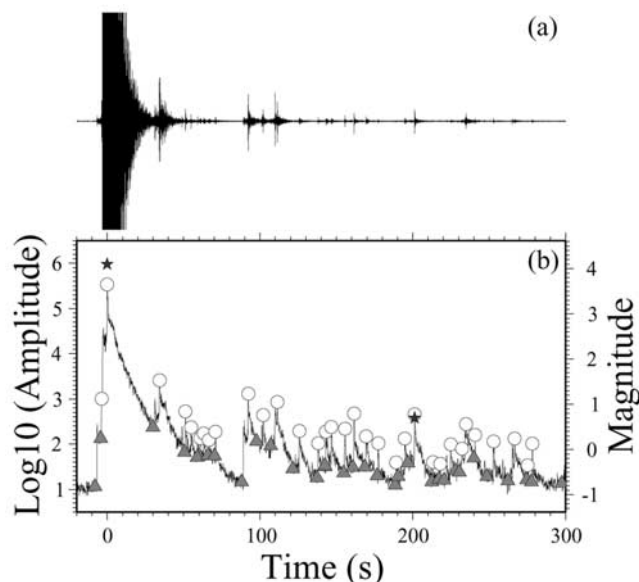


Figure 3. (a) High-pass-filtered vertical component seismogram recorded by station KNHH for event 05021315 with $m = 4.1$. The trace is manually clipped, with the peak amplitude of the main shock off-scale to illustrate small aftershocks. (b) Logarithm of the envelope function obtained by stacking the envelopes of the high-pass-filtered three-component seismograms. Each circle marks a seismic phase arrival (i.e., either P or S wave of an event). A total of 33 events are identified by the handpicking procedure. The gray triangle marks the amplitude level right before each event. The two stars denote the times of event identified in the JMA catalogue. The magnitude for this sequence are calibrated using an empirical relationship $m = \log_{10}(\text{amplitude}) - 1.88$.

[14] We identify an event by searching for clear double peaks in the envelope that correspond to the P and S arrivals. The arrival time of the larger of the two peaks is used as a proxy for the origin time of the corresponding event, and the peak amplitude is used to estimate the event magnitude. An example is shown in Figure 3 for an $m = 4.1$ event. The event location is shown in Figure 2. We note that delays of a few seconds may exist between the high-frequency peaks (circles) in the envelope and the actual origin time of an event given in the JMA catalogue (stars). However, such delays are not relevant in this study when the origin times of the events are measured with respect to the high-frequency peak of the main shock. During the identification process, we check that the $S-P$ time of each handpicked event is close to that of the main shock. This ensures that the events are located near the main shock. We also examine the high-pass-filtered envelopes of the nearby stations to confirm that large events are recorded by other stations besides the best one. Finally, the amplitude immediately before each event is used as a measure for the preevent noise level. Only events with peak amplitude at least 0.3 (in logarithmic scale) larger than the noise level (i.e., a signal-to-noise ratio of ~ 2) are retained for further analysis.

[15] Assuming that a tenfold increase in amplitude corresponds to an increase in one unit of magnitude, we

estimate the magnitude of each handpicked event using its peak amplitude. First, we use events that are both identified by our procedure and are listed in the JMA catalogue to calibrate an amplitude-magnitude relation. This is done by shifting the envelope function so that the logarithmic amplitudes match the magnitudes of small events listed in the JMA catalogue. Only events small enough with their corner frequencies above 30 Hz are used for such calibration. If we assume a circular crack model [Eshelby, 1957] and a stress drop of 3 MPa, the corner frequency of an event with $m = 1.5$ is about 30 Hz. So we use $m < 1.5$ to calibrate the amplitude-magnitude relation in this study. As shown in Figure 4, the magnitudes estimated from the envelope amplitudes match well with the JMA magnitudes for small events. We underestimate magnitudes of events larger than 1.5 because 30 Hz is above their corner frequencies. Next, we derive an empirical relation of $m_{\text{JMA}} = (m_{\text{AMP}} - 1.5) \times 1.6 + 1.5$ ($m_{\text{AMP}} \geq 1.5$), and $m_{\text{JMA}} = m_{\text{AMP}} \times 1.0$ ($m_{\text{AMP}} < 1.5$) from a least squares fitting (Figure 4), and apply it for magnitude correction, where m_{JMA} denotes the JMA magnitude, and m_{AMP} is the magnitude from the envelope amplitude.

[16] We then combine the handpicked events (e.g., Figure 3) with those listed in the JMA catalogue into one catalogue for each sequence (e.g., Figure 5). For consistency, we use the origin times and magnitudes determined from the envelopes for events in the time range of -200 to 900 s relative to the main shock origin time. For those events that are outside this time window, we use the origin times and magnitudes listed in the JMA catalogue. Finally, we examine all 82 sequences together for both foreshocks and aftershocks.

4. Results

4.1. General Features

[17] The stacked aftershock sequences (Figure 6a) show several interesting features. First, we have detected ~ 5 times

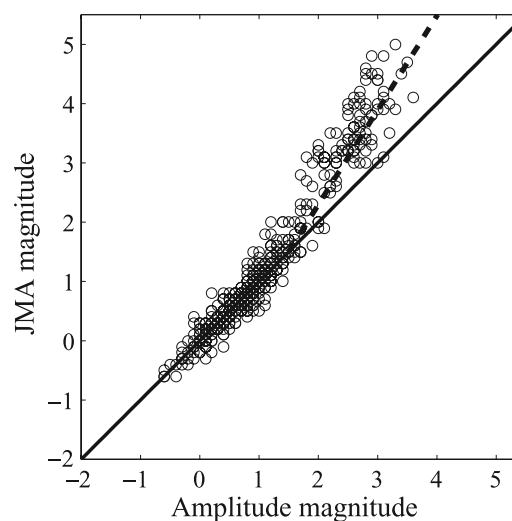


Figure 4. A comparison of magnitudes listed in the JMA catalogue and those derived from the envelope amplitudes for 494 events identified in the 82 sequences. The diagonal line denotes perfect correlation. The dashed line is the best fitting regression line for magnitudes larger than 1.5.

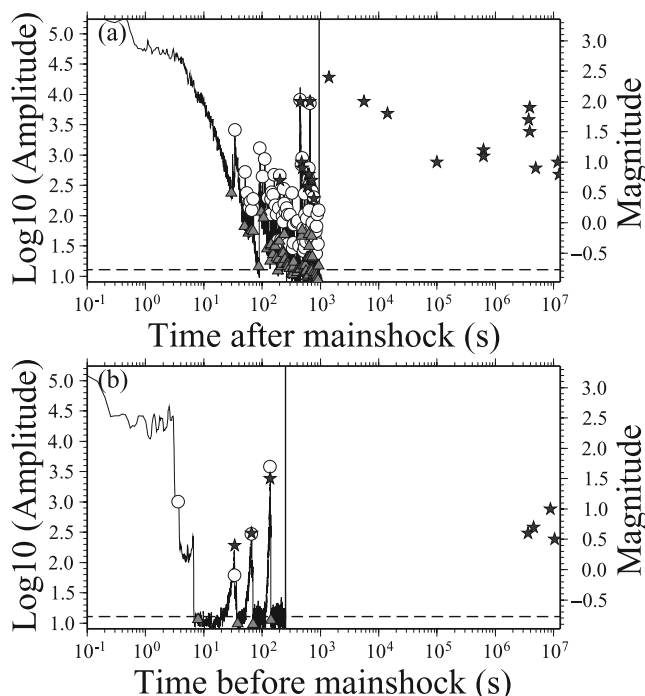


Figure 5. Event magnitudes versus logarithmic times (a) after and (b) before the $m = 4.1$ main shock shown in Figures 1, 2, and 3. Each circle marks an event picked by hand. The star denotes an event listed in the JMA catalogue. The black curve is the high-pass-filtered envelope function for station KNHH. The triangle marks the amplitude level before each event. The horizontal dashed line marks the amplitude level before the main shock. The vertical line denotes the end time of the available waveform data. The increase in amplitude at $t \approx 4$ s before the main shock is due to the main shock first arrival.

as many aftershocks in the first 200 s with $m \geq 0$ as are listed in the JMA catalogue. This indicates that our procedure is more effective in identifying early aftershocks than the conventional methods employed by the JMA. Figure 7 shows the fraction of events listed in the JMA catalogue and identified by our handpicking procedure in each logarithmic time bins. As expected, the fraction of detected events in the conventional catalogue increases with an increase in cutoff magnitude and time elapsed since the main shock [Kagan, 2004; Helmstetter *et al.*, 2006]. We note that some small events are still missing in the JMA catalogue at 900 s after the main shock.

[18] The change in detection threshold with time and magnitude is also evident for our handpicked events. This indicates that our handpicked catalogue is probably incomplete at early times, just like the JMA catalogue. That is, small events escape detection if their amplitudes are below these of the codas of the main shock or large aftershocks [Kagan, 2004; Kilb *et al.*, 2004]. The completeness of our handpicked catalogue within the first 900 s is quantified in section 4.2. We return to examining the seismicity rate around the main shocks in section 4.3.

[19] Figure 6a show that the density (number of events per unit log time) of events becomes greater at all magni-

tude levels as time passes. Because the horizontal axis is in logarithm of time, a constant density of events indicates that the seismicity rate decays with $1/t$, which means $p = 1$. Thus an increase in density at all magnitude levels suggests that the p value is less than 1, especially at times immediately after the main shock.

[20] The stacked foreshock sequences shown in Figure 6b are more scattered than the aftershock sequences. The log-scale densities for both foreshocks and aftershocks are comparable at several tens of days away from the main shock occurrence time. Since foreshocks occur before the main shocks, they are relatively free of the incompleteness caused by the main shock coda. Indeed, we observe two immediate foreshocks within 10 s before the main shock. However, there is a clear lack of small foreshocks in the last 20 s before the main shock.

4.2. Catalogue Completeness

[21] The catalogue completeness is often discussed in terms of magnitude threshold (or cutoff) m_c , the magnitude above which all events are identified in the catalogue. A standard way of estimating m_c is to find the minimum

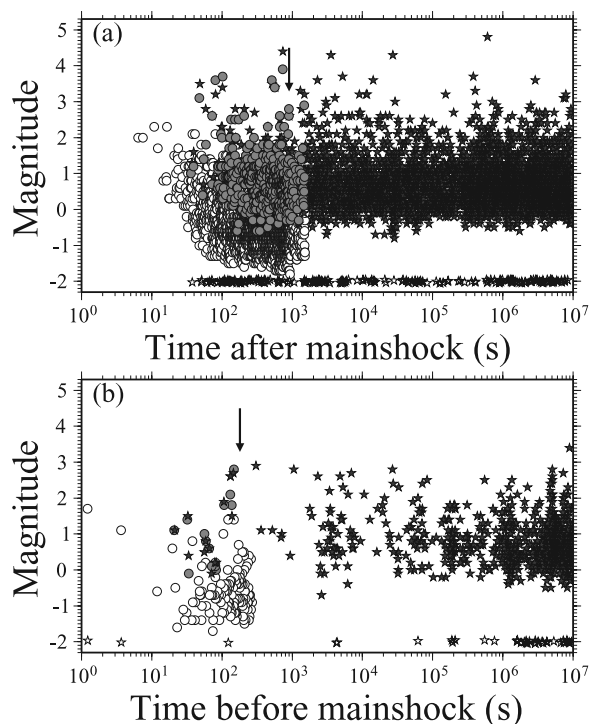


Figure 6. Event magnitude versus logarithm of times (a) after and (b) before the 82 main shocks. The open circles mark events picked by hand. The stars denote events listed in the JMA catalogue. The gray circles mark those that are both picked by hands and listed in the JMA catalogue. Events in the JMA catalogue with no magnitude (open stars) are assigned $m = -2.0$. A random number between -0.05 and 0.05 is added to the magnitudes for plotting purposes. The two arrows mark the time of 900 s (Figure 6a) and -200 s (Figure 6b) from the main shock occurrence time when we have waveform recordings for all 82 sequences.

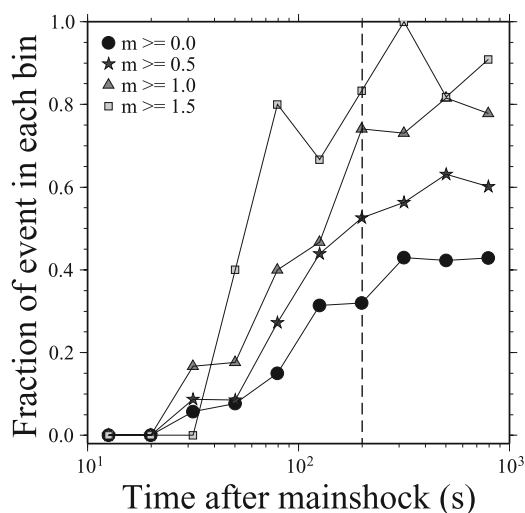


Figure 7. Ratio between events listed in the JMA catalogue and those identified by the handpicking procedure with four different magnitude thresholds in bins of equal width in the logarithm of time. The vertical dashed line marks the time of 200 s after the main shock.

magnitude that fits the *Gutenberg and Richter* [1944] (G-R) frequency-magnitude relation. Assuming self-similarity, *Wiemer and Wyss* [2000] proposed to estimate m_c using the point of the maximum curvature (MAXC) for the noncumulative frequency-magnitude distribution. Since aftershocks listed in the JMA catalogue are not complete in the first few hundred seconds (Figure 7), we use aftershocks between 2000 s and 100 days to compute a noncumulative frequency-magnitude distribution. The number of aftershocks is peaked at $m = 0.7$ (Figure 8a). *Woessner and Wiemer* [2005] have found that m_c values obtained from the MAXC method is typically 0.1–0.2 smaller than those from other methods. However, as will be shown in section 4.5, our results do not vary significantly with the m_c value. For $m_c = 0.7$, we obtain a b value of 0.81 using the discrete G-R model of *Utsu* [1966], with magnitude bands of width $dm = 0.1$ for the JMA catalogue.

[22] Since the maximum curvature of the noncumulative frequency-magnitude distribution for the handpicked events is not well defined, it is difficult to estimate their m_c and the b values even at several hundred seconds after the main shock (Figure 8b). In addition, we only have ~ 300 aftershocks in the first 100 s after the main shock when the m_c is changing dramatically with time (Figure 8c). Since the maximum curvature method typically requires several hundred events in each space-time window for statistically significant result, this method cannot provide an accurate estimation of m_c in the first 100 s [Kagan, 2004]. Recently, *Helmstetter et al.* [2006] have proposed an empirical relation for m_c as a function of the main shock magnitude and the elapsed time. However, that relationship is based on several large ($m \geq 6$) aftershock sequences in southern California. Thus the applicability of this relationship to the 82 sequences in Japan with smaller main shock magnitude is not clear.

[23] Incomplete detection of small aftershocks is most likely due to overlapping with coda of the main shock and large aftershocks immediately after the main shock [Kagan,

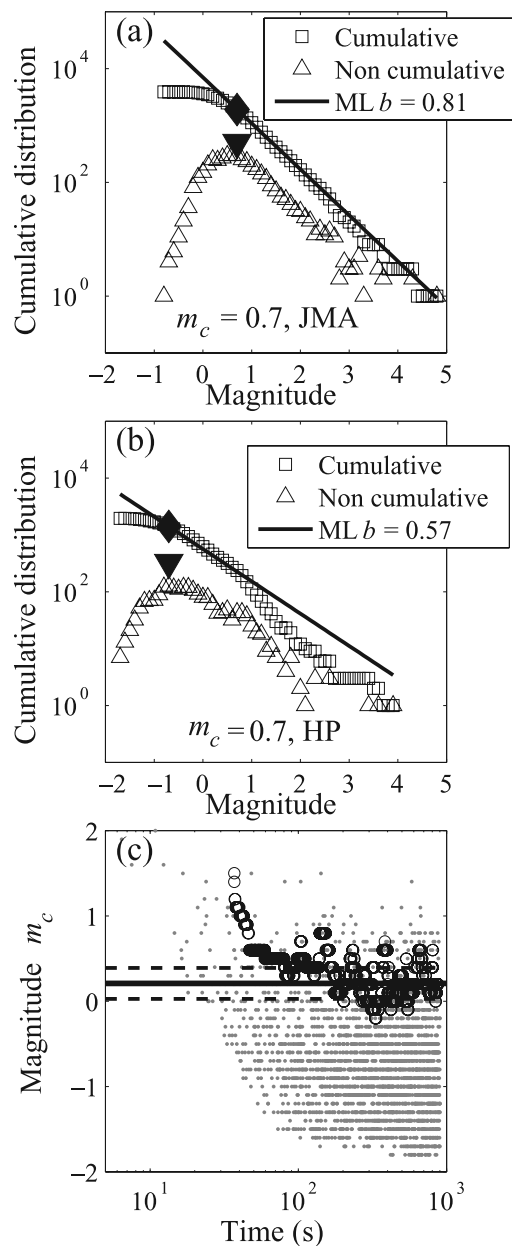


Figure 8. (a) Cumulative (square) and noncumulative (triangle) number of aftershocks versus magnitude for events listed in the JMA catalogue starting 2000 s after the main shock. The solid line marks the maximum-likelihood fit for the frequency-magnitude relationship. The inverted triangle marks the maximum curvature of the noncumulative distribution. The diamond denotes the magnitude of completeness m_c . (b) Same as in Figure 8a except for the handpicked events starting 200 s after the main shock. (c) The m_c value (gray dot) obtained from the preevent noise level of each handpicked aftershock versus logarithmic time after the main shock for all 82 sequences. The open circle denotes the magnitude below which 95% of the m_c values (gray dots), or $m_{c95}(t)$, as a function of time for every 100 points. The average and standard deviation in magnitude for $m_{c95}(t)$ after 200 s, 0.21 ± 0.18 , are marked by the solid and dashed lines.

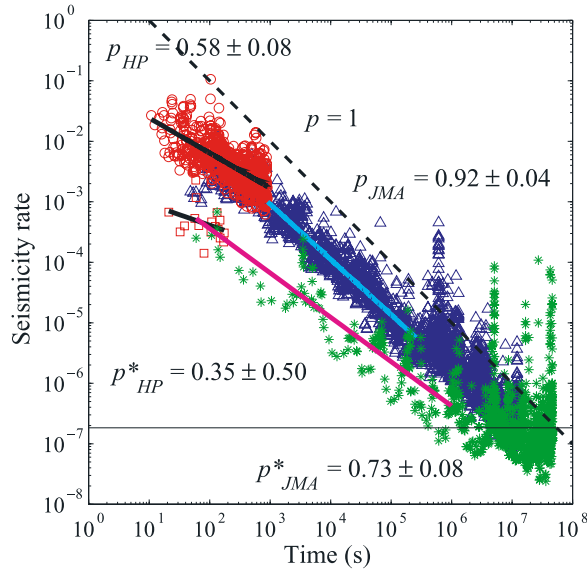


Figure 9. Seismicity rate for the 82 sequences as a function of time relative to the main shock. The triangle denotes the rate measured from the aftershocks in the JMA catalogue. The circle marks the rate measured from the handpicked events after correcting for the m_c . The asterisk and square denote the foreshock seismicity rate from JMA catalogue and handpicked events, respectively. All curves have been shifted so that their $m_{c0} = 1.0$. The solid lines mark the least squares fitting for different seismicity rate. The corresponding slopes (p values) and the 95% confidence levels are marked, where p_{HP} , p_{JMA} , p^*_{HP} , and p^*_{JMA} stand for the p value for the handpicked aftershocks, aftershocks in the JMA catalogue, handpicked foreshocks, and foreshocks in the JMA catalogue, respectively. The dashed line denotes a reference line with $p = 1$. The background seismicity rate, defined as the logarithmic average of the foreshock rate $\geq 10^7$ s before the main shock, is marked by the horizontal line.

2004]. The amplitudes and durations of coda waves depend not only on the main shock magnitude but also on the hypocentral distance, the heterogeneity of the crust, and near-station structures [e.g., *Aki*, 1969; *Sato and Fehler*, 1998]. Thus different aftershock sequences are expected to have different coda durations, and hence different magnitude detection thresholds.

[24] To treat each sequence separately, we use the envelope amplitude right before each event (the noise level discussed in section 2) + 0.3 as a measure of the local m_c at the time of that event. In addition, we compute $m_{c95}(t)$, the magnitude below which 95% of the m_c values exist as a function of time, and use it as a measure of the completeness for the stacked aftershock sequences. The $m_{c95}(t)$ for 100 consecutive events contained in a sliding window (shifted by one event each time) are estimated. As shown in Figure 8c, the $m_{c95}(t)$ values quickly drop from ~ 1.5 in the first 50 s, to around 0.2 after 200 s.

[25] Next, we correct for the changing m_c values and hence incompleteness in the catalogue at early times after the main shock. We weight each aftershock based on the local m_c using a weighting function $w(i) = 10^{[(m_c(i) - m_{c0})/b]}$, where $m_c(i)$ is the m_c value for each event i , b is

exponent for the G-R relationship, and m_{c0} is the minimum m_c value used in the analysis, equal to the long time value of m_c . We use $b = 0.81$ and $m_{c0} = 0.2$ from the results shown in Figure 8. That is, $w(i) = 1$ if $m_c(i) = 0.2$, but $w(i) > 1$ if $m_c(i) > 0.2$. Events with magnitude $m_c < m_c(i)$ are not used in the analysis. By doing so, we assume that the aftershock size distribution follows the G-R relation immediately after the main shock and does not change over time [Kagan, 2004]. Since foreshocks occur before the main shock and are relatively free from overlapping with previous events, we did not apply this correction for the foreshock seismicity rate.

4.3. Seismicity Rate Before and After the Main Shock

[26] Finally, we compute the seismicity rate for the 82 sequences after correcting for the changing m_c value. Previous studies often use a fixed time window to compute seismicity rate based on the aftershock occurrence time [e.g., *Kagan*, 2004]. If the data are nonuniform, especially at times immediately before and after the main shock, this may result in a time window with no data points, causing a gap in the obtained seismicity rate. So we use a moving data window [e.g., *Ziv et al.*, 2003; *Felzer and Brodsky*, 2006], instead of a moving time window to avoid this problem. Each data point corresponds to the time for an event relative to the main shock. We use a fixed window of 5 data points, and the window slides by one data point each time. Using different window sizes produces similar results. Seismicity rate computed from larger-size window is smoother, but the rate is less instantaneous as compared with that from smaller-size window.

[27] The results for both foreshocks and aftershocks are shown in Figure 9. To measure the change in seismicity rate, we fit the early and later aftershock rates separately using the Omori's law ($r(t) \sim 1/t^p$). The p values obtained by a least squares fitting are 0.58 ± 0.08 between 20 and 900 s for the handpicked aftershocks, and 0.92 ± 0.04 between 900 s and 3 days for aftershocks listed in the catalogue. The error is the 95% confidence interval based on 1000 bootstraps of the aftershock time.

[28] In comparison, the seismicity rate for the foreshock sequences follows an inverse Omori's law with a p value of 0.73 ± 0.08 from several tens of days up to several hundreds of seconds before the main shock. The seismicity rates for both foreshocks and aftershocks appear to merge with the background rates around several hundred days. However, they differ by 1 to 2 orders of magnitude around the main shock occurrence time. This implies that the increase rate foreshock is lower than the aftershock decay rate. Finally, the foreshock rate at 6–200 s before the main shock is scattered due to the small amount of data, but the obtained p value of 0.35 ± 0.50 is lower than the foreshock increase rate at larger times. We did not include foreshocks that are within 6 s of the main shock occurrence time (the peak amplitude for the main shock in this study) to avoid overlapping with the P wave arrival of the main shock.

4.4. Statistical Significance of a Low Early Aftershock Rate

[29] The deficiency in seismicity immediately after the main shock is statistically significant. Assuming that an aftershock sequence is a Poisson process that follows the Omori's law, we can directly compare the number of after-

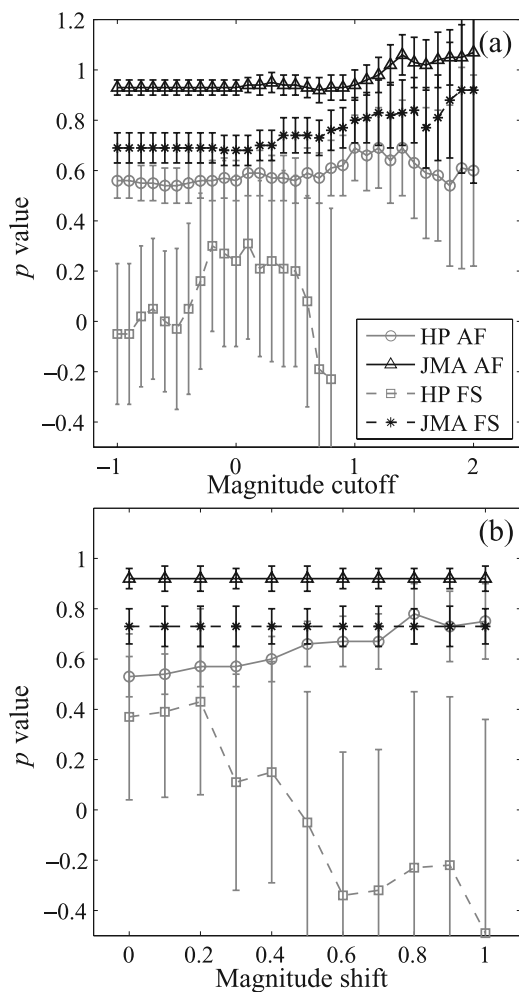


Figure 10. (a) Measured p value and 95% confident level as a function of magnitude of completeness m_{c0} for early aftershocks by handpicking (gray circle), late aftershocks listed in the JMA catalogue (black triangle), immediate foreshocks by handpicking (gray square), and foreshocks listed in the JMA catalogue (black asterisk). (b) Measured p value and 95% confident level as a function of magnitude shift for different categories.

shocks observed at early times with the expected number of events using the Omori's law with the p value from later times. The maximum magnitude of completeness at short times ($20 \leq t \leq 100$ s) is 1.5. Extrapolating the long-term Omori's law \pm uncertainties to this time range for events with $m \geq 1.5$, we expect to have 32–45 events. The observed number in this time and magnitude interval is 16. The probability of having no more than 16 events for a Poisson distribution with an expected number of 32–45 events is 5×10^{-7} to 0.001. Thus the aftershock rate at short times $20 \leq t \leq 100$ s is significantly less than expected from the long-term aftershock rate at the 99.9% confidence level.

4.5. Dependence of the p Values on Different Parameters

[30] The result shown in Figure 9 is based on $m_{c0} = 0.2$ for immediate foreshocks and early aftershocks, and

$m_{c0} = 0.7$ for long-term foreshocks and aftershocks. We also use the preevent noise level + 0.3 magnitude shift as a measure of m_c for each handpicked event. We systematically vary the choosing parameters to test their influences on the p values. As shown in Figure 10, the p values do not depend strongly on m_{c0} ranging from -1 to 2 . However, there is slight increase of p values for early aftershocks if we increase the magnitude shift from 0 to 1. At a magnitude shift of 1, there is no significant difference in the p value for short and large times. However, a value of 1 means that we can only detect an event if its amplitude is about 10 times the preevent noise level, which is well above the detection ability. On the basis of our handpicked experience, a value of 0.3 (close to a signal-to-noise ratio of 2) is close to the true identification ability (i.e., magnitude of completeness) at the time of each handpicked event.

[31] The weighting procedure as described in section 4.2 is employed to correct the changing completeness levels during the coda of the main shock and large aftershocks. However, it does not consider the fact that many aftershocks may occur close in time. When the seismicity rate is high immediately after the main shock, this will result in mixed phase arrivals that are difficult to associate with individual event. Thus some events may not be identified, causing a lower p value in the early aftershock period than that at larger times. In addition, misidentification of events, especially for the smaller events, may cause the proportion of small/large aftershocks (i.e., the b value of the magnitude distribution) to be smaller at early times compared to that at later times. Alternatively, the change in p and b values may be a real effect. For example, Ziv *et al.* [2003] have shown from relocated catalogue of microearthquakes in northern California that the b value within 10^4 s of a previous earthquake is significantly lower than that of the long-term value. Shcherbakov *et al.* [2006] found that the b value increased from 0.60 ± 0.01 after 0.1 day to 0.89 ± 0.01 after 365 days in the aftershock zone of the 2004 Parkfield earthquake. However, these studies use aftershocks listed in the catalogue. So the effect of catalogue incompleteness on the p and b values cannot be ruled out.

[32] To check if mixing of phase arrivals causes a significant number of missing aftershocks in our study, we evaluate the dependence of p value on the aftershock productivity. If mixing of phase arrivals is mainly responsible for the lower p value at early times, we should have higher p values for the less productive sequences. We order the 82 sequences according to the number of aftershocks with $m \geq 0$ observed within 100 days, and divide them into two groups so that each group has roughly the same number of aftershocks. The first group has 12 sequences, and is considered as the productive group, as compared with the second group, which has 70 sequences. Figure 11 shows that early aftershock decay rates are similar for both groups, despite a factor of ~ 6 differences in aftershock productivity. This indicates that we do not miss a significant fraction of aftershocks due to mixed phase arrivals. We also check the dependence of p values with the main shock magnitude by separating the 82 sequences into four magnitude groups with 0.5 unit interval based on the main shock magnitude (Figures 11b and 11c). We do not find a strong dependence of p values for sequences with the main shock magnitude $3 \leq m \leq 5$.

[33] We note that for aftershock sequences with larger main shock magnitude (e.g., $m \geq 6$), a significant fraction of small events could be missing due to mixed phase arrivals, and overlapping with main shock coda. For example, *Kagan* [2004] have shown that the number of small ($m \leq 3$) aftershocks that were missing in the time interval of 0–128 days exceeds those that were listed in the catalogue for the 1992 Landers earthquake. *Peng et al.* [2006] found that a significant fraction of events were missing in the Northern California Seismic Network catalogue in the first hour after

the 2004 M6 Parkfield earthquake. However, we are confident from the above analysis that to the first order, mixed phase arrivals does not cause a significant change in the number of the aftershocks determined in this study.

4.6. Direct Estimation of the Level of Aftershock Activity

[34] Results shown in sections 4.1–4.5 are based on events listed in the JMA catalogue and identified by our handpicking procedure. If we assume that energy radiated at 30 Hz is indicative of the level of seismic activity, the seismicity rate immediately before and after the main shock can be directly estimated from the envelope functions without picking individual event. We compute the stacked envelope functions for all 82 sequences by summing the square of each envelope function, and taking the square root and logarithm (based 10). Since each individual envelope has different noise level before the main shock, directly stacking the envelope functions would bias the result toward those that have higher background noise levels. Thus we shift each envelope using the average values from 180 to 150 s before the main shock (when the foreshock activity is relatively low) to achieve a similar noise level (near zero). The choice of offset is somewhat arbitrary, but helps to separate features that are near the noise level (near zero) and those well above the noise level.

[35] The resulting stacked envelope is shown in Figure 12a in linear timescale. The coda from the main shocks merges with the aftershock signal at ~ 30 s, and the rest of the curve represents the level of the aftershock activity. Figure 12b shows the stacked envelope functions in logarithmic timescale. We smooth the envelope by convolving the log time with a Gaussian kernel of width 0.08 [*Helmstetter and Shaw, 2006*]. The best fitting p value for envelopes between 30 and 900 s after the main shock is 0.69, close to that estimated from the handpicked catalogue. Measuring uncertainty on the obtained p value is difficult, because it depends on how the curve is smoothed.

[36] Figure 12b also shows that a burst with the largest foreshocks is visible between 150 and 120 s before the main shock. This is consistent with a cluster of handpicked events shown in Figure 6b. It is clear that the foreshock activity does not fit the inverse Omori's law in the last 200 s before the main shock. Also there is a relative quiescence in the

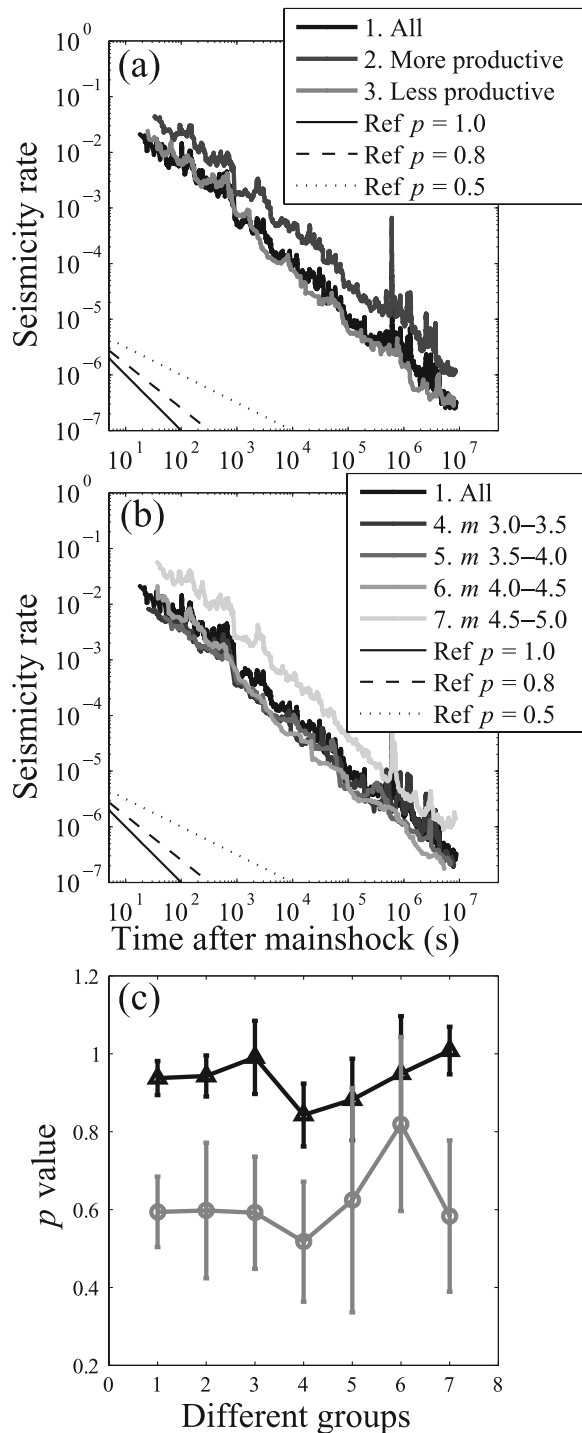


Figure 11. (a) Aftershock rates as a function of time since the main shock for two groups of sequences according to the aftershock productivity (see text for description). To produce smoother curve for comparison, we use a window size of 21 points to compute the seismicity rate. The solid, dashed, and dotted lines show the reference rate with $p = 1$, 0.8, and 0.5, respectively, for comparison of slope. (b) Aftershock rates as a function of the time since the main shock for four magnitude ranges with a 0.5 unit intervals based on the main shock magnitude. (c) Measured p values corresponding to different groups as shown in Figures 11a and 11b for the early aftershocks by handpicking (gray circles) and late aftershocks listed in the JMA catalogue (dark triangles). The vertical bars denote 95% confidence level of the fit.

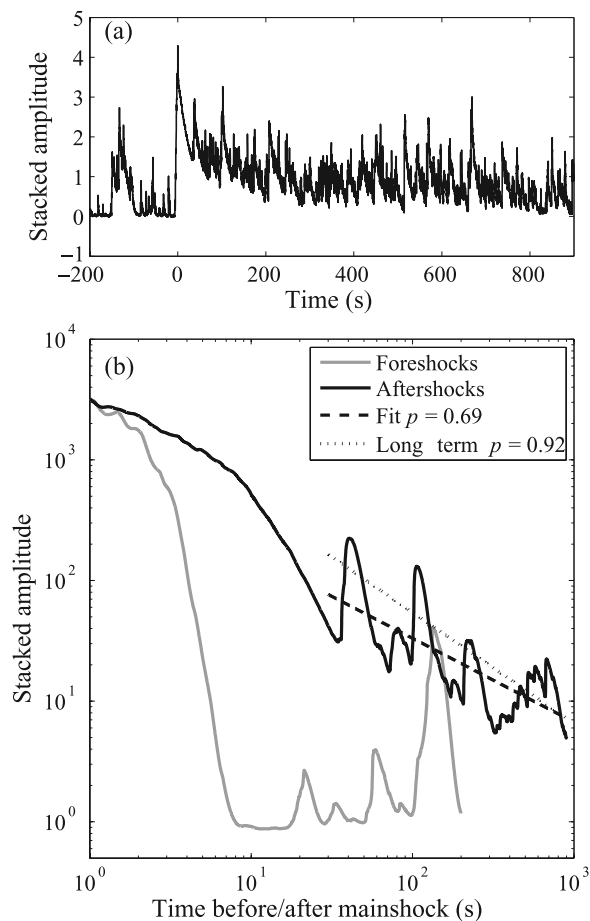


Figure 12. (a) Stacked envelope function for the 82 sequences immediately before and after the main shocks. (b) Smoothed amplitude for foreshocks (solid gray curve) and aftershocks (solid black curve). The dashed black line denotes a least squares fit to the data between 30 and 900 s, with a p value of 0.69. The dotted gray line is generated with $p = 0.92$ (corresponding to the aftershock decay rate at later time, as shown in Figure 9) for comparison. The origin time is chosen as the main shock peak amplitude. An increase in amplitude at $t \approx 6$ s before the main shock is due to the main shock first arrival.

last few tens of seconds, which is evident in the lack of handpicked events in that time period (Figure 6b).

5. Interpretation

[37] Figure 13 compare the p values obtained in this study as a function of time before and after the main shock. The early aftershock decay rate is smaller than the late aftershock decay rate. The foreshock increase rate is smaller than the aftershock decay rate, and the immediate foreshock increase rate is smaller than the long-term foreshock increase rate. Below we will discuss possible interpretations that can explain observations for aftershocks and foreshocks separately.

5.1. Possible Interpretations for Aftershock Decay Rate

[38] The decay of the number of aftershocks with time has been observed for more than a century [Omori, 1894].

However, the underlying physics for such time-dependent phenomena remains unclear. Aftershock triggering is commonly explained by the static stress change induced by the main shock [e.g., Stein, 1999], but dynamic stress changes associated by the main shock have also been invoked [Gomberg et al., 1997; Brodsky et al., 2003; Parsons, 2005]. Many mechanisms have been proposed to explain earthquake triggering by stress changes, such as stress corrosion [Das and Scholz, 1981; Yamashita and Knopoff, 1987; Gomberg, 2001], rate-and-state friction [Dieterich, 1994], and damage rheology [Ben-Zion and Lyakhovskiy, 2006]. Aftershock decay can be also explained by post-seismic relaxation following the main shock, which can result from fluid flow [Nur and Booker, 1972], viscous relaxation [Mikumo and Miyatake, 1979] or aseismic slip [Benioff, 1951; Schaff et al., 1998; Peng et al., 2005; Zöller et al., 2005].

[39] Among them, we focus on the rate-and-state (RS) model [Dieterich, 1994] that has been widely applied to explain the time dependence of aftershock activity. This model assumes that populations of faults around the main shock rupture obey the laboratory-derived RS friction law [Dieterich, 1979]. If the loading rate is increased uniformly due to a main shock, the change in seismicity can be approximated by the Omori's law with $p = 1$ for intermediate times after the main shock. At very short times, i.e., times less than a characteristic time $c = t_a e^{-\tau/A\sigma}$ (where t_a is the aftershock duration, A is a parameter of the RS friction law, τ is the coseismic stress change, and σ is the normal stress), the seismicity rate is constant. The seismicity rate returns to the background seismicity rate at very long times.

[40] Our observation of a small p value immediately after the main shock is consistent with the RS model. However, we did not find a clear gap of activity (or flattening of seismicity rate) since 10 s after the main shock. This indicates that the c value, if exists, is likely to be very small (<10 s) in this study. Recent studies have shown the existence of small, but nonzero c values (on the order of a few minutes) in the early aftershock decay rate for several large earthquakes in Japan and California [Vidale et al., 2003; Peng et al., 2006; Enescu et al., 2007]. The difference might be due to the main shock magnitude, employed techniques, or systematic bias by missing early aftershocks for large main shocks.

[41] The epidemic-type aftershock sequence (ETAS) model [Kagan and Knopoff, 1981; Ogata, 1988; Helmstetter and Sornette, 2002] can also explain our observed foreshock and aftershock rates. This model assumes that each earthquake can trigger other earthquakes according to the modified Omori's law with an exponent p larger than 1. However, due to cascades of earthquake triggering, the observed exponent p in this model is smaller than 1 at short times, and increases slowly with time [Helmstetter and Sornette, 2002; Helmstetter et al., 2003], as observed in our data.

5.2. Mapping Aftershock Rate Into Spatial Stress Distribution

[42] Since the coseismic stress change τ is likely to be different at each aftershock nucleation point, this will result in different characteristic time c overlapping with each other. Thus the observed aftershock rate is a superposition

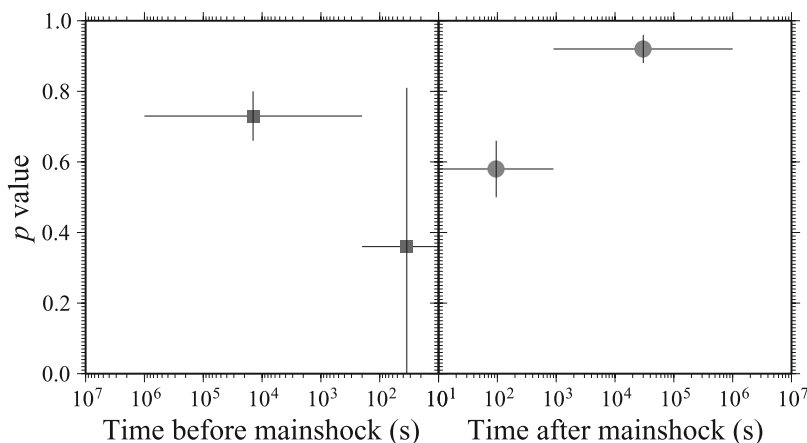


Figure 13. Comparison of the observed foreshock (square) and aftershock (circle) decay rates in this study. The horizontal bars mark the time range in which the least squares fitting with $1/t^p$ is made. The vertical bars denote 95% confidence level of the fit.

of those from different stress values, and the observed rate can be used to infer stress distribution after the main shock.

[43] There are two possible ways to map temporal changes of aftershock rate into spatial stress distribution. *Dieterich* [1994] showed that if the main shock is modeled as a dislocation, aftershock rate within a finite time interval and region decays with the p value of ~ 0.8 , due to a nonuniform stress change around the main shock. Because the stress change decreases with the distance r from the crack tip, this model also gives an apparent Omori exponent that decreases with r [*Dieterich*, 1994].

[44] In this study, we compute seismicity rate using all events within one-rupture length of the main shock and location error. The observed p value for long-term aftershocks is 0.92 ± 0.04 . This value is close to the aftershock decay rate of $p \sim 0.8$ predicted by the RS model for a finite region surrounding a shear crack [*Dieterich*, 1994]. However, this model does not explain our observation of a smaller p value at short times than that for larger times. At very short times in the RS model, the large stress drop resulting from a singularity at the crack tip control the seismicity. This gives a p value that increases progressively toward $p = 1$ as time approaches to the main shock occurrence time. In contrast, we find p value of ~ 0.6 immediately after the main shock.

[45] *Dieterich* [1994] assumed that stress drop is uniform across the main shock rupture area. This model would predict no aftershocks on or very close to the rupture surface. However, aftershocks do occur very close to the rupture surface of the main shock, although the fraction of events occurring on and off the main shock slip plane remains unclear [*Liu et al.*, 2003]. Several studies have shown an anticorrelation between the main shock slip area and the aftershock distribution [e.g., *Mendoza and Hartzell*, 1988; *Schaff et al.*, 2002; *Woessner et al.*, 2006]. These observations suggest that the coseismic stress change on the main shock fault plane is heterogeneous. The heterogeneity may be caused by slip fluctuations [*Herrero and Bernard*, 1994], or the rugosity of the fault [*Dieterich*, 2005]. Additional evidences for a heterogeneous stress field come from both main shock slip inversion [e.g., *Wald and Heaton*, 1994] and a high diversity of aftershock focal

mechanisms [e.g., *Michael et al.*, 1990; *Hardebeck et al.*, 1998].

[46] *Helmstetter and Shaw* [2006] and *Marsan* [2006] proposed a method to estimate distributions of stress heterogeneity from the aftershock rate based on the RS model [*Dieterich*, 1994]. Instead of mapping changes in seismicity onto time-dependent stress changes [*Dieterich et al.*, 2000], they mapped subtle but significant deviations from the pure Omori's law onto measures of stress heterogeneity on the main shock rupture area. *Helmstetter and Shaw* [2006] showed that the p value increases toward 1 with increasing stress heterogeneity. Furthermore, using a scale invariant " k^2 " slip model [*Herrero and Bernard*, 1994], they produced an aftershock rate that is close to the Omori's law with $p \leq 1$. In their stress heterogeneity model, the p value increases slowly with time. This is consistent with our observation.

[47] We have applied their technique to our catalogue. We fit the aftershock rate by the RS model, assuming that the

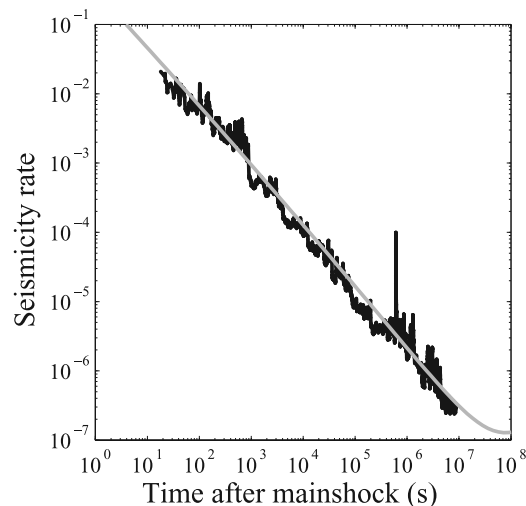


Figure 14. Seismicity rate (black line) observed in this study and a fit (gray line) by the rate-and-state model assuming a Gaussian stress distribution.

Coulomb stress change, in the region where we select aftershocks, has spatial fluctuations that can be modeled by a Gaussian distribution. This stress distribution can be obtained, for instance, by a kinematic “ k^2 ” slip model [Herrero and Bernard, 1994]. We fit the aftershock rate $R(t)$ in the time range 10 s to 100 days by the function [Dieterich, 1994; Helmstetter and Shaw, 2006]

$$R(t) = \int_{-\infty}^{\infty} \left[\frac{R_r}{(e^{-\tau/A\sigma} - 1)e^{-t/t_a} + 1} \right] \left[\frac{1}{\tau^* \sqrt{2\pi}} \right] e^{[-(\tau-\tau_0)^2/2(\tau^*)^2]} d\tau \quad (2)$$

where R_r is the reference seismicity rate, τ_0 is the average stress change, and τ^* is the standard deviation of the stress change. There are three unknown parameters in this function: the normalized average stress change $\tau_0/A\sigma$, the standard deviation $\tau^*/A\sigma$, and the aftershock duration t_a . The reference rate $R_r = 2 \times 10^{-7}$ events/s with $m \geq 1$ is estimated from the average seismicity rate at large times before the main shock (i.e., $\geq 10^7$ s). We use a least squares minimization to invert for $\tau_0/A\sigma$, $\tau^*/A\sigma$, and t_a from the observed aftershock rate. The problem is poorly constrained for a three-parameter inversion, thus we assume that the average stress change is 0. This is justified because we select aftershocks over an area slightly larger than the main shock rupture length. The large stress increase around the rupture area roughly compensates for the average stress decrease on the rupture area. The stress change averaged over an infinite area is equal to zero, while its absolute value is equal to the stress drop if we average over the rupture area. The inversion gives $\tau^*/A\sigma = 10.6$ and $t_a = 0.88$ year (Figure 14). If the average stress change is -1.5 , we obtain $\tau^*/A\sigma = 11.2$ and $t_a = 1.01$ years.

[48] There is a large uncertainty on the value of $A\sigma$ in the crust. Assuming $A = 0.01$, as measured in laboratory friction experiments [Dieterich, 1994], and $\sigma = 100$ MPa (corresponding to the lithostatic pressure at a depth of about 5 km), this gives $A\sigma = 1$ MPa. Using a typical stress drop of 3 MPa, Dieterich [1994] found $A\sigma = 0.15$ MPa from the relation between aftershock duration and main shock recurrence time. Cochran *et al.* [2004] used the RS model to explain tidal triggering of earthquakes, and obtained a preferred value of $A\sigma = 0.064$ MPa, and an acceptable range $0.048 < A\sigma < 0.11$ MPa. Moreover, there are probably large spatial fluctuations of $A\sigma$ in the crust. Heterogeneity of $A\sigma$ also modifies the time decay of aftershocks [Helmstetter and Shaw, 2006] and can bias the value of τ^* obtained by assuming that $A\sigma$ is constant in space and time. We can thus conclude that the RS model, with stress heterogeneity, can explain our observations (Omori exponent smaller than 1, and increasing slowly with time). However, it is difficult, from the observed seismicity rate $R(t)$ alone, to characterize the stress heterogeneity, and to distinguish different sources of heterogeneity (Coulomb coseismic stress change τ , parameter A of the friction law, and normal stress σ).

5.3. Possible Interpretations for Foreshock Increase Rate

[49] Our results indicate that foreshocks follow the inverse Omori’s law, but the foreshock increase rate is smaller than the aftershock decay rate, and there appears

to be a peak of foreshock activity ~ 100 s before the main shock. There are two possible models to explain the relationship between the foreshocks and main shocks. The first is that foreshocks change the stress distribution and the main shock. The second is that foreshocks are caused by a change of stress field related to a main shock.

[50] The first model is compatible with the ETAS model, and models of earthquake triggering such as the RS model [Dieterich, 1995; Ziv, 2003], in which cascade of events trigger future events with the same physical mechanism [Jones *et al.*, 1995; Kilb and Gomberg, 1999; Felzer *et al.*, 2004]. Using this model, the rate of foreshocks is predicted to increase following the inverse Omori’s law [Helmstetter *et al.*, 2003; Helmstetter and Sornette, 2003]. This is consistent with our observation of a power law increase of foreshock occurrences.

[51] The fact that Omori exponent p is smaller for foreshocks than for aftershocks is also in agreement with the predictions of the ETAS model [Helmstetter *et al.*, 2003]. Ziv [2003] used numerical simulations of the RS model, and found that the Omori exponent p is slightly smaller for foreshocks than for aftershocks. In his model, exponents for both foreshocks and aftershocks are smaller or equal to 1, and decrease as the distance from the main shock decreases. An observation that is consistent with this prediction was also made from time distribution of foreshocks of several hundred main shocks with $m \geq 5.0$ listed in the JMA catalogue [Maeda, 1999].

[52] The foreshock increase rate in our data may also be influenced by our selection rule for the main shock. We did not include an earthquake that was preceded by a larger event as a potential main shock. In comparison, an earthquake that was followed by a larger event is considered as a potential main shock. Such rule may lower the average acceleration of the seismicity before the main shock, and hence lower the Omori exponent for foreshocks.

[53] We also observed that there is a slight dependence of foreshock rate with the m_c value. The rate of larger events is higher than that of smaller events (Figure 10). In addition, there is a lack of small events in the last 20 s before the main shock. These results are expected by the triggering model such as ETAS. Because a large foreshock is more likely to trigger the main shock than a smaller one, the magnitude distribution deviates from the average distribution [Michael and Jones, 1998; Helmstetter *et al.*, 2003; Helmstetter and Sornette, 2003]. If the long-term magnitude distribution obeys the G-R law $P(m) \sim 10^{-bm}$, where $P(m)$ is the number of events with magnitude $\geq m$, and aftershock productivity scales as 10^{am} with the main shock magnitude m , the magnitude distribution of foreshocks at short times before the main shock is $\sim 10^{-(b-a)m}$. This results in a small apparent b value.

[54] The peak of foreshock activity at ~ 100 s before the main shock may be due to statistical fluctuations. The Omori’s law decay is generally observed for each individual aftershock sequence, while the inverse Omori’s law is observed only on stacked foreshock sequences, due to larger fluctuations in the foreshock rate than in the aftershock rate. In addition, foreshocks usually occur in clusters, which may explain the peak observed in the data. We have generated synthetic ETAS catalogues, which reproduce qualitatively our results: increase of p value with time for foreshocks and

aftershocks, lower p value for foreshocks than for aftershocks, and large fluctuations of the foreshock rate.

[55] In the second model, foreshocks and main shock nucleation are considered as a consequence of nonseismic stress change associated with the main shock [e.g., Dodge *et al.*, 1996; Olson and Allen, 2005]. Possible candidates for the nonseismic stress changes are episodic aseismic slip [Linde *et al.*, 1996; Nadeau and McEvilly, 2004; McGuire, 2003; McGuire *et al.*, 2005], pore fluid pressure fluctuations [e.g., Ake *et al.*, 2005], dike injection [e.g., Smith *et al.*, 2004], and sometimes a combination of these candidates [Hainzl and Ogata, 2005].

[56] Vidale and Shearer [2006] systematically investigated 71 earthquake bursts in southern California, and found that many bursts of activities are driven by underlying geophysical disturbance such as episodic aseismic slip, or pore fluid pressure fluctuations, rather than simple elastic stress triggering. On the basis of stacked global earthquake sequences, Jones and Molnar [1979] found that foreshock activity rises a few days before the main shock and peaks 1–2 hours before the main shock. Chen *et al.* [1999] showed that foreshocks are separated by a few hours of quiescence from several major earthquakes in China. Zanker *et al.* [2003] identified two clusters of foreshock sequences, one between 20 to 13 hours and the other starting 8 hours before the 1999 Hector Mine earthquake. McGuire *et al.* [2005] found higher rates of foreshocks and lower rates of aftershocks on the oceanic East Pacific Rise transform faults. They inferred that slow slip events, which accommodate most of the aseismic plate motion on those faults, trigger the earthquake sequences. A concentration of foreshock activity in the last 100–1000 s before the main shock observed on these oceanic transform faults [McGuire, 2003] is similar to our observation of a peak in foreshock activity at ~ 100 s before the main shock. It is possible that concentration of foreshock activity may differ for individual sequence, or it is related to the main shock magnitude. Our study involves waveform recording of the last 200 s before the main shock, so the possibility of intense foreshock activity at earlier times cannot be ruled out.

6. Conclusions

[57] We have analyzed foreshock and aftershock rates of 82 shallow earthquake sequences (depth ≤ 30 km) with main shock magnitude $3 \leq m \leq 5$ in Japan using waveforms recorded by the Hi-net borehole array. By scrutinizing high-frequency signals, we have detected ~ 5 times as many aftershocks in the first 200 s as listed in the JMA catalogue. The difference between the aftershocks listed in the JMA catalogue and the additional events identified by our procedure indicates that seismic activity immediately before and after the main shock can be obtained from high-quality continuous waveform data.

[58] After correcting for the changing completeness level immediately after the main shock, we find that early aftershocks decay slower than the later aftershocks. Similarly, the seismicity rate for the foreshock sequences follows an inverse Omori's law with a p value of ~ 0.7 from several hundred days up to near the main shock occurrence time. The seismicity rate in the last 200 s before the main shock appears steady instead of increasing with time.

[59] These observations can be explained by the ETAS model, and the rate-and-state model [Dieterich, 1994] for a heterogeneous stress change along the fault plane [Dieterich, 2005; Marsan, 2006; Helmstetter and Shaw, 2006]. Alternatively, seismicity around the main shock may be caused by nonseismic stress changes around the source region, due, for example, to episodic aseismic slip or pore fluid pressure fluctuations.

[60] **Acknowledgments.** We thank the National Research Institute for Earth Science and Disaster Prevention in Japan for making the Hi-net data available. Yehuda Ben-Zion, Emily Brodsky, Jim Dieterich, Karen Felzer, Heidi Houston, Debi Kilb, Bruce Shaw, and Yan Kagan generously provided commentary. The manuscript benefited from valuable comments by the Editor John C. Mutter and two anonymous referees. The analysis is done using the Seismic Analysis Code [Goldstein *et al.*, 2003]. Most figures are prepared in GMT [Wessel and Smith, 1998]. This work is supported by UC Lawrence Livermore National Laboratory (U.S. DOE) 04-GS-026, and Southern California Earthquake Center (SCEC). SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. This is SCEC contribution 1029.

References

- Abercrombie, R. E., and J. Mori (1996), Occurrence patterns of foreshocks to large earthquakes in the western United States, *Nature*, *381*, 303–307.
- Ake, J., K. Mahrer, D. O'Connell, and L. Block (2005), Deep-injection and closely monitored induced seismicity at Paradox Valley, Colorado, *Bull. Seismol. Soc. Am.*, *95*, 664–683.
- Aki, K. (1969), Analysis of seismic coda of local earthquakes as scattered waves, *J. Geophys. Res.*, *74*, 615–631.
- Benioff, H. (1951), Earthquakes and rock creep. part I: Creep characteristics of rocks and the origin of aftershocks, *Bull. Seismol. Soc. Am.*, *41*, 31–62.
- Ben-Zion, Y., and Y. Lyakhovskiy (2006), Analysis of aftershocks in a lithospheric model with seismogenic zone governed by damage rheology, *Geophys. J. Int.*, *165*, doi:10.1111/j.1365-246X.2006.02878.x.
- Brodsky, E. E., E. Roeloffs, D. Woodcock, I. Gall, and M. Manga (2003), A mechanism for sustained groundwater pressure changes induced by distant earthquakes, *J. Geophys. Res.*, *108*(B8), 2390, doi:10.1029/2002JB002321.
- Chen, Y., J. Liu, and H. Ge (1999), Pattern characteristics of foreshock sequences, *Pure Appl. Geophys.*, *155*, 395–408.
- Cochran, E. S., J. E. Vidale, and S. Tanaka (2004), Earth tides can trigger shallow thrust fault earthquakes, *Science*, *306*, 1164–1166.
- Das, S., and C. Scholz (1981), Off-fault aftershock clusters caused by shear stress increase?, *Bull. Seismol. Soc. Am.*, *71*, 1669–1675.
- Dieterich, J. H. (1979), Modeling of rock friction: 1. Experimental results and constitutive equations, *J. Geophys. Res.*, *84*, 2161–2168.
- Dieterich, J. H. (1994), A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.*, *99*, 2601–2618.
- Dieterich, J. H. (1995), Earthquake simulation with time-dependent nucleation and long-range interactions, *Nonlinear Processes Geophys.*, *2*, 109–120.
- Dieterich, J. H. (2005), Role of stress relaxation in slip of geometrically complex faults, *Eos Trans. AGU*, *86*(52), Fall Meet. Suppl., Abstract NG31A-04.
- Dieterich, J. H., V. Cayol, and P. Okubo (2000), The use of earthquake rate as a stress meter at Kilauea volcano, *Nature*, *408*, 457–460.
- Dodge, D. A., G. C. Beroza, and W. L. Ellsworth (1996), Detailed observations of California foreshock sequences: Implications for the earthquake initiation process, *J. Geophys. Res.*, *101*, 22,371–22,392.
- Enescu, B., J. Mori, and M. Miyazawa (2007), Quantifying early aftershock activity of the 2004 mid-Niigata Prefecture earthquake ($M_w 6.6$), *J. Geophys. Res.*, doi:10.1029/2006JB004629, in press.
- Eshelby, J. D. (1957), The determination of the elastic field of an ellipsoidal inclusion and related problems, *Proc R. Soc. London, Ser. A*, *241*, 376–396.
- Felzer, K. R., and E. E. Brodsky (2006), Decay of aftershock density with distance indicates triggering by dynamic stress, *Nature*, *441*, 735–738.
- Felzer, K. R., R. E. Abercrombie, and G. Ekström (2004), A common origin for aftershocks, foreshocks, and multiplets, *Bull. Seismol. Soc. Am.*, *94*, 88–98.
- Geller, R. J., D. D. Jackson, Y. Y. Kagan, and F. Mulargia (1997), Earthquakes cannot be predicted, *Science*, *275*, 1616.
- Gerstenberger, M. C., S. Wiemer, L. M. Jones, and P. A. Reasenber (2005), Real-time forecasts of tomorrow's earthquakes in California, *Nature*, *435*, 328–331, doi:10.1038/nature03622.

- Goldstein, P., D. Dodge, M. Firpo, and L. Minner (2003), SAC2000: Signal processing and analysis tools for seismologists and engineers, in *The IASPEI International Handbook of Earthquake and Engineering Seismology, Part B*, edited by W. H. K. Lee et al., chapter 85.5, Elsevier, New York.
- Gomberg, J. (2001), The failure of earthquake failure models, *J. Geophys. Res.*, *106*, 16,253–16,264.
- Gomberg, J., M. L. Blanpied, N. M. Beeler, and P. Bodin (1997), Transient triggering of near and distant earthquakes, *Bull. Seismol. Soc. Am.*, *87*, 294–309.
- Gutenberg, R., and C. F. Richter (1944), Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.*, *34*, 185–188.
- Hainzl, S., and Y. Ogata (2005), Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. Geophys. Res.*, *110*, B05S07, doi:10.1029/2004JB003247.
- Hardebeck, J. L., J. J. Nazareth, and E. Hauksson (1998), The static stress change triggering model: Constraints from two southern California aftershock sequences, *J. Geophys. Res.*, *103*, 24,427–24,437.
- Helmstetter, A., and D. Sornette (2002), Subcritical and supercritical regimes in epidemic models of earthquake aftershocks, *J. Geophys. Res.*, *107*(B10), 2237, doi:10.1029/2001JB001580.
- Helmstetter, A., and B. E. Shaw (2006), Relation between stress heterogeneity and aftershock rate in the rate-and-state model, *J. Geophys. Res.*, *111*, B07304, doi:10.1029/2005JB004077.
- Helmstetter, A., and D. Sornette (2003), Foreshocks explained by cascades of triggered seismicity, *J. Geophys. Res.*, *108*(B10), 2457, doi:10.1029/2003JB002409.
- Helmstetter, A., D. Sornette, and J. Grasso (2003), Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws, *J. Geophys. Res.*, *108*(B1), 2046, doi:10.1029/2002JB001991.
- Helmstetter, A., Y. Kagan, and D. Jackson (2005), Importance of small earthquakes for stress transfers and earthquake triggering, *J. Geophys. Res.*, *110*, B05S08, doi:10.1029/2004JB003286.
- Helmstetter, A., Y. Kagan, and D. D. Jackson (2006), Comparison of short-term and long-term earthquake forecast models for southern California, *Bull. Seismol. Soc. Am.*, *76*, 90–106.
- Herrero, A., and P. Bernard (1994), A kinematic self-similar rupture process for earthquakes, *Bull. Seismol. Soc. Am.*, *84*, 1216–1228.
- Jones, L. M., and P. Molnar (1979), Some characteristics of foreshocks and their possible relationship to earthquake prediction and premonitory slip on fault, *J. Geophys. Res.*, *84*, 3596–3608.
- Jones, L. M., R. Console, F. Di Luccio, and M. Murru (1995), Are foreshocks mainshocks whose aftershocks happen to be big? Evidence from California and Italy (abstract), *Eos Trans. AGU*, *76*(46), Fall Meet. Suppl., F388.
- Kagan, Y. Y. (1991), Seismic moment distribution, *Geophys. J. Int.*, *106*, 123–136.
- Kagan, Y. Y. (2004), Short-term properties of earthquake catalogues and models of earthquake source, *Bull. Seismol. Soc. Am.*, *94*, 1207–1228.
- Kagan, Y. Y., and H. Houston (2005), Relation between mainshock rupture process and Omori's law for aftershock moment release rate, *Geophys. J. Int.*, *163*(3), 1039–1048.
- Kagan, Y. Y., and L. Knopoff (1981), Stochastic synthesis of earthquake catalogues, *J. Geophys. Res.*, *86*, 2853–2862.
- Kilb, D., and J. Gomberg (1999), The initial subevent of the 1994 Northridge, California, Earthquake: Is earthquake size predictable?, *J. Seismol.*, *3*, 409–420.
- Kilb, D., V. G. Martynov, and F. L. Vernon (2004), Spatial/Temporal interdependence of aftershocks following the 10/31/2001 M5.1 Anza earthquake, *Eos Trans. AGU*, *85*(47), Fall Meet. Suppl., Abstract S54A-02.
- Kisslinger, C. (1996), Aftershocks and fault-zone properties, *Adv. Geophys.*, *38*, 1–36.
- Kisslinger, C., and L. M. Jones (1991), Properties of aftershock sequences in southern California, *J. Geophys. Res.*, *96*, 11,947–11,958.
- Linde, A. T., M. T. Gladwin, M. J. S. Johnston, R. L. Gwyther, and R. G. Bilham (1996), A slow earthquake sequence on the San Andreas fault, *Nature*, *383*, 65–68.
- Liu, J. K. Sieh, and E. Hauksson (2003), A structural interpretation of the aftershock “cloud” of the 1992 M_w 7.3 Landers earthquake, *Bull. Seismol. Soc. Am.*, *93*, 1333–1344.
- Lolli, B., and P. Gasperini (2006), Comparing different models of aftershock rate decay: The role of catalogue incompleteness in the first times after mainshock, *Tectonophysics*, *423*, 43–59.
- Maeda, K. (1999), Time distribution of immediate foreshocks obtained by a stacking method, *Pure Appl. Geophys.*, *155*, 381–394.
- Marsan, D. (2006), Can coseismic stress variability suppress seismicity shadows? Insights from a rate-and-state friction model, *J. Geophys. Res.*, *111*, B06305, doi:10.1029/2005JB004060.
- McGuire, J. J. (2003), Immediate foreshock sequences of oceanic transform earthquakes on the East Pacific Rise, *Bull. Seismol. Soc. Am.*, *93*, 948–952.
- McGuire, J. J., M. S. Boettcher, and T. H. Jordan (2005), Foreshock sequences and short-term earthquake predictability on East Pacific Rise transform faults, *Nature*, *434*, 457–461, doi:10.1038/nature03377.
- Mendoza, C., and S. H. Hartzell (1988), Aftershock patterns and mainshock faulting, *Bull. Seismol. Soc. Am.*, *78*, 1438–1449.
- Michael, A. J., and L. M. Jones (1998), Seismicity alert probabilities at Parkfield, California, revisited, *Bull. Seismol. Soc. Am.*, *88*, 117–130.
- Michael, A. J., W. L. Ellsworth, and D. Oppenheimer (1990), Co-seismic stress changes induced by the 1989 Loma Prieta, California earthquake, *Geophys. Res. Lett.*, *17*, 1441–1444.
- Mikumo, T., and T. Miyatake (1979), Earthquake sequences on a frictional fault model with non-uniform strengths and relaxation times, *Geophys. J. R. Astron. Soc.*, *59*, 497–522.
- Mogi, K. (1962), Study of the elastic shocks caused by the fracture of heterogeneous materials and its relation to earthquake phenomena, *Bull. Earthquake Res. Inst. Univ. Tokyo*, *40*, 125–173.
- Mori, J. J., and R. E. Abercrombie (1997), Depth dependence of earthquake frequency-magnitude distributions in California: Implications for rupture initiation, *J. Geophys. Res.*, *102*, 15,081–15,090.
- Nadeau, R. M., and T. V. McEvilly (2004), Periodic pulsing of characteristic microearthquakes on the San Andreas Fault, *Science*, *303*, 220–222.
- Nanjo, K., H. Nagahama, and M. Satomura (1998), Rates of aftershock decay and the fractal structure of active fault systems, *Tectonophysics*, *287*, 173–186.
- Narteau, C., P. Shebalin, and M. Holschneider (2002), Temporal limits of the power law aftershock decay rate, *J. Geophys. Res.*, *107*(B12), 2359, doi:10.1029/2002JB001868.
- Nur, A., and J. R. Booker (1972), Aftershocks caused by pore fluid flow?, *Science*, *175*, 885–888.
- Ogata, Y. (1988), Statistical models for earthquake occurrence and residual analysis for point processes, *J. Am. Stat. Assoc.*, *83*, 9–27.
- Okada, Y., K. Kasahara, S. Hori, K. Obara, S. Sekiguchi, H. Fujiwara, and A. Yamamoto (2004), Recent progress of seismic observation networks in Japan –Hi-net, F-net, K-NET and KiK-net, *Earth Planet Space*, *56*, xv-xxviii.
- Olson, E. L., and R. M. Allen (2005), The deterministic nature of earthquake rupture, *Nature*, *438*, 212–215, doi:10.1038/nature04214.
- Omori, F. (1894), On the aftershocks of earthquakes, *J. Coll. Sci. Imp. Univ. Tokyo*, *7*, 111–200.
- Parsons, T. (2005), A hypothesis for delayed dynamic earthquake triggering, *Geophys. Res. Lett.*, *32*, L04302, doi:10.1029/2004GL021811.
- Peng, Z., J. E. Vidale, C. Marone, and A. Rubin (2005), Systematic variations in moment with recurrence interval of repeating aftershocks, *Geophys. Res. Lett.*, *32*, L15301, doi:10.1029/2005GL022626.
- Peng, Z., J. E. Vidale, and H. Houston (2006), Anomalous early aftershock decay rate of the 2004 M_w 6.0 Parkfield, California, earthquake, *Geophys. Res. Lett.*, *33*, L17307, doi:10.1029/2006GL026744.
- Reasenber, P. A. (1999), Foreshock occurrence before large earthquakes, *J. Geophys. Res.*, *104*, 4755–4768.
- Reasenber, P. A., and L. M. Jones (1989), Earthquake hazard after a mainshock in California, *Science*, *243*, 1173–1176.
- Reasenber, P. A., and L. M. Jones (1994), Earthquake aftershocks: Update, *Science*, *265*, 1251–1252.
- Sato, H., and M. Fehler (1998), *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*, Springer, New York.
- Schaff, D. P., G. C. Beroza, and B. E. Shaw (1998), Postseismic response of repeating aftershocks, *Geophys. Res. Lett.*, *25*, 4549–4552.
- Schaff, D. P., G. H. R. Bokelmann, G. C. Beroza, F. Waldhauser, and W. L. Ellsworth (2002), High resolution image of Calaveras Fault seismicity, *J. Geophys. Res.*, *107*(B9), 2186, doi:10.1029/2001JB000633.
- Shcherbakov, R., D. L. Turcotte, and J. B. Rundle (2004), A generalized Omori's law for earthquake aftershock decay, *Geophys. Res. Lett.*, *31*, L11613, doi:10.1029/2004GL019808.
- Shcherbakov, R., D. L. Turcotte, and J. B. Rundle (2006), Scaling properties of the Parkfield aftershock sequence, *Bull. Seismol. Soc. Am.*, *96*, S376–S384.
- Smith, K. D., D. V. Seggern, G. Blewitt, L. Preston, J. G. Anderson, B. P. Wernicke, and J. L. Davis (2004), Evidence for deep magma injection beneath Lake Tahoe, Nevada-California, *Science*, *305*, 1277–1280.
- Stein, R. S. (1999), The role of stress transfer in earthquake occurrence, *Nature*, *402*, 605–609.
- Utsu, T. (1966), A statistical significance test of the difference in b-value between two earthquake groups, *J. Phys. Earth*, *14*, 37–40.
- Utsu, T., Y. Ogata, and R. S. Matsu'ura (1995), The centenary of the Omori formula for a decay law of aftershock activity, *J. Phys. Earth*, *43*, 1–33.
- Vidale, J. E., and P. M. Shearer (2006), A survey of 71 earthquake bursts across southern California: Exploring the role of pore fluid pressure

- fluctuations and aseismic slip as drivers, *J. Geophys. Res.*, *111*, B05312, doi:10.1029/2005JB004034.
- Vidale, J. E., E. S. Cochran, H. Kanamori, and R. W. Clayton (2003), After the lightning and before the thunder; non-Omori behavior of early aftershocks?, *Eos Trans. AGU*, *84*(46), Fall Meet. Suppl., Abstract S31A-08.
- Wald, D. J., and T. H. Heaton (1994), Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake, *Bull. Seismol. Soc. Am.*, *84*, 668–691.
- Wessel, P., and W. H. F. Smith (1998), New, improved version of the Generic Mapping Tools released, *Eos Trans. AGU*, *79*, 579.
- Wiemer, S., and K. Katsumata (1999), Spatial variability of seismicity parameters in aftershock zones, *J. Geophys. Res.*, *104*, 13,135–13,151.
- Wiemer, S., and M. Wyss (2000), Minimum magnitude of completeness in earthquake catalogues: Examples from Alaska, the western United States, and Japan, *Bull. Seismol. Soc. Am.*, *90*, 859–869.
- Woessner, J., and S. Wiemer (2005), Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty, *Bull. Seismol. Soc. Am.*, *95*, 684–698.
- Woessner, J., D. Schorlemmer, S. Wiemer, and P. M. Mai (2006), Spatial correlation of aftershock locations and on-fault main shock properties, *J. Geophys. Res.*, *111*, B08301, doi:10.1029/2005JB003961.
- Working Group on California Earthquake Probabilities (2003), Earthquake probabilities in the San Francisco Bay region: 2002-2031, *U.S. Geol. Surv. Open File Rep.*, 03-214.
- Yamashita, T., and L. Knopoff (1987), Models of aftershock occurrence, *Geophys. J. R. Astron. Soc.*, *91*, 13–26.
- Zanzerkia, E. E., G. C. Beroza, and J. E. Vidale (2003) Waveform analysis of the 1999 Hector Mine, California, foreshock sequence, *Geophys. Res. Lett.*, *30*(8), 1428, doi:10.1029/2002GL016612.
- Ziv, A. (2003), Foreshocks, aftershocks, and remote triggering in quasi-static fault models, *J. Geophys. Res.*, *108*(B10), 2498, doi:10.1029/2002JB002318.
- Ziv, A., A. M. Rubin, and D. Kilb (2003), Spatiotemporal analyses of earthquake productivity and size distribution: Observations and simulations, *Bull. Seismol. Soc. Am.*, *93*, 2069–2081.
- Zöller, G., S. Hainzl, M. Holschneider, and Y. Ben-Zion (2005), Aftershocks resulting from creeping sections in a heterogeneous fault, *Geophys. Res. Lett.*, *32*, L03308, doi:10.1029/2004GL021871.

A. Helmstetter, Observatoire des Sciences de l'Université de Grenoble, Laboratoire de Géophysique Interne et Tectonophysique, BP 53, F-38041 Grenoble Cedex 9, France.

M. Ishii, Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA.

Z. Peng, School of Earth and Atmospheric Sciences, Georgia Institute of Technology, 311 Ferst Drive, Atlanta, GA 30332, USA. (zpeng.seismo@gmail.com)

J. E. Vidale, Department of Earth and Space Sciences, University of California, Los Angeles, CA 90095-1567, USA.