www.sciencemag.org/content/357/6357/1277/suppl/DC1

# Supplementary Materials for

## The hidden simplicity of subduction megathrust earthquakes

M.-A. Meier,* J. P. Ampuero, T. H. Heaton

*Corresponding author. Email: menandrin@gmail.com

Published 22 September 2017, *Science* **357**, 1277 (2017)
DOI: 10.1126/science.aan5643

**This PDF file includes:**

Supplementary Text S1 to S7
Figs. S1 to S13
References

**Other Supplementary Materials for this manuscript include the following:**
(available at www.sciencemag.org/content/357/6357/1277/suppl/DC1)

Data File S1 (iSTF_vs_medSTF.zip): A set of 116 png files showing the individual source time functions and how they compare to the median source time functions in seven magnitude bins

## Supplementary Text

### S1: Kinematic source inversion by Ye et al., 2016a

Ye et al., 2016a have applied a uniform kinematic source inversion method to a large number of large subduction zone thrust earthquakes of the past three decades. They use P-wave (and in a few cases SH-) displacement records in the frequency band 0.005-0.9Hz, and theoretical Greens functions assuming a 1D velocity model. Each earthquake is described as a rectangular plane array of sub-faults. On each sub-fault slip rate is parameterized as a sequence of overlapping triangles whose amplitudes are optimized to best explain the observed waveforms in a least-squares sense. The method imposes the following constraints: slip rate positivity, a higher bound on rupture velocity and rise time, and linear spatial smoothing (*18*). The STF of the 2004 $M_w$ 9.1 Sumatra earthquake is an exception in that it has not been inferred with the same code, but has been taken from (*48*).

### S2: Sampling log-normal distributions

Supplementary Fig. S1 shows how well synthetic input STFs are recovered using different distribution center estimates if residuals are log-normally distributed. We construct 1e4 synthetic STFs for each of the seven magnitude bins using the functional form $y_{fit} = \mu*t*\exp(-1/2*(\lambda t)^2)$. We choose $\mu$=1.25e18Nm/s$^2$ and optimize $\lambda$ to fit the functional form to the median observed STF in each bin. We multiply each STF with a random log-normal error term, $10^{\varepsilon}$, with $\varepsilon = N(0, \sigma^2)$, and where $\sigma^2$ has been measured from the real STF relative to the fitted curve of each bin. We then estimate the distribution center at each point in time by taking the median, the arithmetic mean and the arithmetic mean of the log-transformed amplitudes (Fig. S1). As expected, the median and the mean of the log-transformed values give an unbiased estimate of the distribution center, while the arithmetic mean is strongly biased towards higher amplitudes. This is because log-normal distributions are skewed towards large values, i.e. the deviations above the center are larger then the ones below the center. This bias is stronger for larger magnitude bins, which then results in an apparent magnitude-dependence of STF growth (Fig. S1B). In this study we use median statistics because they are reliable estimates independent of the distribution, and are particularly suitable for log-normal distributions.

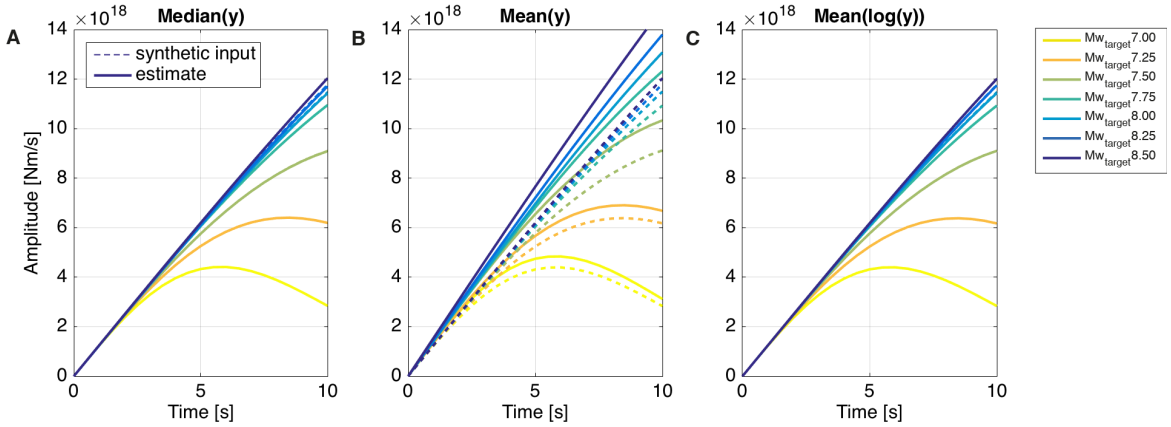


**Fig. S1. Synthetic test for sampling log-normal STF amplitude distributions.** At each point in time the distribution center is estimated using (**A**) medians, (**B**) arithmetic means and (**C**) arithmetic means of log-transformed values (right). While the median and log-means accurately recover the synthetic input function, the arithmetic mean leads to a biased estimate with over-estimated amplitudes and an apparent magnitude dependence of the initial slope.

## S3: Comparison of STF from three independent data sets

Earthquake rupture source inversions are generally ill-posed and non-unique, and source inversion results can be affected by method-specific biases that are difficult to quantify. In order to evaluate if the STF models we use are affected in this way we make use of two alternative data sets that contain sufficiently large numbers of STFs, and that were also obtained by applying a consistent method to a large number of events. The two data sets are those of Hayes 2017, (*20*, in the following referred to as "Hayes17"), and that of Vallée et al., 2011 (*21*, "Vallée11"). We refer to the data set of Ye et al., 2016a as "Ye16".

We first identified matching STFs in the three data sets in order to do a one-to-one STF comparison. We then repeated the main part of our STF analysis with the two alternative data sets, for a systematic robustness analysis of our results. This analysis reveals that the observations we report are stable across data sets and inference methods. In particular, all three data sets have i) linear and magnitude-independent moment rate growth, ii) a simple near-triangular average shape that - after normalization - is virtually identical for all magnitudes in the studied range, and iii) log-normal deviations from a best fit model.

### Data sets

All three STF data sets are derived from teleseismic data, but use different methods to infer the STFs. The Hayes17 STFs were generated with a kinematic source inversion procedure (REF) similar to that applied by Ye et al., 2016a, but Hayes17 used a different slip

parameterization with smaller spatial grid cells and a single cosine slip rate function on each sub-cell, leading to a non-linear inverse problem. This is in contrast to the linear inverse problem based on multi-window source parameterization used for the Ye16 data set. 62 of the events in the Ye16 data set are also contained in Hayes17 data set. The missing events are mostly events with Mw<7.5, which is a range of magnitudes outside the focus of the Hayes17 data set. The STFs from Vallée11, on the other hand, are computed via de-convolution of synthetic Greens functions, and in a fully automated fashion. For each recording station an apparent STF is computed, which is then averaged to obtain the actual STF. Of the 116 STFs in the Ye16 data set we found 99 matching events in the Vallée11 data set.

**Comparison of individual STFs**

Despite fundamental differences in how the STFs are computed, they only differ in terms of second-order details for a majority of cases. Fig. S2 shows all individual 116 STFs of the Ye16 data set, along with the matching STFs from the other two data sets. Apart from a small number of cases with large differences, such as the Mw7.2, 2007 Santa Cruz earthquake, or the failure of the Vallée11 method for the Mw9.1 2004 Sumatra earthquake, the STF from the different methods are very similar. For most earthquakes, the first-order STF shape is consistent between the three methods, and they have comparable amounts of high frequency oscillations. In some cases, the Vallée11 method seems to oversimplify the STF, in that it suggests a very simple STF while the other two methods suggest a more complicated rupture evolution (e.g. Mw7.2 2003 New Zealand, Mw8.6 2005 Sumatra). The Hayes17 method often seems to produce long tails (e.g. Mw8.8 2010 Chile, Mw7.8 2010 N-Sumatra). Since the termination of the ruptures is typically not well constrained (see discussion in main manuscript), these tails are rather unrealistic. Other than that, the STFs are very similar.

The fact that kinematic source inversions lead to highly similar STFs to those from a de-convolution approach is remarkable, and speaks to the general stability of source time function estimates. This supports the general perception that the STF is one of the most robust features of finite source descriptions because it is a relatively direct mapping of the displacement waveforms. This is particularly true for situations where the dip is well constrained, such as in the large subduction earthquakes considered in this study. The fundamental ill-posedness of source inversion problems may manifest itself more strongly in other aspects of source inversions, in particular in the spatial distribution of slip.

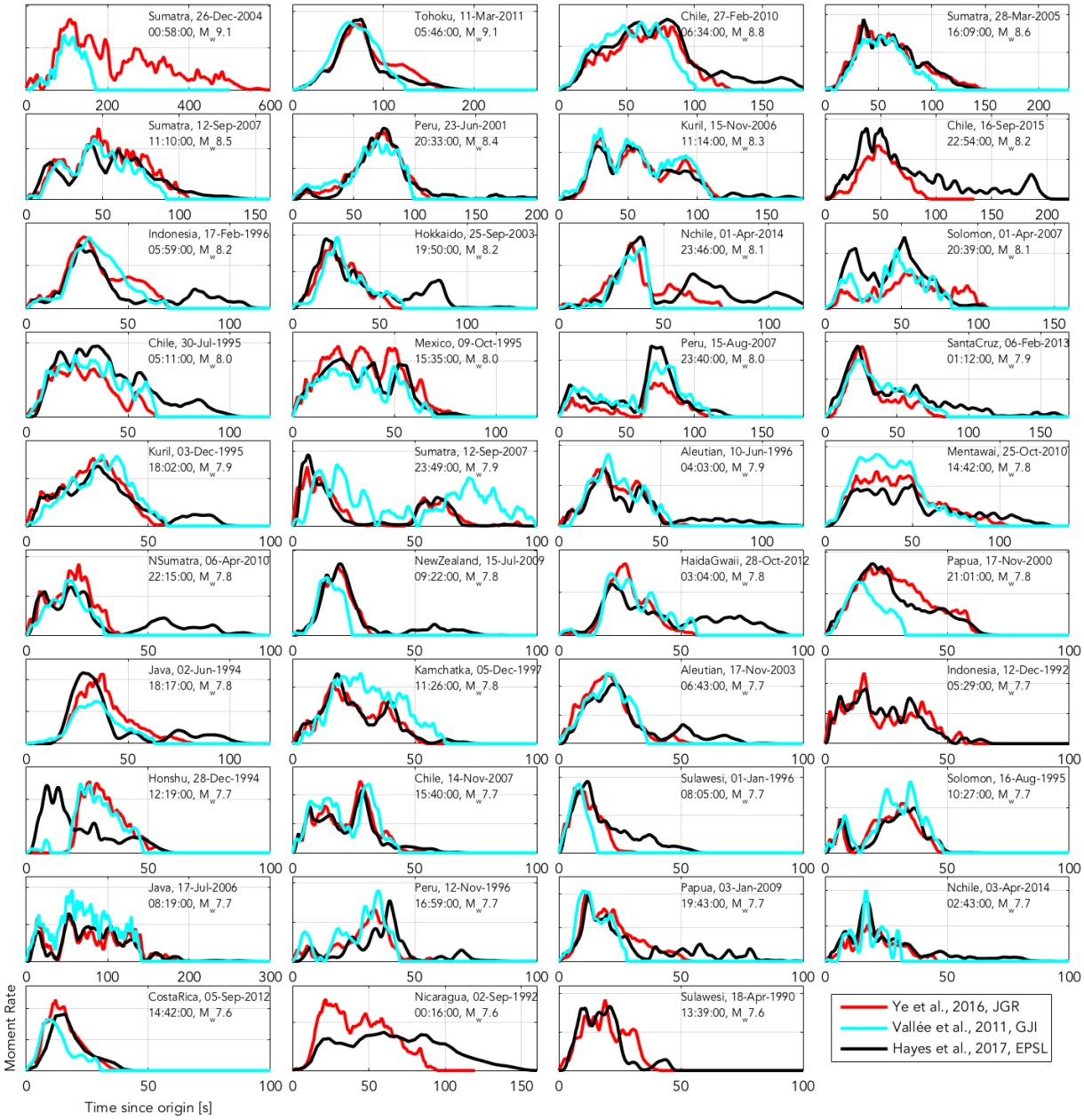


**Fig. S2: Comparison of individual STFs from 3 different methods.** Of the 116 STFs from the Ye16 data set, 99 are also contained in the Vallée11 data set, and 62 in the Hayes17 data set. The STFs generally agree very well across methods and show the same first order shape. No method appears to be producing systematically steeper onsets, or systematically more skewed shapes. The Hayes17 method often produces rather long low amplitude tails that are not supported by the other two methods.

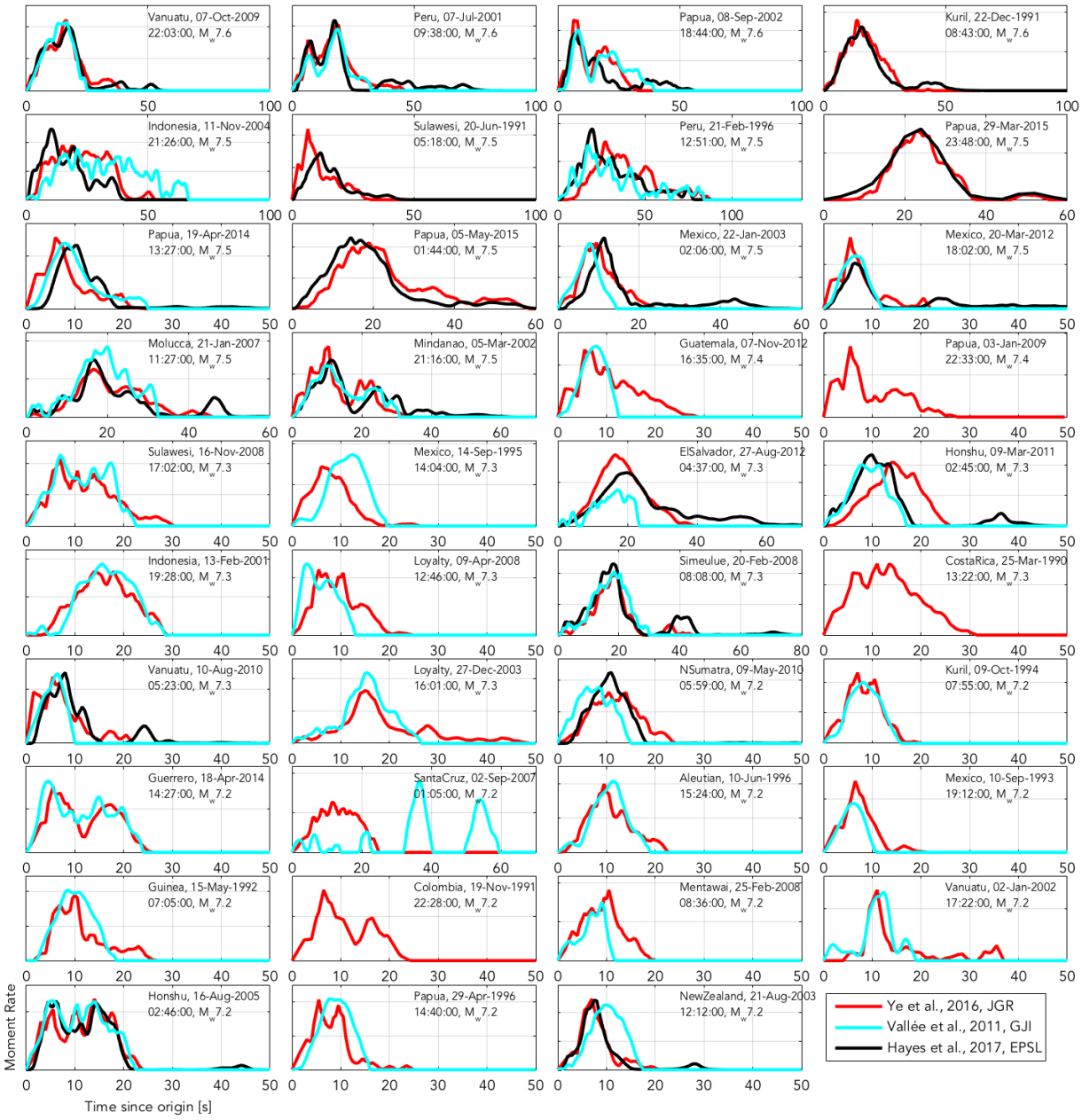


**Fig. S2, continued.**

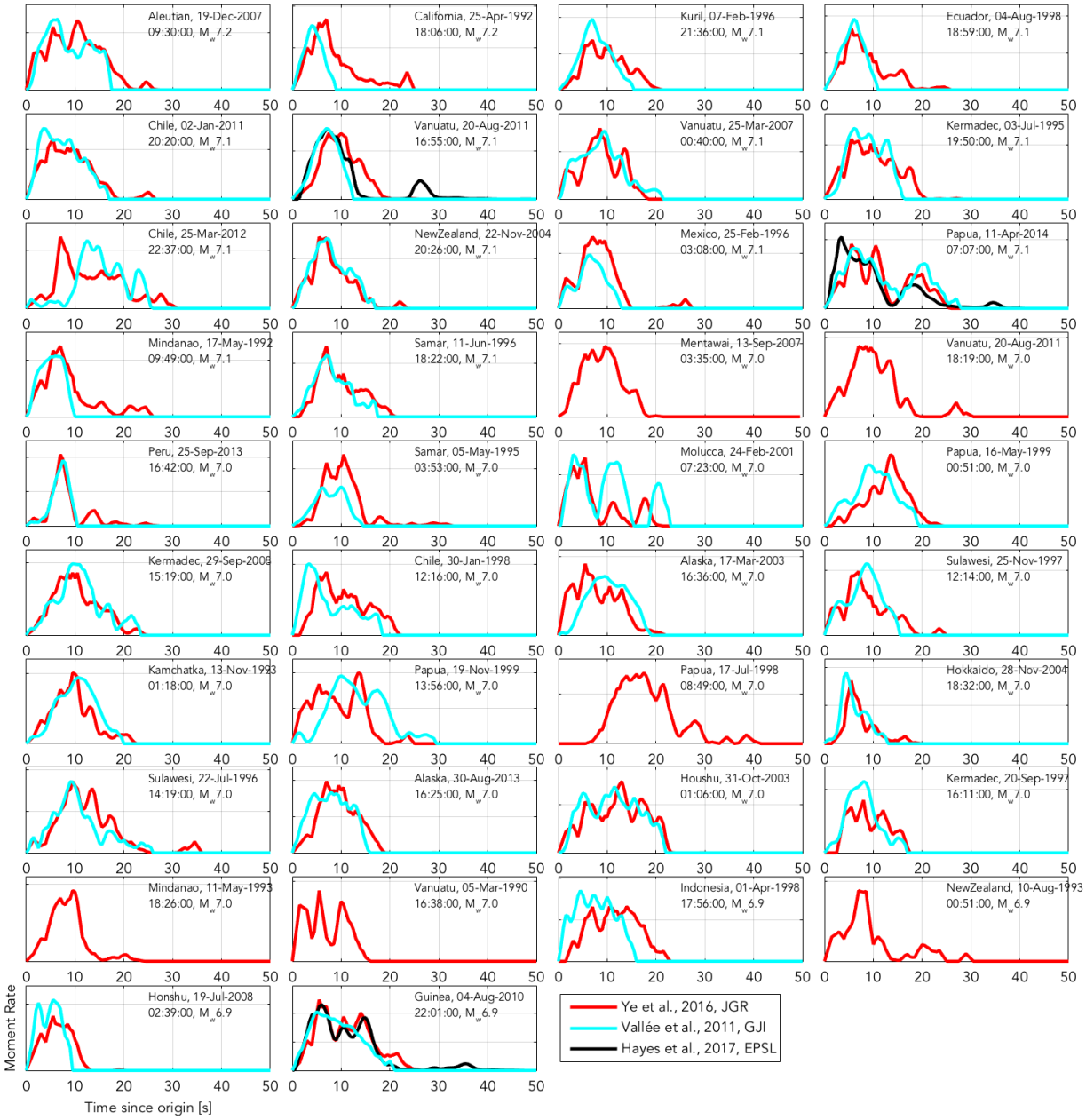


**Fig. S2, continued.**

**Inferring the median STF with the three data sets**

We repeated the main body of our STF analysis with the matching STFs of the Vallée11 (99 STFs) and the Hayes17 (62 STFs) data sets. The median STFs in the seven magnitude bins for all 3 data sets lead to the near-triangular typical STF shape (Fig. S3A-C). The largest difference is that the median Hayes17 STFs (Fig. S3C) are more skewed towards early times and feature longer tails than the other two data sets. Apart from this discrepancy, all three data sets show the same main features that our analysis reports:

- Larger magnitudes start neither more nor less impulsively than smaller ones. Large earthquakes started as smaller ones do, but do not stop as quickly. This means, there will not be strong rupture predictability.
- The moment rate growth is approximately linear all the way until peak moment rate is reached; this is in strong contradiction to often-used standard models, which predict quadratic growth.
- After normalization, the median STF for different magnitudes collapse onto a single near-triangular shape, suggesting scalability between small and large ruptures.

The largest difference is that, while the median STFs from the Ye16 and the Vallée11 data sets are fairly symmetric, those of Hayes17 are somewhat more skewed towards early times and feature longer tails (Fig. S3 E&F). This asymmetry likely stems from the prominent low-amplitude tails we have observed with the individual STFs of Hayes17 (Fig. S2). In order to verify this presumption we truncated the Hayes17 STFs when the corresponding STFs from the other two data sets ended. We measure the time of the last non-zero sample of the Ye16 and, if available, Vallée11 STFs and use the longer of the two for the truncation. In other words, we suppress the tails of the Hayes17 STFs if they are not supported by at least one of the other two methods. Note that for the truncation we use the duration estimate of when the STFs reach amplitude 0, rather than the centroid times, to be conservative and truncate as little as possible.

Indeed, after the truncation, the median STFs (Fig. S3 G&H) are no longer more asymmetric than those of the other data sets, and share the same near-triangular median STF shape. This confirms that it is indeed the badly constrained low amplitude tails that caused the stronger asymmetry. The truncated Hayes17 STF on average reach peak moment rates between 35-55% of total rupture durations, as is the case with the other two data sets. For normalizing the STFs in Fig. S3H we used twice the centroid time, rather than the truncation duration, in order to be consistent across all data sets.

Note that, owing to the different sizes of the three data sets, and their different magnitude ranges, the median STFs cannot always be compared directly. The smaller number of STFs in the Hayes17 data set (roughly half the size of the Ye16 data set) means that the 20 nearest neighbors of our binning scheme cover a wider magnitude range. Furthermore, the lower magnitude bins are not well represented in the Hayes17 data set, because the data set has few events with Mw<7.5. One should consider the median Mw for the bins, rather than the target Mw (both are given in the figure legends).

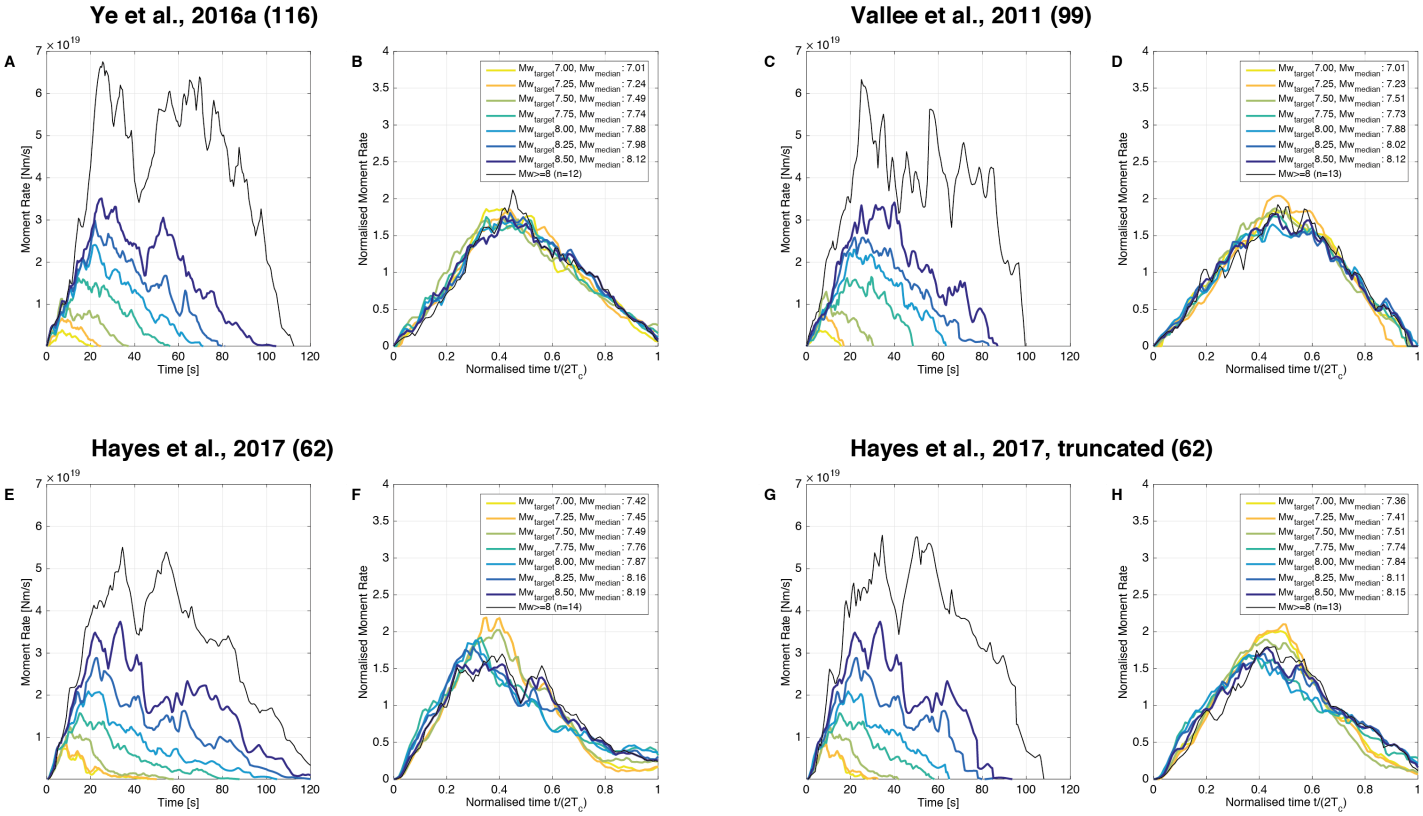


**Fig. S3: Median STFs for the different data sets**. Equivalent plots to Fig. 2 in main manuscript. The number of STFs in each data set is given in title. (A & B) Main data set used in this study, by Ye et al., 2016a. (C & D) Data set of Vallée et al., 2011, GJI. (E & F) Data set of Hayes et al., 2017, EPSL. (G & H). Data set of Hayes et al., 2017, EPSL with badly constrained low-amplitude tails removed. All data sets exhibit linear and magnitude independent STF growth, near-triangular STF shapes, and magnitude-independent normalized shapes. For detailed caption, see Fig. 2 in main manuscript.

### Parameterization of individual STF with the three data sets

When we fit the functional form $y_{fit} = \mu * t * \exp(-1/2 * (\lambda t)^2)$ to each individual STF, we find very similar optimal parameters across the three data sets (Fig. S4). After seeing the general similarity of the individual STFs (Fig. S2) this is not surprising. This exercise shows that the fits between all three data sets are not just similar, but indeed statistically consistent in that the difference between individual bootstrap realizations are larger than the differences between data sets. All three data sets are consistent with the hypothesis that the initial moment rate growth is not different for smaller and larger events. Furthermore, in all data sets the scaling exponent between seismic moment and duration is higher than 1/3 (Ye16: 0.41, Vallée11: 0.46, Hayes17: 0.48, Hayes17, truncated: 0.46).

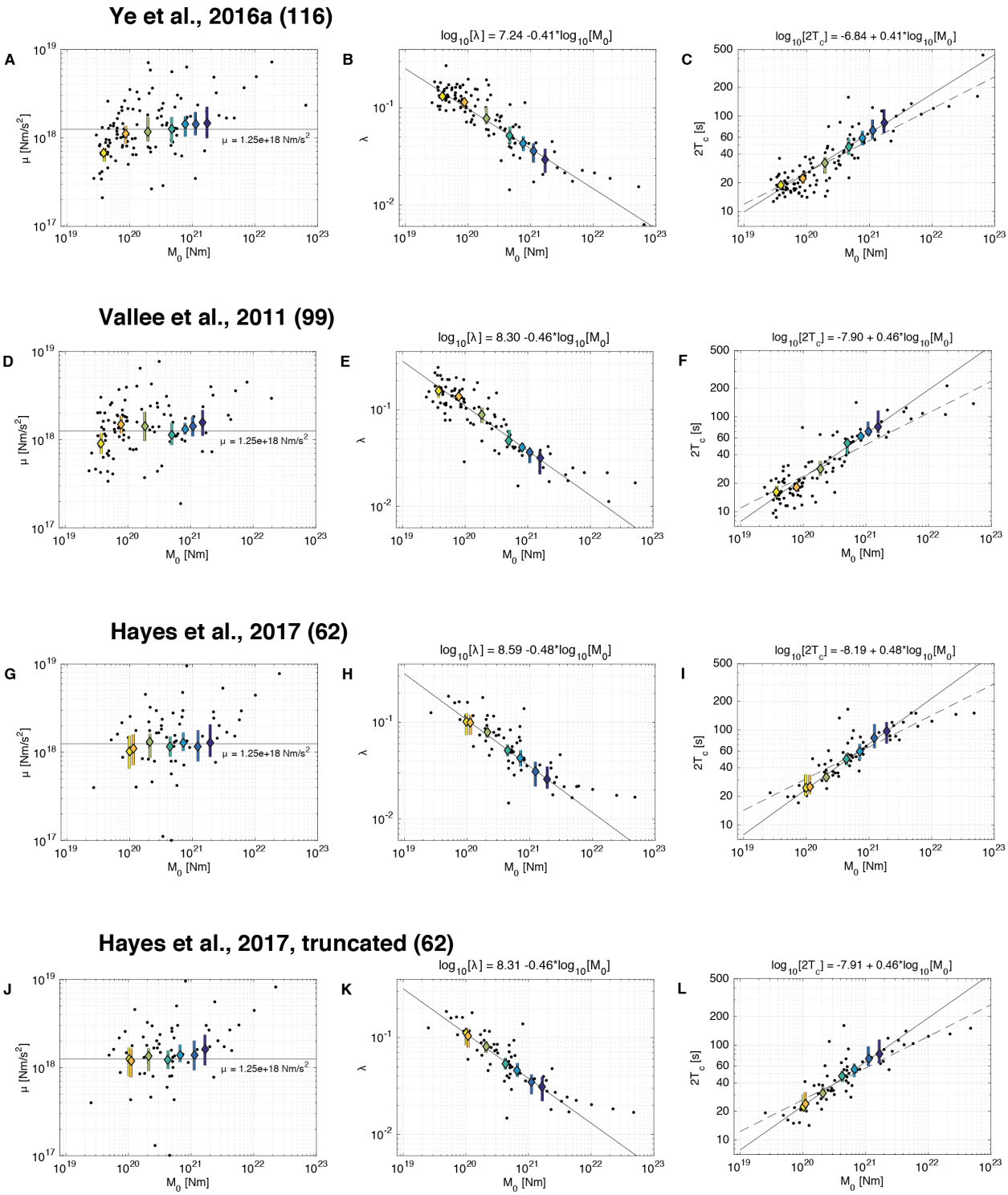


**Fig. S4: Optimized parameters for the different data sets**: Ye16 (A-C), Vallée11 (D-F), Hayes17 (G-I) and Hayes17 with truncated tails (J-L). The trends from all data sets are consistent with each other in that the variability between data sets is smaller than the bootstrap variability for any individual data set. For detailed caption, see Fig. 3 in main manuscript.

**Deviations from best-fit model**

Similarly, if we compare the deviations of the individual STFs from the best-fit model (Fig. S5), the results from the three data sets are consistent. In particular, the deviations are multiplicative and Gaussian, with standard deviations that do not change substantially between magnitude bins. The average standard deviations are slightly larger for the Vallée11 (0.49) and the Hayes17 data (0.42, 0.42) than those of the Ye16 data (0.38).

Most of the Hayes17 STFs have negative residuals at the very beginning, suggesting that the observed STFs have lower amplitudes than the best-fit STF. This likely is a consequence of the cosine slip rate parameterization employed by Hayes17. While this parameterization requires only a single parameter per sub-fault, this comes at the cost of reduced flexibility (*49*). The shape of the slip rate function is prescribed and cannot be adapted to best match the observed shape. The multiple-time-window approach of Ye et al., 2016a, on the other hand, comes with more flexibility to match a wide range of observed shapes. Since both the multiple-time-window approach of Ye16, as well the free form approach of Vallée11 agree well with the initial part of the functional form, we conclude that those descriptions are more reliable for the purpose of analyzing STF onsets.

**Summary**

The systematic comparison of the STFs from the three different data sets shows that the STFs are well constrained, with very similar properties across the data sets. This supports the common perception that the STFs are among the best-resolved aspects of earthquake source descriptions. In particular the fact that fundamentally different methodologies (kinematic source inversion vs. Greens function deconvolution) lead to very similar STF demonstrates the robustness of the estimated STFs. We therefore conclude that the observations we report (near-triangular STF shape, linear and magnitude-independent moment rate growth, and multiplicative Gaussian moment rate deviations) are stable and are not systematically affected by the properties of the inversion methodology of Ye et al. 2016a.

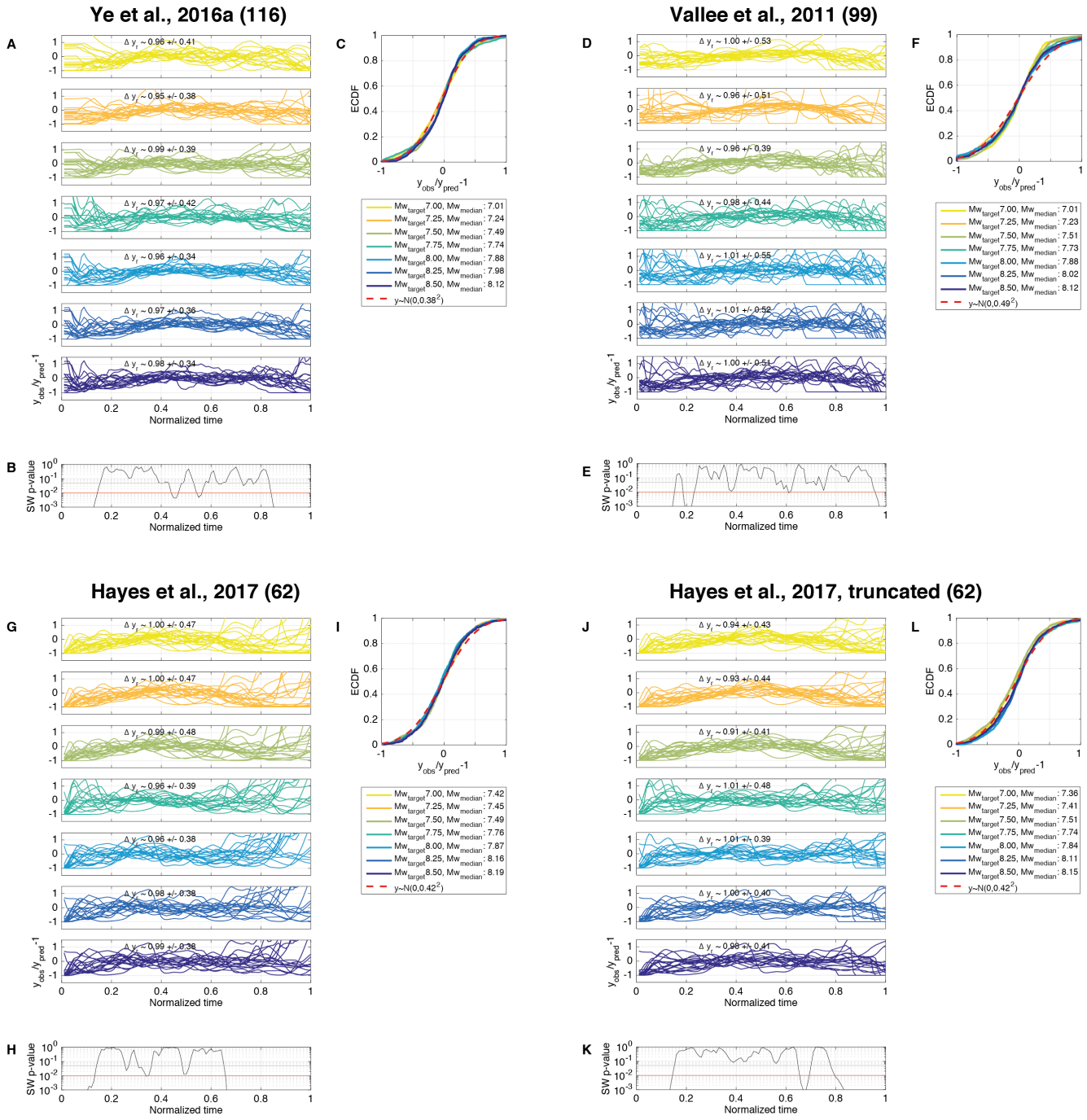


**Fig. S5: Residuals from best-fit STFs for the different data sets**: Ye16 (A-C), Vallée11 (D-F), Hayes17 (G-I) and Hayes17 with truncated tails (J-L). For detailed caption, see Fig. 4 in main manuscript.

## S4: Synthetic resolution tests for inversion slip parameterization

There is the possibility that some of the observed STF patterns are affected by artifacts from the slip parameterization used in the kinematic source inversion procedure. For example, the imposed higher bound on rupture velocity, that defines the earliest point at which any given subfault can start rupturing, could bias the shape of the initial part of the MRF. In order to understand the effects of the imposed fault and slip parameterizations on the STFs we perform synthetic tests with a self-similar pulse as input model.

We test the effect of source parameterization on source recovery. This test does not address the effects introduced by wave propagation and limited observational sampling at the Earth's surface. We build an earthquake source scenario based on the self-similar pulse model by Nielsen and Madariaga (2003). The model provides a continuum space-time distribution of slip rate, $v(\boldsymbol{x}, t)$. In order to avoid the slip rate singularity at the rupture front we replace the amplitudes during the first 0.5s with a linear ramp (Fig. S2A). In the source inversion, the modeled slip rate in each fault cell is spatially uniform and its time-dependence is parameterized as a linear combination $\tilde{v}(t) = \sum \tilde{v}_k b_k(t)$ of triangular temporal basis functions $b_k(t)$. Our goal here is to determine the coefficients $v_k$ that provide the best fit to the true slip rate $v(\boldsymbol{x}, t)$ in a least-squares sense. This can be viewed as an idealized source inversion problem in which slip rate data is available everywhere on the fault at all times, as if the fault surface were covered by a very dense seismic network. This scenario provides the best possible model one could hope for using the assumed source parameterization. Note that this test only assesses the effect of the slip parameterization and does not cover other aspects of the inversion procedure, such as regularization, or the limited spatial coverage of recording stations.

The problem is solved by applying the following procedure in each cell:

1. Compute the cell-averaged slip rate: $\bar{v}(t) = \iint v(\boldsymbol{x}, t) dx^2$.
2. Compute the vector $\bar{\boldsymbol{v}}$ whose *i*-th component is $\bar{v}_i = \int \bar{v}(t) b_i(t) dt$
3. Compute the matrix $A$ whose components are $A_{ij} = \int b_i(t) b_j(t) dt$.
4. Find the vector $\tilde{\boldsymbol{v}}$ of coefficients $\tilde{v}_k$ that minimize the norm of residuals $\chi^2 = \int \left(\sum \tilde{v}_j b_j - \bar{v}\right)^2 dt$ by solving the linear problem $A\tilde{\boldsymbol{v}} = \bar{\boldsymbol{v}}$

Derivation: Minimization of the cost function $\chi^2$ with respect to the coefficients $\tilde{v}_k$ requires $\frac{\partial \chi^2}{\partial \tilde{v}_i} = 0$ for all *i*. This condition can be written as

$$\frac{\partial \chi^2}{\partial \tilde{v}_i} = \int \left(\sum \tilde{v}_j b_j - \bar{v}\right) b_i dt = \sum_j \int b_i b_j dt \times \tilde{v}_j - \int \bar{v} b_i dt = 0$$

That is $A_{ij}\tilde{v}_j - \bar{v}_i = 0$.

Fig. S6A shows how the synthetic local slip rate function on a sub-fault is approximated by the sum of a sequence of overlapping triangles. The initial sharp peak is not well approximated because the width of the triangles is fixed to 4s, which limits the temporal resolution. Fig. S2B shows the global fitted source time function. Although the fitted function is relatively rough, owing to the coarse spatial resolution and to the relatively large width of the triangles, the overall shape of the synthetic slip rate function is well recovered. In particular, there is no apparent linearization, which suggests that the observed linear moment rate growth is a real source characteristic, rather than an artifact from the slip parameterization.

Furthermore, we find that the coarse spatial grid leads to a very poor spatial resolution in the first 5-10s. During the first ~4s only the sub-patch that hosts the hypocenter can be used to explain the observed waveforms. This can be seen from the kink of the STFs in Fig. 1A at ~4s which is when the next closest set of sub-patches is "activated" by the passing rupture front. After that, the number of activated sub-patches grows rapidly and the slip rate parameterization becomes highly flexible. Thus, while the temporal resolution is much better than the spatial resolution, inferred STF amplitudes at <10s should be interpreted with caution. This is particularly important for the smallest events of the data set (Mw~7.0) for which this time span corresponds to a large fraction of the total rupture duration.

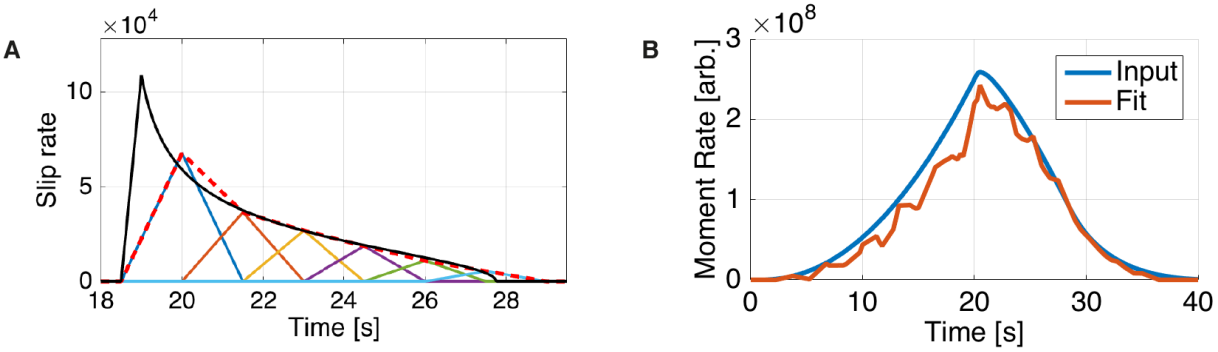


**Fig. S6: Comparison of synthetic input signals and recovered output.** (A) Synthetic input slip rate function for single sub-fault patch with singularity replaced by linear ramp (black line) and approximation by sum of sequence of six isosceles triangles (red dashed line). (B) Synthetic input moment rate function (blue line) and optimized solution (red line) from adding slip rate functions on individual sub-fault patches.

## S5: Source time functions for 12 Mw>=8 events

Arguably the most important earthquakes, with respect to their societal relevance, are the 12 Mw>=8 earthquakes. These events are not accurately represented by any of the magnitude bins (Fig. 1C). The one feature that all of these events have in common is that they, at some point, develop moment growth rates that are much larger than the typical STFs growth rates. During such periods, either the rupture area or the slip rates or both must grow at a rate that is much higher than what is typically observed. The median initial growth rate measured with all 116 STFs is ~1.25x10$^{18}$Nm/s$^2$ (Fig. 3A). This rate also corresponds to the slope of the median STF curves in Fig. 2A. Fig. S7 shows these 12 STFs and highlights periods during which moment acceleration exceeds 4 times the moment acceleration of the median STF. Although there are also smaller events that temporarily exhibit such high growth rates (e.g. 2007 Mw7.9 Sumatra, 2012 Mw7.5 Mexico, 2010 Mw7.3 Vanuatu) this behavior is much more rare in those magnitude ranges. In normalized form, however, these events follow the same shape as the smaller events (Fig. 2), and their residual ratios with respect to the simple base model are on the same order as those of the smaller events.

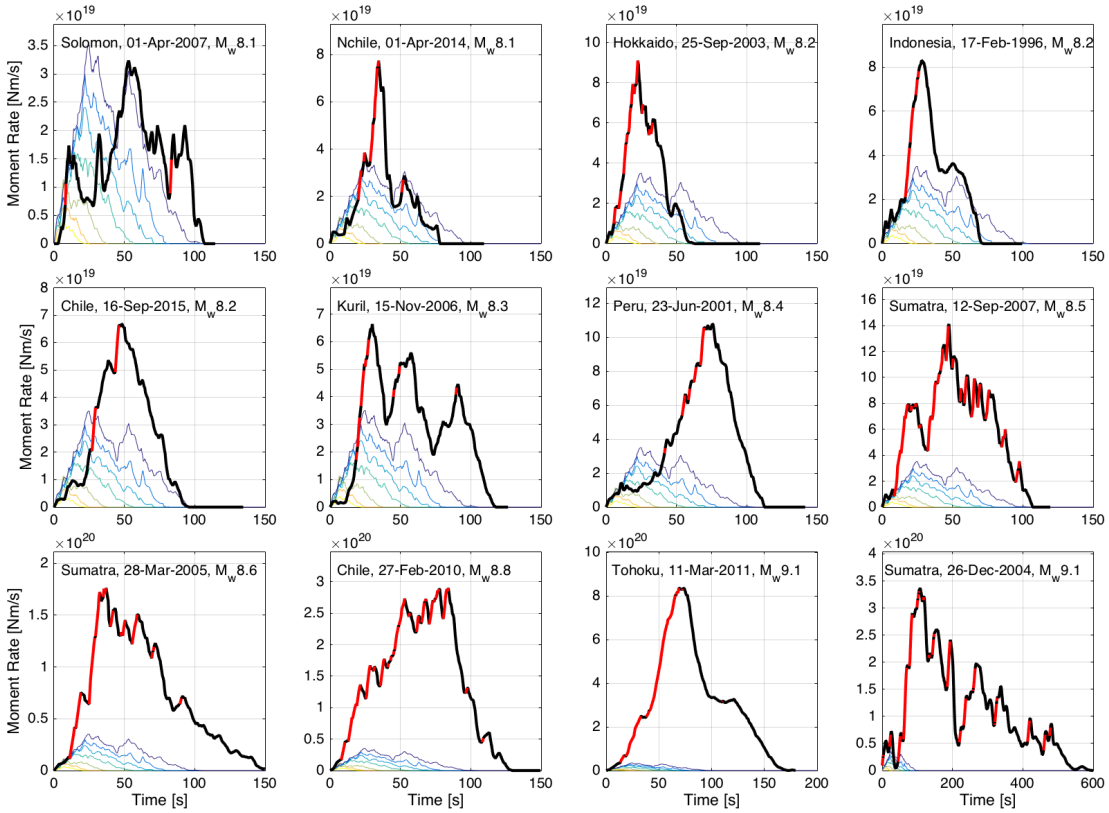


**Fig. S7: STFs of the 12 events with Mw>=8.0 and median STF in 7 magnitude bins.** Red lines mark episodes in which moment acceleration exceeds 5x10^18Nm/s^2, i.e. 4 times the median growth rate of 1.25x10^18Nm/s^2. Identical figures for all 116 earthquakes are provided as electronic supplement.

## S6: Moment-duration scaling from alternative data sets

Given the large scatter in the relation between observed seismic moments and rupture durations, a wide magnitude range is necessary to infer a reliable trend between the two quantities. The source durations of the data set used in this study are consistent with a relation of the form $T \sim M_0^p$ with an exponent p=1/3 ($R^2$=0.76) or p=1/2 ($R^2$=0.70), Fig. S8. While the former scaling is well established with abundant data for smaller magnitude events, it is not clear how well extends to large magnitudes. The near-triangular shape of the median STFs observed in this study (Fig. 2) suggests that the slope should be 1/2.

For resolving the issue a data set with a wider magnitude range is necessary. However, most seismological methods are applicable only within a specific magnitude range. In particular, methods that rely on teleseismic long-period ground motion only work for sufficiently large events (M~>7). Methods that work with local data, on the other hand, can work well for smaller events, but fail for the large events. As a consequence, existing data sets do not constitute a strong and convincing test for the presumed scaling break at M~7.0.

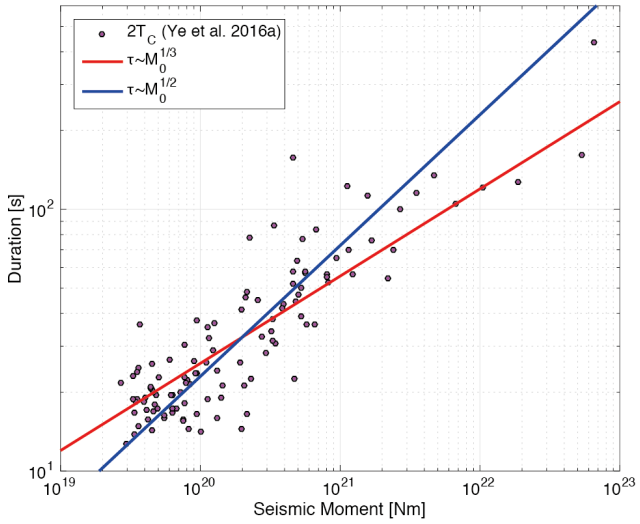


**Fig. S8: Centroid durations and best-fit models with imposed slopes of 1/3 and 1/2.** The data is consistent with either scaling, with corresponding $R^2$ values of 0.76 (slope 1/3) and 0.70 (slope 1/2).

The Global Centroid Moment Tensor (GCMT) data, for example, is based on long period ground motion (>1s) and it can therefore only resolve source durations for events that are substantially longer than a few seconds. In Fig. S9 we compare the centroid times reported in the GCMT catalog to those estimated with the STFs of Ye et al., 2016a. From the GCMT catalog we have selected all shallow events since 1995 with centroid depths between 11 and 40km. We exclude the shallowest events because their depths are notoriously poorly constrained by long period data, and are set to a default value. The centroid times we estimate with the STFs of Ye et

al., 2016a agree very well with those of the GCMT catalog. For smaller magnitudes, $M_w$<7.0, however, the GCMT estimates flatten out at ~10s, suggesting that these estimates are not accurate because they are affected by the above described resolution issues. As a consequence, the scaling of centroid times with seismic moment for events with $M_w$ <7 are not assessed reliably with such data. This demonstrates that existing long period duration estimates cannot be used to test the scaling break hypothesis in a rigorous way.

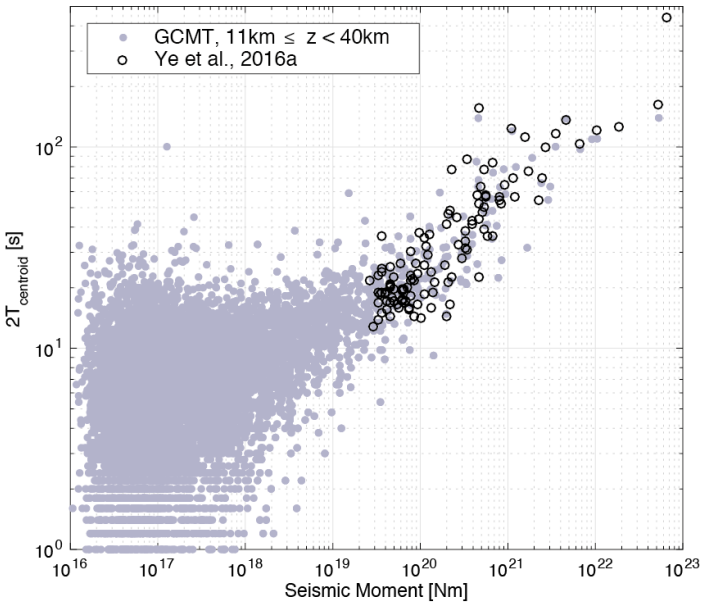


**Fig. S9: Comparison of duration estimates of this study with those of Harvard GCMT catalog**. The centroid durations estimated with the STFs of Ye et al., 2016a are consistent with those reported in the GCMT catalog. For smaller magnitudes than $M_w$ ~7, however, the GCMT estimates become unreliable because they are affected by resolution issues.

Alternatively, we can compare the source duration estimates to those inferred with frequency-domain approaches. Corner frequencies of seismic amplitude spectra are inversely proportional to rupture durations and, for an assumed rupture model, the actual durations can be computed (*50*). In the model assumed by (*33*) and (*35*), the duration is related to the corner frequency $f_c$ as $T = k/(0.9\, f_c)$ with $k = 0.38$. If we compare our source duration estimates, $2T_c$, to those of the two corner frequency studies (Fig. S10) we find that the corner frequency estimates give a scaling exponent of exactly 1/3, but that the durations are systematically offset from our time-domain estimates. Apparently, the frequency-domain estimates are not directly comparable to the time-domain estimates. A plausible candidate explanation for this discrepancy is that the model that relates duration and corner frequencies assumes a self-similar rupture, whereas the linear STF growth reported in this study suggests that this model is not applicable for large subduction earthquakes. By itself, the corner frequency estimate data is not sufficient to test the scaling break either, because it contains very few events above the presumed scaling break at $M_w$ ~7.0. Notably, the study that extends the approach of the two corner frequency studies to large subduction zone earthquakes (*35*) finds the same scaling break that we report

from $M_0^{1/3}$ to $M_0^{1/2}$ at $M_w$ ~7.0. However, owing to limitations of the employed methods they had to change methods at just the critical $M_w$ ~7.0.

For a rigorous and strong test of the proposed scaling break it would be desirable to have a rupture duration data set that is compiled with a consistent method applicable to a wide enough magnitude range on either side of the break. At present, at least two lines of evidence suggest that the scaling break exists, including corner frequency estimates (*36*) and the typical shape of subduction zone STFs (this study).

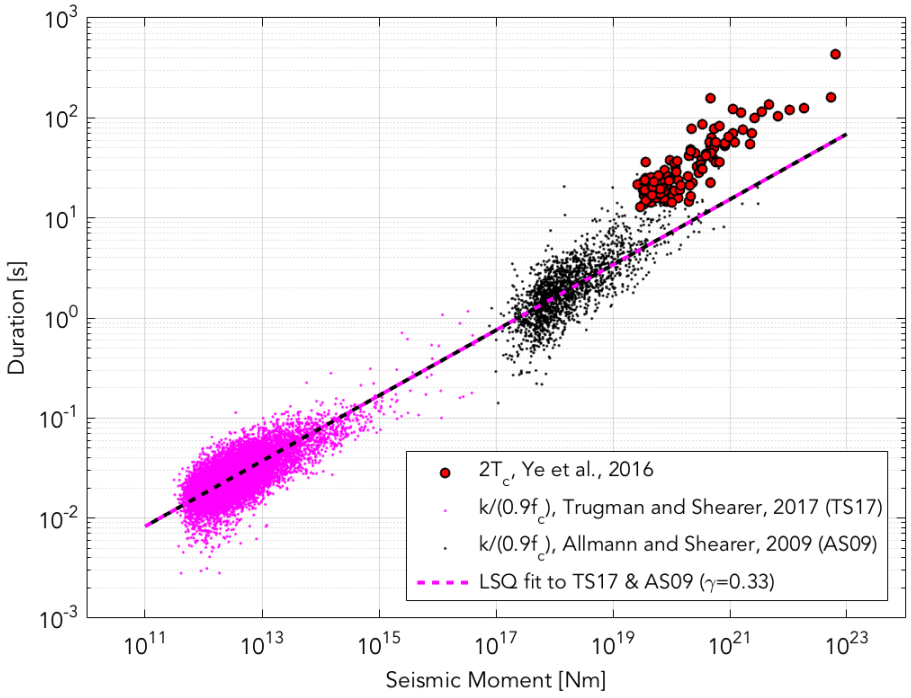


**Fig. S10: Comparison of duration estimates of this study with inverse corner frequencies estimates.** Inverse corner frequencies from Allmann and Shearer, 2009, JGR, and from Trugman and Shearer, 2017, JGR with a least squares fit through the joined data set. The slope of the best-fit curve equals 1/3. The centroid time estimates from this study are offset and systematically larger than the trend observed from the corner frequency estimates.

## S7: Fit and amplitude spectra of Brune and proposed STF model

While the functional form $y_{fit} = \mu*t*\exp(-1/2*(\lambda t)^{\varphi})$, with $\varphi=2$ provides an excellent fit to the median STF, the form used by Brune 1970, i.e. the same form with $\varphi=1$, does not (Fig. S11). In order to reach the observed peak STF amplitudes a much higher initial slope is necessary than what is observed. The bad fit with the model of Brune 1970 is not surprising since it corresponds to the approximate far field solution for an instantaneous spherical pressure source, whereas the earthquakes considered here have extended finite source durations.

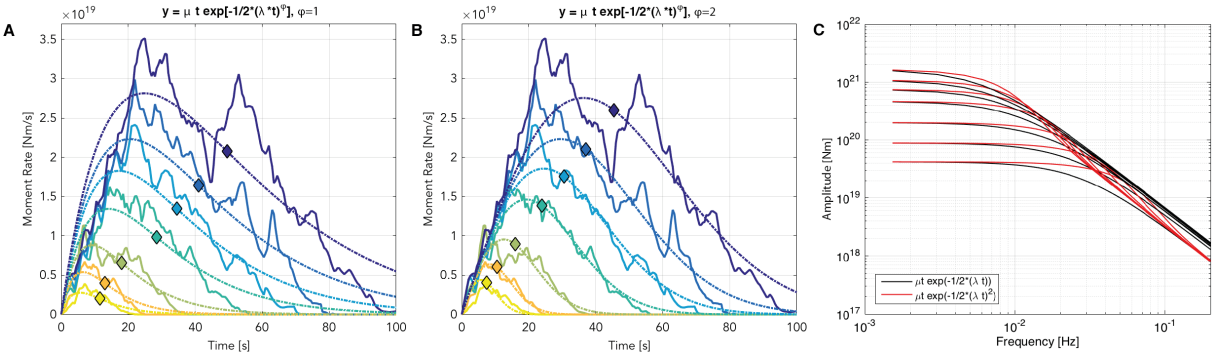


**Fig. S11. Least-squares best fit of the functional form to the median STFs with different values of φ, and corresponding amplitude spectra.** (A) using $\varphi=1$ (Brune 1970, *27*) and (B) $\varphi=2$ (this study, and *25*). While the latter leads to an excellent fit, the Brune model cannot be fitted well. If $\varphi$ is left as a free parameter, together with $\mu$ and $\lambda$, its optimal values for the seven magnitude bins are 2.73, 2.27, 1.54, 2.28, 2.07, 1.93, 1.87, suggesting that the data indeed requires $\varphi\sim2$. (C) The amplitude spectra of both functions fall off with $\omega^{-2}$ at high frequencies but the function with $\varphi=2$ has an intermediate frequency range with a steeper fall-off. Note that, since these median STFs represent strongly smoothed versions of the actual STFs, their spectra lack a substantial fraction of their total signal energy. For a direct comparison with observed spectra, and for considerations of radiated energy, the lacking high frequency energy would have to be added with a noise model.

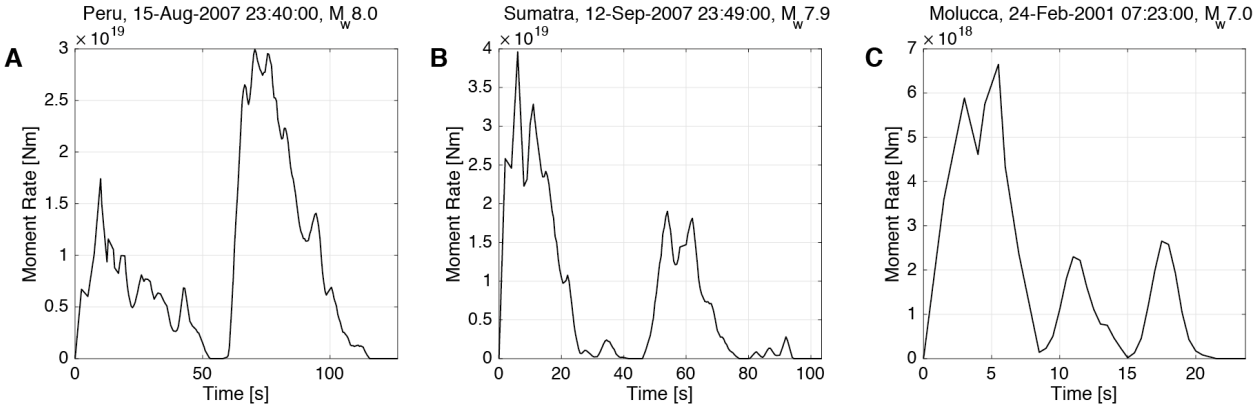


**Fig. S12. STFs of the three events that are excluded as outliers from the residual analysis in Fig. 4.** The three events feature two or more episodes of strong moment rate growth and therefore are poorly approximated with the simple functional form $y_{fit} = \mu * t * \exp(-1/2 * (\lambda t)^2)$.

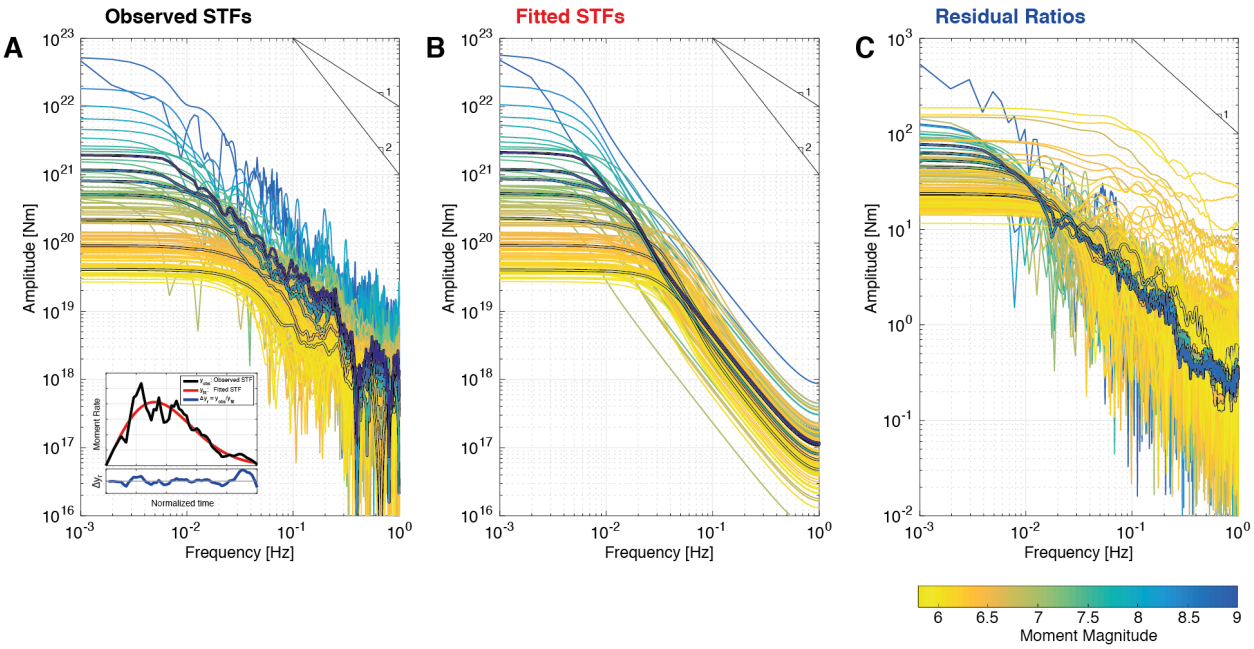


**Fig. S13: STF Fourier amplitude spectra. The amplitude spectra of the individual STFs (thin lines) and the median spectrum in each of the seven magnitude bins (thick lines) are shown for (A) the actual STFs $y_{obs}(t)$, (B) the fitted smooth models $y_{fit}(t)$, and (C) for the residual ratios $\Delta y_r(t) = y_{obs}(t)/y_{fit}(t)$. The spectra of the observed STF fall off with a slope of less than 2, as had previously been noted by (*17, 18*). At all frequencies signal energy increases with magnitude. The smooth models, on the other hand, have approximately equal amplitudes for all magnitudes at higher frequencies (≥0.03Hz). The higher amplitudes for the larger magnitudes in the observed STF spectra are caused by the oscillations around the smooth model, which are higher for larger magnitude events (as demonstrated by constant residual ratios, cf. Fig. 4 in main text). Owing to continuous moment rate accelerations and decelerations, high frequency seismic radiation is emitted throughout the rupture process. The residual ratios for larger magnitude events have higher low-frequency asymptotes because of their longer signal duration. At higher frequencies the amplitude spectra of the residual ratios are characterized by a distinct 1/f trend that is on average near-identical for all magnitudes. This 1/f trend is indicative of a process that lies somewhere between purely random white noise with no temporal correlations (and a constant amplitude spectrum), and a brown noise process in which temporal correlations are strong (with a $1/f^2$ spectrum). In order to compare the STF amplitude spectra of the smooth models to observed source spectra, and for considerations of radiated energy, an appropriate 1/f noise model has to be considered.**

**References and Notes**


1. I. J. Hamling, S. Hreinsdóttir, K. Clark, J. Elliott, C. Liang, E. Fielding, N. Litchfield, P. Villamor, L. Wallace, T. J. Wright, E. D'Anastasio, S. Bannister, D. Burbidge, P. Denys, P. Gentle, J. Howarth, C. Mueller, N. Palmer, C. Pearson, W. Power, P. Barnes, D. J. A. Barrell, R. Van Dissen, R. Langridge, T. Little, A. Nicol, J. Pettinga, J. Rowland, M. Stirling, Complex multifault rupture during the 2016 $M_w$ 7.8 Kaikōura earthquake, New Zealand. *Science* **356**, eaam7194 (2017). doi:10.1126/science.aam7194 Medline

2. L. Ye, T. Lay, H. Kanamori, Z. Zhan, Z. Duputel, Diverse rupture processes in the 2015 Peru deep earthquake doublet. *Sci. Adv.* **2**, e1600581 (2016). doi:10.1126/sciadv.1600581 Medline

3. L. Meng, J. P. Ampuero, J. Stock, Z. Duputel, Y. Luo, V. C. Tsai, Earthquake in a maze: Compressional rupture branching during the 2012 M(w) 8.6 Sumatra earthquake. *Science* **337**, 724–726 (2012). doi:10.1126/science.1224030 Medline

4. K. Aki, Scaling law of seismic spectrum. *J. Geophys. Res.* **72**, 1217–1231 (1967). doi:10.1029/JZ072i004p01217

5. S. Nielsen, R. Madariaga, On the self-healing fracture mode. *Bull. Seismol. Soc. Am.* **93**, 2375–2388 (2003). doi:10.1785/0120020090

6. I. Main, *Nature* **397**, 284 (1999).

7. T. H. Heaton, Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Phys. Earth Planet. Inter.* **64**, 1–20 (1990). doi:10.1016/0031-9201(90)90002-F

8. W. L. Ellsworth, G. C. Beroza, Seismic evidence for an earthquake nucleation phase. *Science* **268**, 851–855 (1995). doi:10.1126/science.268.5212.851 Medline

9. E. L. Olson, R. M. Allen, The deterministic nature of earthquake rupture. *Nature* **438**, 212–215 (2005). doi:10.1038/nature04214 Medline

10. S. Colombelli, A. Zollo, G. Festa, M. Picozzi, Evidence for a difference in rupture initiation between small and large earthquakes. *Nat. Commun.* **5**, 3958 (2014). doi:10.1038/ncomms4958 Medline

11. P. Rydelek, S. Horiuchi, Earth science: Is earthquake rupture deterministic? *Nature* **442**, E5–E6, author reply E6 (2006). doi:10.1038/nature04963 Medline

12. M. A. Meier, T. Heaton, J. Clinton, Evidence for universal earthquake rupture initiation behavior. *Geophys. Res. Lett.* **43**, 7991–7996 (2016). doi:10.1002/2016GL070081

13. T. H. Heaton, A model for a seismic computerized alert network. *Science* **228**, 987–990 (1985). doi:10.1126/science.228.4702.987 Medline

14. M. Böse *et al*., in *Early Warning for Geological Disasters* (Springer, 2014), pp. 49–69.

15. O. Kamigaichi, M. Saito, K. Doi, T. Matsumori, S. Tsukada, K. Takeda, T. Shimoyama, K. Nakamura, M. Kiyomoto, Y. Watanabe, Earthquake early warning in Japan: Warning the general public and future prospects. *Seismol. Res. Lett.* **80**, 717–726 (2009). doi:10.1785/gssrl.80.5.717

16. J. A. Strauss, R. M. Allen, Benefits and costs of earthquake early warning. *Seismol. Res. Lett.* **87**, 765–772 (2016). doi:10.1785/0220150149

17. L. Ye, T. Lay, H. Kanamori, L. Rivera, Rupture characteristics of major and great ($M_w \geq 7.0$) megathrust earthquakes from 1990 to 2015: 1. Source parameter scaling relationships. *J. Geophys. Res. Solid Earth* **121**, 826–844 (2016). doi:10.1002/2015JB012426

18. S. H. Hartzell, T. H. Heaton, Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake. *Bull. Seismol. Soc. Am.* **73** , 1553–1583 (1983).

19. Materials and methods are available as supplementary materials

20. G. P. Hayes, The finite, kinematic rupture properties of great-sized earthquakes since 1990. *Earth Planet. Sci. Lett.* **468**, 94–100 (2017). doi:10.1016/j.epsl.2017.04.003

21. M. Vallée, J. Charléty, A. M. Ferreira, B. Delouis, J. Vergoz, SCARDEC: A new technique for the rapid determination of seismic moment magnitude, focal mechanism and source time functions for large earthquakes using body-wave deconvolution. *Geophys. J. Int.* **184**, 338–358 (2011). doi:10.1111/j.1365-246X.2010.04836.x

22. Y. Ben-Zion, J. R. Rice, Earthquake failure sequences along a cellular fault zone in a three-dimensional elastic solid containing asperity and nonasperity regions. *J. Geophys. Res. Solid Earth* **98**, 14109–14131 (1993). doi:10.1029/93JB01096

23. D. S. Fisher, Collective transport in random media: From superconductors to earthquakes. *Phys. Rep.* **301**, 113–150 (1998). doi:10.1016/S0370-1573(98)00008-8

24. K. A. Dahmen, Y. Ben-Zion, J. T. Uhl, A simple analytic theory for the statistics of avalanches in sheared granular materials. *Nat. Phys.* **7**, 554–557 (2011). doi:10.1038/nphys1957

25. A. P. Mehta, K. A. Dahmen, Y. Ben-Zion, Universal mean moment rate profiles of earthquake ruptures. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **73**, 056104 (2006). doi:10.1103/PhysRevE.73.056104 Medline

26. D. M. Boore, Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. *Bull. Seismol. Soc. Am.* **73**, 1865–1894 (1983).

27. J. N. Brune, Tectonic stress and the spectra of seismic shear waves from earthquakes. *J. Geophys. Res.* **75**, 4997–5009 (1970). doi:10.1029/JB075i026p04997

28. Z. Duputel, V. C. Tsai, L. Rivera, H. Kanamori, Using centroid time-delays to characterize source durations and identify earthquakes with unique characteristics. *Earth Planet. Sci. Lett.* **374**, 92–100 (2013). doi:10.1016/j.epsl.2013.05.024

29. A. M. Dziewonski, T. A. Chou, J. H. Woodhouse, Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *J. Geophys. Res. Solid Earth* **86** , 2825–2852 (1981). doi:10.1029/JB086iB04p02825

30. H. Kanamori, D. L. Anderson, Theoretical basis of some empirical relations in seismology. *Bull. Seismol. Soc. Am.* **65**, 1073–1095 (1975).

31. T. C. Hanks, in *Stress in the Earth* (Birkhäuser, Basel, 1977), pp. 441–458. .

32. H. Kanamori, E. E. Brodsky, The physics of earthquakes. *Rep. Prog. Phys.* **67**, 1429–1496 (2004). doi:10.1088/0034-4885/67/8/R03

33. B. P. Allmann, P. M. Shearer, Global variations of stress drop for moderate to large earthquakes. *J. Geophys. Res. Solid Earth* **114**, B01310 (2009). doi:10.1029/2008JB005821

34. A. Baltay, S. Ide, G. Prieto, G. Beroza, Variability in earthquake stress drop and apparent stress. *Geophys. Res. Lett.* **38**, L06303 (2011). doi:10.1029/2011GL046698

35. D. T. Trugman, P. M. Shearer, Application of an improved spectral decomposition method to examine earthquake source scaling in Southern California. *J. Geophys. Res. Solid Earth* **122**, 2890–2910 (2017). doi:10.1002/2017JB013971

36. M. A. Denolle, P. M. Shearer, New perspectives on self-similarity for shallow thrust earthquakes. *J. Geophys. Res. Solid Earth* **121**, 6533–6565 (2016). doi:10.1002/2016JB013105

37. J. Gomberg, A. Wech, K. Creager, K. Obara, D. Agnew, Reconsidering earthquake scaling. *Geophys. Res. Lett.* **43**, 6243–6251 (2016). doi:10.1002/2016GL069967

38. P. M. Mai, K. K. S. Thingbaijam, SRCMOD: An online database of finite-fault rupture models. *Seismol. Res. Lett.* **85**, 1348–1357 (2014). doi:10.1785/0220140077

39. R. Madariaga, Dynamics of an expanding circular fault. *Bull. Seismol. Soc. Am.* **66**, 639–666 (1976).

40. B. V. Kostrov, Selfsimilar problems of propagation of shear cracks. *J. Appl. Math. Mech.* **28**, 1077–1087 (1964). doi:10.1016/0021-8928(64)90010-3

41. T. Sato, T. Hirasawa, Body wave spectra from propagating shear cracks. *J. Phys. Earth* **21**, 415–431 (1973). doi:10.4294/jpe1952.21.415

42. A. A. Gabriel, J. P. Ampuero, L. A. Dalguer, P. M. Mai, The transition of dynamic rupture styles in elastic media under velocity-weakening friction. *J. Geophys. Res. Solid Earth* **117**, B09311 (2012). doi:10.1029/2012JB009468

43. G. Perrin, J. R. Rice, G. Zheng, Self-healing slip pulse on a frictional surface. *J. Mech. Phys. Solids* **43**, 1461–1495 (1995). doi:10.1016/0022-5096(95)00036-I

44. J. J. McGuire, L. Zhao, T. H. Jordan, Predominance of unilateral rupture for a global catalog of large earthquakes. *Bull. Seismol. Soc. Am.* **92**, 3309–3317 (2002). doi:10.1785/0120010293

45. T. I. Allen, G. P. Hayes, Alternative rupture-scaling relationships for subduction interface and other offshore environments. *Bull. Seismol. Soc. Am.* **107**, 1240–1253 (2017). doi:10.1785/0120160255

46. B. W. Crowell, D. Melgar, Y. Bock, J. S. Haase, J. Geng, Earthquake magnitude scaling using seismogeodetic data. *Geophys. Res. Lett.* **40**, 6089–6094 (2013). doi:10.1002/2013GL058391

47. N. A. Abrahamson, W. J. Silva, R. Kamai, Summary of the ASK14 Ground Motion Relation for Active Crustal Regions. *Earthq. Spectra* **30**, 1025–1055 (2014). doi:10.1193/070913EQS198M

48. C. J. Ammon, C. Ji, H. K. Thio, D. Robinson, S. Ni, V. Hjorleifsdottir, H. Kanamori, T. Lay, S. Das, D. Helmberger, G. Ichinose, J. Polet, D. Wald, Rupture process of the 2004 Sumatra-Andaman earthquake. *Science* **308**, 1133–1139 (2005). doi:10.1126/science.1112260 Medline

49. C. Ji, D. J. Wald, D. V. Helmberger, Source description of the 1999 Hector Mine, California, earthquake, Part I: Wavelet domain inversion theory and resolution analysis. *Bull. Seismol. Soc. Am.* **92**, 1192–1207 (2002). doi:10.1785/0120000916

50. Y. Kaneko, P. M. Shearer, Variability of seismic source spectra, estimated stress drop, and radiated energy, derived from cohesive-zone models of symmetrical and asymmetrical circular and elliptical ruptures. *J. Geophys. Res. Solid Earth* **120**, 1053–1079 (2015). doi:10.1002/2014JB011642