Supporting Information for "Earthquakes Within Earthquakes: Patterns in Rupture Complexity"

Philippe Danré[1,2] , Jiuxun Yin[1], Bradley Lipovsky[1], Marine A. Denolle[1] (corresponding author)

[1]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

[2]Département de Géosciences, École Normale Supérieure, PSL Res. Univ. Paris, France

Contents of this file

Additional Supporting Information (Files uploaded separately)

1. Caption for Movie S1

Introduction In the supplementary materials, we expand on our proposed algorithm for earthquake early warning using source time functions (Text S1). This algorithm is illustrated in Movie S1. After justifying our choice of kernel for the subevent function (Text S2), we then expand on the dynamic rupture modeling parameters (Text S3), and show the pre-stress distribution (Figure S7) and an example of a single example of simulations (Figure S8). We also add supplementary figures to show the cross validation against the USGS database (Figure S1), the sensitivity to the threshold in subevent size (Figure S2), the application of the EEW algorithm to all events in the SCARDEC database (Figure S3), the properties of the first and the largest subevent to support the distribution of Figure S3 (Figure S4), the number of subevents and scaling of subevent moments to main event moment obtained from triangle kernels (Figure S5) and the ratio between moment obtained from Gaussian or from Triangle functions (Figure S6).

Text S1: Earthquake early warning algorithm

In the context of a real-time application of our findings, we do not have knowledge of the maximum value of the STF, which is known a posteriori. Thus, we modify the de-

tection algorithm, and show one example in main manuscript Figure 4. In order to avoid fitting residuals, we proceed as follow:

1. Detect a first peak.
   (a) Compute its moment M_S and duration $T_S = 4\sigma$, if $T_S < 1$ s, go back.
   (b) Use the scaling relation to get M_0 from M_S and combine with magnitude-moment relation: $M_W = 0.84log_{10}(M_S) - 8.7$.
2. Detect the next peak (number i) if it has at least 0.25 times the amplitude of the first peak, or of the median of the previous subevents peaks:
   (a) Compute its moment M_S^i and duration T_S^i.
   (b) Use the scaling relation to get M_0^i from M_S^i and combine with magnitude-moment relation: $M_W^i = 0.84log_{10}(M_S^i) - 8.7$
   (c) Take the median of the individual magnitude estimates $M_W^i = median(M_W)$.
3. Go back to step 2.

Because this algorithm necessitates a different criterion for the amplitude threshold, we re-calculate the scaling between subevent and main event moments (Figure S3a). The coefficients of the regression in the scaling change slightly: $log_{10}(M_S) = 0.79log_{10}M_0 + 3.12$. This is the method implemented for Figure 4 of main manuscript. We apply this to all events in SCARDEC database. We show in Figure S3b the difference between the ground truth magnitude and our prediction against the rupture time normalized to duration time. In Figure S3b, each dot is a magnitude estimate at the peak time of a subevent $M_W^i - M_W$. First estimates can come in as early as 5% of the source duration with an underprediction of about 0.5 magnitude increment. Between 30-50%, our method overpredicts the magnitude by an increment of 0.5. We show in Figure S4(c) that the small events have few subevents (mostly a single one) and thus the time at which that subevent occurs tends to be half way in the rupture. This is also visible in Figure 4 of the main manuscript. Biases could be corrected for in future development of this approach.

Text S2: Kernel choice for subevent function

Our algorithm of subevent detection uses Gaussian functions. We choose their width with a grid search for the best-fitting solution that minimizes the L_2 norm of the residuals of the STF that we construct from progressively stripping down from the subevent

Gaussian functions (in a similar fashion to *Kikuchi and Kanamori* [1982]). The choice of a Gaussian kernel is motivated by the observed shape of the STF, and is preferred to other kernels because it leads to greater accuracy in the reconstructed STF and in the total moment (Figure S6a). We also try the triangular function as a choice of kernel. We find in general that the uncertainty in the total moment from the original STF and that reconstructed with the subevents is biased and contain greater uncertainty,(Figure S6b). Besides these inaccuracies, the moment of the subevents scales with the moment of the main events, as shown in Figure S5, which supports the robustness of our results.

Text S3: Dynamic rupture simulation set up Our simulations solve the equations of 2D mode III elastodynamics coupled to a frictional fault using the spectral boundary integral methods (SBIEMLAB, code developed by Jean-Paul Ampuero, `http://web.gps.caltech.edu/~ampuero/software.html`, last accessed 11/27/2018). The ingredients for our simulations are listed below:

- **Basic parameters in the simulations**: Model domain is 400 km long, with 200 km long fault domain and 200 km boundary domain, and is uniformly discretized by 4096 grid points with a grid size of Δx = 97.7 m. Other fundamental parameters are listed in Table S1.
- **Constitutive relation**: We use a linear slip weakening law to describe the friction on the fault:

$$\mu = \begin{cases} \frac{\mu_d - \mu_s}{D_c} + \mu_s, s \leq D_c, \\ \mu_d, s > D_c, \end{cases} \tag{1}$$

  where s is the slip, μ_d and μ_s are the dynamic and static friction coefficients, respectively, D_c is the critical slip distance, which we choose to be $D_c = 0.8\ m$ in the main text (we also tested $D_c = 0.4\ m$ and $1.2\ m$). The frictional parameters μ_d and μ_s are chosen to be uniform along the fault.
- **Heterogeneous pre-stress distribution**: We produce a heterogeneous pre-stress distribution with the scaled Power Spectral Density (PSD) function:

$$P_m(k) = C|k|^{-\gamma}, \tag{2}$$

  where k is the wavenumber, and C is chosen such that $P_m(k)$ is normalized to its absolute maximum in space domain and scaled to $0.9(\mu_s - \mu_d)\sigma_0$. We choose $\gamma = 0.8$. The phase of the pre-stress distribution is randomly generated from a uniform distribution in $[0, 2\pi]$. Combining $P_m(k)$ with the random phases, we can

produce a large ensemble of statistically similar distributions. We pad the negative part of these distributions to zero and normalize non-negative distributions to the range between $[\mu_d\sigma_0,\ 0.9\mu_s\sigma_0]$. We further add a random noise between $[-0.05(\mu_s-\mu_d)\sigma_0,\ 0.05(\mu_s-\mu_d)\sigma_0]$ to these distribution, to mimic the background heterogeneity on the fault. Finally, we apply a Tukey-window to taper the 100 km on either end of the 400 km pre-stress distributions to avoid the artifacts in STF from abruptly stopping of rupture at the fault boundary in the spectral boundary integral solutions (Fig. S7 a). To make sure the series of post-processing on the pre-stress distribution does not distort the original PSD, we also compare the spectra of unprocessed (pure self-affine) and processed (non-negative constraint) pre-stress distributions in the frequency domain (Fig. S8 b) to confirm that the spectral slope is preserved.

- **2D spontaneous dynamic rupture model**: To nucleate the dynamic rupture, we randomly and artificially set an over-stressed nucleation area. We only keep the pre-stress distributions whose pre-stress peak is within $x = [-30, 30]$ km on the fault and nucleate rupture at that peak pre-stress location. This setting helps to avoid the boundary effects of pre-stress distribution and thus ensure that most of the simulated events evolve on a stress condition that is statistically similar during the event. The nucleation length is based on the relation from *Uenishi and Rice* [2003]:

$$L_c = \frac{1.158\mu D_c}{(\mu_s-\mu_d)\sigma_0}. \tag{3}$$

For $D_c = 0.8m$, the corresponding $L_c = 1626.9\ m \approx 17\Delta x$, which also ensures the sufficient spatial resolution of our simulations. We set the size of nucleation zone to be $1.5L_c = 2.44\ km$ and this is to guarantee the ending of quasi-static stage and beginning of dynamic unstable slip, after the over-stress nucleation [*Uenishi and Rice*, 2003]. For each D_c, we produce 600 pre-stress distributions that qualify the aforementioned condition, and finally get the simulated STFs with different event sizes.

Movie S1. Example of potential application to Earthquake Early Warning. The algorithm follows that described in the supplementary materials (Text S1). The example is for the M_w 7.5 2018 Palu Earthquake. Top panel: progressive subevent detection (picks in red crosses from the peaks of the residual STF) and Gaussian fitting (shaded blue Gaussian curves) against the evolution of the moment-rate function. Bottom panel: predicted mag-

nitude estimates from each subevent (black circles), from the median of the previous magnitude estimates (orange circles), and against the magnitude of the previously released moment (green curve).

References

Hayes, G. P. (2017), The finite, kinematic rupture properties of great-sized earthquakes since 1990, *Earth and Planetary Science Letters*, *468*, 94–100.

Iio, Y. (1995), Observations of the slow initial phase generated by microearthquakes: Implications for earthquake nucleation and propagation, *Journal of Geophysical Research: Solid Earth*, *100*(B8), 15,333–15,349.

Kikuchi, M., and H. Kanamori (1982), Inversion of complex body waves, *Bulletin of the Seismological Society of America*, *72*(2), 491–506.

Shearer, P. M., G. A. Prieto, and E. Hauksson (2006), Comprehensive analysis of earthquake source spectra in southern california, *Journal of Geophysical Research: Solid Earth*, *111*(B6).

Uenishi, K., and J. R. Rice (2003), Universal nucleation length for slip-weakening rupture instability under nonuniform fault loading, *Journal of Geophysical Research: Solid Earth*, *108*(B1), 2042, doi:10.1029/2001JB001681.

Parameter	Value
S wave velocity V_s (km/s)	3.46
Density ρ (kg/m^3)	2.670
Shear modulus μ (GPa)	32
Normal stress σ_0 (MPa)	120
Dynamic friction μ_d	0.525
Static friction μ_s	0.677
D_c (m)	0.8
Power law exponent γ	0.8

Table S1. Parameters used in the simulations.

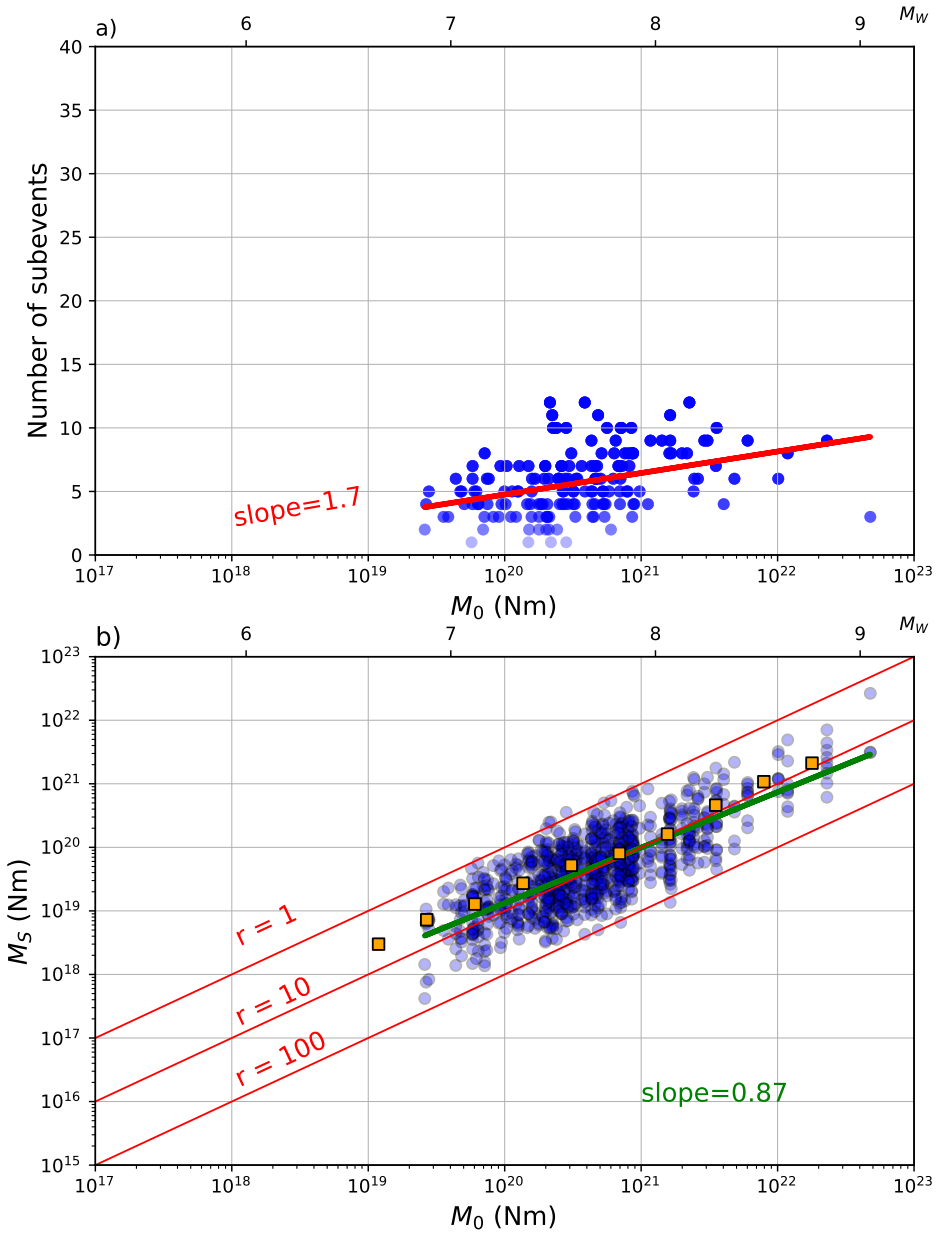

Figure S1. Results for the 167 USGS STFs of *Hayes* [2017] that had detectable subevents (out of 180). (a) Number of subevents as a function of main event moment. There is an increase of complexity with main moment (slope of 1.7, R-value of 0.39). (b) Subevent moment M_S as a function of main event moment M_0 for a peak detection with a 0.1 threshold. The green line represents the result of the regression performed over the whole dataset in a log-log space. Red lines represent the r=1, 10, and 100 ratios of main event moment to subevent moment. Results still show a scaling similar to the SCARDEC results: $M_s \approx M_0^{0.87}$ (R-value of 0.77). Moment magnitudes M_W are labeled on the top of the plots.

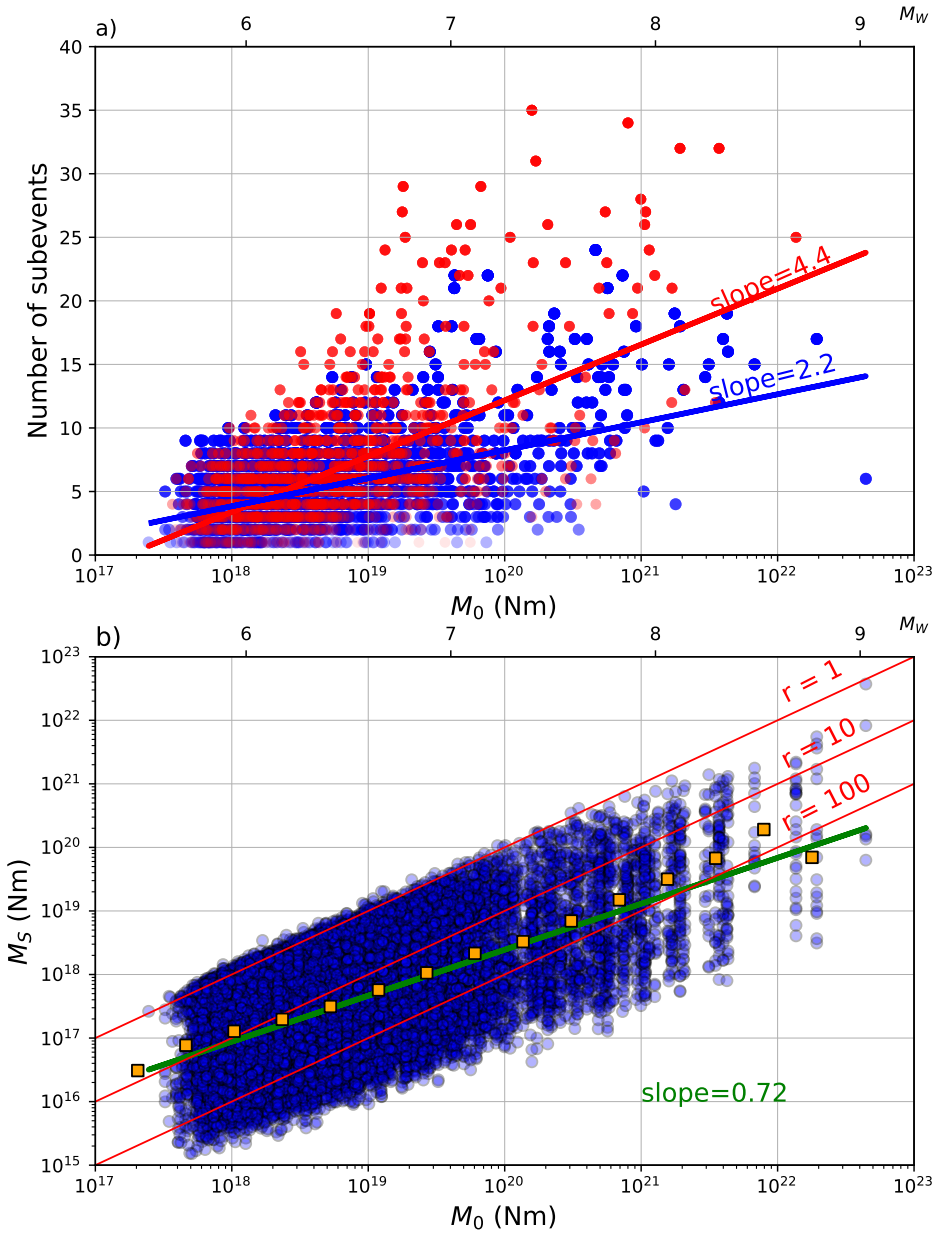

Figure S2. Results of subevent decomposition using a threhold of 0.01. (a) Number of subevents as a function of main event moment (similar to Figure 2 of main manuscript) with red dots strike slip events, blue dots dip-slip events and linear regressions with their slope selected from the focal mechanism parameter defined in *Shearer et al.* [2006]. (b) Subevent moment M_S against main event M_0. The green line is the linear regression performed over the whole dataset in a log-log space. Red lines represent the r=1, 10, and 100 ratios moments. The scaling still holds with a slope of 0.72. There are smaller subevents that could be spurious.

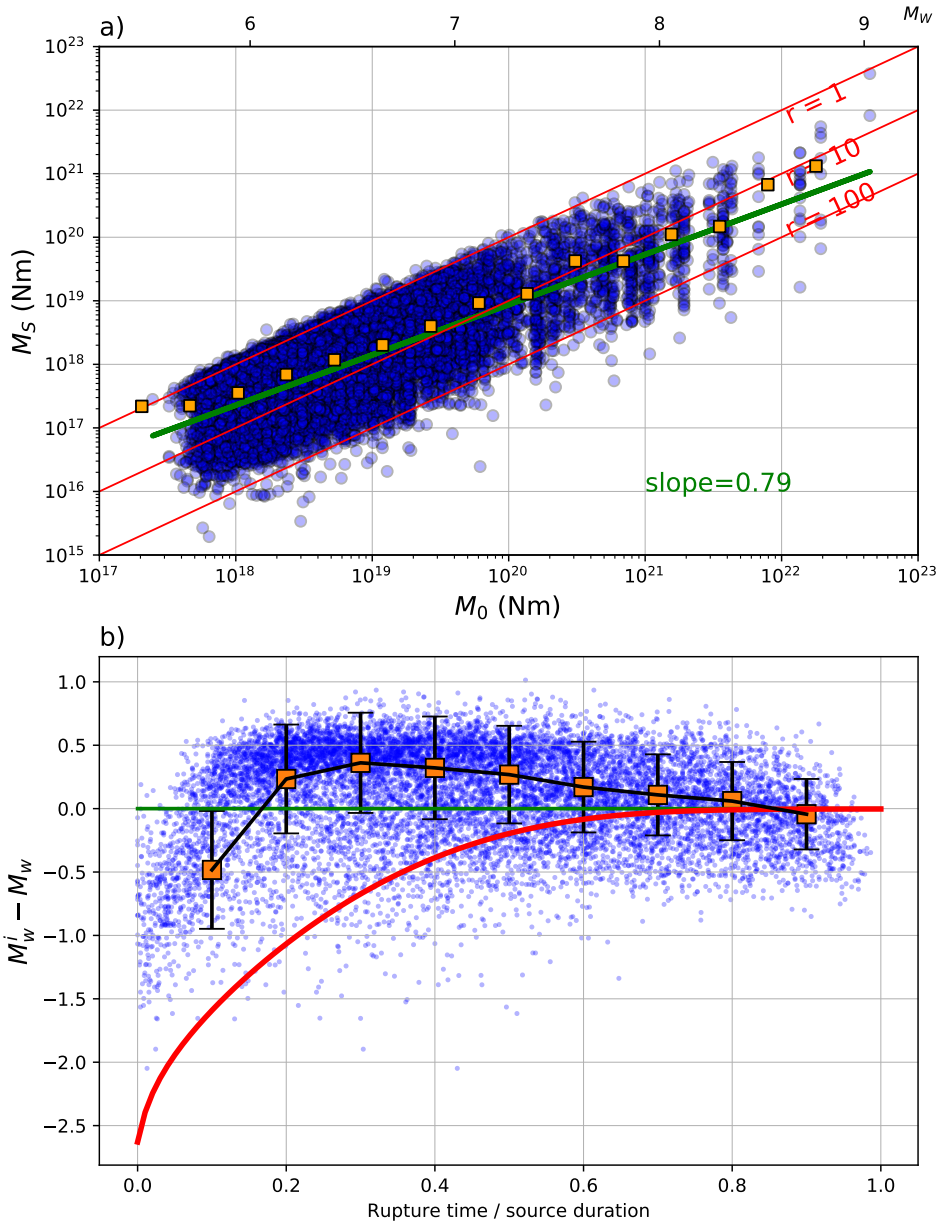

Figure S3. EEW application on the entire SCARDEC data set. (a) Subevent and main event with the new threshold criterion (blue dots), the new regression (green line), the new moment-binned median (orange square), similar to Figure 3(a) of main manuscript. (b) Residuals between magnitudes estimated at each subevent M_W^i and true main event moment M_W, similar to Figure 4 of the main manuscript for the entire SCARDEC dataset. The median in each time bin and 2 standard deviation tall error bars are shown. Moment scaling is that found in (a) Between 20% and 40% of STF duration, our algorithm overpredicts the magnitude by 0.5 unit, we explain this because the largest subevents occurs generally early in the rupture while the first subevent tends to be smaller (see Figure S4 a).

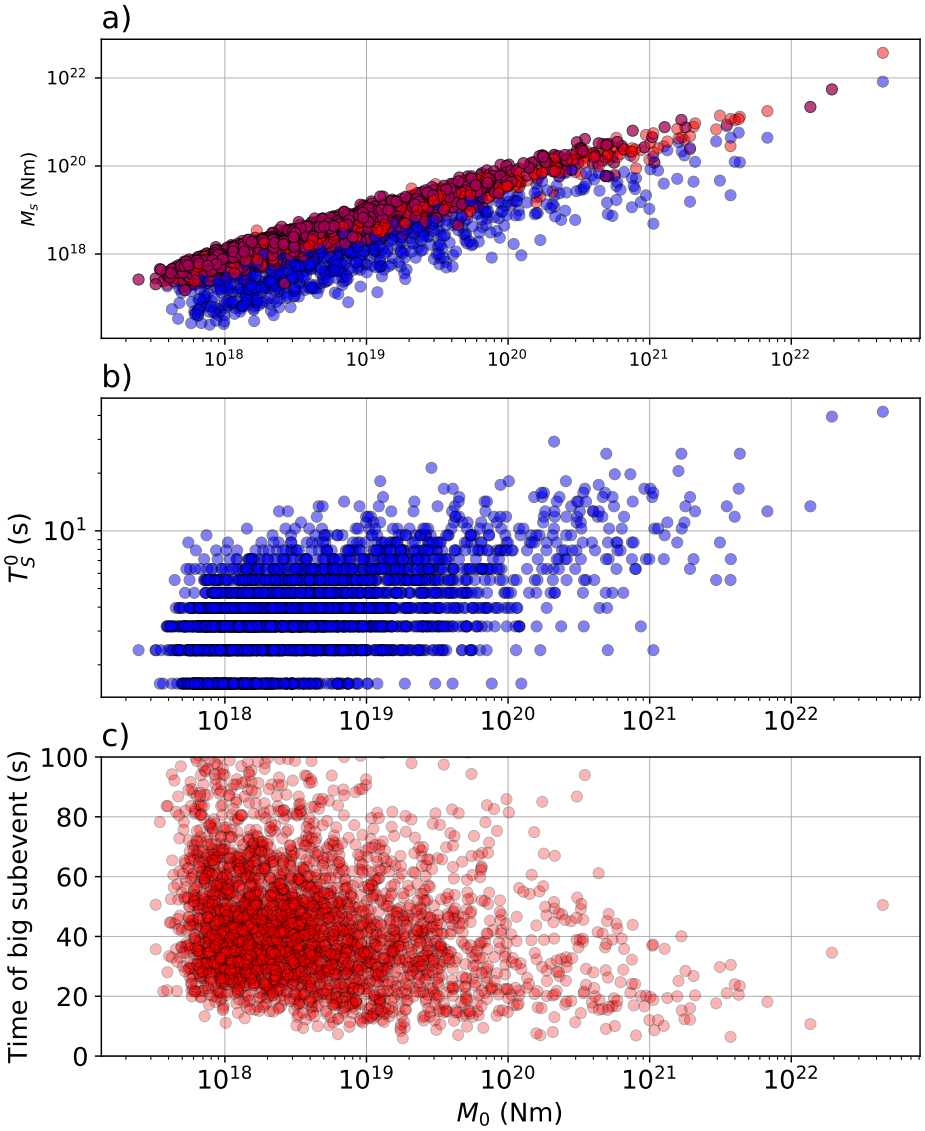

Figure S4. (a) Subevent moment against main event moment of the largest (red) and first (blue) subevent. The first subevents shows are not the largest, which explains why the estimates of magnitude from the first subevent is slightly lower. (b) Duration of the first subevent correlates with main event moment, which suggests that the duration of the first subevent could be used for early magnitude estimates as well, in relation to the results of *Iio* [1995]. (c) Time at which the largest subevent occurs: the small earthquakes have their biggest subevent in about the middle of the earthquake, probably because they may only have one or two subevents, but early on during the rupture for larger earthquakes.

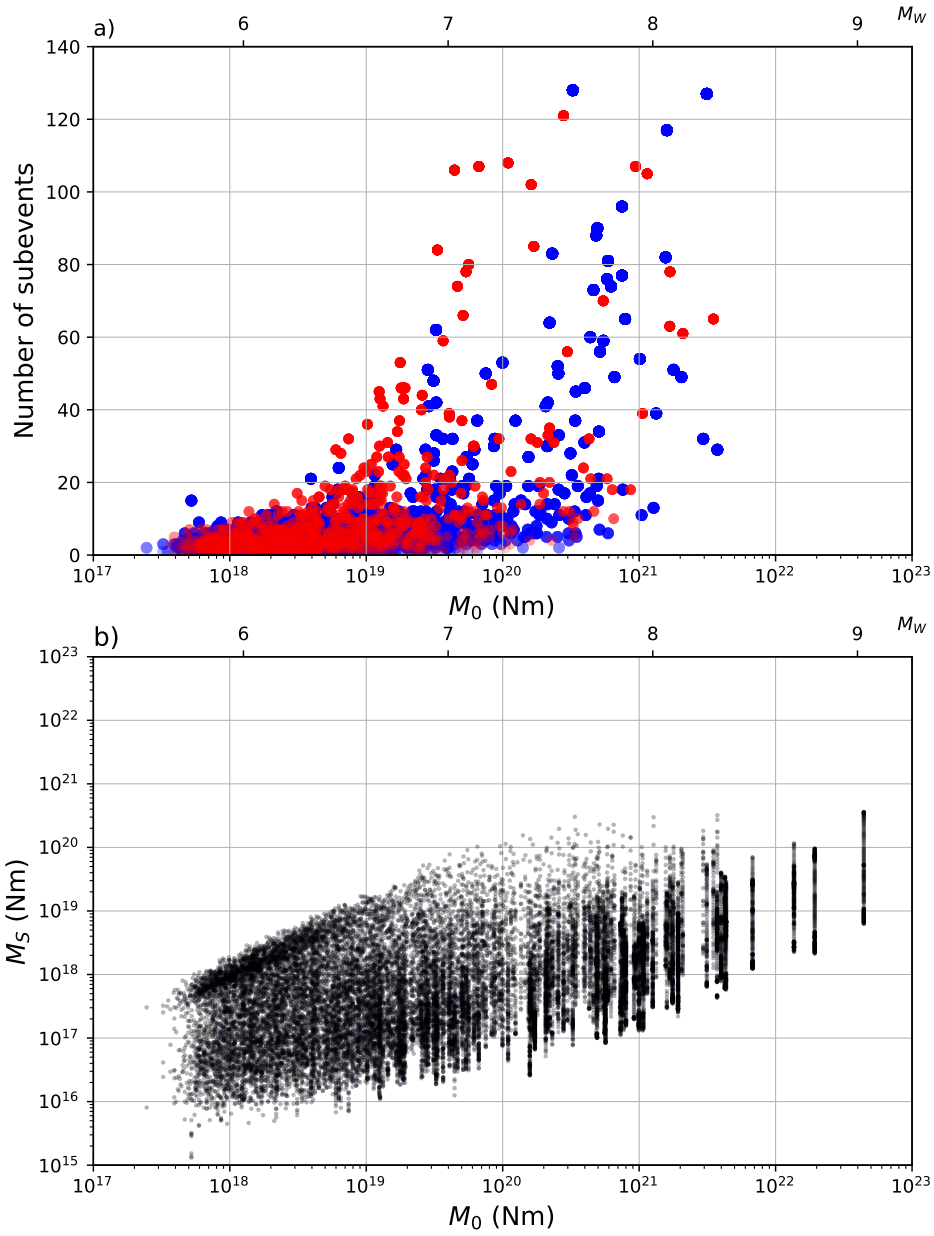

Figure S5. Analysis with triangle kernels with a similar style and notation as in Figure 2(a) and Figure 3(a) of main manuscript. (a) Number of subevents against main event moment for strike-slip (red) and dip-slip (blue) earthquakes. (b) Subevent–main event moments. We see: i) still a growth of the number of subevents with main earthquake moment, ii) the fit to the largest magnitude becomes likely spurious because there are too many subevents (probably not well separated in time and thus over fitting the STF). Corresponding moment magnitude shown in the top axis label.

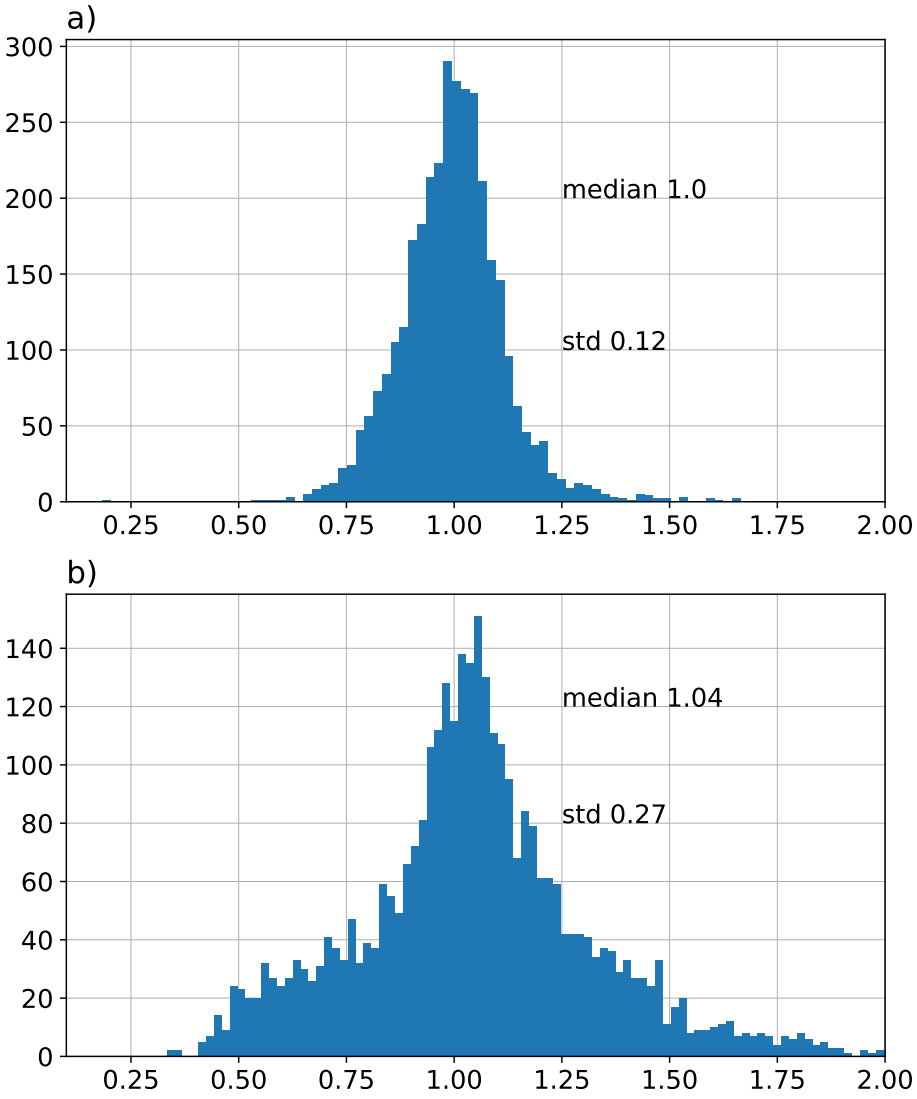

Figure S6. Histograms of the ratios between the moments of our reconstructed STFs and the SCARDEC moments for (a) Gaussian kernels and (b) triangle kernels. Median show no systematic bias, but the standard deviation show greater uncertainty (and thus overall poorer fit).

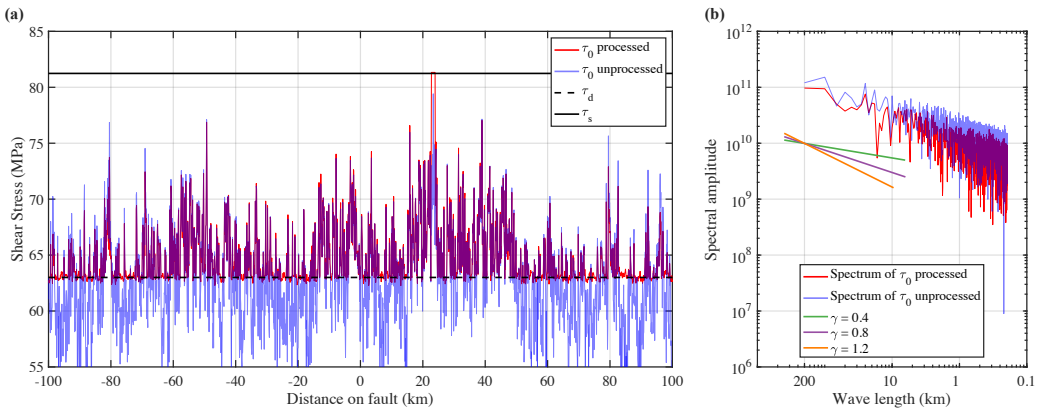

Figure S7. (a) One example of the stochastically generated pre-stress distribution. Blue line shows the pre-stress distribution without any post-processing, while red line is the pre-stress after processing (non-negativity constraint) and used in the dynamic simulation. Black solid and dashed lines are the uniform fault strength and dynamic friction, respectively. (b) Comparison between the processed (red) and unprocessed (blue) pre-stress distributions in the frequency domain. Green, purple and orange lines show the references of different γ exponents of 0.4, 0.8 (used in this study) and 1.2, respectively.

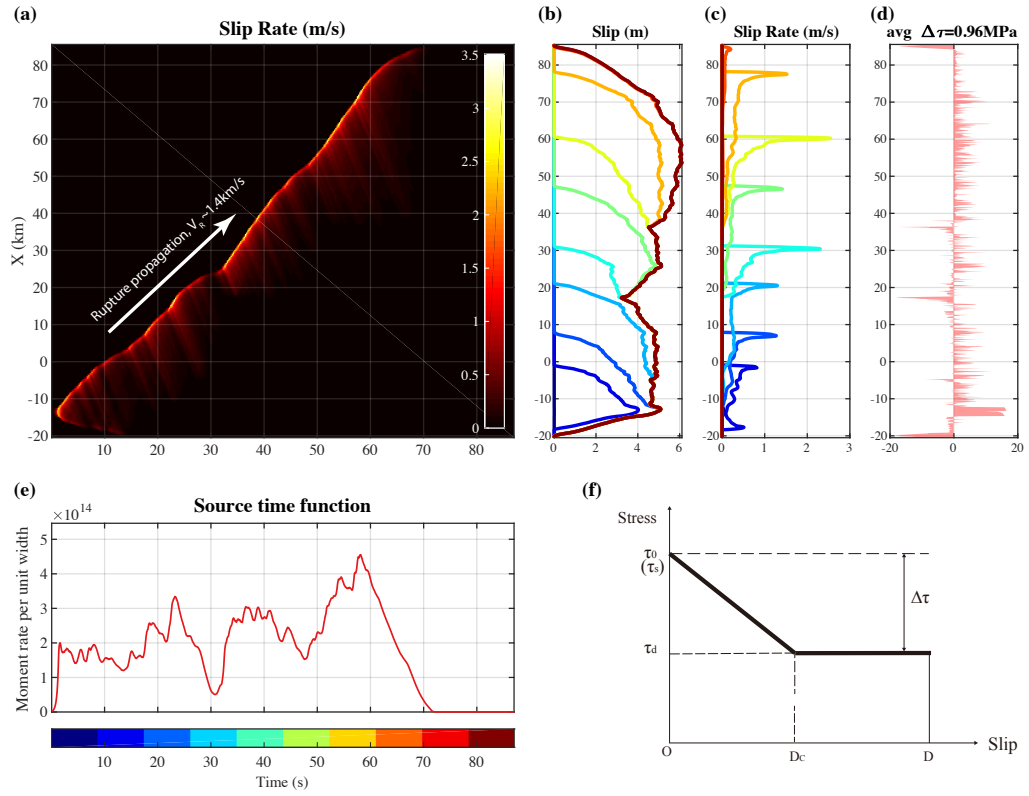

Figure S8. (a) Shows the space-time evolution of the on-fault slip rate, (b) the slip profile across the fault, (c) the slip rate as a function of time, (d) the stress change on the fault, (e) the final STF, (f) the linear slip weakening.