*Geophysical Research Letters*

Supporting Information for

# Larger aftershocks happen farther away: non-separability of magnitude and spatial distribution of aftershocks

Nicholas J. van der Elst[1,2] and Bruce E. Shaw[2]

1. U.S. Geological Survey Earthquake Science Center, Pasadena, Calif. 2. Lamont-Doherty Earth Observatory, Columbia University, Palisades, N.Y.

## Contents of this File



## Introduction

This document contains the following supporting information: (1) Maps and descriptions of the subset of aftershock sequences for which the largest aftershock is larger than the mainshock (Table S1, Figure S1); (2) Plots of aftershock ranked distance as a function of aftershock-mainshock magnitude difference, for alternative Catalogs and reference frames (Figures S2 – S4); and (3) Descriptions of the procedures used to fit parameters (Text S1) and establish statistical significance (Text S2).

## Text S1. Fitting the power-law kernel to the aftershock distributions

We fit the parameters of the spatial kernel (Eqs. 4-6) using a grid search for maximum likelihood (Fig. S5). We use the population of aftershocks with $\Delta M \leq 0$, $r \leq 10$ km, and $t \leq 3.2$ days.

The maximum likelihood estimate (MLE) should weight the aftershocks of each magnitude bin approximately evenly, since the combined mainshocks of each magnitude bin tend to produce the same total number of aftershocks [*Helmstetter et al.*, 2005]. However, due to secondary triggering and other factors, the population of aftershocks may not be entirely independent and identically distributed. This makes it difficult to rigorously compare the likelihoods of the best-fit and the theoretically constrained model. Nevertheless, we use the Akaike Information Criterion (AIC) [*Akaike*, 1973] to adjust for the difference in free parameters and calculate rough confidence bounds on the best-fit parameters. The difference in AIC between the two models is defined as

$$\Delta AIC = 2k_{th} - 2\log\left(L_{th}\right) - 2k_{mle} + 2\log\left(L_{mle}\right). \qquad \text{(S1)}$$

Here $k_{th}$ and $k_{mle}$ are the number of free parameters, and $L_{th}$ and $L_{mle}$ are the maximum likelihoods computed for the theoretical and MLE parameter sets, respectively. The 'plausibility' of a candidate model relative to the best-fit model under the AIC framework can be expressed as $P = \exp(-\Delta AIC/2)$ [*Burnham and Anderson*, 2002] (equivalent to a likelihood ratio corrected for number of parameters). Contours of $\Delta AIC$ (Eq. S1) are given in Fig. S5, along with the $P = 5\%$ ($\Delta AIC \approx 6$) boundary. The difference in AIC score between the MLE best-fit and the theoretically constrained parameter set is 7.4, putting the theoretical parameters near the boundary of what is considered to be plausibly supported by the data.

The AIC also justifies the use of the geometrical correction given by Equation 5. The maximum likelihood score for a model with no geometrical correction, i.e. $N(r)$ = 1, gives a power-law decay constant $\hat{\gamma}$ = 3.21, and a log-likelihood 3714 units smaller than with Eq. 5. Using a simpler geometrical correction of the form $N(r) = r$ gives $\hat{\gamma}$ = 3.67 and a likelihood score 179 log-likelihood units smaller. The reduced number of free parameters in these alternative models can only account for a difference of 2 log-likelihood units, meaning the AIC dramatically prefers the geometrical correction given by Eq. 5.

For comparison with previous studies, we also directly fit the linear density using least squares, following Felzer and Brodsky [2006]. This approach yields a geometrical increase of $r^{1.05\pm0.02}$ for distances less than ½ the mainshock rupture length (from Eq. 6), and a power-law decay of $r^{-1.77\pm0.007}$ for distances greater than the mainshock rupture length. The uncertainty estimates given are 95% confidence bounds.

### S1.1. Fitting the internal length scale for individual magnitude bins

In the inset of Figure 4, we plot the internal length-scale $d$ for individual magnitude bins. To get these individual $d$ values, we solve for the expected median $r_{50}$ of the power law kernel (Eq. 4), and use this formula to compute $d$ from the observed median of the aftershock population. The median of Eq. 4 is given by the formula

$$r_{50} = d\left[\left(1-0.5D\right)^{1/(1-\gamma/2)} - 1\right]^{1/2}, \qquad \text{(S2)}$$

where

$$D = 1 - \left[\left(\frac{R}{d}\right)^2 + 1\right]^{1-\gamma/2}$$

The constant $R$ is the maximum bound on the aftershock distance. For fitting the median we use a maximum bound $R = 0.75w$, where $w$ = 7.8 km, and use the approximation $N(r) \sim r$, valid for $r \ll w$ (Eq. 5).

## Text S2. Comparison with the ETAS model

This study rests on the interpretation of combined aftershock sequences from a catalog in which the actual triggering relationships between foreshocks, mainshocks, and aftershocks are ambiguous. Whenever such a stacking method is employed, there is a possibility that unanticipated signals will show up due to unconsidered interactions between overlapping sequences. We therefore compare the results from the earthquake catalogs to results from simulated earthquake catalogs constructed with Epidemic-Type Aftershock sequence (ETAS) model.

### S2.1. ETAS ingredients

The ETAS model combines well-established statistical laws of aftershocks to produce realistic simulated earthquake catalogs. We start with an initial catalog of spontaneous or "background" earthquakes, with random magnitudes assigned from the Gutenberg-Richter distribution. To each of these earthquakes, we assign a number of aftershocks (offspring) that depends exponentially on the parent magnitude, and distribute these offspring in time according to Omori's law. Magnitudes are then assigned to the offspring, and the process repeats until a generation arises with no new earthquakes. The ETAS simulation procedure is well described elsewhere [*Helmstetter and Sornette*, 2003; *Ogata*, 1998; 2011; *Zhuang et al.*, 2002]. We generate an initial catalog of background earthquakes proportional to a gridded estimate of background rate for the California catalog, and set the ETAS parameters to produce catalogs with on average the same total number of events as in the observed catalog. For reproducibility, we report the parameters in Table S2. See the aforementioned references for a more elaborate description of the function of each parameter in the simulation.

We assign parent-offspring distances using a kernel of the form

$$p(r) \propto r\left(r^2 + d_{MS}^2\right)^{-\gamma/2} \qquad \text{(S3)}$$

where $d_{MS}$ is the length scale of the parent earthquake from Eq. (6), and $\gamma = 3.0$.

Depths for the initial seed catalog of 'background' earthquakes are drawn from the observed depth distribution of the catalog mainshocks, but aftershock depths are allowed to range from 0 to 30 km. In the real catalog, the aftershock depths are more restricted, leading to the geometrical correction $N(r)$ given as Eq. 5 in the main text. In the simulations we use the approximation $N(r) \sim r$ (Eq. S3) to permit the efficient numerical generation of random aftershock distances. The ETAS simulations generated with Eq. S3 produce median aftershock distances that are slightly larger than those in the catalog (~2.1 km for the simulations vs. ~1.4 km for the real catalog). Since the goal of the simulations is to look for differences in the distributions of large and small aftershocks, and not to perfectly reproduce the observed catalog, these small discrepancies should not be a problem. To be safe we compare the results as a function of ranked distance, which normalizes each aftershock zone by the distribution of smaller events and corrects for any small systematic differences between the observation and simulations.

### S2.2. ETAS results

The results are summarized in Fig. S7. While the scatter in the distribution of ranked distance increases with ΔM (as the sample size diminshes), the median ranked distance stays centered on 0.5, and the ETAS simulations do not reproduce the observed tendency for the larger aftershocks to occur farther away (Fig S7). The difference between observations and simulations is significant above the 95% level for ΔM ≥ -0.5, as measured by the fraction of simulations that produce a mean or median ranked distance as large as that observed. While the mean ranked distance in the observational data continues to grow with ΔM, the sample size becomes too small for a meaningful comparison above ΔM = 1.

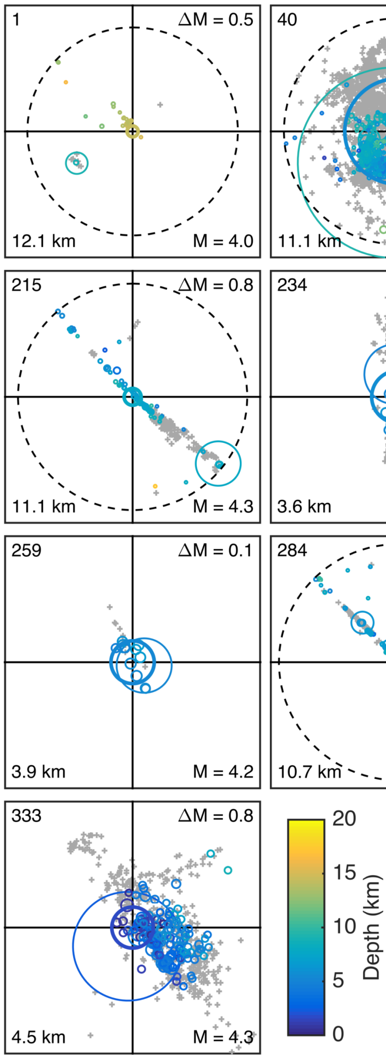


**Fig S1a.** Northern California mainshock-aftershock sequences for which the largest aftershock is as large or larger than the mainshock in map view. Mainshock is the thick circle centered on the crosshairs. Other colored circles are aftershocks up to and including the first aftershock larger than the mainshock. Circles have radius equal to the rupture radius (Eq. 6). Grey crosses are aftershocks in the 10 days after the first aftershock larger than the mainshock. Numbers in corners are: NW - sequence number (Fig. 1); NE - magnitude difference between largest aftershock and mainshock; SE - mainshock magnitude; SW - edge length of one quadrant of each plot for scale. The dashed circle shows the 10 km limit for collecting aftershocks.

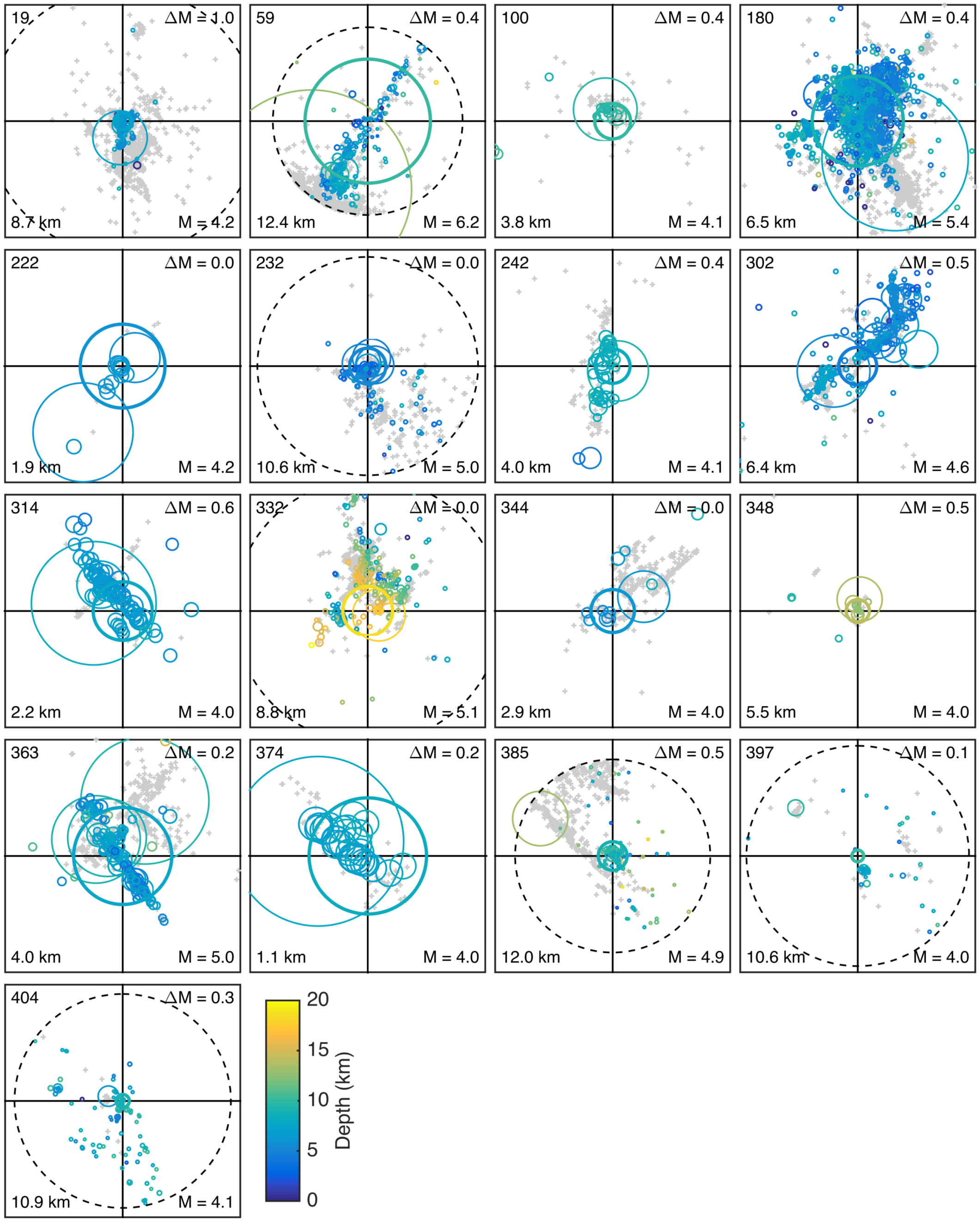


**Fig S1b.** Sequence plots for Southern California. See Fig. S1a caption for additional details.

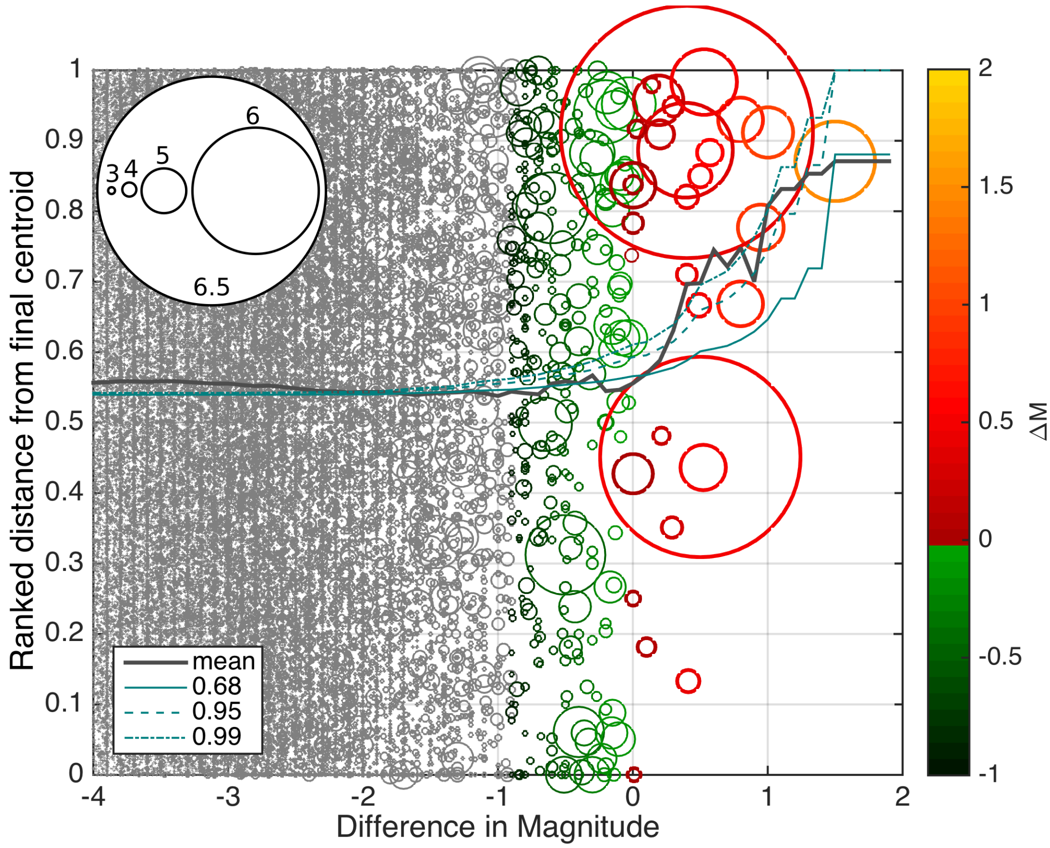


**Fig S2.** Ranked distance of aftershocks with respect to the *final centroid* location of all aftershocks, up to and including the first aftershock with magnitude as large or larger than the mainshock. Compare to Fig. 3 in the main text. Circles are scaled to aftershock magnitude. Solid black line is the running average over one magnitude unit. Green lines are 68, 95, and 99% confidence bounds on the mean from repeated random sampling of the smaller aftershocks ($-2 \leq \Delta M < -1$).

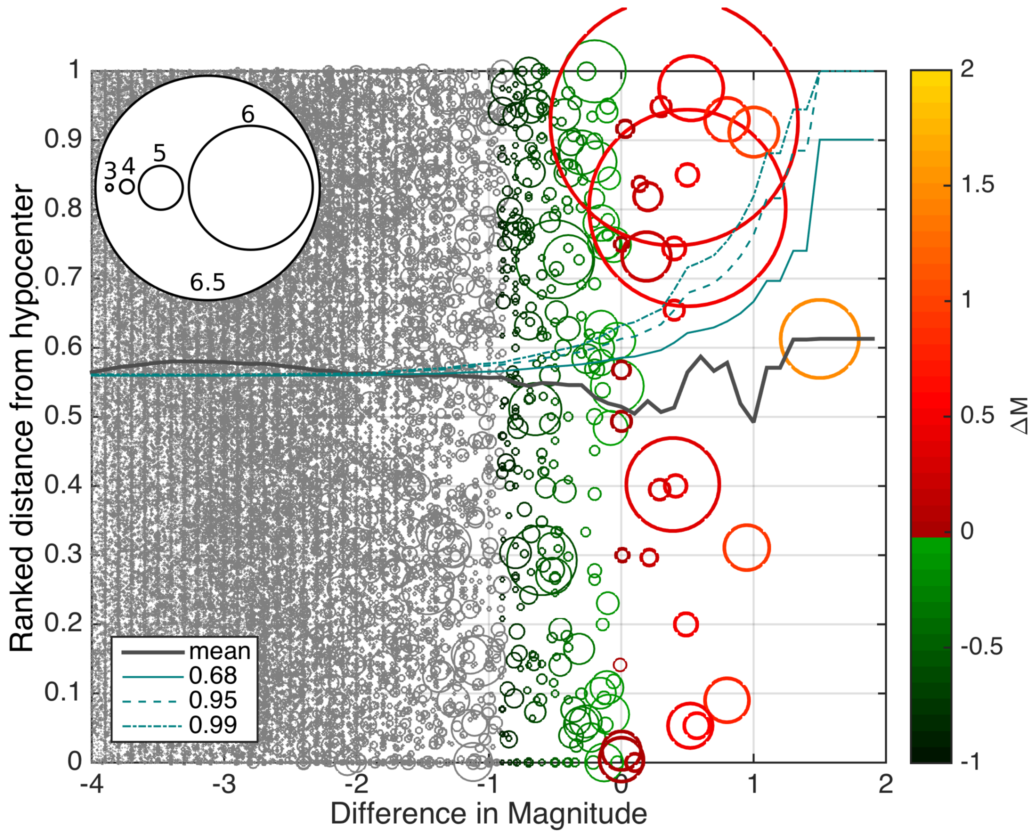


**Fig S3.** Ranked distance to aftershocks with respect to the mainshock *hypocenter*, as a function of aftershock-mainshock magnitude difference. Compare to Fig. 3 in the main text. Circles are scaled to aftershock magnitude. Solid black line is the running average over one magnitude unit. Green lines are 68, 95, and 99% confidence bounds on the mean from repeated random sampling of the smaller aftershocks ($-2 \leq \Delta M < -1$).

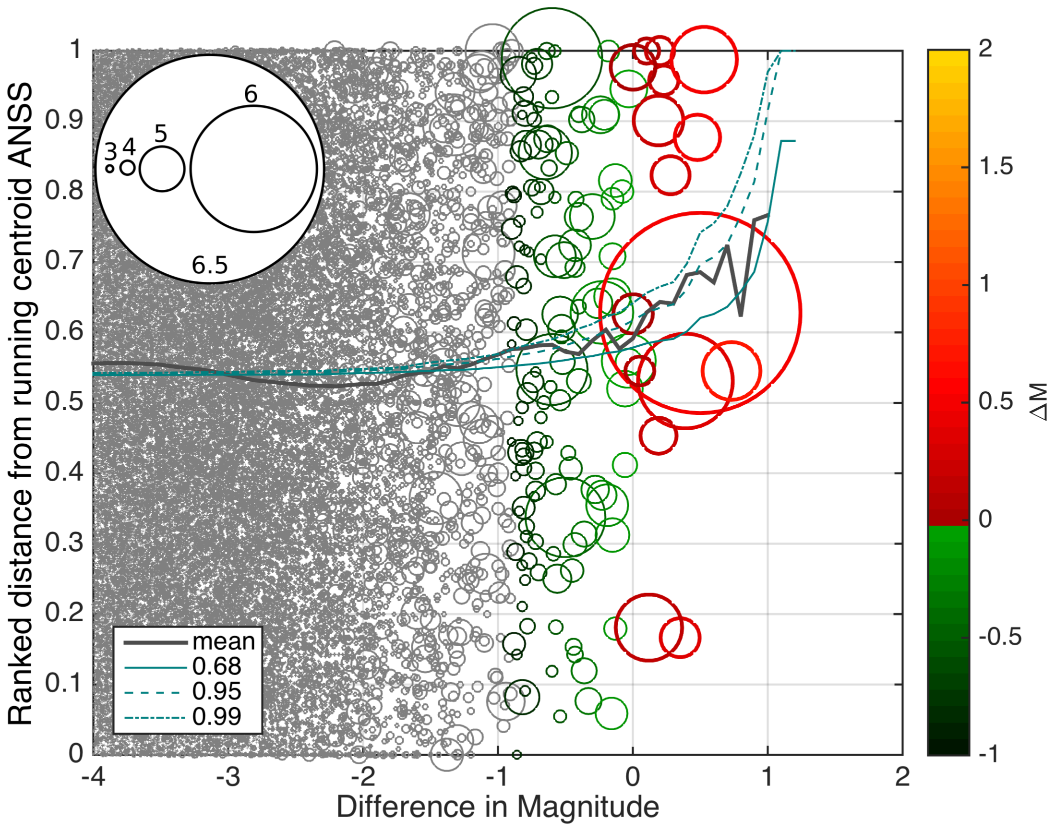


**Fig S4.** Ranked aftershock distance in the ANSS catalog, with respect to the running centroid location of all aftershocks, up to and including the first aftershock with magnitude as large or larger than the mainshock. Compare to Fig. 3 in the main text. Mainshock magnitudes in the ANSS catalog are limited to M4.5 and above (~1 km source radius), to account for the greater location uncertainty. Circles are scaled to aftershock magnitude. Solid black line is the running average over one magnitude unit. Green lines are 68, 95, and 99% confidence bounds on the mean from repeated random sampling of the smaller aftershocks ($-2 \leq \Delta M < -1$).

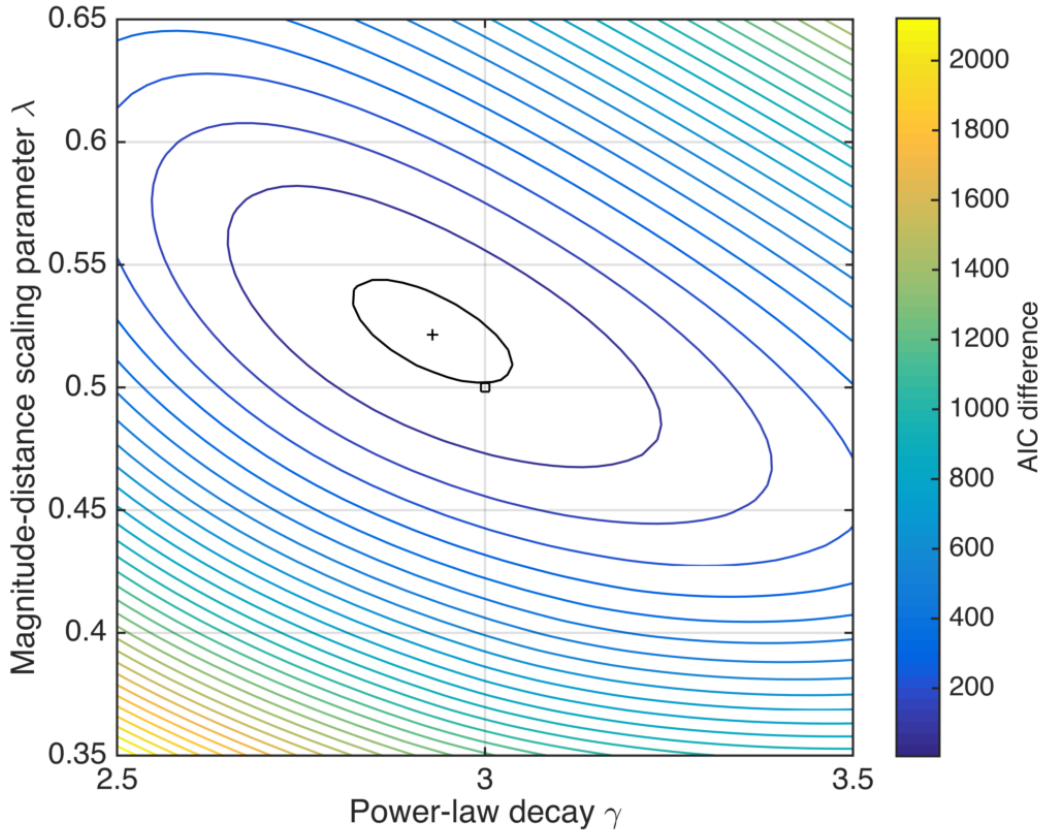


**Fig S5.** Maximum likelihood estimate (MLE) of the spatial kernel parameters (+). Contours give the difference in Akaike Information Criterion (AIC), relative to the MLE (Eq. S1) The black contour encloses the parameter space with greater than 5% relative likelihood compared to the MLE. The theoretically-motivated parameters are marked by a square.

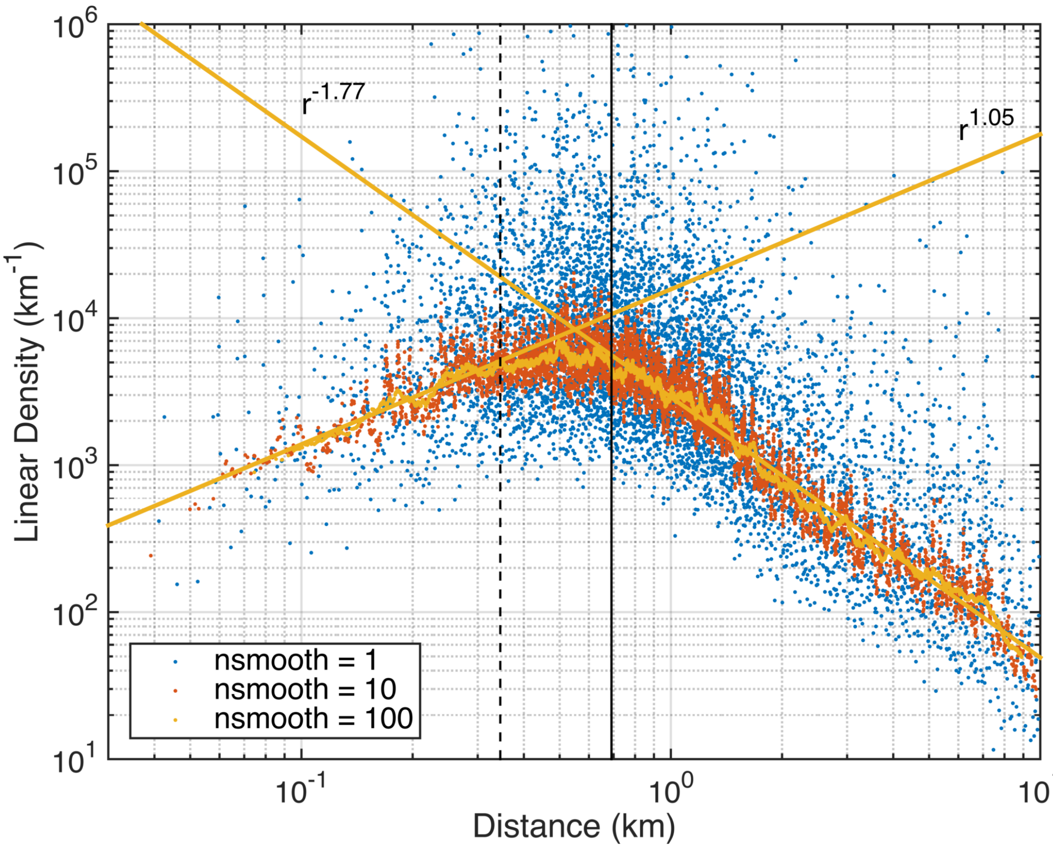


**Fig S6.** Least-squares fit to linear density for M4 – 4.5 mainshocks, using aftershocks with $\Delta M < 0$, up to 3.2 days after the mainshock. The decay is fit to points outside the estimated rupture length (solid black line), and the geometrical increase is fit to points inside ½ the rupture length (dashed black line). The legend gives the number of points over which the data are smoothed. The three smoothing windows yield spatial decays that agree within 0.01. The 95% confidence bounds for the two exponents are 1.05±0.02 and -1.77±0.007.

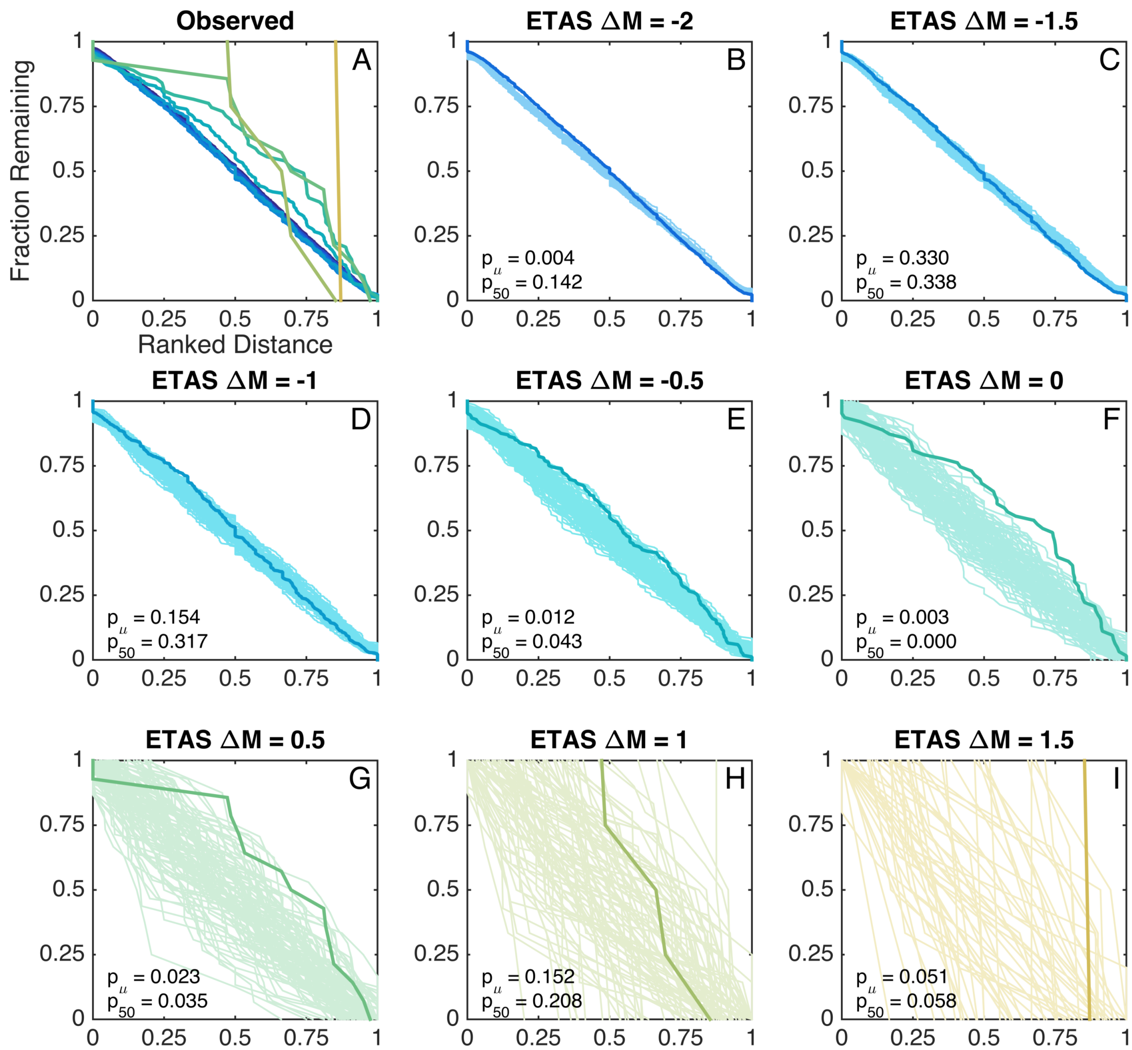


**Fig S7. (A)** Ranked distance as a function of the difference between aftershock and mainshock magnitude ΔM (Fig 5C, main text). **(B-I)** Distribution of ranked distance curves from 100 simulated ETAS catalogs in which distance depends only on mainshock magnitude, for varying ΔM. Thick lines are observed distributions from (A). Reported *p*-values are the fraction of 1000 ETAS simulations that produce mean ($p_\mu$) or median ($p_{50}$) ranked distances larger than observed.

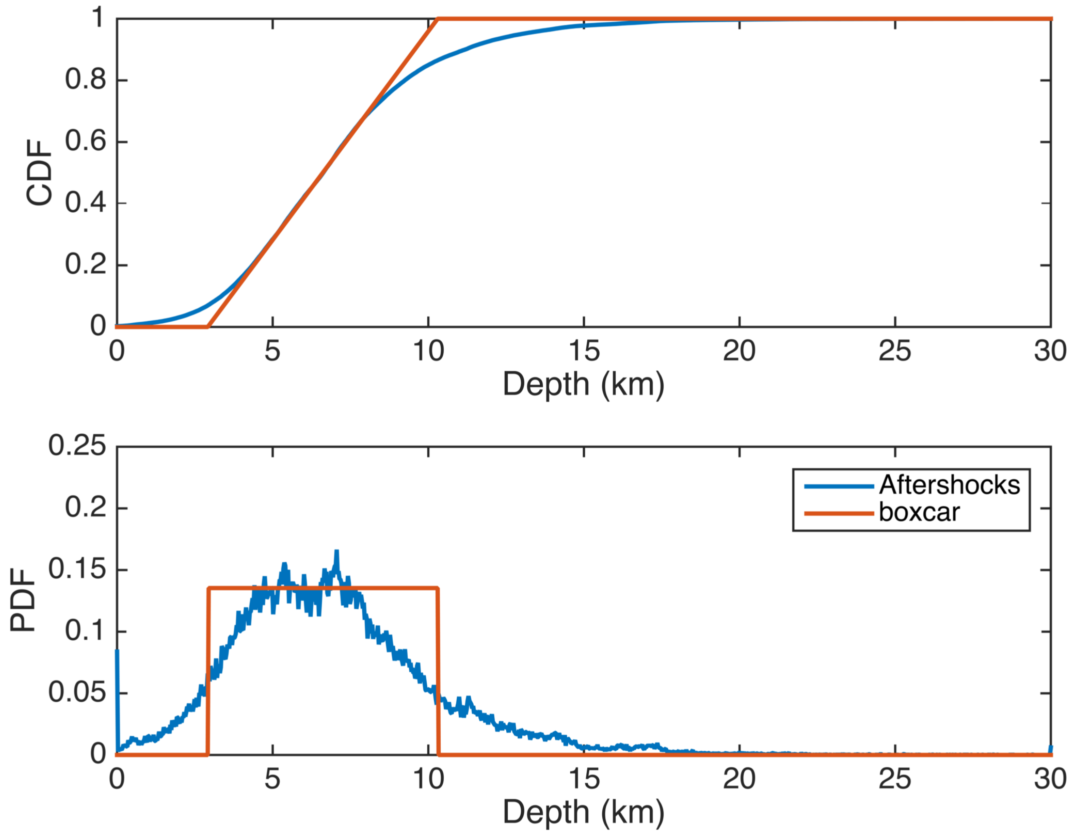


**Fig. S8.** Distribution of aftershock depths compared to the boxcar used to derive the geometrical correction $N(r)$ (Eq. 5). The width of the boxcar (7.39 km) was independently constrained by a maximum likelihood fit to the distance decay of aftershocks through Eq. 5, and is centered on the median of the data for comparison.

**Table S1.** Sequences with aftershocks larger than the mainshock

| Date | Time (UTC) | Mag | Lat. | Lon. | Depth (km) | Max ΔM | Seq. # | Cat. |
|---|---|---|---|---|---|---|---|---|
| **27-Sep-1982** | 18:21 | 4.2 | 35.75 | -117.75 | 7.2 | 1.0 | 19 | SC |
| **12-Jan-1984** | 3:11 | 4.0 | 37.41 | -118.53 | 14.5 | 0.5 | 1 | NC |
| **18-Jul-1986** | 14:29 | 5.9 | 37.56 | -118.44 | 5.5 | 0.5 | 40 | NC |
| **22-Nov-1987** | 1:53 | 6.2 | 33.08 | -115.78 | 10.1 | 0.4 | 59 | SC |
| **5-Apr-1990** | 2:39 | 4.4 | 37.86 | -121.99 | 4.8 | 0.0 | 94 | NC |
| **29-Aug-1990** | 1:06 | 4.1 | 33.24 | -116.05 | 9.9 | 0.4 | 100 | SC |
| **6-Sep-1990** | 12:48 | 4.1 | 36.68 | -121.31 | 6.1 | 0.3 | 105 | NC |
| **15-Aug-1995** | 22:39 | 5.4 | 35.77 | -117.66 | 9.1 | 0.4 | 180 | SC |
| **26-May-1998** | 20:31 | 4.3 | 36.81 | -121.54 | 7.8 | 0.8 | 215 | NC |
| **17-Feb-1999** | 3:08 | 4.2 | 32.60 | -116.16 | 6.0 | 0.0 | 222 | SC |
| **19-Oct-1999** | 1:53 | 5.0 | 34.86 | -116.40 | 5.3 | 0.0 | 232 | SC |
| **8-Jan-2000** | 21:41 | 4.3 | 38.76 | -122.92 | 5.6 | 0.0 | 234 | NC |
| **12-Jun-2000** | 19:00 | 4.1 | 32.89 | -115.50 | 8.2 | 0.4 | 242 | SC |
| **30-Jun-2001** | 17:34 | 4.1 | 36.70 | -121.33 | 4.8 | 0.0 | 253 | NC |
| **12-Jul-2001** | 17:31 | 4.2 | 36.03 | -117.87 | 7.5 | 1.0 | 255 | NC |
| **5-Dec-2001** | 14:29 | 4.2 | 39.05 | -123.12 | 5.2 | 0.1 | 259 | NC |
| **28-Aug-2004** | 4:30 | 4.0 | 36.58 | -121.18 | 7.4 | 0.4 | 284 | NC |
| **16-Sep-2004** | 7:07 | 4.1 | 38.01 | -118.67 | 3.8 | 1.5 | 285 | NC |
| **29-Aug-2005** | 22:48 | 4.6 | 33.16 | -115.62 | 4.6 | 0.5 | 302 | SC |
| **12-Sep-2006** | 0:11 | 4.1 | 32.71 | -116.04 | 7.8 | 0.6 | 314 | SC |
| **17-Jan-2008** | 17:18 | 4.5 | 40.16 | -122.76 | 10.3 | 0.2 | 330 | NC |
| **7-Feb-2008** | 7:12 | 5.1 | 32.41 | -115.31 | 18.7 | 0.0 | 332 | SC |
| **22-Apr-2008** | 22:55 | 4.3 | 39.53 | -119.93 | 1.6 | 0.8 | 333 | NC |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **27-Nov-2008** | 21:14 | 4.0 | 35.97 | -117.31 | 6.4 | 0.0 | 344 | SC |
| **12-Dec-2008** | 8:41 | 4.0 | 32.53 | -115.52 | 14.1 | 0.5 | 348 | SC |
| **29-Sep-2009** | 10:01 | 5.0 | 36.39 | -117.86 | 7.4 | 0.2 | 363 | SC |
| **7-Mar-2010** | 4:17 | 4.0 | 33.00 | -116.34 | 8.2 | 0.2 | 374 | SC |
| **11-Jun-2010** | 3:08 | 4.9 | 33.39 | -116.40 | 8.8 | 0.5 | 385 | SC |
| **7-Dec-2010** | 9:02 | 4.0 | 32.63 | -115.77 | 10.2 | 0.1 | 397 | SC |
| **7-Apr-2011** | 8:58 | 4.1 | 32.64 | -115.73 | 9.1 | 0.3 | 404 | SC |

**Table S2.** ETAS parameters

| Parameter | Value | Description |
|---|---|---|
| $\mu$ | 13.2508 $day^{-1}$ | Background rate |
| $c$ | 0.0055 days | Omori time constant |
| $p$ | 1.1 | Omori decay exponent |
| $k$ | 0.0085 | Aftershock productivity |
| $\alpha$ | 0.9 | Productivity-magnitude scaling |
| $M_c$ | 1.0 | Minimum simulated magnitude |
| $b$ | 1.0 | Gutenberg-Richter parameter |

## References

Akaike, H. (1973), lnformation theory and an extension of the maximum likelihood principle, in *Second International Symposium on In/ormation Theory*, edited by B. N. Petrov and F. Csak, pp. 267-281, Akademiai Kiado, Budapest.

Burnham, K. P., and D. R. Anderson (2002), *Model selection and multimodel inference: a practical information-theoretic approach*, Springer Science & Business Media, New York.

Helmstetter, A. S., Y. Y. Kagan, and D. D. Jackson (2005), Importance of small earthquakes for stress transfers and earthquake triggering, *Journal of Geophysical Research-Solid Earth*, *110*(B5), doi:10.1029/2004jb003286.

Helmstetter, A. S., and D. Sornette (2003), Importance of direct and indirect triggered seismicity in the ETAS model of seismicity, *Geophysical Research Letters*, *30*(11), doi:10.1029/2003gl017670.

Ogata, Y. (1998), Space-time point-process models for earthquake occurrences, *Ann. Inst. Stat. Math.*, *50*(2), 379-402.

Ogata, Y. (2011), Significant improvements of the space-time ETAS model for forecasting of accurate baseline seismicity, *Earth Planets and Space*, *63*(3), doi:10.5047/eps.2010.09.001.

Zhuang, J. C., Y. Ogata, and D. Vere-Jones (2002), Stochastic declustering of space-time earthquake occurrences, *Journal of the American Statistical Association*, *97*(458), 369-380, doi:10.1198/016214502760046925.