LETTERS

Decay of aftershock density with distance indicates triggering by dynamic stress

K. R. Felzer[1] & E. E. Brodsky[2]

The majority of earthquakes are aftershocks[1], yet aftershock physics is not well understood. Many studies suggest that static stress changes[2,3] trigger aftershocks, but recent work suggests that shaking (dynamic stresses) may also play a role[4,5]. Here we measure the decay of aftershocks as a function of distance from magnitude 2–6 mainshocks in order to clarify the aftershock triggering process. We find that for short times after the mainshock, when low background seismicity rates allow for good aftershock detection, the decay is well fitted by a single inverse power law over distances of 0.2–50 km. The consistency of the trend indicates that the same triggering mechanism is working over the entire range. As static stress changes at the more distant aftershocks are negligible, this suggests that dynamic stresses may be triggering all of these aftershocks. We infer that the observed aftershock density is consistent with the probability of triggering aftershocks being nearly proportional to seismic wave amplitude. The data are not fitted well by models that combine static stress change with the evolution of frictionally locked faults[3].

Previous studies of how aftershock density decays with distance from the mainshock have found a range of functions, including power laws and combinations of power laws with constants and exponentials[6–8]. The ability to study the decay with improved clarity has recently been provided by the publication of large catalogues with precise earthquake locations. We use the Shearer *et al.* relocated 1984–2002 Southern California catalogue[9]. Standard location error is of the order of tens of metres, similar to that obtained in more localized studies using similar relocation techniques[10,11]. For most of the analysis we use magnitude $M = 2$–4 mainshocks and $M \geq 2$ aftershocks. We prefer small mainshocks because their large number allows for good statistical averaging and for the use of a small difference between mainshock and aftershock magnitude, which improves catalogue completeness[12]. For near field measurements, where larger mainshocks are necessary for appropriate range and precision, we use $M = 5$–6 mainshocks and $M \geq 3$ aftershocks. Earthquakes are used as mainshocks if they are sufficiently isolated from larger earthquakes in time and space (see Methods). To improve statistics, we combine the aftershocks of each unit magnitude range of mainshocks (Fig. 1).

The spatial density of a point process can be measured in any dimension. For instance, a density could be the number of points per length, per area, or per volume. We choose to measure linear density, that is, the number of aftershocks per length. This is done by collapsing all of the aftershocks onto a single line with their position on the line equal to their distance from the mainshock. The number of aftershocks per unit length can then be measured at different positions using standard statistical tools (see Methods).

We first study earthquakes occurring within 5 min and 50 km of $M = 2$–4 mainshocks. The short time window minimizes the amount of background seismicity[13], that is, earthquakes not aftershocks of the designated mainshocks. We approximate the mainshocks as point sources and measure the distance, r, between mainshock and aftershock hypocentres. From 0.2 to 50 km, the data are well fitted by

$$\rho(r) = cr^{-n} \quad (1)$$

where $n = 1.37 \pm 0.1$ for the $3 \leq M < 4$ mainshocks and 1.35 ± 0.12 for the $2 \leq M < 3$ mainshocks, and c is a constant that varies with the number of aftershocks (Fig. 2). The error is the 98% confidence interval based on 5,000 bootstraps. For $r < 0.2$ km the point source approximation is no longer accurate (Supplementary Figs 1 and 2).

We also check the applicability of equation (1) to longer times. For 30 minutes of post-mainshock data the time window is still short, and an inverse power law, with $n = 1.36 \pm 0.07$, fits the $3 \leq M < 4$ mainshocks. For $2 \leq M < 3$ mainshocks background seismicity begins to level the decay around 16 km, but a clear inverse power law with $n = 1.37 \pm 0.15$ is evident at shorter distances (Supplementary Fig. 3). Background interference is worse for smaller mainshocks because of lower aftershock productivity per mainshock. For 30 minutes to 25 days of post-mainshock earthquakes there is universal deviation from a pure power law, but we find that a combined power law/background function fits the data well, at 95% and 65% confidence (Supplementary Fig. 4). This suggests that the aftershock decay with distance does not change with time.

To study the density of aftershocks within one fault length of the

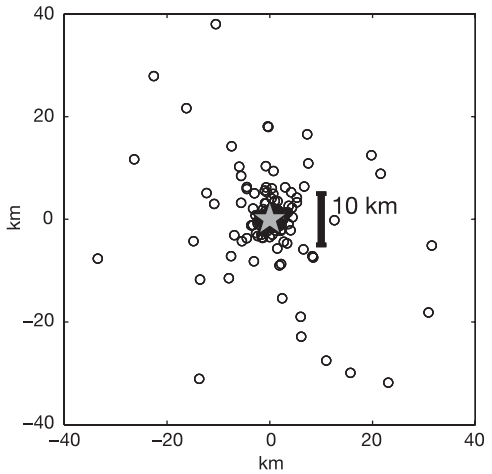

Figure 1 | Combined aftershocks of $M = 3$–4 mainshocks. To create a composite aftershock data set, we move all of the mainshocks to the origin in space and time and move their aftershocks with them. Data here are for the first 30 minutes of aftershocks of $M = 3$–4 mainshocks. The grey star gives the locations of the mainshocks, at the origin.

[1]US Geological Survey, 525 S. Wilson, Pasadena, California 91106, USA. [2]Department of Earth Science, University of California Santa Cruz, 1156 High Street, Santa Cruz, California 95060, USA.

mainshock, we use $M = 5$–6 mainshocks that have a Harvard CMT (centroid moment tensor) focal mechanism solution and aftershock distributions that clearly delineate the preferred fault plane. We estimate the length and width of the faults from empirical relationships[14], and centre the fault planes at the median aftershock location. We measure aftershock density from 0.2 km to 12 km from the fault plane (~0.05 fault lengths of an $M = 5$ earthquake to one fault length of an $M = 6$ earthquake). As before, we recover an inverse power law, with $n = 1.34 \pm 0.25$ (Fig. 3). The decay levels out at $r < 0.2$ km owing to error in fault plane location and catalogue incompleteness in the very near field.

To verify that the decay we observe is due to aftershock physics, we repeat the analysis for the $M = 2$–4 mainshocks with a time-randomized catalogue. This produces a large scatter of points (Fig. 4), indicating that the pattern in Fig. 2 is aftershock-related. To verify that a pure inverse power law is a good functional fit to the aftershocks, we use the Kolmogorov–Smirnoff test (see Methods). We also test the fit of a composite power law of the form $\rho(r) = a_1 r^{-n_1} + a_2 r^{-n_2}$. The bayesian information criterion[15,16] prefers the single power law fit. To check if our results are catalogue dependent, we verify that aftershocks in unrelocated Japanese and Northern California catalogues also follow an inverse power law decay (Supplementary Fig. 6).

The consistent aftershock decay relationship observed from distances of 0.2 km to 50 km—from within 0.05 fault lengths of $M = 5$ mainshocks to over 100 fault lengths of $M = 2$–3 mainshocks—implies that static stress change is not triggering the aftershocks. Static stresses decay rapidly. At 3 km from a $M = 4$ earthquake the static stress change is at most 4 kPa, comparable to tidal stresses that have not been found to trigger earthquakes[17]; at 10 km from a $M = 3$ earthquake the static stress change is three orders of magnitude lower. Triggering by static stress in the near field and dynamic stress in the far field would require a discontinuity in the aftershock decay. Only uniform triggering by dynamic stress matches the observation of a single, consistent decay that traverses a wide range of distances.

The hypothesis of aftershock triggering by static stress change has received strong support in part because there is a physical model, rate and state friction[3], which explains how static stress changes could result in the power law distribution of aftershock times[18]. Our observations indicate, however, that this model does not fit the distribution of aftershocks in space. At very short times, using a point source approximation for the static stress in a whole space, the model predicts that the density of aftershocks from an earthquake of moment M_0 is

$$\rho(r) = B(r)e^{c_2 r^{-3}} \quad (2)$$

where $B(r)$ is the background seismicity per kilometre per unit time as a function of distance from the mainshock and $c_2 = M_0/4\pi A\sigma$, where σ is normal stress and A is a frictional constant. At a given distance from a set of mainshocks, the observed density in a combined data set is the sum of the individual ones.

The functional form of $B(r)$ can be estimated from distance measurements between random earthquakes (Fig. 4). The scatter is large, but the data can be fitted with the same functional form used by others[19,20]:

$$B(r) = c_1 r^{(D-1)} \quad (3)$$

This function describes points randomly scattered on a structure with effective dimension D. We substitute equation (3) into equation (2) and use a grid search to find the best fit to the 30 min/16 km aftershocks of $M = 3$–4 mainshocks over a wide parameter range (see Methods). For the best fit case, with $D = 0.1$, the summed squared error is 1.4 times worse than for the best inverse power law fit, and the correlation of the data residuals is high ($r = 0.45$; 601 data points) (Supplementary Fig. 7). This correlation has less than a 0.01% chance of occurring for a good functional fit. For the residuals of the inverse power law, $r = 0.0079$. For more realistic values of $D = 1$–2, the rate and state model fit is much worse (Supplementary Fig. 7). The point source approximation may produce an inaccurate representation of the static stress change within one fault length of the source (for example, within 1 km for $M = 4$ mainshocks and 0.1 km for $M = 2$ mainshocks), but the functional shape and associated misfit is problematic at further distances as well.

The poor fit of the data to equation (2) indicates that the aftershocks are not triggered by static stress change coupled with rate and state friction, at least at distances beyond one fault length. We also find more model-dependent evidence that the number of aftershocks triggered varies linearly with dynamic stress change amplitude.

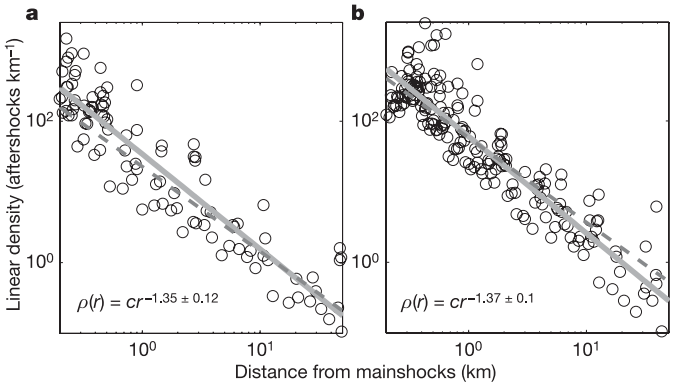

Figure 2 | Distance from the mainshock hypocentre versus aftershock linear density. Aftershocks are $M > 2$ and occur in the first 5 minutes. **a**, $2 \leq M < 3$ mainshocks. The plot uses 7,396 mainshocks and 104 aftershocks. **b**, $3 \leq M < 4$ mainshocks. The plot uses 2,355 mainshocks and 199 aftershocks. The data are fitted with an inverse power law with an exponent of -1.35 ± 0.12 for the $2 \leq M < 3$ data and -1.37 ± 0.1 for the $3 \leq M < 4$ data (solid lines). The fit is made from 0.2 to 50 km for both plots. For comparison, dashed lines give the decay of maximum seismic wave amplitude, a proxy for dynamic stress, as derived from the standard Richter relationship[22]. Data over a wider distance range are given in Supplementary Fig. 1.

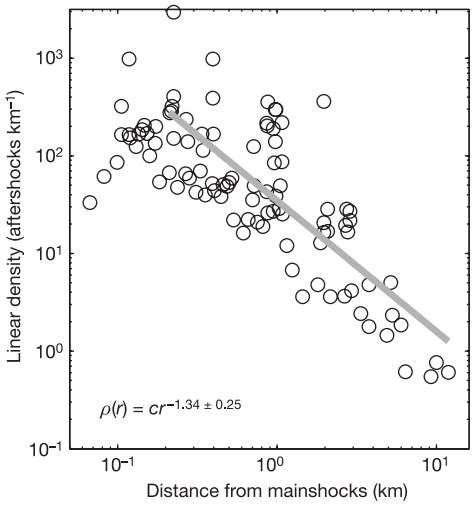

Figure 3 | Aftershock density versus distance from the closest point on the fault planes of $M = 5$–6 mainshocks. The plot uses nine mainshocks and 104 $M > 3$ aftershocks that occurred within 2 days of the mainshock. A fit to the data is made from 0.2 to 12 km, or from 0.05 fault lengths of an $M = 5$ earthquake to one fault length of an $M = 6$ earthquake. The decay rate shows good agreement with the far field decay found for $M = 2$–4 mainshocks. Far field aftershocks of these mainshocks are given in Supplementary Fig. 5.

Linear aftershock density, $\rho(r)$, can be separated into geometric and physical terms

$$\rho(r) = \frac{N_{\mathrm{aft}}}{\Delta r} = \frac{N_{\mathrm{hyp}}}{\Delta r} \times \frac{N_{\mathrm{aft}}}{N_{\mathrm{hyp}}} = cr^{-1.4} \quad (4)$$

where N_{aft} is the number of aftershocks in a shell of width Δr centred at distance r from the mainshock and N_{hyp} is the number of potential hypocentres, or places where aftershocks could be triggered, in the same shell. If the background seismicity is roughly evenly distributed over the active fault population, the functional form of the background earthquake distribution (equation (3)) can be substituted for the geometric term:

$$\rho(r) = c_1 r^{D-1} c_3 r^m = cr^{-1.4} \quad (5)$$

The trade-off between D and m prevents a direct solution for D. We instead estimate D by fitting the data in Fig. 4. We find a better fit with $D = 1$ than with $D = 2$ or 3; that is, the linear density is independent of distance. A D value close to 1 is also suggested by geometric considerations (see Methods).

If $D \approx 1$ the fraction of potential hypocentres triggered decays as $r^{-1.4}$, that is, at a rate slightly stronger than $1/r$. The maximum amplitude of seismic waves[21] also decay somewhat faster than $1/r$. The standard Richter relationship for maximum short period seismic wave amplitude[22,23] is well fitted by a combination of $r^{-1.2}$ decay and attenuation with $Q = 300$. Attenuation is less important at the depth of the aftershocks, so the peak amplitude of shaking at depth falls off as $r^{-1.2}$ (Fig. 2). The power law decay of local seismic amplitude is primarily due to the geometrical spreading of the S waves, wave focusing, and surface wave amplitude decay. Given the uncertainty in D, the data are consistent with the probability of triggering an aftershock being proportional to the amplitude of the seismic wave. The probability of aftershock triggering scaling with the amplitude of the mainshock wave also predicts that the number of aftershocks increases by a factor of ten with each mainshock magnitude unit, as has been observed[20,24,25].

In summary, the decay of aftershock linear density with distance from $M = 2$–6 mainshocks is well fitted by an inverse power law. The trend exists at least from 0.2 to 50 km for the first 5 minutes of the aftershock sequence. At longer times background seismicity makes distant aftershocks more difficult to detect, but we can trace the decay trend to at least 16 km for the first 30 minutes, and to at least 10 km for the first 2 days. If the linear density of faults is independent of distance ($D \approx 1$), then the data indicate that the probability of triggering an aftershock is directly proportional to the amplitude of seismic shaking. Whether or not $D = 1$, dynamic triggering is preferred by the data. The similarity of aftershock decay from distances of 0.05 to over 100 fault lengths implies a single physical triggering mechanism, and dynamic stress change is the only plausible agent over most of this range.

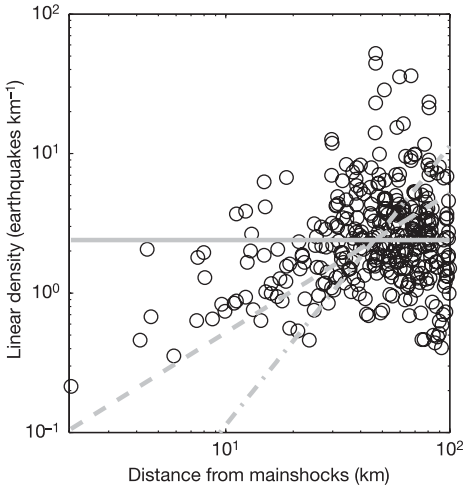

Figure 4 | Distance versus earthquake linear density for a time-randomized catalogue. Distances are measured between ten random earthquakes from the mainshock data set and random catalogue earthquakes, producing 346 earthquake pairs. Unlike the aftershock data, there is no systematic decay of density with distance. The sum of squared residuals is lowest when the data are fitted with the relationship that linear earthquake density is constant (solid line), as opposed to linear density equals r (dashed line), which would correspond to hypocentres being located randomly in two dimensions, or r^2 (dashed-dotted line), which would correspond to random locations in a volume (see equation (3)). The fits are done from 0 to 100 km, which completely covers the range of aftershock data analysed in this paper.

METHODS

Mainshock and aftershock selection. Earthquakes are used as mainshocks if they are separated from larger earthquakes by at least L km or by t_1 days if the larger earthquake comes first, and t_2 days if it comes after. This separation minimizes contamination from aftershocks of larger earthquakes. L is set at 100 km; larger values (at least up to 500 km) produce the same results. The results are insensitive to the values of t_1 and t_2 as long as $t_1 \ll t < t_2$, where t is the time after the mainshock for which we use aftershock data. For $t = 5$ min and $t = 30$ min we obtain the same aftershock decay for values of t_1 between 3 and 100 days. We use $t_1 = 3$ days, which, within this range, maximizes the number of qualified mainshocks, and $t_2 = 0.5$ days. For Supplementary Fig. 4, where t can be as large as 25 days, we set $t_1 = 100$ days and $t_2 = 26$ days. For Fig. 3, where $t = 2$ days, we use $t_1 = 30$ days and $t_2 = 2$ days.

We use a uniform distance cut-off for measuring aftershocks of all mainshocks, even though the mainshock magnitudes vary. This is justified by our observation, in agreement with previous authors[8,26], that the distribution of aftershock distances is independent of mainshock magnitude (Supplementary Fig. 8).

Linear density measurement. To measure linear density we use the nearest neighbour method[27], in which densities are estimated by taking the inverse of the width of the box required to contain k neighbouring points. The edges of sequential boxes meet between data points. Smoothing is controlled by k. We find that our results are constant for $k = 1$–20, although the fitting error increases with k (Supplementary Fig. 9). We use $k = 1$, which produces the smallest error and minimum bias[27]. The advantages of the nearest neighbour method are that the number and location of measurements are determined by the location of the data points, smoothing is uniform, and there are no empty bins.

The definition of linear density averages over all azimuths and is therefore insensitive to radiation pattern.

Catalogue completeness. We check catalogue completeness by fitting the Gutenberg-Richter magnitude frequency relationship[28], with $b = 1$. The aftershocks of $M = 2$–3 mainshocks are complete to $M = 2$. For $M = 3$–4 mainshocks, 10–15% of the smallest aftershocks ($M = 2.0$–2.2) measured over the first 30 minutes may be missing, but there is no systematic loss with distance and thus no expected effect. Over the first 5 minutes, closer to 30–40% of the smallest aftershocks of the $M = 3$–4 mainshocks may be missing, but again the loss is not significantly systematic with distance. Note that this incompleteness results in aftershock productivity appearing to vary as a factor of six with mainshock magnitude in Fig. 2, rather than the factor of ten noted in the literature[20,24,25]. The $M \geq 3$ aftershocks used for the $M = 5$–6 mainshocks are complete.

Goodness of fit. If the proposed single inverse power law fit is adequate, the residuals should have a distribution similar to that of data drawn from a pure power law distribution via Monte Carlo simulation. We test for similarity between the residuals of simulated data and the largest robust data set in this study (Southern California 30 min/16 km aftershocks of $M = 3$–4 mainshocks) with the Kolmogorov-Smirnoff test. At 95% and 65% confidence the test indicates that the null hypothesis—that the observed and simulated residuals come from the same distribution—cannot be rejected.

We test the composite power law $\rho(r) = a_1 r^{-n_1} + a_2 r^{-n_2}$ by fitting a_1 and a_2 with a nonlinear least-squares algorithm for every value of n_1 and n_2 between 0.5 and 3. We find that the single power law is preferred (bayesian information criterion).

Testing rate and state friction. To fit the rate and state friction equation to the aftershock density data, we try the parameters: $c_1 = 10^{-5.72}$–$10^{-1.172}$; stress drop = 0.1–10 MPa; normal stress = 10–1,000 MPa; $A = 0.005$–0.012 (ref. 3); $D = 0.1$–2. The resulting range of c_2 is 0.008–$200 \times (M_0/4\pi)$. We achieve a minimum least-squared error at $D = 0.1$, $c_1 = 0.0034$ and $c_2 = 0.33(M_0/4\pi)$.

Potential hypocentre distribution effective dimension, *D*. Because of the limited seismogenic depth of Southern California, at length scales over ~10 km the system

would effectively be two-dimensional if potential hypocentres were randomly scattered throughout the crust. In reality, earthquakes concentrate on planar faults, whose width is also limited by the seismogenic depth. At distances longer than ~10–20 km, effective D for earthquakes randomly scattered on a fault tends towards 1. Multiple faults increase D, but the concentration of earthquakes in streaks and clusters on faults decreases it[11]. Thus we expect $D \approx 1$ at large distances, and the lack of any break in the slope of the aftershock density decay, from 200 m to 100 km, for a wide range of mainshock magnitudes, suggests that $D \approx 1$ throughout.

Received 7 October 2005; accepted 10 April 2006.

1. Gardner, J. K. & Knopoff, L. Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? *Bull. Seismol. Soc. Am.* **64,** 1363–1367 (1974).
2. Stein, R. S. The role of stress transfer in earthquake occurrence. *Nature* **402,** 605–609 (1999).
3. Dieterich, J. A. Constitutive law for the rate of earthquake production and its application to earthquake clustering. *J. Geophys. Rev.* **99,** 2601–2618 (1994).
4. Gomberg, J., Bodin, P. & Reasenberg, P. A. Observing earthquakes triggered in the near field by dynamic deformations. *Bull. Seismol. Soc. Am.* **93,** 118–138 (2003).
5. Johnson, P. A. & Jia, X. Nonlinear dynamics, granular media and dynamic earthquake triggering. *Nature* **437,** 871–874 (2005).
6. Ichinose, G. A., Anderson, J. G. & Smith, K. D. Static stress change caused by the 1978 Diamond Valley, California and 1994 Double Spring Flat Nevada earthquakes. *Eos* **78**, abstr. S22B-04 (1997).
7. Ogata, Y. Statistical models for earthquake occurrence and residual analysis for point processes. *J. Am. Stat. Assoc.* **83,** 9–27 (1988).
8. Huc, M. & Main, I. G. Anomalous stress diffusion in earthquake triggering: Correlation length, time-dependence, and directionality. *J. Geophys. Res.* **108,** 2324, doi:10.1029/2001JB001645 (2003).
9. Shearer, P., Hauksson, E. & Lin, G. Southern California hypocenter relocation with waveform cross-correlation, part 2: Results using source-specific station terms and cluster analysis. *Bull. Seismol. Soc. Am.* **95,** 904–915 (2005).
10. Dodge, D. A., Beroza, G. C. & Ellsworth, W. L. Foreshock sequence of the 1992 Landers, California earthquake and its implications for earthquake nucleation. *J. Geophys. Res.* **100,** 9865–9880 (1995).
11. Rubin, A. M., Gillard, D. & Got, J. Streaks of microearthquakes along creeping faults. *Nature* **400,** 635–641 (1999).
12. Kagan, Y. Y. Short-term properties of earthquake catalogs and models of earthquake source. *Bull. Seismol. Soc. Am.* **94,** 1207–1228 (2004).
13. Bak, P., Christensen, K., Danon, L. & Scanlon, T. Unified scaling law for earthquakes. *Phys. Rev. Lett.* **88,** 109901 (2002).
14. Wells, D. L. & Coppersmith, K. J. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bull. Seismol. Soc. Am.* **84,** 974–1002 (1995).
15. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **6,** 461–464 (1978).
16. Main, I. G., Leonard, T., Papasouliotis, O., Hatton, C. G. & Meredith, P. G. One slope or two? detecting statistically significant breaks of slope in geophysical data, with application to fracture scaling relationships. *Geophys. Res. Lett.* **26,** 2801–2804 (1999).
17. Vidale, J. E., Agnew, D. C., Johnston, M. J. S. & Oppenheimer, D. H. Absence of earthquake correlation with earth tides: An indication of high preseismic fault stress rate. *J. Geophys. Res.* **103,** 24567–24572 (1998).
18. Omori, F. On the aftershocks of earthquakes. *J. Coll. Sci. Imp. Univ. Tokyo* **7,** 111–200 (1894).
19. Kagan, Y. Y. & Knopoff, L. Spatial distribution of earthquakes: the two-point correlation function. *Geophys. J. R. Astron. Soc.* **62,** 303–320 (1980).
20. Helmstetter, A., Kagan, Y. Y. & Jackson, D. D. Importance of small earthquakes for stress transfers and earthquake triggering. *J. Geophys. Res.* **110,** B05S08, doi:10.1029/2004JB003286 (2005).
21. Campbell, K. W. *Strong Motion Attenuation Relations* (Academic, London, 2003).
22. Richter, C. F. An instrumental earthquake-magnitude scale. *Bull. Seismol. Soc. Am.* **25,** 1–32 (1935).
23. Kanamori, H. *et al.* Determination of earthquake energy release and M_L using terrascope. *Bull Seismol. Soc. Am.* **83,** 330–346 (1993).
24. Michael, A. J. & Jones, L. M. Seismicity alert probabilities at Parkfield, California, revisited. *Bull. Seismol. Soc. Am.* **88,** 117–130 (1998).
25. Felzer, K. R., Abercrombie, R. E. & Ekström, G. A common origin for aftershocks, foreshocks, and multiplets. *Bull. Seismol. Soc. Am.* **94,** 88–98 (2004).
26. Davidsen, J. & Paczuski, M. Analysis of the spatial distribution between successive earthquakes. *Phys. Rev. Lett.* **94,** 048501 (2005).
27. Silverman, B. W. *Density Estimation for Statistics and Data Analysis* (Chapman and Hall, New York, 1986).
28. Gutenberg, B. & Richter, C. F. Frequency of earthquakes in California. *Bull. Seismol. Soc. Am.* **4,** 185–188 (1944).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature

Acknowledgements We thank R. Abercrombie, A. Felzer, N. Field, M. Gerstenberger, J. Gomberg, S. Gross, J. Hardebeck, M. Harrington, A. Helmstetter, S. Hough, L. Jones, Y. Kagan, H. Kanamori, I. Main, S. Prejean, P. Shearer, R. Stein, J. Vidale, K. Richards-Dinger and A. Yong for comments. This work was supported in part by the National Science Foundation.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to K.R.F. (kfelzer@gps.caltech.edu).