Title
The spatial density of foreshocks

Permalink
https://escholarship.org/uc/item/8dk263fz

Journal
Geophysical Research Letters, 38(10)

Author
Brodsky, Emily E.

Publication Date
2011-05-21

Peer reviewed
The spatial density of foreshocks
Emily E. Brodsky

Received 2 March 2011; revised 14 April 2011; accepted 14 April 2011; published 21 May 2011.

[1] Aftershocks follow a well-defined spatial decay pattern in the intermediate field. Here I investigate the same pattern for foreshocks. Foreshock linear density decays as $r^{-1.501}$ over distances $r$ of 0.1–30 km for 15 minutes before magnitude 3–4 mainshocks. This trend is the same as that of the aftershocks within the error of the measurement. This consistency of spatial decay can be explained by the clustering inherent in earthquake interactions. No additional preparatory process beyond earthquake triggering is necessary to explain the spatial decay. Citation: Brodsky, E. E. (2011), The spatial density of foreshocks, Geophys. Res. Lett., 38, L10305, doi:10.1029/2011GL047253.

1. Introduction

[2] Earthquake triggering is a possible route into studying earthquake initiation and, by extension, predictability. Earthquake interactions are generally dissected by looking at relationships between earthquakes that are close in either time or space. For instance, Felzer and Brodsky [2006] examined the spatial pattern of smaller earthquakes following main magnitude mainshocks within 5 minutes. The study found that the linear density $\rho$ of aftershocks followed a well-defined pattern of $\rho \propto r^{-1.5}$ where $r$ is the distance from the mainshock. The persistence of this spatial trend to large distances was interpreted as indicative of a consistent triggering process across the entire range of distances. Since seismic waves are thought to be significant triggers at great distances, the inference was that seismic waves are an important part of triggering at all distances [Hill and Prejean, 2007; van der Elst and Brodsky, 2010].

[3] Foreshocks are the best-known predictor of earthquakes. Early thinking attributed foreshocks to a stress accumulation process that ultimately culminated in a large earthquake [e.g., Jones and Molnar, 1979]. In this scenario, the spatial distribution of foreshocks is controlled by the stress distribution and fatigue on the fault plane. This line of thinking has most recently been supported by a connection between fluid flow and foreshocks [Lucente et al., 2010; Terakawa et al., 2010]. Other work has entertained alternative possibilities such as the foreshocks being symptomatic of an earthquake cascade. Earthquakes are constantly triggering each other. When seismicity is tightly clustered, the likelihood of a large earthquake increases and, after the fact, the seismicity cluster is interpreted as foreshocks [Helmstetter et al., 2003; Felzer et al., 2004]. This type of clustering results in distinct patterns, like the Inverse Omori’s Law [Helmstetter et al., 2003]. Still other work has suggested that both kinds of foreshocks exist in earthquake catalogs [McGuire et al., 2005].

[4] This paper starts by establishing that the foreshock spatial trend is similar to the aftershock one. At first, this observation is disconcerting. If the spatial decay of $r^{-1.5}$ is generated by aftershocks, then why is it present before the mainshock? I will follow up the observation by showing that a statistical seismicity model in which aftershocks follow the spatial decay law of $r^{-1.5}$ automatically generates foreshocks with the same spatial decay. The trend can be a natural consequence of the clustering provided by the mutual triggering of the earthquake cascade.

2. Observation

[5] I begin by using the well-located Lin-Shear-Hauksson (LSH) catalog of Southern California earthquakes 1981–2005 to compare the spatial distribution of aftershocks and foreshocks [Lin et al., 2007]. Only magnitudes >2 are included to ensure completeness. For the purpose of comparison, I use a similar windowing criteria as Felzer and Brodsky [2006] to isolate sequences. Mainshocks are defined as earthquakes that are at least 4 days after and 0.5 day before any larger earthquake at any distance. Aftershocks are defined to be earthquakes that follow a larger identified mainshock within the specified time period $\Delta t$. Foreshocks are earthquakes that are followed by a larger identified mainshock within the specified time period $\Delta t$. When comparing foreshocks and aftershocks for a particular sequence, the same value of $\Delta t$ is used to define both types of seismicity.

[6] These windowing criteria are not meant to imply a physical limit to interaction. The insensitivity of the aftershock spatial decay to the windowing details was established in previous work [Felzer and Brodsky, 2006]. The point of the present study is simply to compare foreshock decays to aftershock decays given the same conditioning on the data.

[7] To measure density, the mainshocks (and their accompanying sequences) for the entire catalog are combined to improve the statistical sampling. A single, ordered vector of aftershock-mainshock distances is created, $\vec{r}$. The linear density of aftershocks (and foreshocks) is estimated at the midpoints between each element of the combined aftershock distance vector $\vec{r}$ by

$$\rho(r_i + r_{i+1}) = \frac{1}{2} \frac{r_{i+1} - r_i}{r_{i+1} - r_i}$$

For mainshocks with magnitude <4, the fault rupture length is less than the spatial precision of the data (~0.1 km) and therefore the point approximation embedded in equation (1) is appropriate. Larger magnitude mainshocks require a better geometrical model for earthquake density estimation.

[8] The aftershocks within 15 minutes of magnitude 3–4 mainshocks yield a fit of $\gamma = 1.5 \pm 0.08$ for data from
differs slightly from for aftershocks and foreshocks of 100 time lines connect the measured density (equation (1)) for each earthquake generates its own aftershocks. Forequarters to facilitate comparison with the 2-D ETAS model below. For \( \Delta t = 5 \) minutes, a similar trend appears, but fewer quakes are recorded (Figure 1b). For close distances, the largest mainshocks in this range can no longer be treated as point sources and so the distance \( r \) is likely inaccurate. For large distances, the number of triggered events is sufficiently small that unrelated seismicity overwhelms the signal. The background level can be seen by shuffling the event times randomly for the same earthquake locations and re-measuring the density. The thin light lines in Figure 1 show 100 such reshirflings and the thick light lines show the mode values of the time-randomized set. As an unusually closely spaced pair of events can easily dominate the median or mean of the density, the mode is the best representative of the ordinary background level of seismicity. This background level demonstrates that indeed the background interferes with recording the after-shock signal at large distances, but an after-shock signal is visible up to at least 30 km distance from the mainshock.

Interestingly, the foreshocks follow almost the same trend (Figure 1). For 15 minutes before, the decay of the composite foreshock sequence from the future mainshock site follows \( r^{-1.5 \pm 0.1} \). For the 5 minutes before the mainshocks, the foreshock linear density is \( r^{-1.6 \pm 0.2} \). This observation is similar to that of Richards-Dinger et al. [2010]. The foreshock decay is also truncated by the background seismicity at large distances.

The specific cases of \( \Delta t = 5 \) and 15 minutes provide insight into the behavior at short times when unrelated background earthquakes are rare. The recovered value of \( \gamma \) for a full suite of values of \( \Delta t \) ranging from 2–40 minutes shows that these results are representative of short times when there are enough earthquakes for the measurement to be made (Figure 2). At extremely short times (<2 minutes), too few events are cataloged to measure \( \gamma \). For \( \Delta t \) up to 15 minutes, the foreshock and aftershock values are indistinguishable. Beyond 15 minutes, the foreshock sequences have a slightly smaller value of \( \gamma \).

3. Clustering Model

In order to recreate the observed foreshock trend, I will employ an Epidemic Triggering Aftershock Sequence (ETAS) model that combines empirical statistical laws of seismicity to generate self-consistent synthetic earthquake catalogs [Ogata, 1999]. An earthquake triggers successive earthquakes with a probability determined by the magnitude of the mainshock. Omori’s Law and a spatial decay relationship determine the time and distance separating the aftershock from the mainshock. Successive earthquake magnitudes are then determined by Gutenberg-Richter statistics and then each earthquake generates its own aftershocks.

The ETAS formulation has been explored extensively in the literature [Helmstetter et al., 2003, 2005; Felzer et al., 2004; McGuire et al., 2005; Zhuang et al., 2008]. Windowing and selection criteria often result in behaviors that are challenging to evaluate analytically. I therefore use a numerical model directly for this work. The parameters used here are meant to illustrate the relationship between the foreshock and aftershock spatial decay. They are certainly not exhaustive,
but are realistic and suffice to elucidate the role of clustering in generating a foreshock spatial pattern. Published versions of the basic ETAS model utilize variants of the observational laws. As the permutations of the model are important and the practicalities of implementation are sometimes difficult, I will go into some detail on the exact forms used in the numerical calculations done here.

3.1. ETAS Implementation

[14] Background seismicity is imposed as a Poissonian process with rate $\lambda$ over the duration of the observed catalog (24 years) and the magnitude of each earthquake determined by Gutenberg–Richter. The background rate is chosen so that the total number of synthesized events (including aftershocks) is approximately equal to the observed number. Each background earthquake generates aftershocks, which in turn spawn their own sequences through four empirical equations cast in terms of probability distributions.

3.1.1. Gutenberg–Richter

[15] The number of aftershocks with magnitudes greater than or equal to $M$ is

$$\log N(M) = -bM + a \quad (2)$$

where $a$ and $b$ are constants. To obtain a cumulative density function (CDF), $N(M)$ is normalized by the total number of earthquakes in the catalog, i.e., $N(M_{\text{min}})$ where $M_{\text{min}}$ the minimum magnitude of the simulated catalog. A random number $p_1$ is then selected between 0 and 1 for each aftershock and inverted using the CDF into a magnitude, i.e.,

$$M = -\log(p_1(10^{-bM_{\text{min}}} - 10^{-bM_{\text{max}}})) / b \quad (3)$$

where $p_1$ is a random variable between 0 and 1 and $M_{\text{max}}$ is the maximum magnitude size to be generated. The extra term in brackets is a normalization that accounts for discarding values of $p_1$ that result in $M > M_{\text{max}}$ [Sornette and Werner, 2005].

3.1.2. Omori’s Law

[16] The rate of seismicity as a function of time $t$ from a mainshock is

$$dN_{AS}/dt = K/(t + C_{\text{omori}})^{\beta} \quad (4)$$

where $\beta$ and $C_{\text{omori}}$ are constants and $K$ is a function of mainshock magnitude to be discussed below. The constant $C_{\text{omori}}$ is difficult to measure as it is often biased by completeness problems at short times after an earthquake. Recent work has suggested that $C_{\text{omori}}$ is significantly shorter than previously imagined [Peng et al., 2007]. In probabilistic terms,

$$t = \left[ (t_{\text{max}} + C_{\text{omori}})^{1-\beta} p_2 + (1 - p_2) C_{\text{omori}} (1 - \beta) \right]^{1/(1-\beta)} - C_{\text{omori}} \quad (5)$$

where $p_2$ is a random variable between 0 and 1.

3.1.3. Aftershock Productivity

[17] The numerator of Omori’s Law determines the number of aftershocks from a mainshock of magnitude $M$

$$K = C10^{\phi(M - M_{\text{min}})} \quad (6)$$

where $C$ and $\alpha$ are constants. Studies agree that $\alpha = 1$ for Southern California seismicity [Felzer et al., 2004; Helmstetter et al., 2005]. The parameter $C$ is defined for the instantaneous rate in Omori’s Law. In order to use equation (6), the total number of aftershocks must be computed for a sequence and so in practice what is required is the integral of equation (4) over the finite time, $t_{\text{max}}$ that is the duration of the simulated direct aftershock sequence. The practical equation is therefore,

$$N_{AS} = C' 10^{\phi(M - M_{\text{min}})}$$

where

$$C' = C \left[ (t_{\text{max}} + C_{\text{omori}})^{1-\beta} - C_{\text{omori}} (1-1/\beta) \right] / (1 - \beta) \quad (7)$$

[18] Equation (7) is implemented deterministically based on aftershock magnitudes. For non-integer values of $N_{AS}$, the fractional part is interpreted as a probability. For each sequence, a uniform random variable $p$ is generated between 0 and 1 and if the value is less than the non-integer fraction, an extra aftershock is added. Table 1 reports a value of $C'$ that fits the observed productivity in the LSH data set with the aftershock identification criteria used here (Figure S1 of the auxiliary material).\footnote{Auxiliary materials are available in the HTML. doi:10.1029/2011GL047253.}

3.1.4. Distance Decay

[19] The distance fall-off is implemented by normalizing the distribution to form a CDF assuming that the aftershocks are distributed over distances greater than $d_{\text{min}}$ from the
mainshock. Since $g > 1$, no upper bound to the distance fall-off is necessary. The probabilistic form for $r$ for each aftershock is

$$r = d_{\text{min}} (1 - p_3)^{1/(1-g)}$$  \hspace{1cm} (8)$$

where $p_3$ is a random variable between 0 and 1. As there are no observational constraints on the angular distribution, $\theta$ is selected with uniform probability from 0 to $2\pi$.

[20] In the implementation of ETAS, it is helpful to note the order of calculation. First the total number of aftershocks must be calculated for each extent earthquake using equation (7), then each aftershock is assigned a magnitude, time and location using equations (3), (5), and (8). Then the aftershocks generate their own aftershocks by the same procedure until the cascade comes to an end.

### 3.2. Simulation Results

[21] The ETAS catalog was made to mimic the observed catalog using the parameters of Table 1. The procedures used to isolate mainshocks and identify 15 and 5-minute aftershock and foreshocks are identical to those used on the observational data in Figure 1.

The overall trends of the Figure 1 are reproduced in the simulated catalog (Figure 3). Both the aftershock and foreshock densities have a best-fit spatial decay consistent with the imposed exponent. This consistency in foreshock and aftershock behavior exists even though there was no specific foreshock preparatory process in the ETAS model. The utility of the statistical simulation is that it shows that non-intuitive behavior, like that of Figure 1, emerges naturally from the earthquake sequences. Like the Inverse Omori’s Law, the spatial decay of the foreshocks is symptomatic of the type of increased seismicity that is likely to trigger a large earthquake.

[22] The behavior of the best-fit values of $\gamma$ over a range of time windows reproduces the major trends of the observational data (Figure 4). Like in the real catalog, $\gamma$ is unmeasurable at extremely short times ($\Delta t < C_{\text{Omori}}$). The decay of the seismicity becomes more gradual (best-fit $\gamma$ decreases) for both aftershock and foreshock sequences as time increases due to the random walk of secondary sequences [Helmstetter et al., 2003] combined with interference from unrelated background events. (Superposition of sequences with offset origins result in a composite fit with reduced values of $\gamma$). In the simulations, both the decrease in $\gamma$ and the variability of realizations is more pronounced for the foreshock sequences because of the relatively small number of events. Multiple simulations are required to identify the expected behavior. The mean of the best-fit values of $\gamma$ for foreshocks over 50 simulations differs by more than 1 standard deviation from that of the aftershocks for time windows $\geq 14.5$ minutes (Figure 4b). These simulation results are consistent with the real data showing significant departures between the aftershock and foreshock decay for $\Delta t > 15$ minutes (Figure 2).

[23] At very large distances, the simulated catalog has clusters due to unisolated ongoing sequences. The isolation

**Figure 4.** The best-fit value of $\gamma$ for simulated aftershocks and foreshocks for a range of time windows. ETAS parameters are as in Figure 1. (a) Best-fit values from a single simulation with error ranges on fit shown by the shaded region as in Figure 2. (b) Mean of the best-fit values of $\gamma$ for 50 simulations with the standard deviations over the suite of simulations shown by the error bars. As in Figure 3, mainshocks are magnitude 3–4 and the linear density decay is fit over 0.1–30 km. Vertical dashed line shows the value of $C_{\text{Omori}}$ in the simulations.
distance of 100 km is insufficient for the simulation because of the more homogeneous distribution of epicenters, i.e., there are no simulated faults.

[25] One conspicuous difference between the simulations and the real catalog is that the ratio of identified aftershocks to foreshocks is much higher in the simulation than the observation. In the actual catalog, the ratio is 2.2 for $\Delta t = 5$ min and 2.8 for $\Delta t = 15$ min, while the simulations result in ratios between 7 and 8 for the parameters used here. Aftershock/foreshock ratios have been analyzed extensively elsewhere [e.g., McGuire et al., 2005] and are not the focus of this study. Nonetheless, some assessment of the source of the discrepancy is in order.

[26] One possibility is that the observed catalog artificially depresses the aftershock/foreshock ratio because of incompleteness. It is more difficult to detect aftershocks than foreshocks because the mainshock and temporally high seismicity rate can obscure individual events. Therefore, the magnitude of catalog completeness for aftershocks can be significantly higher than normal [e.g., Peng et al., 2007].

[27] To initially evaluate the role of catalog incompleteness, I restrict the data to magnitudes $>2.5$ (Figure S2). The observed aftershock to foreshock ratio for $\Delta t = 15$ min increases slightly to 2.9. The same restriction on the simulated data decreases the aftershock/foreshock ratio because the ratio depends on the identified mainshock magnitude range [McGuire et al., 2005]. In a suite of 100 ETAS simulations with the values in Table 1 and $M_{\text{th}} = 2.5$, for $\Delta t = 15$ min the mean ratio is 4 with a standard deviation of 0.8. This improved consistency between simulation and observation is expected based on previous work on foreshock rates in the region [Felzer et al., 2004] and indicates that the lack of aftershock detection is an important issue in the observations. However, the simulated ratios still do not exactly match the observations. It is possible that the completeness threshold needs to be raised still further to ensure aftershock detection, but further restriction limits the dataset to an unreasonably small size.

[28] As an additional probe of time-variable completeness, I exclude a short window immediately before and following the mainshock in measuring the aftershock and foreshock densities (Figure S3). This strategy has the advantage of preserving the number of events at larger times when the completeness threshold drops to the overall catalog value. Excluding a window of 1 minute on either side of the mainshock increases the aftershock/foreshock ratio to 3.3 for $\Delta t = 15$ min. This ratio is still smaller than the simulated ratio, but it does again indicate that early-time aftershock detection is likely a problem in the observations.

[29] The aftershock-rate/foreshock-rate ratio can be adjusted in the simulations by reducing the minimum magnitude of simulation $M_{\text{min}}$ or increasing the productivity constant $C$, both of which have the effect of increasing the average number $n$ of aftershocks per mainshock [Sornette and Werner, 2005]. As $n$ increases, the system approaches a critical state and the foreshock rate should increase relative to the aftershock rate [McGuire et al., 2005]. The calculations also become prohibitively expensive. Thus far, the observed foreshock to aftershock ratio has not been simulated. One possible explanation is that there is a genuine difference in the ETAS process from the observations. If so, the aftershock/foreshock ratio is the only evidence of such a process. Another explanation is that aftershock completeness is still the overwhelming observational problem.

[30] More importantly for this study, the spatial decay of the foreshock density in both restricted datasets remains consistent with the aftershock decay. This consistency of Figure S3 is particularly useful in light of the commentary of Richards-Dinger et al. [2010] that the observed aftershock density decay may be controlled by events that occurred before the arrival of the seismic waves. No such events are included in the fits of Figure S3 as P waves travelling at 6 km/s pass through 360 km in the first minute.

4. Conclusions

[31] Earthquakes are preceded by a halo of earthquakes that decays with distance much like aftershocks. These foreshocks are an expected consequence of earthquake interaction and clustering. Their existence and location near the eventual rupture does not in itself indicate a preparatory strain accumulation process beyond ordinary earthquake triggering.

[32] A foreshock preparatory process could exist in the Earth beyond the clustering process, but it is not required by the observations. The only potential evidence seen here for such a preparatory process is in the ratio of total number of aftershocks to foreshocks. However, the data suggest that catalog completeness is a serious problem that could artificially depress the aftershock/foreshock ratio and therefore this ratio is not the most reliable indicator of the underlying physics. The spatial decay of both the aftershocks and foreshocks is well explained by the clustering model.

[33] Acknowledgments. Many thanks to Karen Felzer, Agnes Helmstetter, Keith Richards-Dinger, Peter Shearer, and Ross Stein for helpful conversations and ideas. This work was funded in part by the Southern California Earthquake Center. SCEC is funded by NSF Cooperative Agreement EAR–0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for this paper is 1481.

[34] The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References


Lin, G., P. M. Shearer, and E. Hauksson (2007), Applying a three-dimensional velocity model, waveform cross correlation, and cluster


E. E. Brodsky, Department of Earth and Planetary Sciences, University of California, Santa Cruz, CA 95060, USA. (brodsky@es.ucsc.edu)