**Supporting Information for**

**"Foreshocks and Mainshock Nucleation of the 1999 $M_w$ 7.1 Hector Mine, California, Earthquake"**

**Clara E. Yoon[1], Nana Yoshimitsu[1,2], William L. Ellsworth[1], and Gregory C. Beroza[1]**

[1]Department of Geophysics, Stanford University, Stanford, California, USA.

[2]Earthquake Research Institute, The University of Tokyo, Tokyo, Japan.

**Contents**

1. Sections S1 to S4

2. Figures S1 to S10

3. Tables S1 to S9

**Additional Supporting Information (Files uploaded separately)**

1. Captions for Datasets S1 to S3

Corresponding author: Clara E. Yoon, `claraeyoon@gmail.com`

**Section S1: Magnitude estimation for 32 uncataloged foreshocks**

We estimate magnitudes for the 32 newly detected foreshocks following the procedure in *Shelly et al.* [2016], by measuring relative amplitudes between new events (with unknown magnitude) and catalog events (with known local magnitudes from the SCSN catalog). The new event magnitude $M_{new}$ is computed from the catalog event magnitude $M_{cat}$ as:

$$M_{new} = M_{cat} + c \log_{10}(\alpha), \tag{S1}$$

where $\alpha$ is the ratio of the newly detected and catalog event amplitudes, and $c$ is a calibration constant describing the amplitude-magnitude scaling. The relative amplitude $\alpha$ is calculated with a principal component fit:

$$\alpha = \frac{v(2)}{v(1)} \tag{S2}$$

where $v(1)$ and $v(2)$ are elements of the singular vector $\mathbf{v}$ for the largest singular value of the covariance matrix $cov(\mathbf{x}, \mathbf{y})$ of the catalog waveform vector $\mathbf{x}$ and the newly detected waveform vector $\mathbf{y}$ [*Shelly et al.*, 2016]. The waveforms $\mathbf{x}$ and $\mathbf{y}$ need to be aligned to their maximum correlation before computing their covariance. In our magnitude estimate, we use $P$ and $S$ waveforms with the same time windows, filter bands, and channel selection used in the cross-correlation differential time calculations (Section S2), at the same 14 stations used for relocation (Table S5); however, the correlation coefficient (CC) threshold was lowered to 0.6 for magnitude calculation. This magnitude estimation method has lower bias compared to a typical least-squares relative magnitude measurement for small events [*Shelly et al.*, 2016].

First, we use the 18 catalog foreshocks to estimate the calibration constant $c$. For each unique pair of catalog events with magnitudes $M_{cat1}$ and $M_{cat2}$ respectively, we estimate their relative amplitude $\alpha_{(cat2,cat1)}$. We first use Equation S2 to calculate a separate estimate of $\alpha$ for the $P$ and $S$ waveform at each station, where $\mathbf{x}$ is the waveform for event $cat1$ and $\mathbf{y}$ is the waveform for event $cat2$, then take their median value as $\alpha_{(cat2,cat1)}$. For the calibration, we can write Equation S2 as:

$$M_{cat2} = M_{cat1} + c \log_{10}(\alpha_{(cat2,cat1)}), \tag{S3}$$

and we estimate $c = 1.085$ with a linear least-squares fit to the catalog event-pair data (Figure S6). This is close to the $c = 1$ value expected for local magnitudes calculated from the log of the peak amplitude on a Wood-Anderson seismometer.

For each new event, we estimate a separate magnitude $M_{new}$ using a different catalog event $M_{cat}$ in Equation S1 with $c = 1.085$. The relative amplitude $\alpha_{(new,cat)}$ is again separately calculated for the $P$ and $S$ waveform at every station, and the median value is selected for the $M_{new}$ calculation.

The final magnitude for the new event is then the median of all 18 $M_{new}$ estimates, one from every catalog event.

**Section S2: Foreshock and mainshock relocation details**

We relocate the 42 largest foreshocks and the mainshock (Data Set S1 and S2) using the double-difference location (hypoDD) algorithm [*Waldhauser and Ellsworth*, 2000]. These 42 foreshocks were also detected by *Zanzerkia et al.* [2003]. We do not relocate the 8 smallest foreshocks (Data Set S3), which are detected reliably on only 2 stations, since they can make the inversion procedure unstable. We only use differential travel times calculated from cross-correlation. We do not include catalog differential times, given the significant uncertainty in the absolute locations from different alternative catalogs (Figure S8a-d); poor quality starting locations, and the lack of nearby (< 5 km) stations, would result in an unstable solution.

To determine the *P* and *S* arrivals at each foreshock, we manually picked *P* and *S* arrival times at all 14 stations used for location (Table S5) for one catalog event: SCSN ID 9108545, with origin time 1999-10-15 16:35:54.16 UTC (Figure S1, red rectangle). For each *P* and *S* pick at the other 13 stations for this event, we subtracted the *P* arrival time at station RMM, in order to get a Δt time difference for each pick at each station relative to the *P* arrival time at RMM. Then, for the rest of the foreshocks, we manually picked the *P* arrival time at RMM, and used the Δt time differences to predict *P* and *S* phase arrival times at the other 13 stations. If the *P* phase at RMM was too noisy to allow us to place a manual pick, we used the *S* phase at station HEC. Since the foreshock waveforms are highly similar (Figures S1, S2, S5), we assume that they have about the same travel time to each station, and let hypoDD sort out the exact travel time differences.

To obtain cross-correlation differential times for hypoDD, we cross-correlate short windows around these predicted *P* and *S* arrivals at each station for every foreshock pair, keeping only differential times with CC ≥ 0.8. We use a 1.28 s window centered around the *P* or *S* arrival (0.64 s before and 0.64 s after), consistent with *Zanzerkia et al.* [2003]. For the *P* arrival, we only use vertical component data. For the *S* arrival, we cross-correlate on both horizontal components and keep the differential time from the component that results in a higher CC, if the station has 3 components; otherwise, we use the only available vertical component. These time windows are tapered and filtered according to the station-specific band in Table S5 before cross-correlation in the time domain. We can get subsample precision of the differential time by fitting a parabola near the

peak of the cross-correlation function and interpolating the time at the peak [e.g. *Deichmann et al.*, 1992; *Schaff et al.*, 2004].

It is important to precisely locate the mainshock relative to the foreshocks, in order to determine whether the mainshock initiated at a location of increased stress, or decreased stress, from the foreshock ruptures. To do this, we calculate cross-correlation differential times between the *P* arrival at each of the 18 catalog foreshocks, and the initial onset rise (first 0.08 seconds) of the mainshock *P* arrival (Figure S7), at each of the 11 stations sampled at 100 Hz (Table S5), keeping only differential times with CC ≥ 0.9. *P* arrival times for the mainshock are obtained from the SCSN. Compared to the foreshock-foreshock cross-correlations, here we use a shorter 0.3 s window centered around the *P* arrival (0.22 s before and 0.08 s after). We oversample each time window by a factor of 5, to a 500 Hz sampling rate, before performing the cross-correlation on these short time windows, in order to get precise subsample differential time estimates. Since the predicted arrival time for the foreshock may differ by more than 0.08 s relative to the mainshock, we shift the foreshock window between -0.4 and 0.4 s with time steps of 0.01 s, cross-correlate with the mainshock time window at every time step, and select the time shift with the highest CC.

We assume the same starting location for all 42 foreshocks and the mainshock for input to hypoDD. For the initial epicenter, we use the mean latitude (34.592°) and longitude (-116.271°) from the 18 catalog foreshocks. The absolute depths of the foreshocks and mainshock are not well constrained, because the distance to the nearest station HEC is 26 km, which is greater than the expected depths of these events. Therefore, we test 5 different trial depths, all equally plausible, for the starting location used in hypoDD: 3 km, 5 km, 8 km, 10 km, 15 km, which are representative of the range of depths for these earthquakes reported in the literature (Figure S9). For each initial trial depth, we also experimented with shifting the starting depth of the mainshock 1 km lower and 1 km higher than the rest of the foreshocks before running hypoDD, and confirmed that the mainshock converged to the same final location each time.

Existing catalogs and previous studies of these foreshocks indicate significant location and depth uncertainties. The SCSN catalog reports 14.9 km for the mainshock depth and 0 - 6.8 km for the 18 foreshock depths (Figure S8a). *Ji et al.* [2002] found that a mainshock hypocenter depth of 14.8 km provided the best fit to a finite fault inversion for the mainshock slip distribution constrained by strong motion data, GPS measurements, and surface offsets. *Zanzerkia et al.* [2003] report 4-5 km depth for most foreshocks, while *Hauksson et al.* [2002] lists both 2.0 km and 5 ± 4 km for the mainshock depth and 2-3 km for the foreshock depths. The *Hauksson et al.* [2012] catalog, where southern

California earthquakes between 1981 and 2011 were relocated with differential travel times obtained by waveform cross-correlation, reports 7.7 km for the mainshock depth and 2.1 - 7.6 km for the 18 foreshock depths (Figure S8b). The *Yang et al.* [2012] "YHS" focal mechanism catalog reports 1.2 km for the mainshock depth and 2.6 - 4.4 km for the depths of the 7 largest foreshocks (Figure S8c). *Chen and Shearer* [2013] relocated the foreshocks using differential times from cross-correlation, and used a STA/LTA (short-term-average/long-term-average) picker to compute differential times to relocate the mainshock relative to the foreshocks; they report 1.6 km for the mainshock depth and 1.0 - 1.6 km for the 18 foreshock depths (Figure S8d).

To locate the 42 largest foreshocks and the mainshock, we ran hypoDD in LSQR mode with parameters from Table S6 and weights from Table S7 with 24 iterations; damping was necessary for convergence to stable final locations. To estimate relative location uncertainties (error bars in Figures 3, 5, S10), we ran hypoDD in singular value decomposition (SVD) mode with 20 iterations (Table S7) for a subset of 31 (out of the total 42) located foreshocks, including all 18 catalog foreshocks, and the mainshock. We did not estimate uncertainties for 10 foreshocks (marked in Data Set S2), because including them when running hypoDD in SVD mode resulted in an unstable solution with singular values. Location uncertainties are within the range of 10-120 m horizontally and 40-700 m vertically for most events.

We also relocated the 42 largest foreshocks and mainshock with an alternative method called GrowClust [*Trugman and Shearer*, 2017], which uses the same differential travel times obtained from waveform cross-correlation that were input to hypoDD, but instead minimizes the L1-norm of the differential travel time residuals (rather than the L2-norm as done in hypoDD), and groups earthquakes with similar waveforms into clusters (rather than matrix inversion as done in hypoDD). We assume the same starting location (34.592°, -116.271°) for all 42 foreshocks and the mainshock, and a starting depth of 8 km, running GrowClust with parameters *delmax* = 200 (maximum station distance between an event pair in km) and *rmsmax* = 0.02 (maximum root-mean-square differential time residual for merging clusters). The resulting GrowClust locations (Figure S8e), are slightly more compact, but still closely resemble the hypoDD locations (Figure S8f, Figure 3a).

**Section S3: Source parameter calculation details**

The source parameter estimates of the foreshocks are limited by the lack of usable data in 1999. Event waveforms for the largest 8 foreshocks were clipped at the nearer short-period, vertical-component stations used for detection (Table S1, all stations except for HEC), so these

stations were unsuitable for the spectral ratio method. We also tried using smaller foreshocks as eGf events, but the resulting spectral ratios were too noisy to estimate corner frequencies reliably. Therefore, we were unable to average over a large number of spectral ratio estimates from different stations, with different combinations of master and eGf events. Only the four 3-component SCSN broadband stations in Table S9 (Figure 1, red triangles) had on-scale recordings of waveforms at a high enough sampling rate of 100 Hz that could yield reliable spectral ratio estimates for the largest 8 foreshocks. A 20 Hz sampling rate does not provide enough bandwidth to estimate both corner frequencies.

For both the master and eGf events, we cut a time window of length 10.24 s, starting 1.28 seconds before the *S* arrival, and applied a 0.5 - 40 Hz bandpass filter. We also experimented with using *P* arrival and coda time windows, but the resulting spectral ratios were too noisy. After applying a taper to the time window, we took the multi-window approach [*Imanishi and Ellsworth*, 2006], dividing the 10.24 s time window into 3 sub-windows, each 5.12 s long, with 50% overlap (2.56 s lag between adjacent sub-windows). We calculate the spectrum for each sub-window by taking the magnitude of its Fourier transform, then compute their average spectrum, and smooth it with a running average for every five samples to improve its stability [*Huang et al.*, 2016]. We also compute the frequency spectrum of noise from the same channel: we cut a 10.24 s time window containing only noise (before the *P* arrival), and follow the same procedure we did for the *S* window (bandpass filter, taper, multi-window spectrum, smooth with running average). We then divide the spectrum of the master event by the spectrum of the eGf event to get the spectral ratio, but we discard the spectral ratio measurement at this component as too noisy if the spectral signal-to-noise ratio (snr) is lower than 5 at any frequency between 0.8 and 10 Hz. We compute the logarithmic mean of the spectral ratio from each component at the same station. We then compute a logarithmic average of the spectral ratios from all stations, and resample this average spectral ratio at equally spaced log(frequency) intervals between 0.5 and 30 Hz. At this point, we can fit the average spectral ratio to the Boatwright spectral model (Equation 1), using a grid search to estimate values for the three unknown parameters (master event corner frequency $f_{c1}$, eGf event corner frequency $f_{c2}$, and seismic moment ratio $M_{01}/M_{02}$) that minimize the residual between data and model. We also tried fitting the average spectral ratio to the Brune spectral model [*Brune*, 1970], with $\gamma = 1$ in Equation 1; however, the Boatwright spectral model, with a sharper corner shape in its spectrum, yielded slightly lower residuals.

We used the largest foreshock (SCSN id 9108606, *M* 3.7) as the master event, and each of the other 7 foreshocks as an eGf event, to compute the spectral ratio and estimate corner frequencies (Figure 4) by fitting the Boatwright spectral model (Equation 1). As a result, we had 7 separate

estimates of the lower corner frequency $f_{c1}$, so we used their median value as the corner frequency

for the $M$ 3.7 largest foreshock. For the 7 other foreshocks, we used the single estimate of its corner

frequency, from the higher corner frequency $f_{c2}$. Some of the spectral ratio estimates were averaged

over only 1 or 2 stations (Figure 4), since they were too noisy at the other stations. We also computed

spectral ratio estimates between all possible pairs of the largest 8 foreshocks, where the master (larger

magnitude) and eGf (smaller magnitude) were at least 0.5 magnitude units apart, but they were all

discarded because they were below our snr threshold. Column 3 in Table 1 lists corner frequency

estimates for the 8 largest foreshocks.

Figure 4 (lower right) compares catalog magnitudes for each eGf (red) with their $M_w$ magnitudes

calculated from the moment ratio estimates (blue), which overestimate the catalog magnitudes by

0.1 to 0.4 magnitude units. Here we assumed $M_w = 3.7$ for the master event (largest foreshock) from

the catalog, and calculated its seismic moment $M_{01}$:

$$M_{01} = 10^{\frac{3}{2}(M_w + 6.07)}, \tag{S4}$$

then for the other 7 foreshocks (eGf events), used the moment ratio $M_r = M_{01}/M_{02}$ estimates (Table 1,

column 4) to calculate the seismic moment $M_{02}$ (Table 1, column 5):

$$M_{02} = \frac{M_{01}}{M_r} \tag{S5}$$

then obtained $M_w$ for each eGf (Table 1, column 6):

$$M_w = \frac{2}{3} \log_{10} M_0 - 6.07 \tag{S6}$$

$M_{01}$ (for the largest foreshock) and $M_{02}$ (for the other 7 foreshocks), displayed in column 5 of Table 1,

are used to calculate the average slip $D$ (Equation 3).

**Section S4: Source parameter uncertainty estimates**

We fit the average spectral ratio from the data (Figure 4, thick gray line) to the Boatwright

spectral model (Equation 1; Figure 4, thick red line), using a grid search to estimate values for the

three unknown parameters (master event corner frequency $f_{c1}$, eGf event corner frequency $f_{c2}$, and

seismic moment ratio $M_r = M_{01}/M_{02}$) that minimize the residual between data and model.

We first performed a coarse-grained grid search to narrow down the range of values for $f_{c1}$,

$f_{c2}$, and $M_r$. The parameter search space for both $f_{c1}$ and $f_{c2}$ was 60 equally spaced grid cells on

a logarithmic scale between the interval $10^{-1} = 0.1$ Hz and $10^{1.6} \approx 40$ Hz. The parameter search

space for $M_r$ was 30 equally spaced grid cells on a logarithmic scale between the interval $10^0 = 1$

and $10^2 = 100$.

We then performed a fine-grained grid search to get more precise estimates of $f_{c1}$, $f_{c2}$, and $M_r$. The parameter search space for both $f_{c1}$ and $f_{c2}$ was 90 equally spaced grid cells on a logarithmic scale between the interval $10^{-0.3} \approx 0.5$ Hz and $10^{1.3} \approx 20$ Hz. The parameter search space for $M_r$ was 45 equally spaced grid cells on a logarithmic scale between the interval $10^{0.7} \approx 5$ and $10^2 = 100$.

We report uncertainties for $f_c$ and $M_r$ from the fine-grained grid search. We report the optimum values of $f_c$ (Table 1, column 3) and $M_r$ (Table 1, column 4) that minimize the residual between data and model for each spectral ratio. From the fine-grained grid search, we report $\delta f_c$ (the uncertainty in $f_c$, Table 1, column 3) as the grid cell spacing between the optimum value of $f_c$ and the adjacent $f_c$ in the next higher grid cell, and $\delta M_r$ (the uncertainty in $M_r$, Table 1, column 4) as the grid cell spacing between the optimum value of $M_r$ and the adjacent $M_r$ in the next higher grid cell.

We propagate the uncertainties in the estimated quantities from the fine-grained grid search ($\delta f_c$ and $\delta M_r$) to the remaining calculated source parameters using this general error propagation function, where $q$ is a function of several variables $x$, ..., $z$ [e.g. *Taylor*, 1997]:

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x}\delta x\right)^2 + \ldots + \left(\frac{\partial q}{\partial z}\delta z\right)^2} \tag{S7}$$

The uncertainty in the seismic moment $M_{02}$ for the 7 eGf foreshocks, in column 5 of Table 1, is (substituting Equation S5 into the partial derivative):

$$\delta M_{02} = \frac{\partial M_{02}}{\partial M_r}\delta M_r = \frac{\partial}{\partial M_r}\left(\frac{M_{01}}{M_r}\right)\delta M_r = \frac{M_{01}}{M_r^2}\delta M_r \tag{S8}$$

where $\delta M_r$ is the uncertainty in the moment ratio $M_r$ from the fine-grained grid search.

The uncertainty in the magnitude estimates $M_w$ for the 7 eGf foreshocks as derived from the seismic moment $M_{02}$, in column 6 of Table 1, is (substituting Equation S6 into the partial derivative):

$$\delta M_w = \frac{\partial M_w}{\partial M_{02}}\delta M_{02} = \frac{\partial}{\partial M_{02}}\left(\frac{2}{3}\log_{10} M_0 - 6.07\right)\delta M_{02}$$
$$\Rightarrow \delta M_w = \frac{2}{3\ln 10}\frac{\delta M_{02}}{M_{02}} \tag{S9}$$

The uncertainty in the source radius $R$, in column 7 of Table 1, is (substituting Equation 2 into the partial derivatives):

$$\delta R = \sqrt{\left(\frac{\partial R}{\partial v_s}\delta v_s\right)^2 + \left(\frac{\partial R}{\partial f_c}\delta f_c\right)^2} = \sqrt{\left(\frac{\partial}{\partial v_s}\left(\frac{k v_s}{f_c}\right)\delta v_s\right)^2 + \left(\frac{\partial}{\partial f_c}\left(\frac{k v_s}{f_c}\right)\delta f_c\right)^2}$$
$$\Rightarrow \delta R = \sqrt{\left(\frac{k}{f_c}\delta v_s\right)^2 + \left(\frac{k v_s}{f_c^2}\delta f_c\right)^2} \tag{S10}$$

where $\delta f_c$ is the uncertainty in the corner frequency $f_c$ from the fine-grained grid search. We use $v_s = (3.54 \pm 0.25)$ km/s for the $S$-wave velocity at the 8-km foreshock depth (Table S8), and the

uncertainty $\delta v_s = 0.25$ km/s comes from the range of possible $v_s$ values in the velocity model at the trial depths, from 3 km ($v_s = 3.29$ km/s) to 15 km ($v_s = 3.78$ km/s).

The uncertainty in the average slip $D$, in column 8 of Table 1, is (substituting Equation 3 into the partial derivatives):

$$\delta D = \sqrt{\left(\frac{\partial D}{\partial M_{02}}\delta M_{02}\right)^2 + \left(\frac{\partial D}{\partial R}\delta R\right)^2} = \sqrt{\left(\frac{\partial}{\partial M_{02}}\left(\frac{M_{02}}{\mu\pi R^2}\right)\delta M_{02}\right)^2 + \left(\frac{\partial}{\partial R}\left(\frac{M_{02}}{\mu\pi R^2}\right)\delta R\right)^2}$$

$$\Rightarrow \delta D = \sqrt{\left(\frac{1}{\mu\pi R^2}\delta M_{02}\right)^2 + \left(\frac{2M_{02}}{\mu\pi R^3}\delta R\right)^2} \tag{S11}$$

The uncertainty in the stress drop $\Delta\sigma$, in column 9 of Table 1, is (substituting Equation 4 into the partial derivatives):

$$\delta\Delta\sigma = \sqrt{\left(\frac{\partial\Delta\sigma}{\partial M_{02}}\delta M_{02}\right)^2 + \left(\frac{\partial\Delta\sigma}{\partial R}\delta R\right)^2} = \sqrt{\left(\frac{\partial}{\partial M_{02}}\left(\frac{7}{16}\frac{M_0}{R^3}\right)\delta M_{02}\right)^2 + \left(\frac{\partial}{\partial R}\left(\frac{7}{16}\frac{M_0}{R^3}\right)\delta R\right)^2}$$

$$\Rightarrow \delta\Delta\sigma = \sqrt{\left(\frac{7}{16R^3}\delta M_{02}\right)^2 + \left(\frac{21M_{02}}{16R^4}\delta R\right)^2}$$
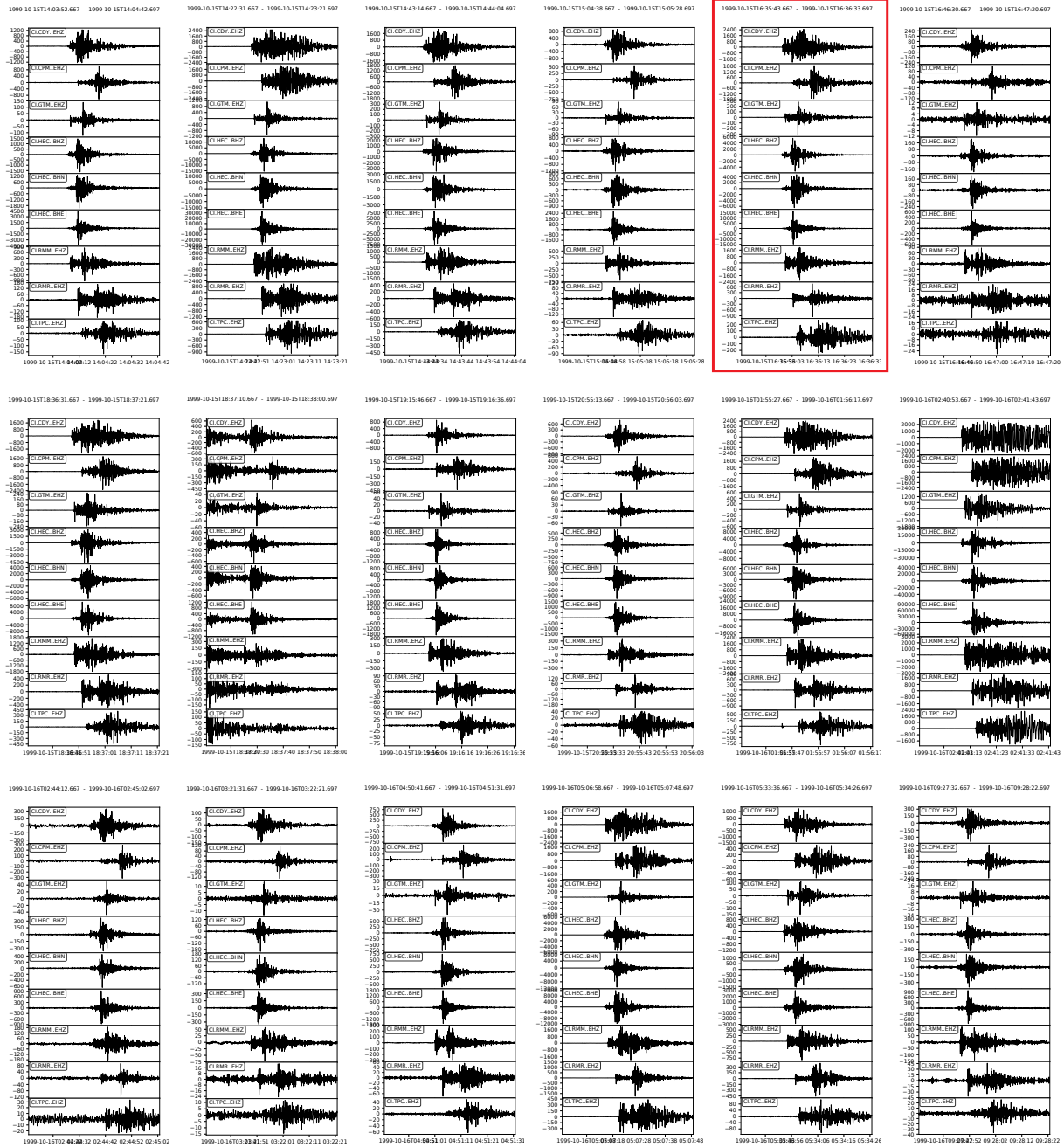
$$\tag{S12}$$

**Figures S1 to S10**



**Figure S1.** 50-second waveforms of the 18 catalog foreshocks at the 7 stations (9 channels) used for detection (Data Set S1), filtered according to Table S1, and ordered by time. The red rectangle indicates the reference event with SCSN ID 9108545 where *P* and *S* arrivals were manually picked at all 14 stations used for location (Section S2).
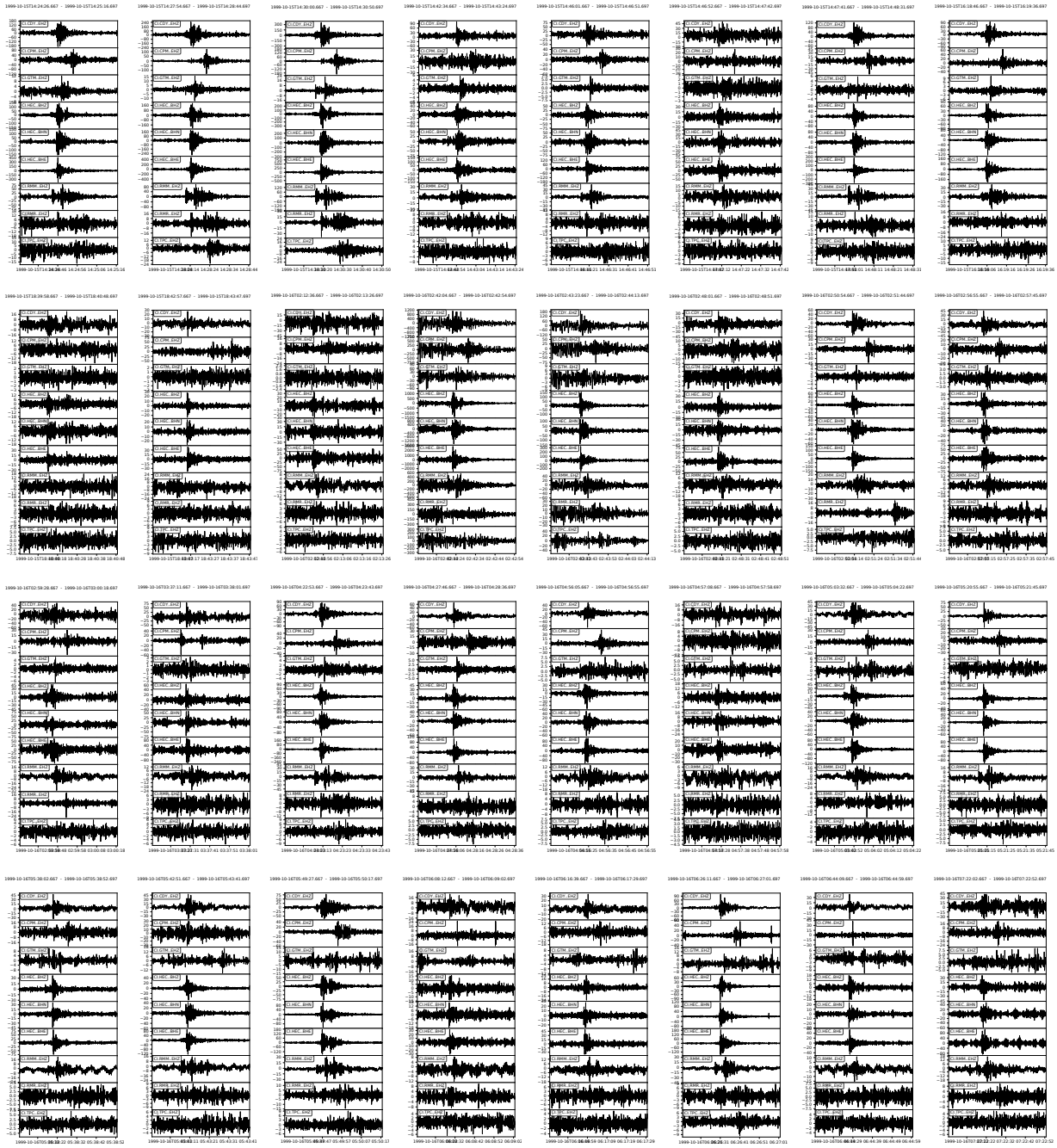
**Figure S2.** 50-second waveforms of the 32 new uncataloged foreshocks at the 7 stations (9 channels) used

for detection (Data Set S2 and S3), filtered according to Table S1, and ordered by time.
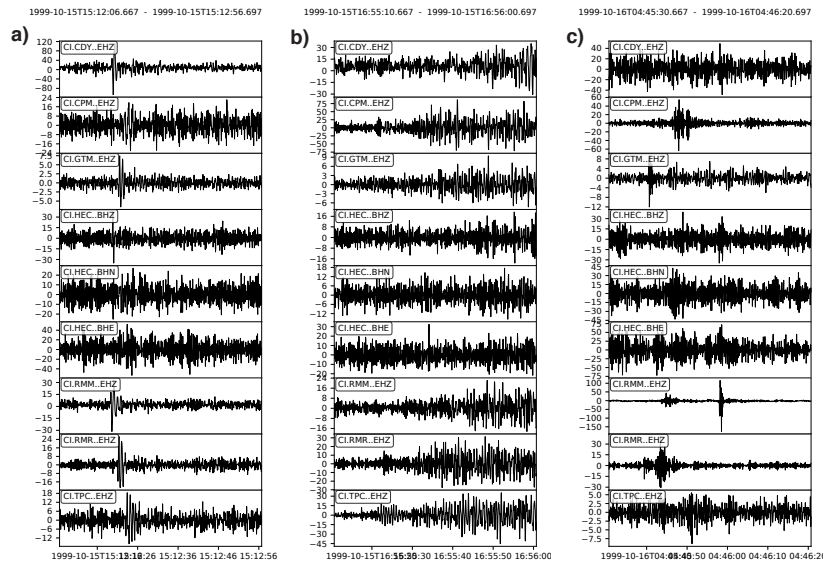
**Figure S3.** 50-second waveforms of the 3 other earthquakes (not foreshocks of the Hector Mine earthquake) at the 7 stations (9 channels) used for detection, filtered according to Table S1, and ordered by time. (a) Deep teleseismic earthquake, Kuril Islands (depth 163.9 km, origin time 1999-10-15 15:01:30 UTC, *M* 4.9): note the impulsive low-frequency waveform. (b) Teleseismic earthquake in Luzon, Philippines (depth 44.5 km, origin time 1999-10-15 16:39:35 UTC, *M* 4.5): the waveform is more emergent and longer in duration. (c) Uncataloged local earthquake.
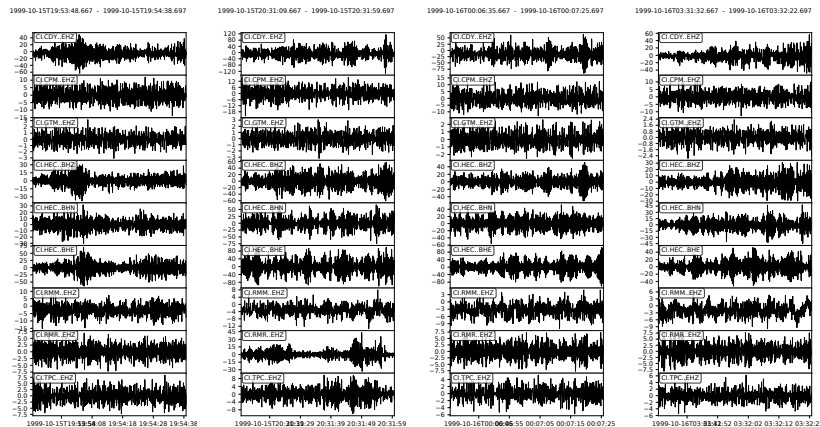
**Figure S4.** 50-second waveforms of 4 false detections (noise, not earthquakes) at the 7 stations (9 channels) used for detection, filtered according to Table S1, and ordered by time.
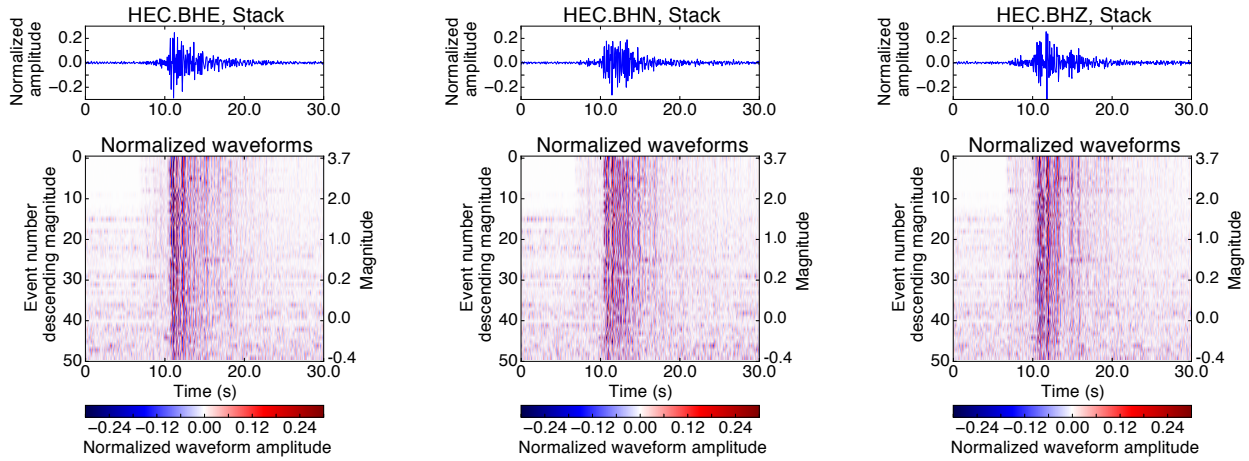
**Figure S5.** Stack waveform (top) and normalized waveforms aligned with cross-correlation (bottom) of all 50 foreshocks on each component at the nearest station HEC.



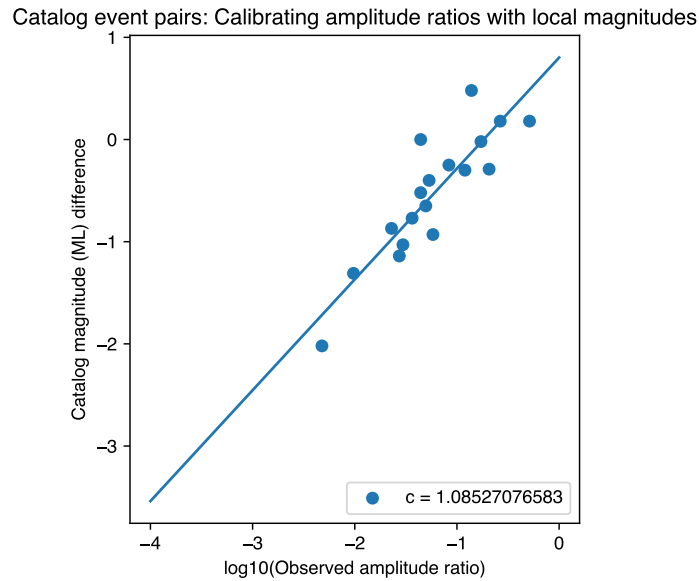**Figure S6.** Calibration using catalog events for estimating magnitudes of newly detected events. Each data point represents a catalog event pair, with the relative amplitude $\log_{10}\left(\alpha_{(cat2,cat1)}\right)$ on the horizontal axis, and magnitude difference $M_{cat2} - M_{cat1}$ on the vertical axis. A linear least-squares fit (Equation S3) to the data is used to calculate the calibration constant $c = 1.085$.
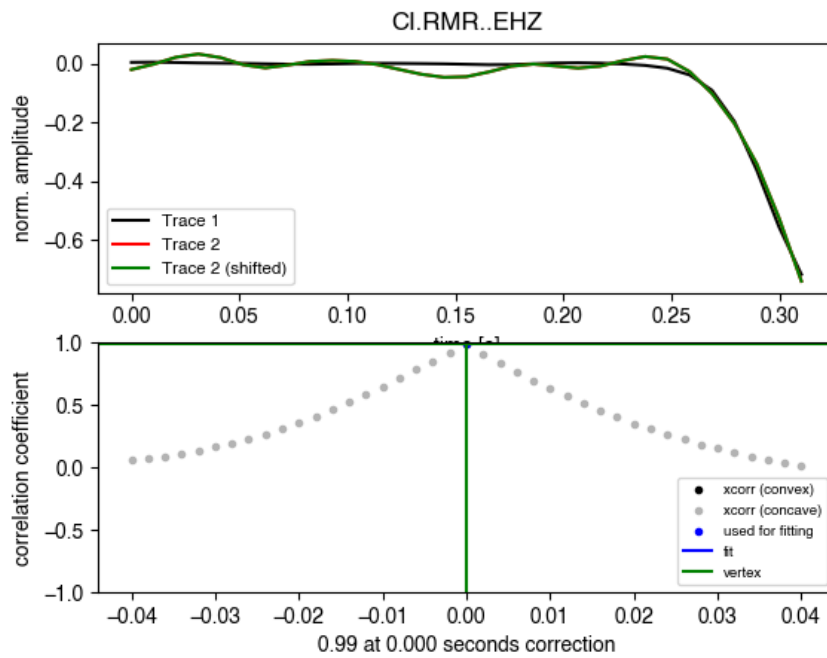
**Figure S7.** Example of cross-correlation differential time calculation between the $M_w$ 7.1 mainshock (black) and $M_w$ 2.2 foreshock (green) at station RMR (0.3 time window from the mainshock and foreshock, 0.22 seconds before and 0.08 seconds after the *P* arrival, top plot) with a correlation coefficient of 0.99 (bottom plot). The initial onset rise (first 0.08 seconds) of the $M_w$ 7.1 mainshock is highly similar that of the $M_w$ 2.2 foreshock, even if their entire waveforms are not similar.

**Figure S8.** Alternative locations from different catalogs for foreshocks (circles sized by relative magnitude, colored by time) of the Hector Mine earth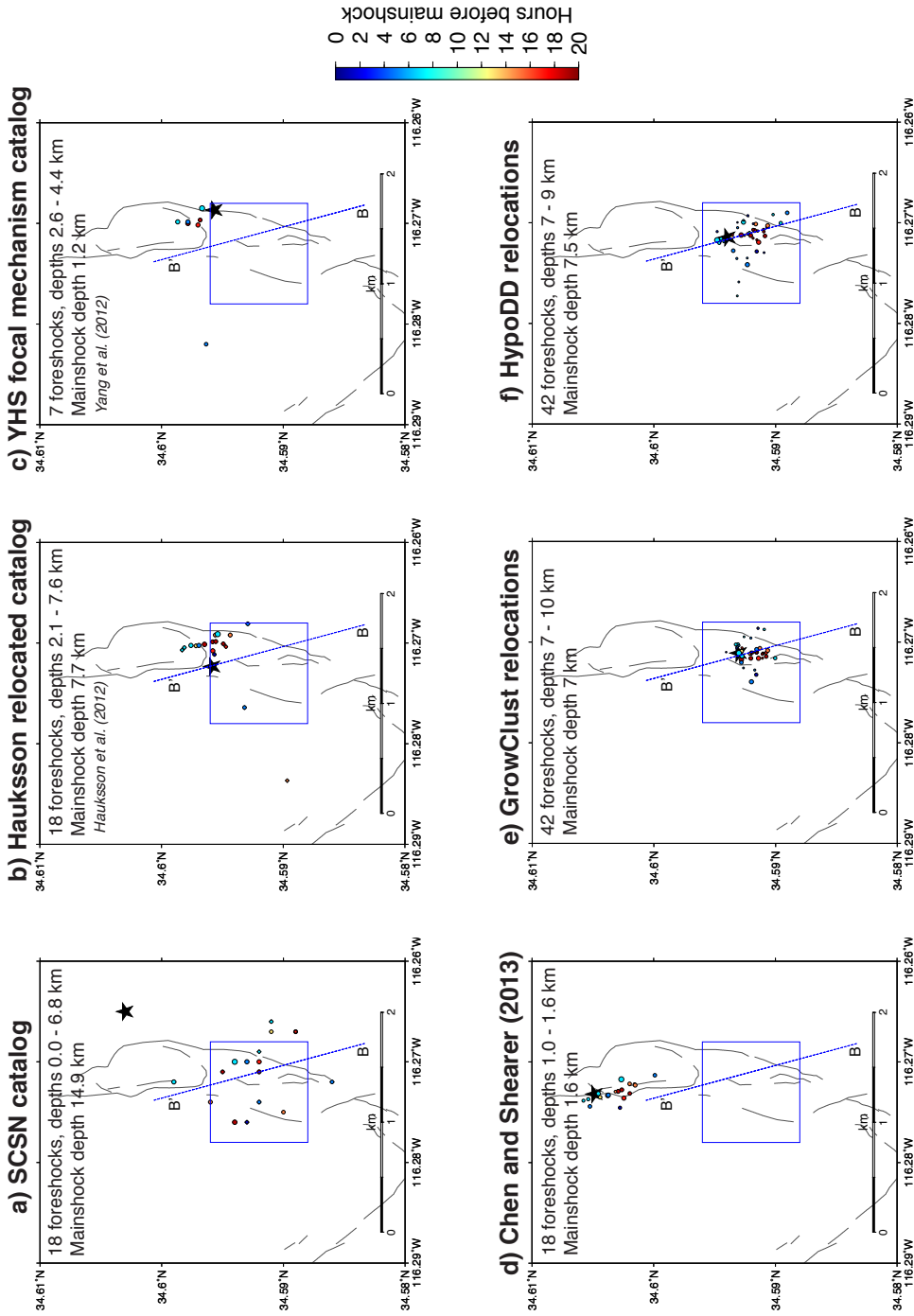quake (mainshock hypocenter denoted by star), which demonstrate the range of location and depth uncertainties. The blue inset box and profile B-B' are the same as those in Figures 2b, 3a, and S9. (a) 18 foreshocks and mainshock from SCSN catalog. (b) 18 foreshocks from *Hauksson et al.* [2012] waveform relocated catalog. (c) 7 largest foreshocks from *Yang et al.* [2012] focal mechanism catalog. (d) 18 foreshocks from *Chen and Shearer* [2013] relocations. (e) 42 foreshocks relocated using GrowClust [*Trugman and Shearer*, 2017] in this study, with same starting location (34.592°, -116.271°) and starting depth 8 km. (f) 42 foreshocks relocated using hypoDD in this study, with same starting location (34.592°, -116.271°) and starting depth 8 km; these locations are reported in Figures S9 and Figure 3a.
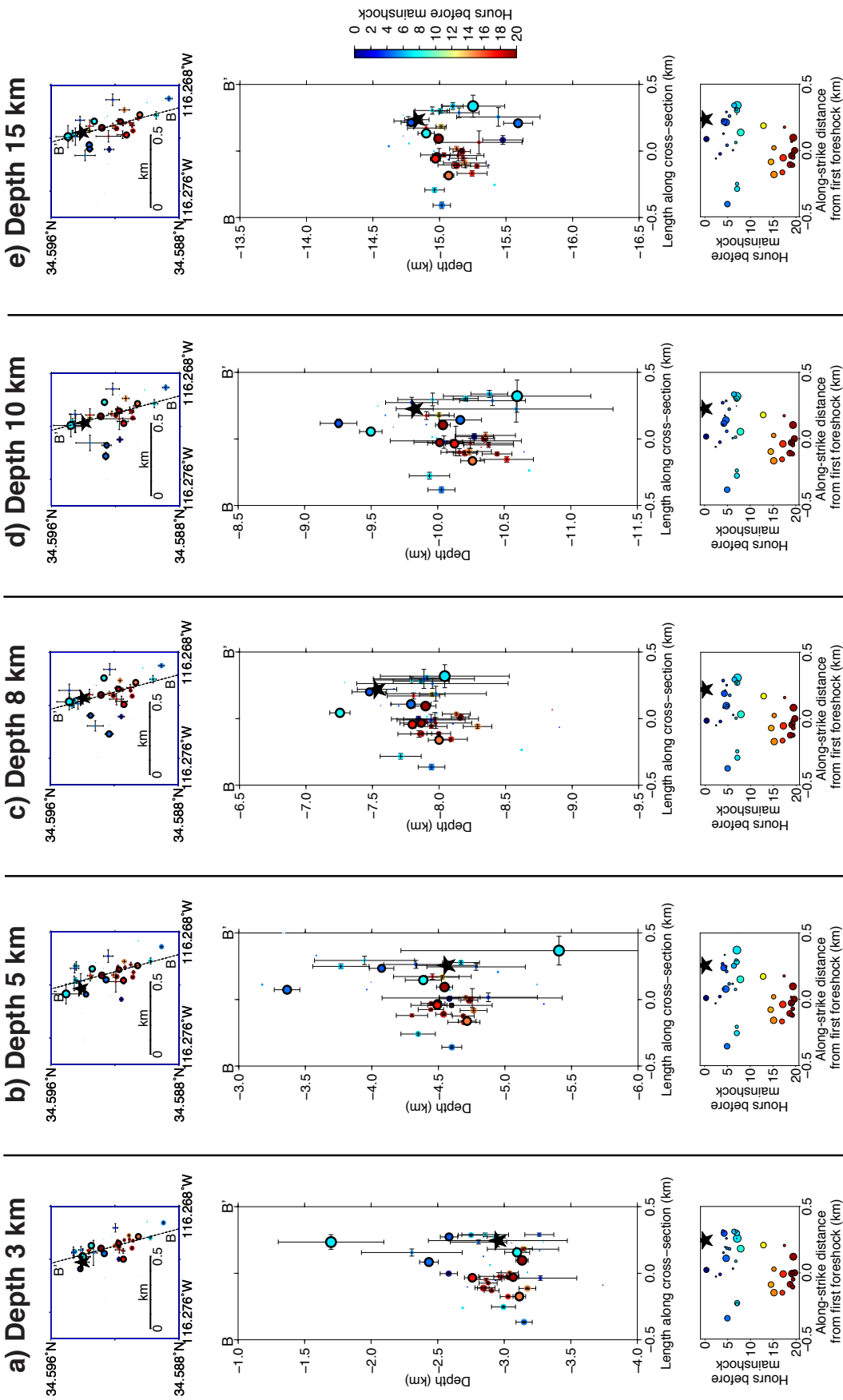
**Figure S9.** Uncertainty analysis for the space-time evolution of 42 located foreshocks (circles sized by relative magnitude, colored by time) of the Hector Mine earthquake (mainshock hypocenter denoted by star): map view (top row), depth cross-section along southeast-to-northwest profile B-B' parallel to fault strike (center row), along-strike time evolution of foreshocks leading up to mainshock (bottom row). Circles with thick edges indicate the 8 largest foreshocks for which source parameters and rupture dimensions are calculated in Figure S10. Each column indicates the trial starting depth for all earthquakes: (a) 3 km, (b) 5 km, (c) 8 km, (d) 10 km, (e) 15 km.
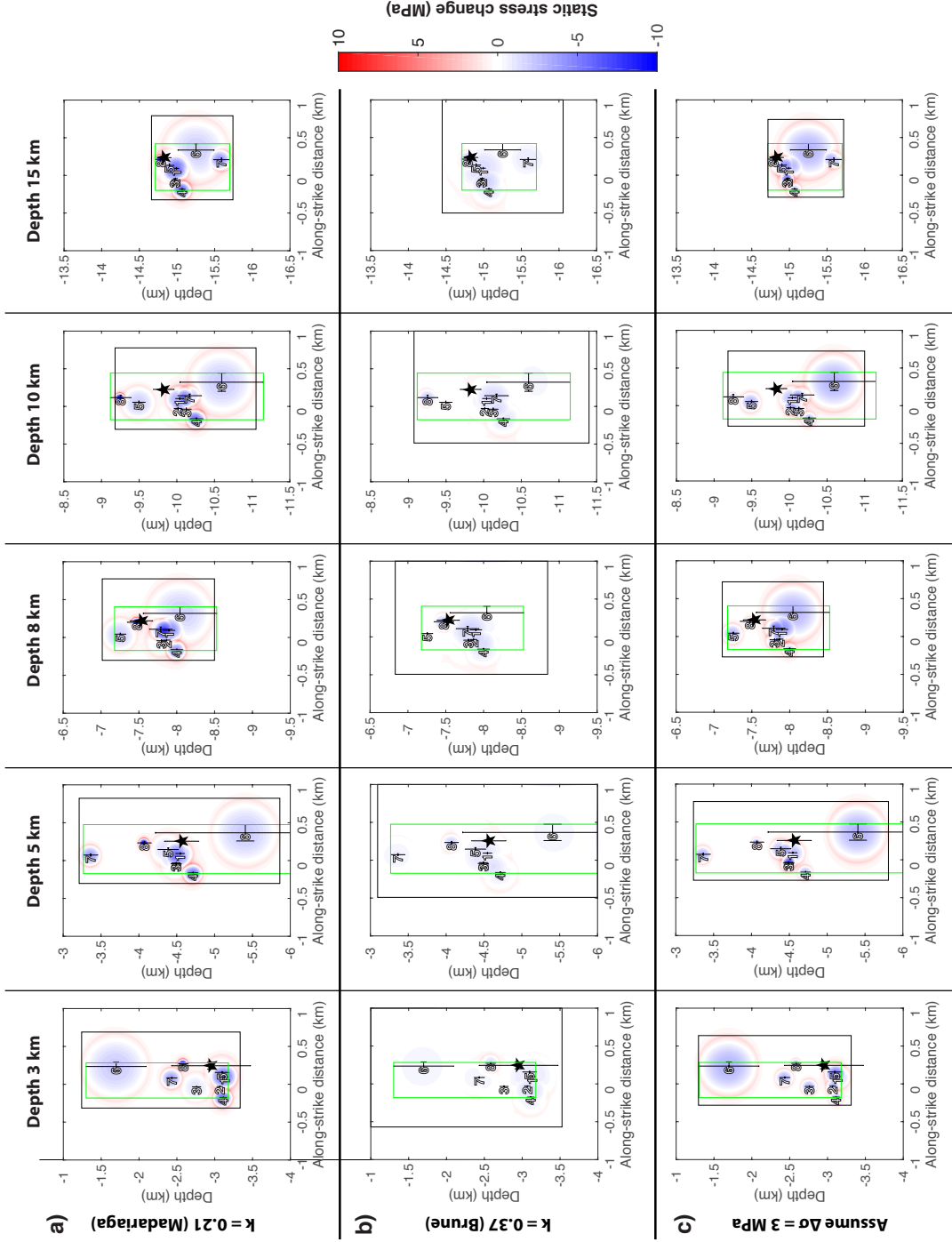
**Figure S10.** Uncertainty analysis for the cumulative static stress change (MPa) on fault plane from 8 largest foreshocks, and mainshock hypocenter (black star), for locations resulting from different trial starting depths (columns: 3 km, 5 km, 8 km, 10 km, 15 km, with locations and uncertainties from center row of Figure S9), and with source parameters (Table 1) calculated using three different model assumptions (rows): (a) $k = 0.21$ [*Madariaga*, 1976], (b) $k = 0.37$ [*Brune*, 1970], (c) assume stress drop $\Delta\sigma = 3$ MPa. The green rectangle is the smallest box containing all foreshock hypocenters including their uncertainties, while the black rectangle is the smallest box containing all foreshock rupture areas along the fault.

**Tables S1 to S9**

**Table S1.** 7 SCSN stations with continuous seismic data used for foreshock detection, spanning the >20 hour time period from 1999-10-15 13:00:00 to 1999-10-16 09:46:44 UTC (cut off just before the mainshock). Stations CDY and HEC are located only 0.2 km apart. Before detecting earthquakes using FAST, we applied a station-specific filter to each channel of data to remove frequencies with repeating noise. Data at the 3-component broadband station HEC was originally sampled at 20 Hz. Data at the other 6 short-period stations (vertical component only) was originally sampled at 100 Hz, but decimated to 20 Hz after the bandpass filter. Therefore, the Nyquist frequency for all channels input to FAST was 10 Hz.

| Station.Channel | Latitude (deg) | Longitude (deg) | Filter frequencies (Hz) |
| --- | --- | --- | --- |
| CDY.EHZ | 34.83007 | -116.33717 | no filter |
| CPM.EHZ | 34.15442 | -116.19771 | highpass 1 Hz |
| GTM.EHZ | 34.2946 | -116.356 | bandpass 1-6 Hz |
| HEC.BHE | 34.8294 | -116.335 | highpass 2 Hz |
| HEC.BHN | 34.8294 | -116.335 | highpass 2 Hz |
| HEC.BHZ | 34.8294 | -116.335 | highpass 2 Hz |
| RMM.EHZ | 34.64384 | -116.62438 | no filter |
| RMR.EHZ | 34.21283 | -116.5763 | bandpass 1-6 Hz |
| TPC.EHZ | 34.10564 | -116.04939 | bandpass 1-5 Hz |

**Table S2.** FAST input parameters [*Yoon et al.*, 2015; *Bergen et al.*, 2016; *Bergen and Beroza*, 2018b; *Rong et al.*, 2018] used to detect foreshocks in each component of continuous seismic data (sampled at 20 Hz) at the 7 stations listed in Table S1. FAST first converts seismic waveforms into compact, discriminative, binary features called "fingerprints", then organizes the fingerprints into a database to efficiently search for similar fingerprints with high probability.

| FAST parameter: Fingerprint | Value |
| --- | --- |
| Time window length (s) for spectrogram | 6 s (120 samples) |
| Time window lag (s) for spectrogram | 0.2 s (4 samples) |
| Spectral image length (samples) | 32 samples (12.4 s) |
| Spectral image lag (samples) = fingerprint sampling period | 5 samples (1 s) |
| Final spectral image width = number of frequency bins | 32 samples |
| Number of wavelet coefficients to keep | 200 (out of 1024) |

| FAST parameter: Similarity Search | Value |
| --- | --- |
| LSH: number of hash functions per hash table $r$ | 4 |
| LSH: number of hash tables $b$ | 100 |
| Initial pair threshold: number $v$ (fraction) of tables, pair in same bucket | 2 (2/100 = 0.02) |
| Similarity search: near-repeat exclusion parameter | 5 samples (5 s) |

**Table S3.** Input parameters for pair-wise association and network detection over the 7-station network [*Bergen and Beroza*, 2018a]. The input to this algorithm is the output of FAST from each station: a list of [*time1, time2, sim*] values, where *time1* and *time2* are times in the continuous data with similar fingerprints (and therefore similar earthquake signals), and *sim* is their FAST similarity, defined by the number of hash tables with this fingerprint pair in the same bucket, with $sim \geq v$, where $v = 2$ is the initial pair threshold (Table S2). For the 3-component station HEC, we added the *sim* values at the same (*time1*, *time2*) pair across all components, and set the station-pair threshold to $\tau_0 = (v = 2)*(3 \text{ components}) = 6$. For the remaining 1-component stations in Table S2, we multiplied every *sim* value by 3 so that the FAST similarity values would be weighted equally as those from station HEC.

| Event-pair Extraction Parameter | Value |
| --- | --- |
| Time gap (along diagonal), $g_L$ | 2 samples (2 s) |
| Time gap (adjacent diagonal), $g_W$ | 1 sample (1 s) |
| Adjacent diagonal merge iterations, $p$ | 2 |
| **Event-pair Pruning Parameter** | **Value** |
| Number of votes (station-pair threshold), $\tau_0$ | 6 |
| Minimum fingerprint-pairs, $|C|_{min}$ | 4 |
| Minimum total similarity, $v_{min}^{(C)}$ | 24 |
| Maximum bounding box width | 8 samples (8 s) |
| **Pseudo-association Parameter** | **Value** |
| Minimum number of stations for detection | 2 (out of 7) |
| Arrival time constraint: maximum time gap, $g_N$ | 3 samples (3 s) |

**Table S4.** Final thresholds applied to network detection parameters *nsta* (number of stations that detected event) and *peaksum* (total FAST similarity over all stations, for the strongest similarity value involving this event) to determine list of earthquakes [*Bergen and Beroza*, 2018a], set empirically after visual inspection. For each value of *nsta*, a different threshold for *peaksum* can be applied. 57 detections exceeded these final thresholds: 50 foreshocks (Figure S1, Figure S2), 3 other earthquakes (Figure S3), and 4 false detections (Figure S4).

| *nsta* | *peaksum* |
| --- | --- |
| $\geq 5$ | all events |
| 4 | 100 |
| 3 | 75 |
| 2 | 110 |

**Table S5.** 14 SCSN stations (20 channels) with continuous seismic data used for foreshock location; the first 7 stations were also used to detect foreshocks (Table S1). At the 7 additional stations, we extracted event waveforms from the continuous data at times of the newly detected events. We apply a station-specific bandpass filter to each channel of data before computing cross-correlation differential travel times for double-difference location. (Note that these filter bands are sometimes different for location than for detection.) We use data at the full sampling rate for this cross-correlation, which is 100 Hz at some stations and 20 Hz at others.

| Station.Channel | Latitude (deg) | Longitude (deg) | Sampling rate (Hz) | Filter frequencies (Hz) |
|---|---|---|---|---|
| CDY.EHZ | 34.83007 | -116.33717 | 100 | 1-10 |
| CPM.EHZ | 34.15442 | -116.19771 | 100 | 1-10 |
| GTM.EHZ | 34.2946 | -116.356 | 100 | 1-10 |
| HEC.BHE | 34.8294 | -116.335 | 20 | 1-9 |
| HEC.BHN | 34.8294 | -116.335 | 20 | 1-9 |
| HEC.BHZ | 34.8294 | -116.335 | 20 | 1-9 |
| RMM.EHZ | 34.64384 | -116.62438 | 100 | 1-10 |
| RMR.EHZ | 34.21283 | -116.5763 | 100 | 1-10 |
| TPC.EHZ | 34.10564 | -116.04939 | 100 | 1-10 |
| EW2.EHZ | 33.94059 | -116.40839 | 100 | 1-10 |
| FRG.EHZ | 33.75722 | -116.06228 | 100 | 1-10 |
| GRP.EHZ | 34.80481 | -115.60705 | 100 | 1-10 |
| GSC.BHE | 35.30177 | -116.80574 | 20 | 1-9 |
| GSC.BHN | 35.30177 | -116.80574 | 20 | 1-9 |
| GSC.BHZ | 35.30177 | -116.80574 | 20 | 1-9 |
| PNM.EHZ | 33.97738 | -115.80161 | 100 | 1-10 |
| SBPX.BHE | 34.2324 | -117.23484 | 20 | 1-9 |
| SBPX.BHN | 34.2324 | -117.23484 | 20 | 1-9 |
| SBPX.BHZ | 34.2324 | -117.23484 | 20 | 1-9 |
| SIL.EHZ | 34.34802 | -116.82746 | 100 | 1-10 |

**Table S6.** Input parameters for hypoDD program (version 2.1b) to relocate earthquakes [*Waldhauser and Ellsworth*, 2000]. For calculating location uncertainties, we ran hypoDD again in SVD mode (ISOLV=1) with NSET=5 (omitting the last 4 iterations in Table S7) for 31 out of the 42 located foreshocks.

| Parameter name | Value | Description |
|---|---|---|
| IDAT | 1 | Cross-correlation only |
| IPHA | 3 | P and S phases |
| DIST | 400 | Maximum distance (km) between cluster centroid and station |
| OBSCC | 0 | Minimum number of obs/pair for crosstime data |
| MINDS | 0 | Minimum distance between individual event pairs and stations |
| MAXDS | 200 | Maximum distance between individual event pairs and stations |
| MAXGAP | -999 | Maximum azimuthal gap between individual event pairs and stations (-999: not used) |
| ISTART | 2 | From network sources (though all initial locations were the same) |
| ISOLV | 2 | LSQR mode |
| IAQ | 0 | Keep air-quakes, reset depths to those of previous (successful) iteration |
| NSET | 6 | Number of sets of iteration |

**Table S7.** hypoDD data weighting parameters [*Waldhauser and Ellsworth*, 2000], LSQR mode, 24 iterations.

| NITER | WTCCP | WTCCS | WRCC | WDCC | DAMP |
|---|---|---|---|---|---|
| 4 | 1 | 1 | -9 | -9 | 12 |
| 4 | 1 | 1 | -9 | 8 | 12 |
| 4 | 1 | 1 | 6 | 8 | 12 |
| 4 | 1 | 1 | 6 | 5 | 12 |
| 4 | 1 | 1 | 6 | 3 | 12 |
| 4 | 1 | 1 | 6 | 2 | 12 |

**Table S8.** 1D velocity model for Southern California [*Kamer et al.*, 2017] used to locate earthquakes in this study. $V_p/V_s = \sqrt{3}$.

| Depth to top of layer (km) | P wave (km/s) |
|---|---|
| 0 | 5.28 |
| 2 | 5.68 |
| 4 | 6.05 |
| 6 | 6.13 |
| 8 | 6.21 |
| 10 | 6.30 |
| 12 | 6.36 |
| 15 | 6.53 |
| 18 | 6.59 |
| 23 | 6.83 |
| 30 | 7.84 |
| 50 | 8.20 |

**Table S9.** 4 broadband SCSN stations with triggered event waveforms used to calculate spectral ratios and corner frequencies. Each station has 3 components (HHE, HHN, HHZ) sampled at 100 Hz.

| Station | Latitude (deg) | Longitude (deg) |
|---|---|---|
| BC3 | 33.6548 | -115.4531 |
| DAN | 34.6375 | -115.3812 |
| GSC | 35.30177 | -116.80574 |
| HEC | 34.8294 | -116.335 |

**Data Set S1.**

List of 18 foreshocks from the SCSN catalog, and Hector Mine mainshock (last event in list), in time order. UTC origin time (columns 2-5) is from the SCSN catalog.

| Column | Description |
| --- | --- |
| 1 | FAST detection time (not origin time) in seconds since UTC 1999-10-15T13:00:00.676 |
| 2 | Origin time (UTC), year, month, day (yyyymmdd) |
| 3 | Origin time (UTC), hour |
| 4 | Origin time (UTC), minute |
| 5 | Origin time (UTC), second |
| 6 | Magnitude from catalog |
| 7 | SCSN event ID |
| 8 | Number of stations where FAST detected this event: *nsta* (Table S4) |
| 9 | Total FAST similarity over all stations, for the strongest similarity value involving this event: *peaksum* (Table S4) |

**Data Set S2.**

List of 24 newly detected and located foreshocks, in time order.

| Column | Description |
| --- | --- |
| 1 | FAST detection time (not origin time) in seconds since UTC 1999-10-15T13:00:00.676 |
| 2 | Number of stations where FAST detected this event: *nsta* (Table S4) |
| 3 | Total FAST similarity over all stations, for the strongest similarity value involving this event: *peaksum* (Table S4) |
| 4 | Magnitude (calculated with procedure in Section S1) |

**Data Set S3.**

List of 8 newly detected foreshocks too small to be located, in time order.

| Column | Description |
| --- | --- |
| 1 | FAST detection time (not origin time) in seconds since UTC 1999-10-15T13:00:00.676 |
| 2 | Number of stations where FAST detected this event: *nsta* (Table S4) |
| 3 | Total FAST similarity over all stations, for the strongest similarity value involving this event: *peaksum* (Table S4) |
| 4 | Magnitude (calculated with procedure in Section S1) |