Source time function clustering reveals patterns in earthquake dynamics

Jiuxun Yin[1,1], Zefeng Li[2,2], and Marine Denolle[1,1]

[1]Harvard University
[2]California Institute of Technology

November 30, 2022

Abstract

We cluster a global data base of 3529 M>5.5 earthquakes in 1995-2018 based on a dynamic time warping dissimilarity of their source time functions (STFs). The clustering exhibits different degrees of STF shape complexity and suggests an association between STF complexity and earthquake source parameters. Thrust events are in large proportion with simple STF shapes and at all depths. In contrast, earthquakes with complex STF shapes tend to be located at shallow depth in complicated tectonic regions with preferentially strike slip mechanism and relatively longer duration. With 2D dynamic modeling of earthquake ruptures on heterogeneous pre-stress and linear slip-weakening friction, we find a systematic variation of the simulated STF complexity with frictional properties. Comparison between the observed and synthetic clustering distributions provides useful constraints on elements of the frictional properties. In particular, the characteristic slip-weakening distance could be constrained to be generally short (< 0.1 m) and depth dependent.

Source time function clustering reveals patterns in earthquake dynamics

Jiuxun Yin[1], Zefeng Li[2], Marine Denolle[1]

[1]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA
[2]Seismological Laboratory, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

Key Points:

- We cluster earthquakes based on the dynamic time warping distance of their source time function (STF) shapes.
- The patterns of complexity correlate with source parameters such as depth, mechanism, and radiation.
- Simulations of dynamic rupture indicate a correlation between the STF complexity and frictional properties.

Corresponding author: Zefeng Li, `zefengli@caltech.edu`

Abstract

We cluster a global data base of 3529 M>5.5 earthquakes in 1995-2018 based on a dynamic time warping dissimilarity of their source time functions (STFs). The clustering exhibits different degrees of STF shape complexity and suggests an association between STF complexity and earthquake source parameters. Thrust events are in large proportion with simple STF shapes and at all depths. In contrast, earthquakes with complex STF shapes tend to be located at shallow depth in complicated tectonic regions with preferentially strike slip mechanism and relatively longer duration. With 2D dynamic modeling of earthquake ruptures on heterogeneous pre-stress and linear slip-weakening friction, we find a systematic variation of the simulated STF complexity with frictional properties. Comparison between the observed and synthetic clustering distributions provides useful constraints on elements of the frictional properties. In particular, the characteristic slip-weakening distance could be constrained to be generally short (<0.1 m) and depth dependent.

Plain Language Summary

Seismic waves carry a signature about the earthquake source process. Earthquake source time functions (STFs), which are directly recovered from seismic waves, reflect the temporal history of earthquake rupture. However, it is often hard to directly compare STFs due to the large differences among earthquakes in terms of amplitude and duration. In this study, we perform a cluster analysis of STFs using a technique called dynamic time warping (DTW). DTW is commonly used in speech recognition to handle with various speeds of elocution. DTW allows us to dynamically stretch the seismic signals and provides a new way to quantify earthquake similarity through analyzing the shapes of their source time functions (STFs). We apply this to a large database of STFs. Our results show that the shape complexity of STFs is correlated with the earthquake source parameters such as the earthquake depth, focal mechanism, and energy radiation. Our numerical simulations further show that those correlations may indicate a spatial heterogeneity of frictional properties.

1 Introduction

Earthquakes are known to break in diverse manners: some events rupture on a geometrically simple fault with a relatively smooth slip distribution (e.g., Yagi & Fukahata, 2011), while others break a network of faults and/or have heterogeneous slip distribution (Li et al., 1994; Ammon et al., 2005; Meng et al., 2012; Cesca et al., 2017). Although the complexity of earthquakes can be directly observed, in some cases, from surface fault trace (Massonnet et al., 1993; Li et al., 1994; Kaneko et al., 2017), many ruptures are buried at depth so that seismic waves are the only observations available to infer the source process. Derived from seismic waves through waveform deconvolution or kinematic inversion, the earthquake Source Time Function (STF) is a foremost important seismic observation that describes the time history of moment release during a rupture. Moreover, the shape of the STF directly controls the variability and uncertainty in the strength and duration of strong ground motion.

Observations of global earthquake STFs and source spectra have shown significant inter-event variability among earthquakes (Allmann & Shearer, 2009; Atik et al., 2010; Denolle, 2019). Such variability may partly come from differences in data processing strategy (Ide & Beroza, 2001). Therefore, large catalogs of STFs (or their spectra) obtained from a uniform approach is preferable to analyze relative differences among earthquakes (Allmann & Shearer, 2009; Convers & Newman, 2011; Denolle & Shearer, 2016; Vallée & Douet, 2016).

Recently, such catalogs of STFs (or of their spectra) have enabled multiple discoveries about earthquake source processes. For example, the total seismic moment M_0 (the time integral of the STF) scales with source duration T^3 (the duration of the STF) for most small to moderate size earthquakes, which implies that the earthquake stress drop is roughly invariant with earthquake magnitudes. At larger magnitudes, this scaling may differ (e.g. $M_0 \sim T^2$ from Denolle and Shearer (2016)). Their properties also have indicated that the ratio of the radiated energy E_R over the moment, also referred to as the scaled energy E_R/M_0, varies spatially and with depth but remains invariant with earthquake magnitude (Convers & Newman, 2011; Baltay et al., 2014; Denolle & Shearer, 2016).

However, both the amplitude and the source duration of the STF vary by orders of magnitude. This requires careful strategies of amplitude and time scaling for across-magnitude visualization and comparison. One approach is to scale the time axis to a duration metric and normalize the amplitude to seismic moment (i.e. the integral of the STF). However, source duration is difficult to measure because near-source and near-site scattering of seismic waves may interfere with waves radiating from the end of the seismic rupture. Therefore previous studies have proposed several metrics of duration: moment-based duration (Houston, 2001), threshold-based duration (Vallée, 2013; Denolle, 2019), and centroid-based duration (Meier et al., 2017). Because these measures are not strictly equivalent, the shapes of the scaled and stretched STFs differ as well. For instance, Meier et al. (2017) find that average STFs have rather a triangle shape whereas Denolle (2019) suggests a rather skewed-Gaussian functional form.

Here, we propose to weaken the assumption of a particular definition of source duration and instead use dynamic time warping (DTW) to compare the shapes of the STFs. DTW has been widely used in speech recognition (Berndt & Clifford, 1994; "Dynamic Time Warping", 2007). The DTW algorithm performs a non-uniform stretching of time and amplitude to match the shape of two time series via the optimal warping path with minimum distance (Figure S1). We measure the similarity between STFs with DTW distance and cluster the STFs accordingly. We apply this to the global SCARDEC catalog of STFs (Vallée & Douet, 2016, available at `http://scardec.projects.sismo.ipgp.fr/`, last accessed 01/20/2020) that contains 3529 earthquakes of magnitude greater than 5.5 from 1/1/1992 and until 12/31/2018. The analysis shows that the STF overall shape is correlated with several earthquake source parameters, such as focal mechanisms, depth, and scaled energy.

To test whether the current physical understanding of earthquake processes reproduces the clustering patterns, we perform dynamic simulations of earthquake ruptures with linear slip-weakening friction to construct synthetic STFs. We find a strong correlation between the grouping distribution of STF shapes and frictional parameters, such as the characteristic slip-weakening distance D_c. Furthermore, we find that the grouping pattern of the SCARDEC STF shapes are most similar to those simulated STFs with small values of D_c, thus the grouping patterns of a large number of STFs can potentially provide observational constraints to earthquake dynamics.

2 Dynamic time warping and clustering analysis

DTW measures the similarity between two time series that may not share the same frequency content or the same sampling rate. The series are "warped" (or stretched) non-uniformly in the time dimensions to optimally match two series (Figure S1). This algorithm is widely used in automated speech recognition in which different audio sequences may have different speaking speeds (Berndt & Clifford, 1994; "Dynamic Time Warping", 2007). One important advantage of DTW is its ability to preserve topological structures of the time series by assimilating their temporal elongation or compression. Once stretched, the DTW distance is taken as a new metric for STF similarity, which can be used for clustering. Our approach follows four steps: 1) STF pre-conditioning, 2) DTW distance calculation, 3) clustering, 4) re-grouping around a centroid event.

We first perform minimal pre-conditioning of the STF shapes. The STFs are built from the deconvolution of teleiseismic P waves that are relatively well constrained at frequencies below 1 Hz (Vallée & Douet, 2016). Given that the maximum duration of the STF in the catalog is about 100 s, we re-sample the data to 100 points giving a minimum sampling rate of 1 point per second. We then normalize the amplitude STFs to the event seismic moment. These two processing steps improve the stability of the warping. We have tested various strategies to resample and normalize the STFs, which did not affect the conclusions of this analysis.

Second, we apply the DTW to each pair of STFs. The DTW distance is the Euclidean distance between two STFs warped along the optimal warping path, and is chosen here as the measure of similarity between two STFs (see Figure S1 (a)-(b)).

Then, the STF shapes are clustered based on their DTW distance with a single-linkage hierarchical clustering analysis that provides the flexibility to form clusters at any desired level (Text S1, Figure S1 (c)). Here, we constrain the number of clusters to be 20, which is about equivalent of DTW distance threshold of 0.4. For each of these clusters, we choose a representative STF (defined as the centroid event) that has the minimum median distance with all of the other members of the cluster. It is similar to the stack of all stretched STF within each cluster (Figure 1), which, in turn, exhibits the common features of all cluster members.

Furthermore, we parameterize the characteristic STF shape for each of these clusters by calculating the number of prominent peaks of each centroid event. The number of prominent peaks is commonly used for topographic relief analysis and is defined as the amplitude of the peak (hill summit) relative to the lowest amplitude point (valley) that does not contain a higher peak. This metric differs from the calculation of Gaussian subevents that Danré et al. (2019) use. One hyper-parameter we tune is a threshold for peak amplitude of the prominent peak, which we choose to be 10% of the global maximum of the STF amplitude. The raw and stretched STFs have a lot fewer prominent peaks than individual peaks from the Gaussian decomposition by Danré et al. (2019) (Figure S2). Furthermore, the stretched STFs have fewer prominent peaks than the raw STFs, but in general the same number of prominent peaks as the centroid event (Figure S3). For instance, a STF may have multiple separated amplitude peaks, but only one single prominent peak (Figure 1 (a)-(b)).

Finally, we group the clusters based on the number of prominent peaks of the centroid event, where G1 is the group where the centroid event has 1 prominent peak, G2 is the group where the centroid event has 2 prominent peaks, ... (Figure 1 (c)). G4 is the group where the centroid event has at least 4 prominent peaks. Examples of detected prominent peaks are found in Figure 1 (a)-(b) (see Figure S4 for the unstretched STFs). In this study, we define the STFs to be "complex" if their DTW stretched STFs have multiple prominent peak. The first order result from the grouping is that most events have a single prominent peak whereas about 20% events are more complex.

3 Correlations between shape complexity and source parameters

We now explore the correlation between grouping and several source parameters such as depth, focal mechanism, moment, duration, energy, and location.

The first property we investigate is the source depth. Complex STFs (groups G2-G4) are mostly shallow crustal events ($\leq$ 20 km) whereas the simple STFs (group G1) can be found at all depths (Figure 2 (a)). Because co-located events have various degrees

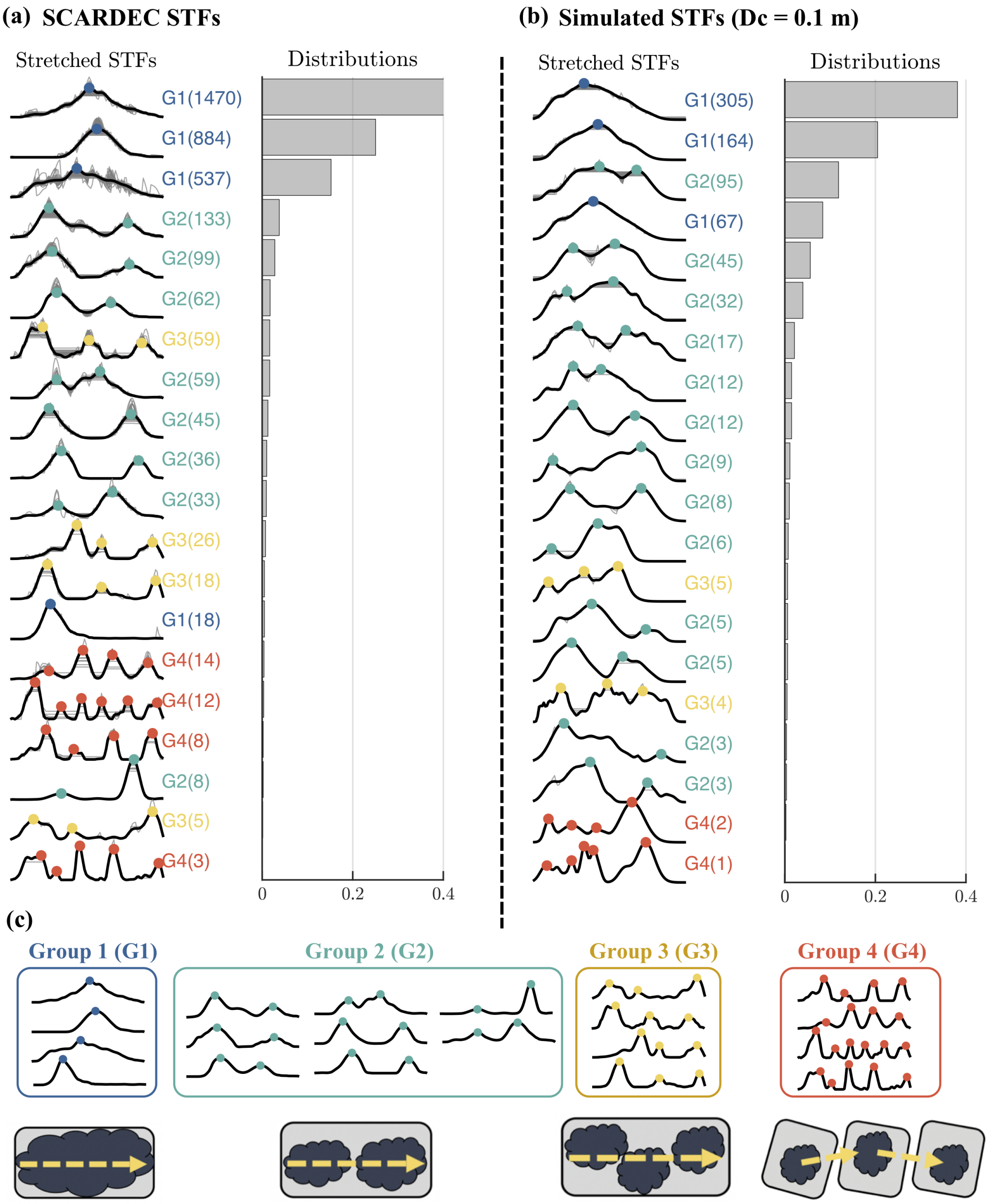

Figure 1. Source time function clustering, grouping, and conceptual interpretation. (a) Individual STFs after dynamic time warping and clustering are shown by gray thin lines. Black thick lines are the STFs of the centroid event of each cluster. Colored dots indicate the prominent peaks of the centroid STF as well as the associated group. Numbers in the parentheses are the number of STFs in each cluster. The corresponding population proportion of each cluster is shown in the right histograms. (b) Same as (a) but for the STFs from our dynamic simulations. (c) Cluster centroid STF shapes and conceptual models for G1-G4. In the model diagram, dark blocks represent major rupture asperities and the arrow indicates the rupture direction.

of complexity (Figure 2 (d), Figure 3), inaccuracy in the Green's function does not strongly bias these specific results.

The second property we investigate is the focal mechanism (Figure 2 (b)). The focal mechanisms are solved simultaneously by the SCARDEC method (Vallée et al., 2011). Most of the thrust earthquakes have simple STFs (G1 and G2), whereas the strike-slip earthquakes are dominated by complex STFs (G3 and G4). There are too few normal events in the database (only 17.5 %) to give any significant conclusion regarding this mechanism.

There is no clear relation between earthquake size (moment) and this metric of complexity (see Figure 2 (d) and Figure S5). For example in Figure 2 (d), we see that the largest events in SCARDEC database may only have one prominent peak in their stretched STF, while the events with smaller moments can be in any of those complexity groups.

We find a clear pattern that G3-G4 events have an abnormally longer duration with respect to other events of similar magnitudes and relative to events of the other groups (Figure 2 (d)). It is illustrated in Figure 2 (d) by visualization of two STFs of co-located events and of similar magnitudes. For the same earthquake moment (or the STF integral), it is intuitive to understand that STFs in G4 have multiple low amplitude prominent peaks and overall extended duration, compared to the G1 STFs that have a single high amplitude and short duration peak. Simple models of crack ruptures yield a relation between moment, source duration, and stress drop that could indicate low stress drops for the G4 events (Figure S6 (a)-(c)) (Brune, 1971; Eshelby, 1957).

We now explore the clustering results against the earthquake scaled energy. Here we calculate radiated energy from the squared time derivative of the STF (moment acceleration function $\ddot{M}_0(t)$) using the relation $E_R = (\frac{1}{15\pi\rho V_p^5} + \frac{1}{10\pi\rho V_s^5}) \int_0^\infty \left(\ddot{M}_0(t)\right)^2 dt$. We select depth-dependent bulk properties (V_p, V_s, ρ) from PREM (Dziewonski & Anderson, 1981). Radiated energy scales almost linearly with seismic moment and look at the scaled energy, the ratio of both radiated energy and seismic moment. Figure 2 (c) shows the distribution of the scaled energy with respect to each group. G3 and G4 events have systematically larger scaled energy as G1 and G2 events. This is consistent with intuition that G3 and G4 events generally have rougher STFs.

The correlations between STF complexity and source depths and focal mechanism are consistent with the findings from previous studies (Houston, 2001; Vallée, 2013; Danré et al., 2019). In particular, shallow strike slip earthquakes are constrained geometrically by the Earth surface on the top and the seismogenic depth on the bottom. They also tend to be composed of segmented faults (Klinger, 2010). These geometrical settings control the evolution of rupture that tends to operate with moving energetic slip pulses (Kaneko & Lapusta, 2010) with repeated rupture acceleration and deceleration as they travel across segments (e.g., Kanamori et al., 1992; Peyrat et al., 2001; Cesca et al., 2017).

Since earthquake source parameters are closely related to the local tectonic regime, we also find that our observations from the clustering and grouping results (G1 - G4) are consistent to the marked variation of tectonic environments (Figure 3). Many of the major subduction zones are dominated by the simpler types of events (G1 and G2) and lack of more complex ones, likely because they are dominated by thrust events located along/within the subducting slabs at various depths. For example, since 1992, there have been only two events ($M_W > 5.5$) belonging to the G3 group along the Southern American and Aleutian subduction zones, respectively (Figure 4 (a)-(b)). Similarly, other subduction zone regions like in Japan and in Sumatra, the Indian-Eurasian collision zone are also dominated by simple-type earthquakes (Figure 4 (c)-(d)). In contrast, the complex group (G3 and G4) events are located mostly along the boundaries around the junction region of the Indo-Australian, western Pacific, Philippine plates and Eurasian plates (Figure 3 and Figure 4 (e)). Bird (2003) explored and documented the kinematics at plate boundaries and found that this region is characterized by a particularly extensive number of micro plates, whose boundaries exhibit varied relative motions and kinematics (their

Figure 6). Therefore, we propose that the complexity in the STF may reflect the complexity in the regional stress field.

4 Modeling STF complexity

Simulations of dynamic ruptures using stochastic distributions of fault-interface parameters are popular in the investigations of complex kinematic source models, realistic fault geometry and roughness models, and to simulate high-frequency ground motions (Mai & Beroza, 2002; Ripperger et al., 2007; Trugman & Dunham, 2014; Graves & Pitarka, 2016; Mai et al., 2017). In order to investigate possible factors that control the STF complexity patterns, we perform a large number of 2-dimensional dynamic rupture simulations with stochastic distributions of pre-stress, and apply the same clustering analysis to the resulting synthetic STFs as to the SCARDEC STFs.

In this study, synthetic dynamic sources are generated in a 2-dimensional medium in an anti-plane setting. Pre-stress on the fault is constrained to follow a power-law amplitude distribution that approximates the scenario caused by natural fault roughness (Candela et al., 2012, Text S2 for more details). We assume a constant normal stress of 120 MPa and linear slip weakening friction law (Andrews, 1976). Linear slip weakening requires three parameters: the static friction coefficient (here chosen as $\mu_s = 0.677$), the dynamic friction coefficient (here chosen as $\mu_d = 0.525$), and the characteristic slip-weakening distance D_c. We set up the experiments so that the fault-average stress drop is about 1 MPa (Figure S7).

Danré et al. (2019) find that heterogeneity is necessary to reproduce realistically rough STFs. Here, we focus on varying D_c, yet aware of the trade-off between strength excess and D_c in controlling rupture velocity and the resulting ground motions (Guatteri & Spudich, 2000). While we keep D_c constant within a single set of simulations, we carry several sets of experiments with values of D_c at various levels 0.05, 0.1, 0.2, 0.4, 0.8, and 1.6 m that are within bounds found in the literature.

For each D_c, we first generate a set of pre-stress distributions that we use in each simulations. The dynamic rupture is solved by 2D boundary integral method SBIEM-LAB (`http://web.gps.caltech.edu/~ampuero/software.html`, last accessed 11/27/2018). We discard the rupture models that unsuccessfully nucleated with a source dimension less than 20 km, or rupture beyond the zone of heterogeneous pre-stress, and obtain 800 qualified simulations for each D_c value. Finally, the STFs are calculated from the integral of the moment-density-rate functions over the fault surface (more details in Text S2).

We perform the hierarchical clustering and group the simulated STFs for each D_c, following the same procedures as for the SCARDEC STFs (Figure 1 (b), Figures S8 - S12). Because our modeling is not three dimensional and does not include the free surface, we are not matching observations such as the focal mechanism and depth. However, our results can match the proportion of the STFs relative to each group: 80% of the STFs belong to the G1 group, 15% belong to the G2, and the rest in higher indexed groups. Comparison of the relative proportion between groups for each set of simulations suggests that an increasing D_c value yield an increase in STF complexity (e.g. proportion of G3-G4 events). This shows that D_c, or more generally, the frictional parameters can impact the complexity of STFs. Compared with the observed global variability in SCARDEC STFs, small value of D_c (< 0.1 m) is preferred in this particular metric of complexity. In contrast, models with large value of D_c tend to generate proportionally more STFs belonging to G3 beyond (Figures S10 - S12).

Our results indicate that the small values of $D_c < 0.1$ m are necessary to produce the general level of complexity of the SCARDEC STFs (Figure 5 (a)). When binning these relative contributions with source depths, we find that crustal events (h $\leq$ 40 km), which show a higher degree of complexity, could be explained by a larger D_c value than the deeper events (Figure 2 (a), Figure 5 (b)). This is more pronounced with the upper-crustal depths (h $\leq$ 20 km).

Depth variations in D_c have been reported in earlier studies. Wibberley and Shimamoto (2005) perform laboratory experiments on samples from the Median Tectonic Line in southwestern Japan, and estimate that D_c ought to vary with depth, with a deeper (6 km) values being systematically 30% smaller than the shallow (2 km) values. Kinematic source inversions also find a systematic depth variation of rise time, which they attribute to a systematic dependence in D_c (Ide & Takeo, 1997). Our results may provide a supporting evidence that the characteristic slip-weakening distance varies at depth over crustal scales.

5 Discussion and Conclusion

We apply a dynamic time warping methodology to cluster a large number of earthquake source time functions based on similarity of their general shapes. We find patterns between source parameters and the STF shape, which we now compare with previous work Danré et al. (2019) that analyzed the same SCARDEC database. Although the definition of complexity in Danré et al. (2019) is different, this study confirms the correlation between STF complexity with focal depth and mechanisms. This study adds to the Danré et al. (2019) in three ways. First, there is no correlation between this particular metric of complexity and earthquake magnitude. This means that the shape of the individual prominent peaks does not systematically change with earthquake magnitude, while the number of individual and separated peaks does. Second, we analyze in this study the relation between degree of complexity and other source parameters, such as the scaling between duration and moment (sometimes used to estimate earthquake stress drop) and the ratio between radiated energy and moment. Taken together, it is reasonable to infer that the complex STFs exhibit large radiation ratio (proportion of radiated energy over available energy).

Finally, the modeled STFs exhibit different degrees of complexity depending on the frictional properties. We find that small values of characteristic slip weakening distance are required to reproduce the variability in complexity measured in the SCARDEC database. Furthermore, we find that the variability in STF complexity of shallow earthquakes is better explained by a larger value of characteristic distance compared to the deeper sources.

There are several limitations to our approaches. First, the database we use is constructed from a Green's function in a radially symmetric Earth. Although this is unlikely to affect the overall results, Green's functions that account for laterally varying structure would improve the temporal resolution of the shallowest events. This requires better understanding of near surface scattering and attenuation. Second, our modeling approach is unable to characterize the correlation between focal mechanisms and STF complexity. Indeed, these parameters could be tested using a 3-dimensional dynamic rupture simulation framework, which however is impractical to implement due to high computational expense and the employed statistical approaches. Nevertheless, because fault geometry and fault properties seem to play a dominant role in shaping the source and the resulting strong ground motions, further 3-dimensional modeling and observations are necessary.

Acknowledgments

We sincerely thank Martin Vallée for his insightful suggestions. All the source time functions are downloaded from SCARDEC source time function database (`http://scardec.projects.sismo.ipgp.fr/`). The dynamic rupture simulation code SBIEMLAB is developed by Jean-Paul Ampuero (available on `http://web.gps.caltech.edu/~ampuero/`

software.html). The Matlab scripts to reproduce the results and figures can be obtained on the Github (https://github.com/yinjiuxun/STF_DTW). Global maps are made by GMT (Wessel et al., 2013, available at http://gmt.soest.hawaii.edu/)).

References

Allmann, B. P., & Shearer, P. M. (2009). Global variations of stress drop for moderate to large earthquakes. *Journal of Geophysical Research: Solid Earth*, *114*(B1), B01310. doi: 10.1029/2008JB005821

Ammon, C. J., Ji, C., Thio, H.-K., Robinson, D., Ni, S., Hjorleifsdottir, V., … Wald, D. (2005). Rupture Process of the 2004 Sumatra-Andaman Earthquake. *Science*, *308*(5725), 1133–1139. doi: 10.1126/science.1112260

Andrews, D. J. (1976). Rupture propagation with finite stress in antiplane strain. *Journal of Geophysical Research*, *81*(20), 3575–3582. doi: 10.1029/JB081i020p03575

Atik, L. A., Abrahamson, N., Bommer, J. J., Scherbaum, F., Cotton, F., & Kuehn, N. (2010). The Variability of Ground-Motion Prediction Models and Its Components. *Seismological Research Letters*, *81*(5), 794–801. doi: 10.1785/gssrl.81.5.794

Baltay, A. S., Beroza, G. C., & Ide, S. (2014). Radiated Energy of Great Earthquakes from Teleseismic Empirical Green's Function Deconvolution. *Pure and Applied Geophysics*, *171*(10), 2841–2862. doi: 10.1007/s00024-014-0804-0

Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *KDD workshop* (Vol. 10, pp. 359–370). Seattle, WA.

Bird, P. (2003). An updated digital model of plate boundaries. *Geochemistry, Geophysics, Geosystems*, *4*(3). doi: 10.1029/2001GC000252

Brune, J. N. (1971). Correction (to Brune, 1970). *J. geophys. Res*, *76*, 5002.

Candela, T., Renard, F., Klinger, Y., Mair, K., Schmittbuhl, J., & Brodsky, E. E. (2012). Roughness of fault surfaces over nine decades of length scales. *Journal of Geophysical Research: Solid Earth*, *117*(B8). doi: 10.1029/2011JB009041

Cesca, S., Zhang, Y., Mouslopoulou, V., Wang, R., Saul, J., Savage, M., … Dahm, T. (2017). Complex rupture process of the Mw 7.8, 2016, Kaikoura earthquake, New Zealand, and its aftershock sequence. *Earth and Planetary Science Letters*, *478*, 110–120. doi: 10.1016/j.epsl.2017.08.024

Convers, J. A., & Newman, A. V. (2011). Global evaluation of large earthquake energy from 1997 through mid-2010. *J. Geophys. Res. Solid Earth*. doi: 10.1029/2010JB007928

Danré, P., Yin, J., Lipovsky, B. P., & Denolle, M. A. (2019). Earthquakes Within Earthquakes: Patterns in Rupture Complexity. *Geophysical Research Letters*, *46*(13), 7352–7360. doi: 10.1029/2019GL083093

Denolle, M. A. (2019). Energetic Onset of Earthquakes. *Geophysical Research Letters*, *46*(5), 2458–2466. doi: 10.1029/2018GL080687

Denolle, M. A., & Shearer, P. M. (2016). New perspectives on self-similarity for shallow thrust earthquakes. *Journal of Geophysical Research: Solid Earth*, *121*(9), 2016JB013105. doi: 10.1002/2016JB013105

Dynamic Time Warping. (2007). In M. Müller (Ed.), *Information Retrieval for Music and Motion* (pp. 69–84). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-540-74048-3_4

Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, *25*(4), 297–356. doi: 10.1016/0031-9201(81)90046-7

Eshelby, J. D. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 241, pp. 376–396). The Royal Society.

Graves, R., & Pitarka, A. (2016). Kinematic Ground-Motion Simulations on Rough Faults Including Effects of 3d Stochastic Velocity PerturbationsKinematic Ground-Motion Simulations on Rough Faults. *Bulletin of the Seismological Society of America*, *106*(5), 2136–2153. doi: 10.1785/0120160088

Guatteri, M., & Spudich, P. (2000). What Can Strong-Motion Data Tell Us about Slip-Weakening Fault-Friction Laws? *Bulletin of the Seismological Society of America*, *90*(1), 98–116. doi: 10.1785/0119990053

Houston, H. (2001). Influence of depth, focal mechanism, and tectonic setting on the shape and duration of earthquake source time functions. *Journal of Geophysical Research: Solid Earth*, *106*(B6), 11137–11150. doi: 10.1029/2000JB900468

Ide, S., & Beroza, G. C. (2001). Does apparent stress vary with earthquake size. *Geophys. Res. Lett*, *28*(17), 3349–3352.

Ide, S., & Takeo, M. (1997). Determination of constitutive relations of fault slip based on seismic wave analysis. *Journal of Geophysical Research: Solid Earth*, *102*(B12), 27379–27391. doi: 10.1029/97JB02675

Kanamori, H., Hong-Kie, T., Doug, D., Egill, H., & Heaton, T. (1992). Initial investigation of the Landers, California, Earthquake of 28 June 1992 using TERRAscope. *Geophysical Research Letters*, *19*(22), 2267–2270. doi: 10.1029/92GL02320

Kaneko, Y., Fukuyama, E., & Hamling, I. J. (2017). Slip-weakening distance and energy budget inferred from near-fault ground deformation during the 2016 Mw7.8 Kaikōura earthquake. *Geophysical Research Letters*, *44*(10), 4765–4773. doi: 10.1002/2017GL073681

Kaneko, Y., & Lapusta, N. (2010). Supershear transition due to a free surface in 3-D simulations of spontaneous dynamic rupture on vertical strike-slip faults. *Tectonophysics*, *493*(3), 272–284. doi: 10.1016/j.tecto.2010.06.015

Klinger, Y. (2010). Relation between continental strike-slip earthquake segmentation and thickness of the crust. *Journal of Geophysical Research: Solid Earth*, *115*(B7). doi: 10.1029/2009JB006550

Li, Y.-G., Aki, K., Vidale, J. E., Lee, W. H. K., & Marone, C. J. (1994). Fine Structure of the Landers Fault Zone: Segmentation and the Rupture Process. *Science*, *265*(5170), 367–370. doi: 10.1126/science.265.5170.367

Mai, P. M., & Beroza, G. C. (2002). A spatial random field model to characterize complexity in earthquake slip. *Journal of Geophysical Research: Solid Earth*, *107*(B11), ESE 10–1. doi: 10.1029/2001JB000588

Mai, P. M., Galis, M., Thingbaijam, K. K. S., Vyas, J. C., & Dunham, E. M. (2017). Accounting for Fault Roughness in Pseudo-Dynamic Ground-Motion Simulations. *Pure and Applied Geophysics*, *174*(9), 3419–3450. doi: 10.1007/s00024-017-1536-8

Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigl, K., & Rabaute, T. (1993). The displacement field of the Landers earthquake mapped by radar interferometry. *Nature*, *364*(6433), 138–142. doi: 10.1038/364138a0

Meier, M.-A., Ampuero, J. P., & Heaton, T. H. (2017). The hidden simplicity of subduction megathrust earthquakes. *Science*, *357*(6357), 1277–1281. doi: 10.1126/science.aan5643

Meng, L., Ampuero, J.-P., Stock, J., Duputel, Z., Luo, Y., & Tsai, V. C. (2012). Earthquake in a maze: Compressional rupture branching during the 2012 Mw 8.6 Sumatra earthquake. *Science*, *337*(6095), 724–726.

Peyrat, S., Olsen, K., & Madariaga, R. (2001, November). Dynamic modeling of the 1992 Landers earthquake. *Journal of Geophysical Research: Solid Earth*, *106*(B11), 26467–26482. doi: 10.1029/2001JB000205

Ripperger, J., Ampuero, J.-P., Mai, P. M., & Giardini, D. (2007). Earthquake source characteristics from dynamic rupture with constrained stochastic fault stress. *Journal of Geophysical Research: Solid Earth*, *112*(B4). doi:

10.1029/2006JB004515

Shearer, P. M., Prieto, G. A., & Hauksson, E. (2006). Comprehensive analysis of earthquake source spectra in southern California. *Journal of Geophysical Research: Solid Earth*, *111*(B6). doi: 10.1029/2005JB003979

Trugman, D. T., & Dunham, E. M. (2014). A 2d Pseudodynamic Rupture Model Generator for Earthquakes on Geometrically Complex FaultsA 2d Pseudodynamic Rupture Model Generator for Earthquakes on Geometrically Complex Faults. *Bulletin of the Seismological Society of America*, *104*(1), 95–112. doi: 10.1785/0120130138

Vallée, M. (2013). Source time function properties indicate a strain drop independent of earthquake depth and magnitude. *Nature Communications*, *4*, 2606. doi: 10.1038/ncomms3606

Vallée, M., Charléty, J., Ferreira, A. M. G., Delouis, B., & Vergoz, J. (2011). 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. *Geophysical Journal International*, *184*(1), 338–358. doi: 10.1111/j.1365-246X.2010.04836.x

Vallée, M., & Douet, V. (2016). A new database of source time functions (STFs) extracted from the SCARDEC method. *Physics of the Earth and Planetary Interiors*, *257*, 149–157. doi: 10.1016/j.pepi.2016.05.012

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., & Wobbe, F. (2013). Generic Mapping Tools: Improved Version Released. *Eos, Transactions American Geophysical Union*, *94*(45), 409–410. doi: 10.1002/2013EO450001

Wibberley, C. A. J., & Shimamoto, T. (2005). Earthquake slip weakening and asperities explained by thermal pressurization. *Nature*, *436*(7051), 689–692. doi: 10.1038/nature03901

Yagi, Y., & Fukahata, Y. (2011). Rupture process of the 2011 Tohoku-oki earthquake and absolute elastic strain release. *Geophysical Research Letters*, *38*(19). doi: 10.1029/2011GL048701

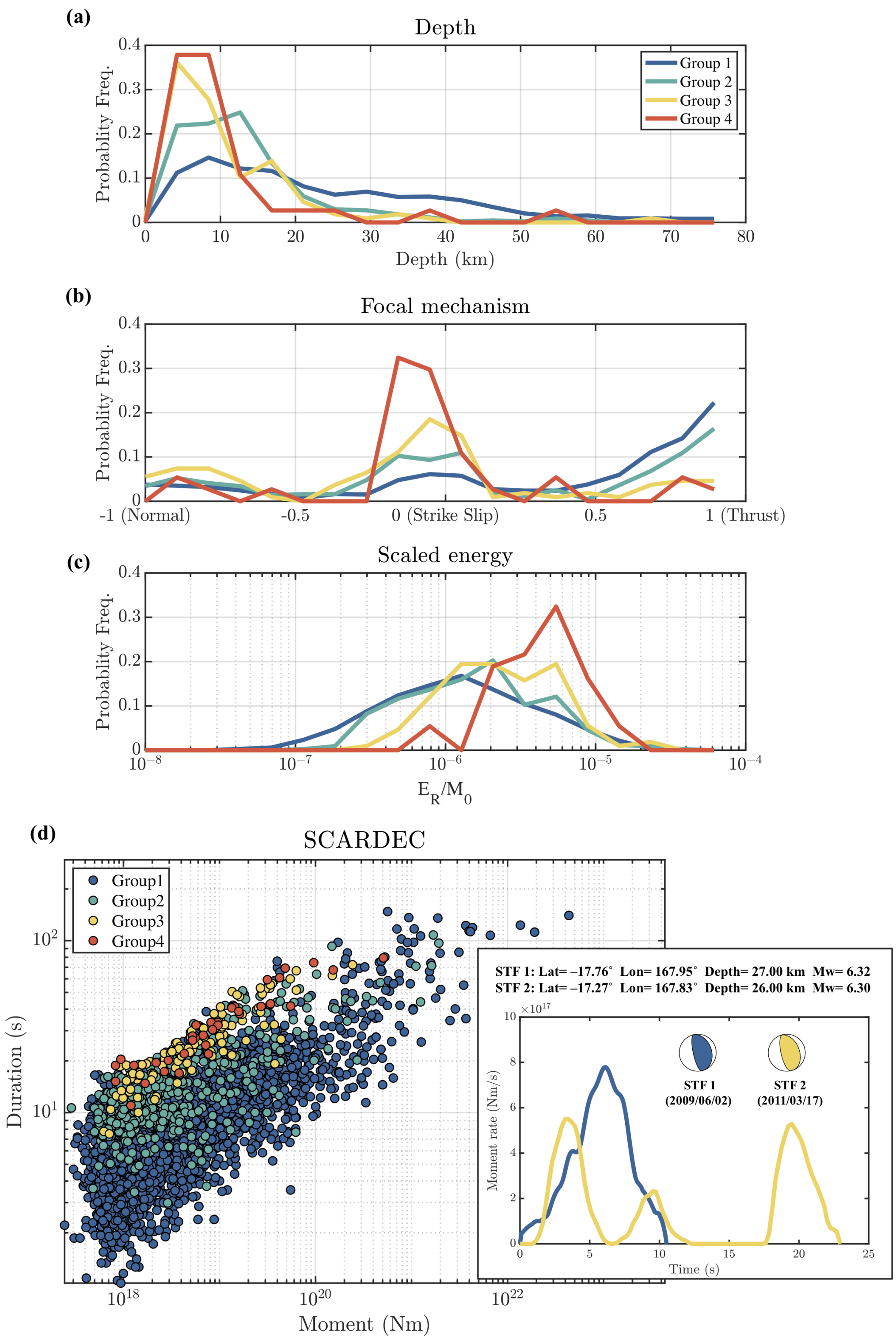

Figure 2. Population distribution of four complexity groups and correlation with different source parameters: (a) centroid depth, (b) focal mechanism (scalar defined by Shearer et al. (2006) that varies from -1 (normal), 0 (strike-slip) to 1 (reverse)), (c) and scaled radiated energy $e = E_R/M_0$. Panel (d) shows the earthquake duration against earthquake moment, colored with the respective group labels. One pair of co-located events with different complexity are also shown in the inset.

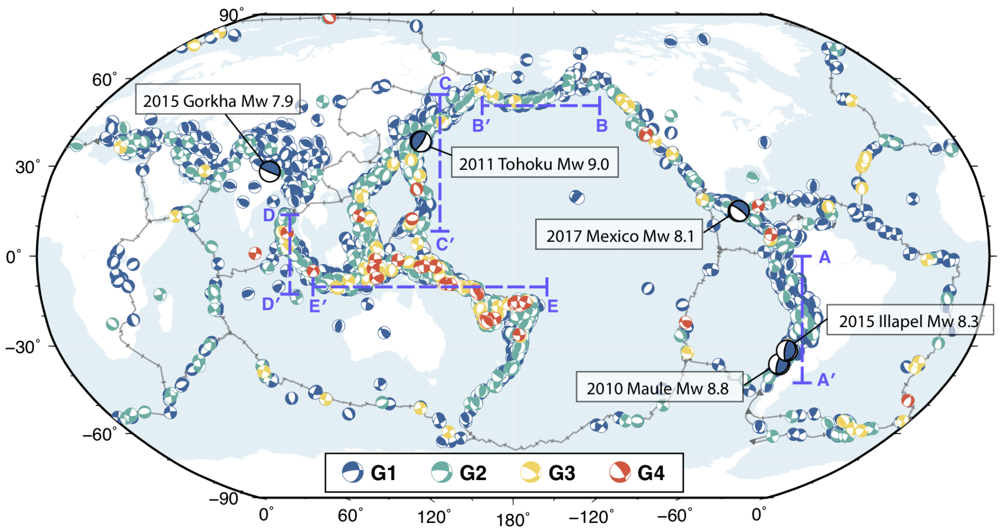

Figure 3. Map of focal mechanisms colored by their group label and overlay of the plate boundaries (gray thin lines). Several recent large megathrust earthquakes are highlighted. Blue dashed lines shown the locations of profiles in Figure 4. Bottom panels show the center STFs in each groups (same as those in Figure 1 (a)) as well as the corresponding schematic rupture propagation.

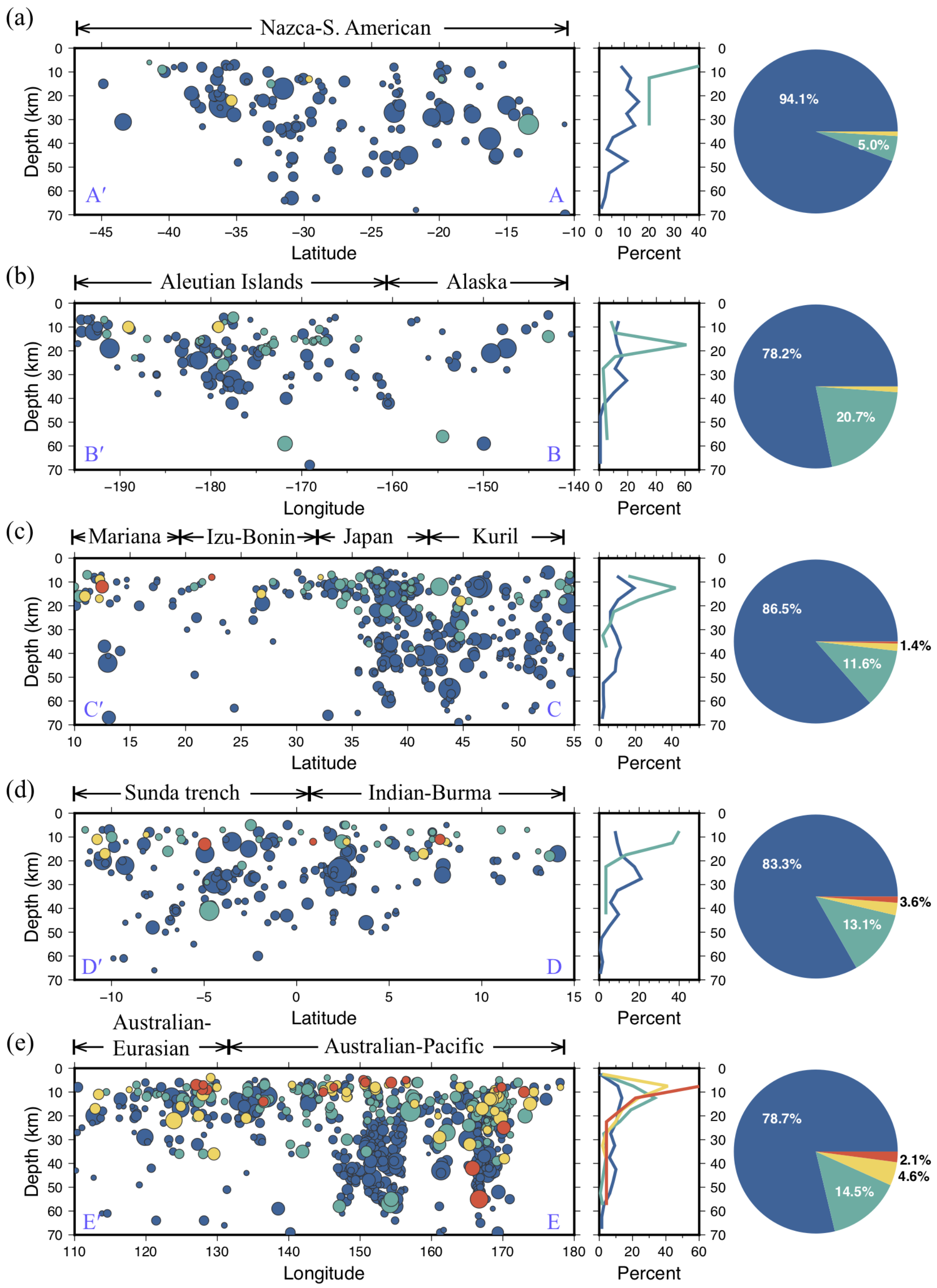

Figure 4. Earthquake distributions of different complexity groups on the vertical profiles (from 0–70 km, locations are indicated by blue dashed lines in Figure 3). The regional along-depth and total group distributions are also shown to the right.

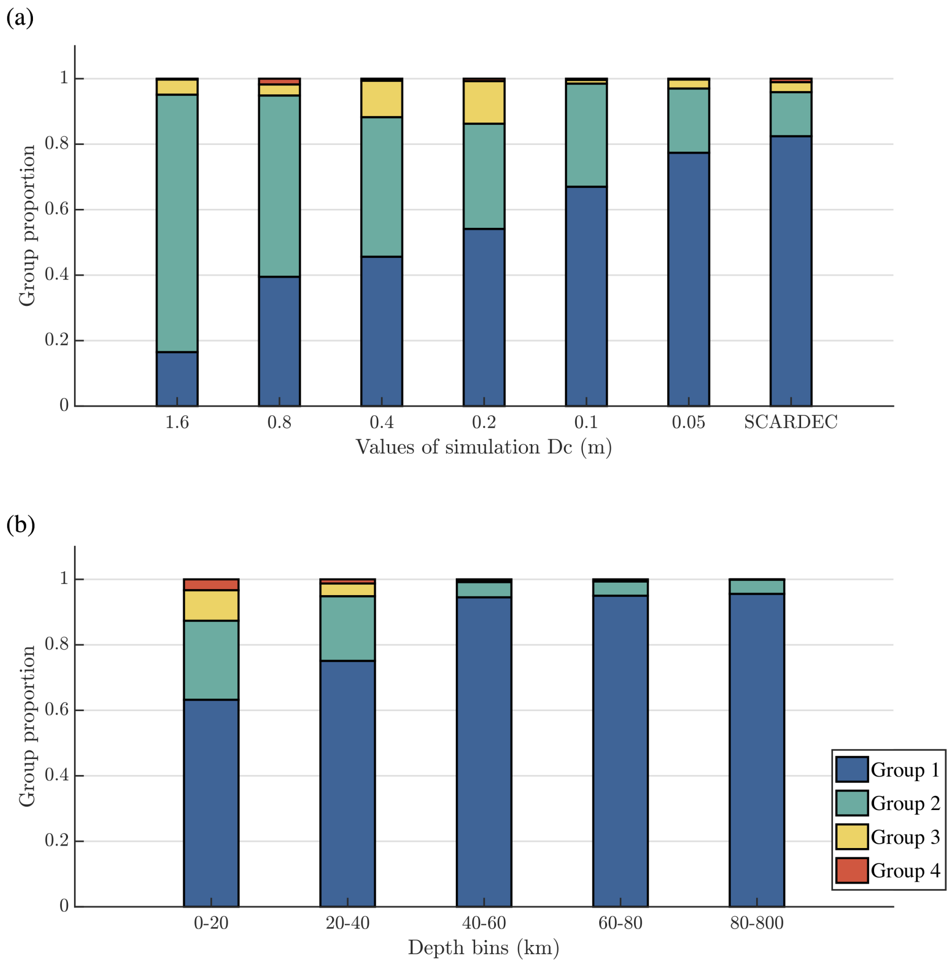

Figure 5. Group proportion distributions: (a) simulated STFs clustering with different values of D_c, compared with the group proportions of real STFs (SCARDEC); (b) Group proportions of real STFs (SCARDEC) within different depth bins.

Supporting Information for "Source time function clustering reveals patterns in earthquake dynamics"

Jiuxun Yin[1], Zefeng Li[2], Marine Denolle[1]

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

[2]Seismological Laboratory, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

Contents of this file

Introduction

In the supporting information for "Source time function clustering reveals patterns in earthquake dynamics", we present additional information on the methods and results. First we provide supplementary information about the clustering of SCARDEC STFs (Text S1, Figures S1 – S6). Second we provide detailed information about the dynamic simulation (Text S2, Figures S7-S12).

Text S1. Dynamic Time Warping clustering of STFs

We downloaded the global catalog of STFs from 3529 $M_W \geq 5.5$ earthquakes from SCARDEC source time function database (`http://scardec.projects.sismo.ipgp.fr`, last accessed 01/20/2020). In this database, there are two types of STFs, average and optimal. In this study, we use the average STFs because their time derivative are not discontinuous. All STFs are resampled over 100 points. The purpose of this step is to retain signals at periods as short at 1 s, while it is not required for the DTW stretching. We also have tested resampling at 200 and 500 points, but our results are insensitive to the number of points. Finally, all STFs are normalized by the seismic moment.

DTW searches for the best point-to-point match between two STFs (Figure S1 (a)) to match their general shapes. The best corresponding relation (white line in Figure S1 (b)) provides an optimal warping/stretching path, along which two STFs can be stretched to the best similarity. The DTW distance is the Euclidian norm once both STFs are warped. Single linkage hierarchical clustering is applied to the DTW distances to build the "family tree" for the entire STF database. 20 clusters are finally determined to keep the rich variations of STF complexity without diving into numerous individual shapes (Figure S1 (c)).

Text S2. Dynamic rupture simulations

Our dynamic simulations are similar with those in (Danré et al., 2019), with slight modification in the pre-stress distribution and in the range of values of frictional parameters. We solve the elastodynamic equations of a mode III fracture with linear slip-weakening friction in a homogeneous infinite medium 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 total length of simulation domain is fixed as 400 km, but the length of the "fault" where rupture can occur is 200 km. Basic material properties are: P wave velocity V_p = 6.00 km/s; shear wave velocity V_s=3.46 km/s; density ρ=2.67 kg/m^3; shear modulus G=32 GPa and we fix the normal stress σ_0=120 MPa. The linear slip weakening friction is used as a simple but general constitutive relation:

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

where d is slip, dynamic friction μ_d=0.525, static friction μ_s=0.677. We vary the characteristic slip weakening distance D_c = 0.05, 0.1, 0.2, 0.4, 0.8, 1.6 m. The nucleation length L_c relates to D_c:

$$L_c = \frac{1.158GD_c}{(\mu_s - \mu_d)\sigma_0}, \tag{2}$$

which varies from 101.6 to 3250.5 m (Uenishi & Rice, 2003). In our simulations, we set the nucleation patch to be 10 km in extent, which is at least $3L_c$ to guarantee the successful nucleation. The cohesive zone size is:

$$\Lambda_0 = \frac{9\pi}{32}\frac{G}{1-\nu}\frac{D_c}{(\mu_s - \mu_d)\sigma_0}, \tag{3}$$

which varies from 103.3 to 3306.9 m and where ν = 0.25 is the Poisson ratio (Day et al., 2005). To guarantee sufficient spatial resolution, we require spatial sampling along the fault axis x of $\Delta x \leq \Lambda_0/2$ at least for each D_c value.

To generate diverse dynamic ruptures, and their corresponding STFs, we generate statistically similar shear pre-stress $\tau_0(x)$ distributions on the fault plane. To a constant level of shear stress, which equals to the dynamic friction $\mu_d \sigma_0$, we add a perturbation $d\tau_0(x)$, such that the pre-stress is:

$$\tau_0(x) = \mu_d \sigma_0 + d\tau_0(x). \tag{4}$$

The power spectral density (PSD) of $d\tau_0(x)$, $d\mathrm{T}_0(k)$, follows power-law decay in the wavenumber domain,

$$d\mathrm{T}_0(k) = C|k|^{-\gamma}, \tag{5}$$

where $\gamma = 0.8$ is based on observational constraints on the self-afine fault roughness (Dunham et al., 2011; Candela et al., 2012), and C is a normalization factor. Combining the PSD $d\mathrm{T}_0(k)$ with the random phases $\phi(k)$, which are taken from a uniform distribution in $[0, 2\pi]$, we can generate various pre-stress distributions. For each realization of a pre-stress perturbation, we further scale the pre-stress perturbation amplitude to vary within the range from $-0.6(\mu_s - \mu_d)\sigma_0$ to $0.8(\mu_s - \mu_d)\sigma_0$. Finally, we apply a Tukey-window to taper the 100 km on either end of the 400 km pre-stress distributions; this avoids the artifacts in STF from abruptly stopping of rupture at the fault boundary in the spectral boundary integral solutions (Figure S7).

To nucleate spontaneous dynamic ruptures, we apply a weakening nucleation. For each pre-stress distribution, we first perform a peak detection of $\tau_0(x)$ to find its absolute maximum τ_0^{max} within within [-50 50] km. Then, we reduce the fault strength $\tau_s = \tau_0^{max} - 4$MPa within a 10 km nucleation region centered at this point, and set $D_c = 0.1$ m in this nucleation region.

For simulations with different D_c values, we keep the identical nucleation processes by fixing the $D_c = 0.1$ m within the nucleation zone. This is to minimize the effects from nucleation on the STFs. Once nucleated, slip and stresses evolve according to elastodynamics.

We repeat the workflow above to generate diverse ruptures. We remove those that unsuccessfully nucleated (L $\leq$ 20 km) or over-ruptured (those that ruptured over the heterogeneous area at $\pm$ 100 km in Figure S7). We run a sufficient number of ruptures in order to keep 800 qualified dynamic rupture models for each D_c value. We then apply the same approach as in the case of the observations (Figures S8-S12) to cluster those synthetic STFs based on their complexities.

References

Brune, J. N. (1971). Correction (to Brune, 1970). *J. geophys. Res*, *76*, 5002.

Candela, T., Renard, F., Klinger, Y., Mair, K., Schmittbuhl, J., & Brodsky, E. E. (2012). Roughness of fault surfaces over nine decades of length scales. *Journal of Geophysical Research: Solid Earth*, *117*(B8). doi: 10.1029/2011JB009041

Danré, P., Yin, J., Lipovsky, B. P., & Denolle, M. A. (2019). Earthquakes Within Earthquakes: Patterns in Rupture Complexity. *Geophysical Research Letters*, *46*(13), 7352–7360. doi: 10.1029/2019GL083093

Day, S. M., Dalguer, L. A., Lapusta, N., & Liu, Y. (2005). Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture. *Journal of Geophysical Research: Solid Earth*, *110*(B12). doi: 10.1029/2005JB003813

Dunham, E. M., Belanger, D., Cong, L., & Kozdon, J. E. (2011). Earthquake Ruptures with Strongly Rate-Weakening Friction and Off-Fault Plasticity, Part 2: Nonplanar FaultsEarthquake Ruptures with Rate-Weakening Friction and Off-Fault Plasticity, Part 2: Nonplanar Faults. *Bulletin of the Seismological Society of America*, *101*(5), 2308–2322. doi: 10.1785/0120100076

Eshelby, J. D. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 241, pp. 376–396). The Royal Society.

Noda, H., Lapusta, N., & Kanamori, H. (2013). Comparison of average stress drop measures for ruptures with heterogeneous stress change and implications for earthquake physics. *Geophysical Journal International*, *193*(3), 1691–1712. doi: 10.1093/gji/ggt074

Uenishi, K., & Rice, J. R. (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

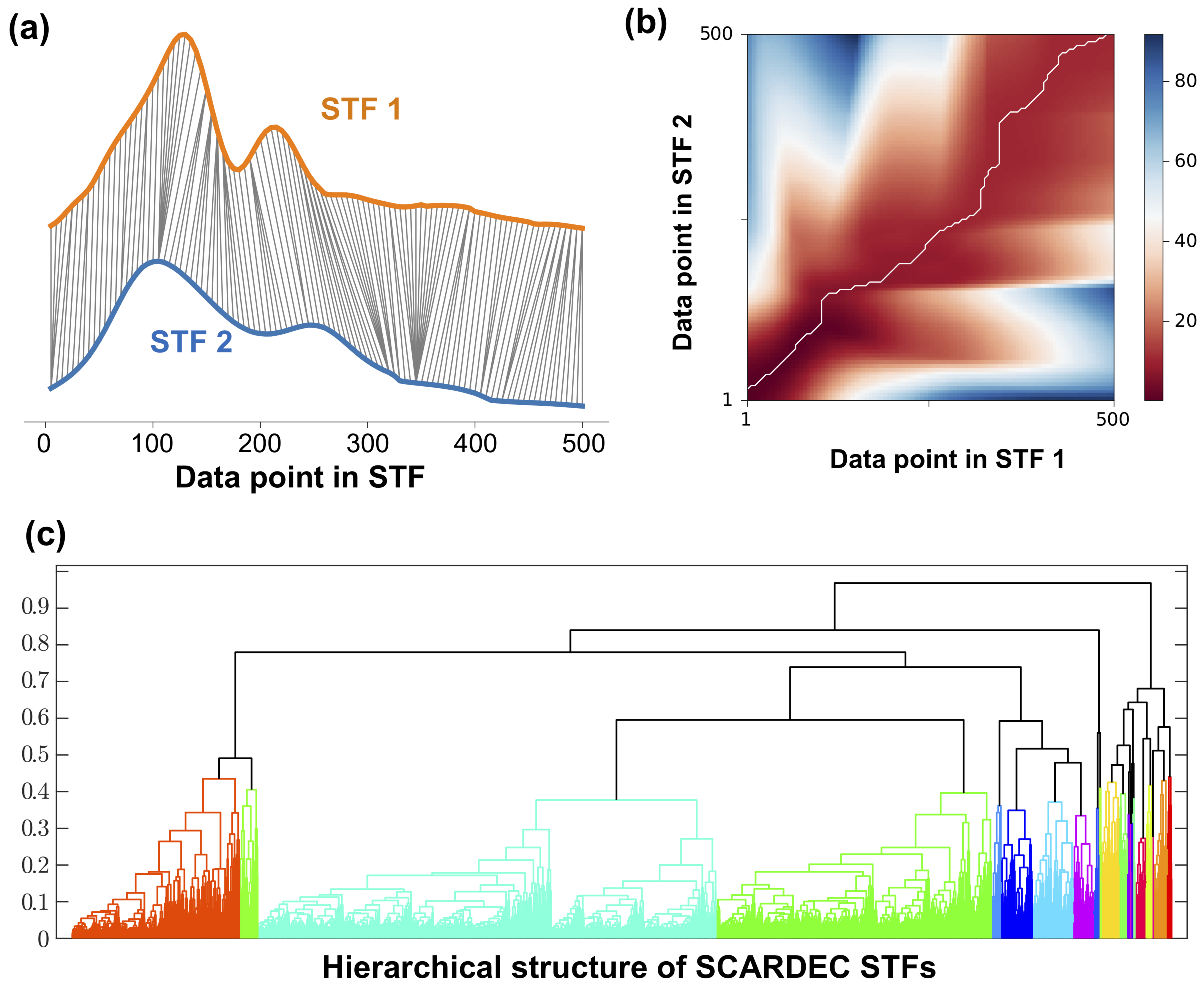

Figure S1. Dynamic time warping (DTW) clustering of earthquake source time functions (STFs). (a) Point-to-point correspondence between two example STFs. (b) Optimal stretching path (white line) from the minimum differences for the two example STFs. (c) Hierarchical structure of all SCARDEC STFs from the DTW clustering.

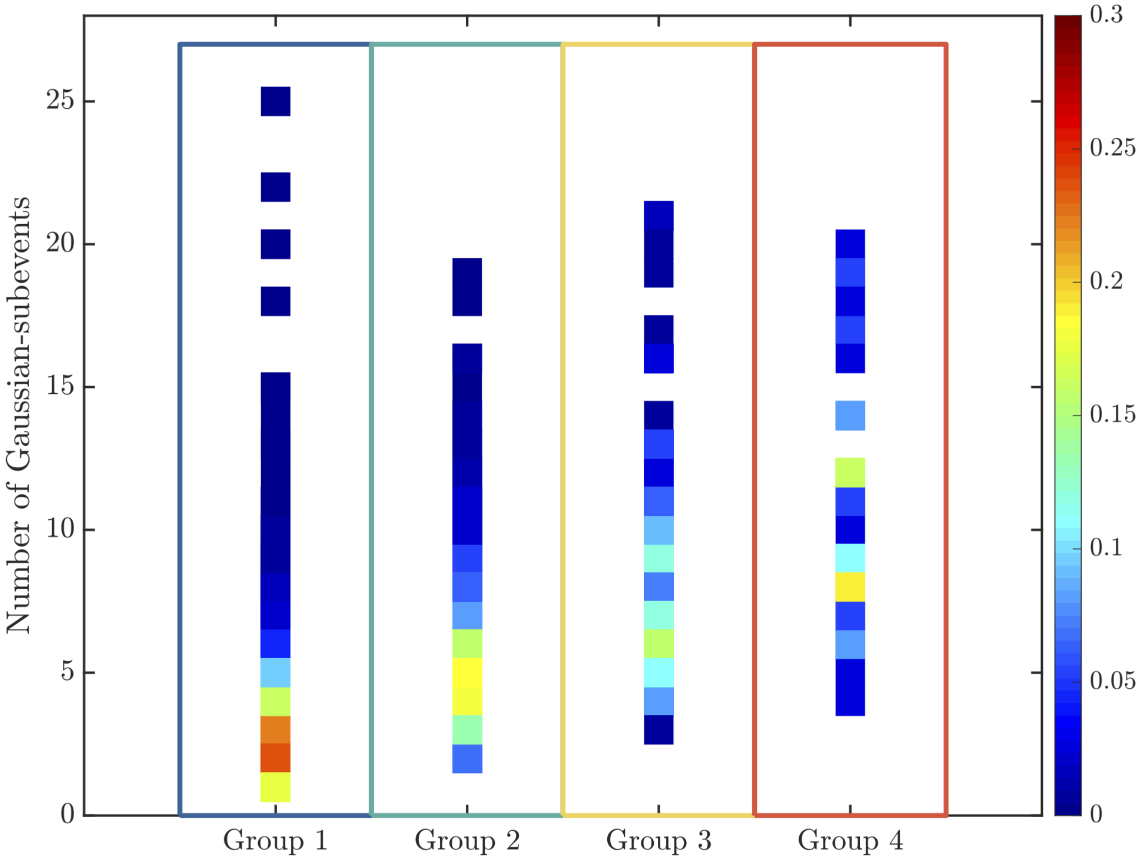

Figure S2. Comparison between the DTW complexity groups and number of Gaussian subevents (Danré et al., 2019). The color indicates the frequency of occurrence within each group.

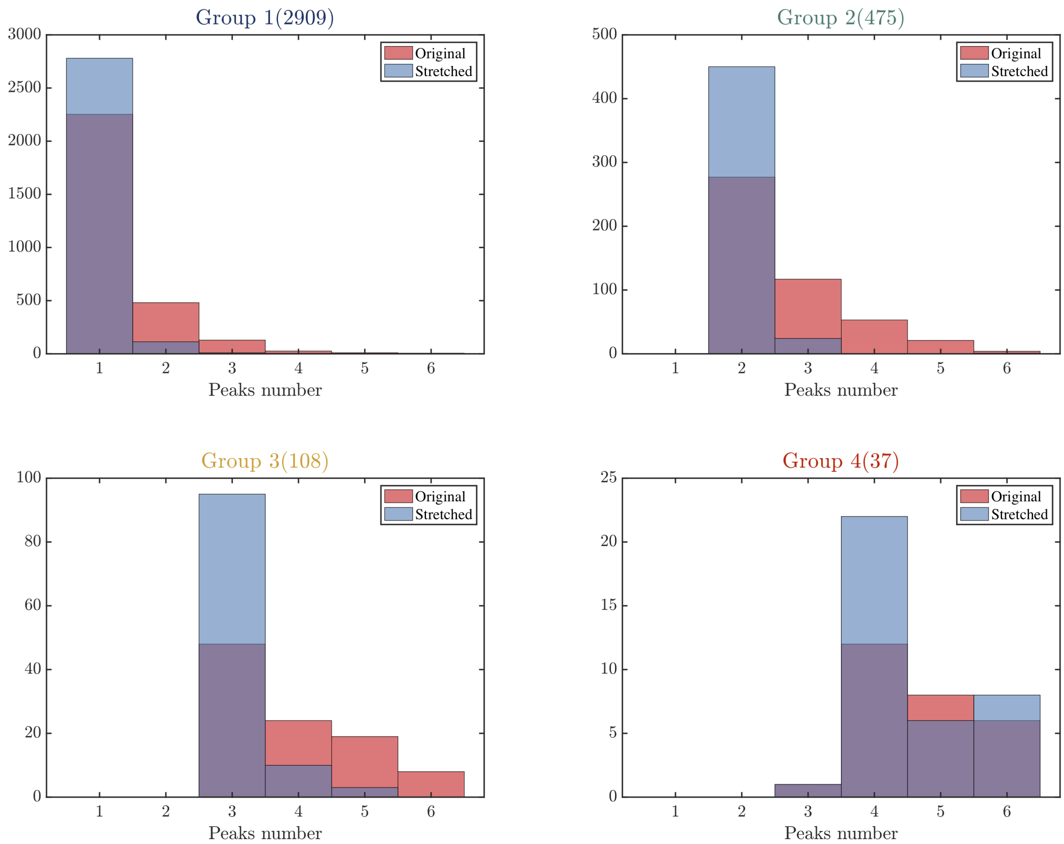

Figure S3. Comparisons between prominent peak (0.1 of STF global maximum) number distributions of original raw STFs (red histograms) and DTW stretched STFs (blue histograms) in each group. Group numbers are also the prominent peak numbers of the centroid event within each group.

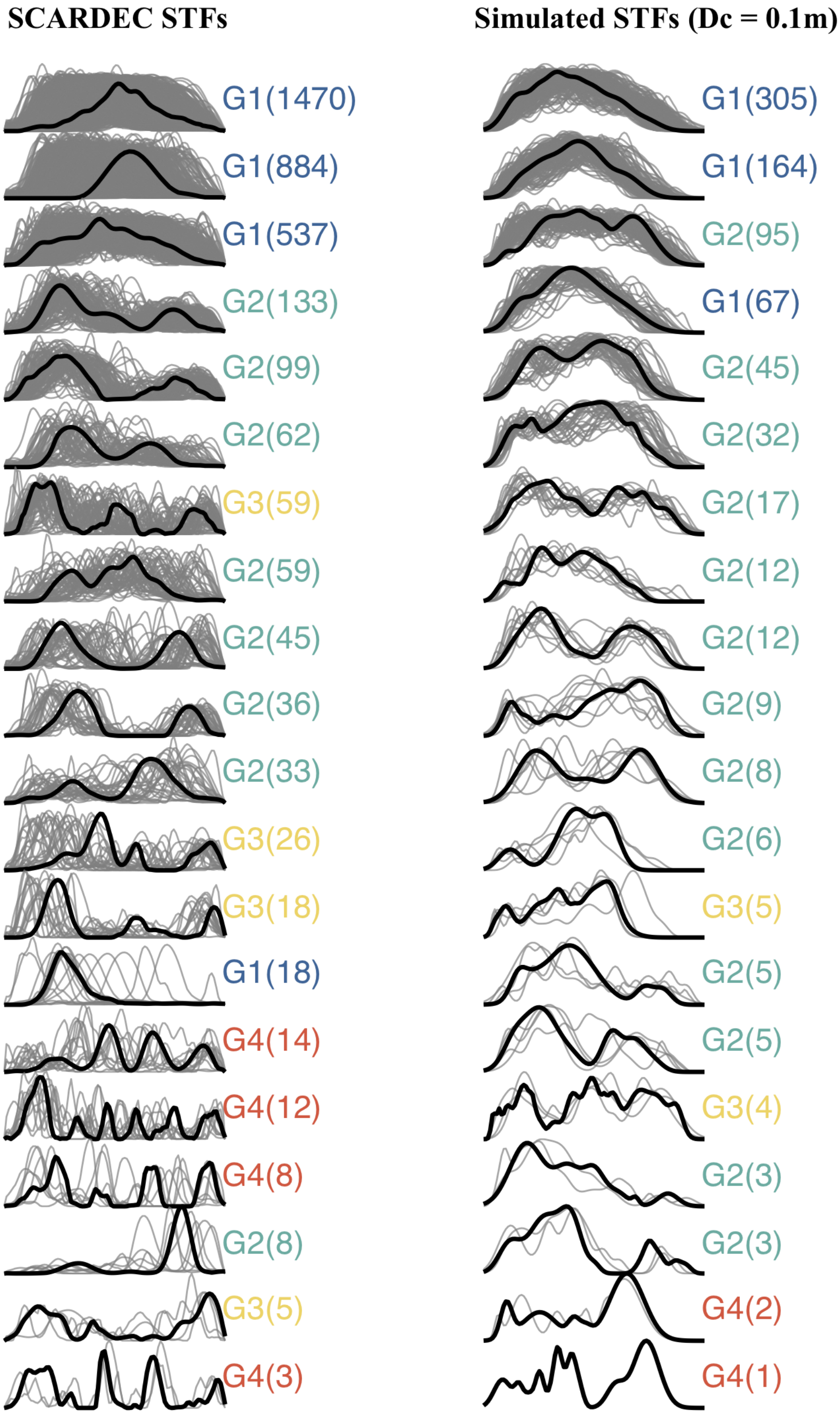

Figure S4. All the STFs before DTW stretching (gray thin lines) compared with the centroid STF (Black thick lines). Other symbols are the same as Figure 1. SCARDEC STFs are shown to the left and simulated STFs ($D_c = 0.1$ m) are shown to the right.

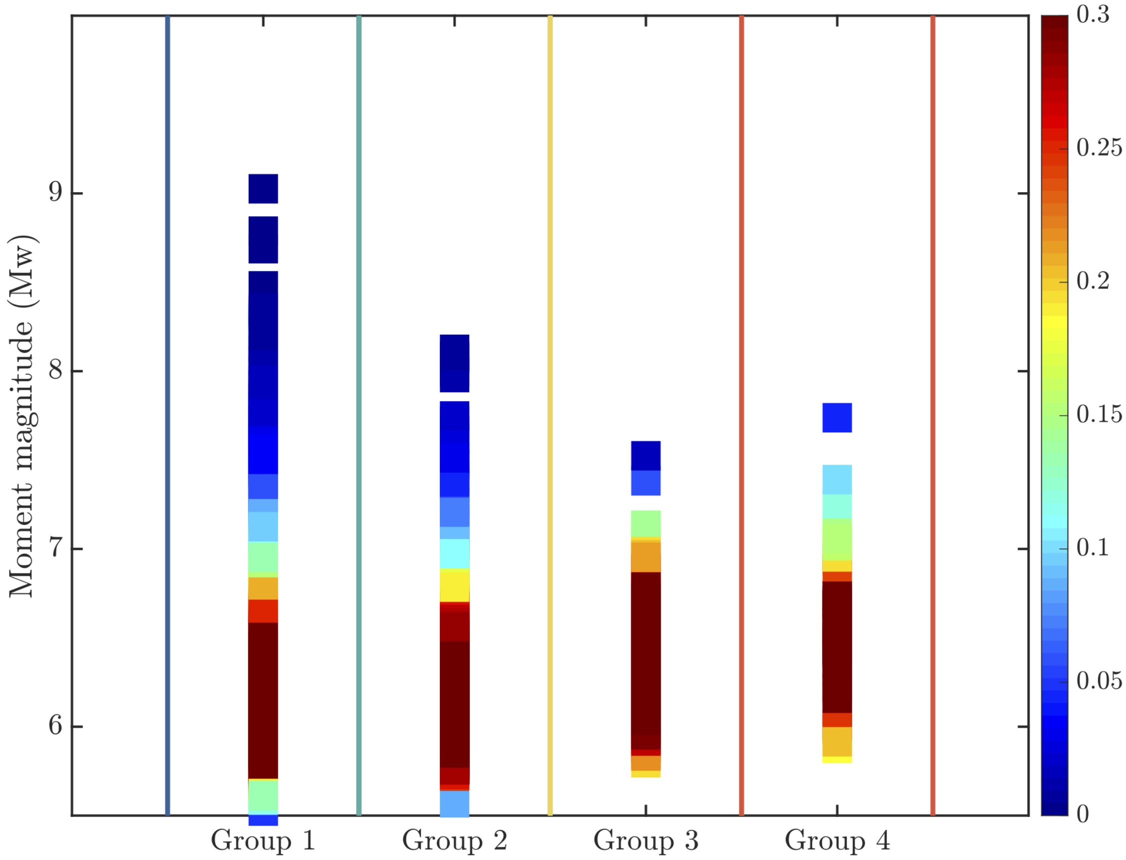

Figure S5. Moment magnitude distributions of the STFs in each group. The color indicates the frequency of occurrence within each group.

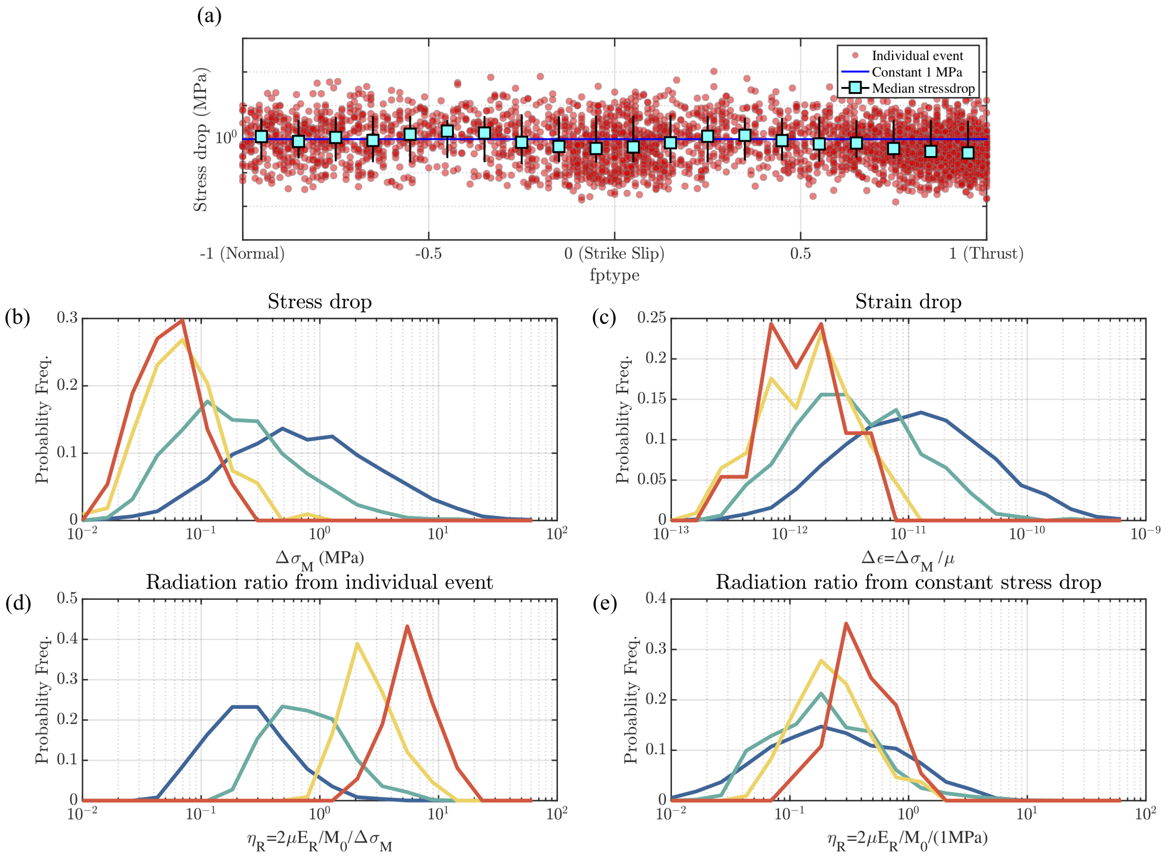

Figure S6. Seismic stress drop estimation using the earthquake duration T_D for each individual earthquake in the SCARDEC database: Panel (a) shows the stress drop variations with focal mechanisms parameters, the stress drop is calculated as $\Delta\tau = 7/16 M_0/(0.32 V_s T_D)^3$ (Eshelby, 1957; Brune, 1971). In the dynamic simulation, the average stress drop of all models is approximately 1 MPa. Panel (b) shows the group distributions of estimated stress drop based on event duration. (c) and (d) show the group distributions of corresponding strain drop and radiation ratio calculated from stress drop, respectively. Note that the stress drop estimation based on duration is model-dependent and may be underestimated for the very heterogeneous earthquake rupture (Noda et al., 2013), such as the complex Group 3 and Group 4, thus leads to very high radiation efficiency. Panel (e) also shows the group distributions of radiation ratio, but estimated based on the assumption that stress drop is a constant value of 1 MPa, for comparison.

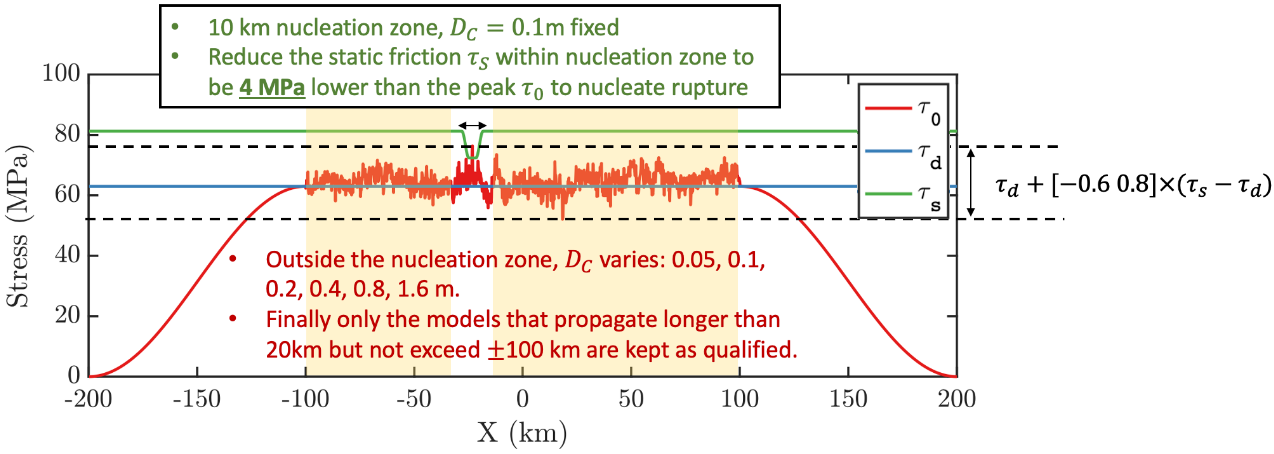

Figure S7. Pre-stress (red curve) and frictional strengths (green curve: static friction; blue curve: dynamic friction) settings of the dynamic rupture simulations. Dashed lines indicate range of values of the randomly generated pre-stress. Finally, only the rupture models terminates within the yellow shadow regions are kept as the qualified models.

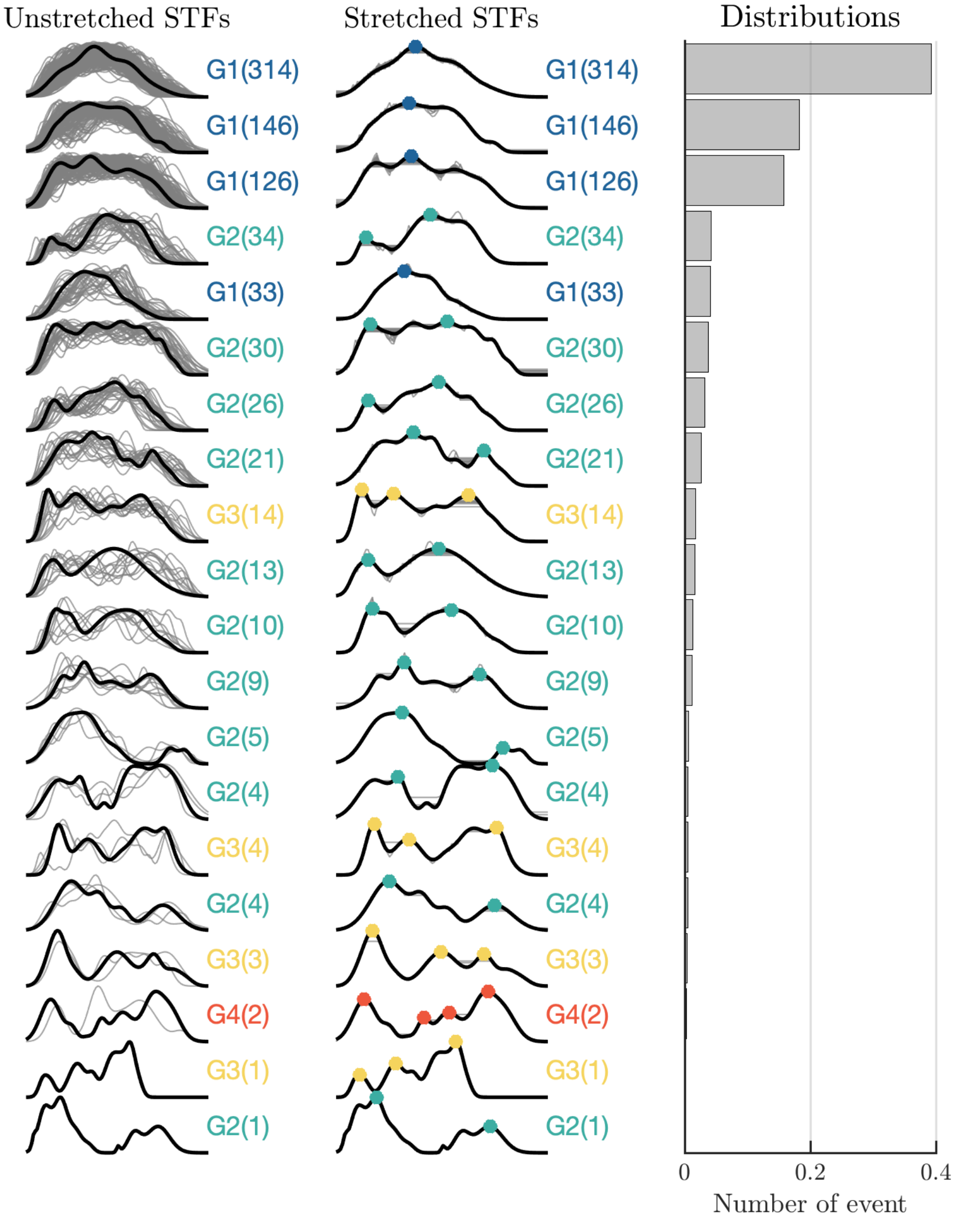

Figure S8. DTW clustering results for the simulated STFs with $D_c = 0.05$ m.

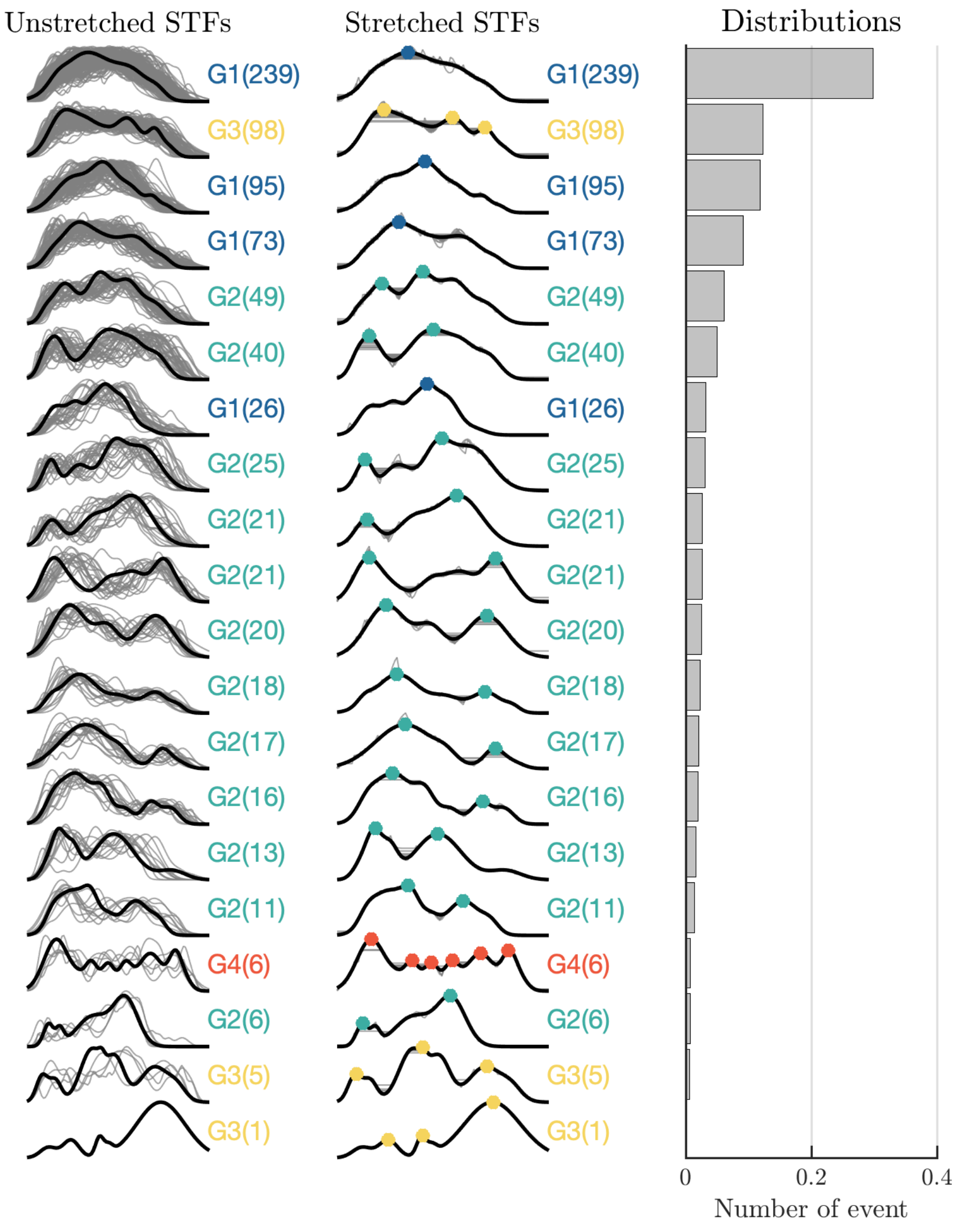

Figure S9. DTW clustering results for the simulated STFs with $D_c = 0.2$ m.

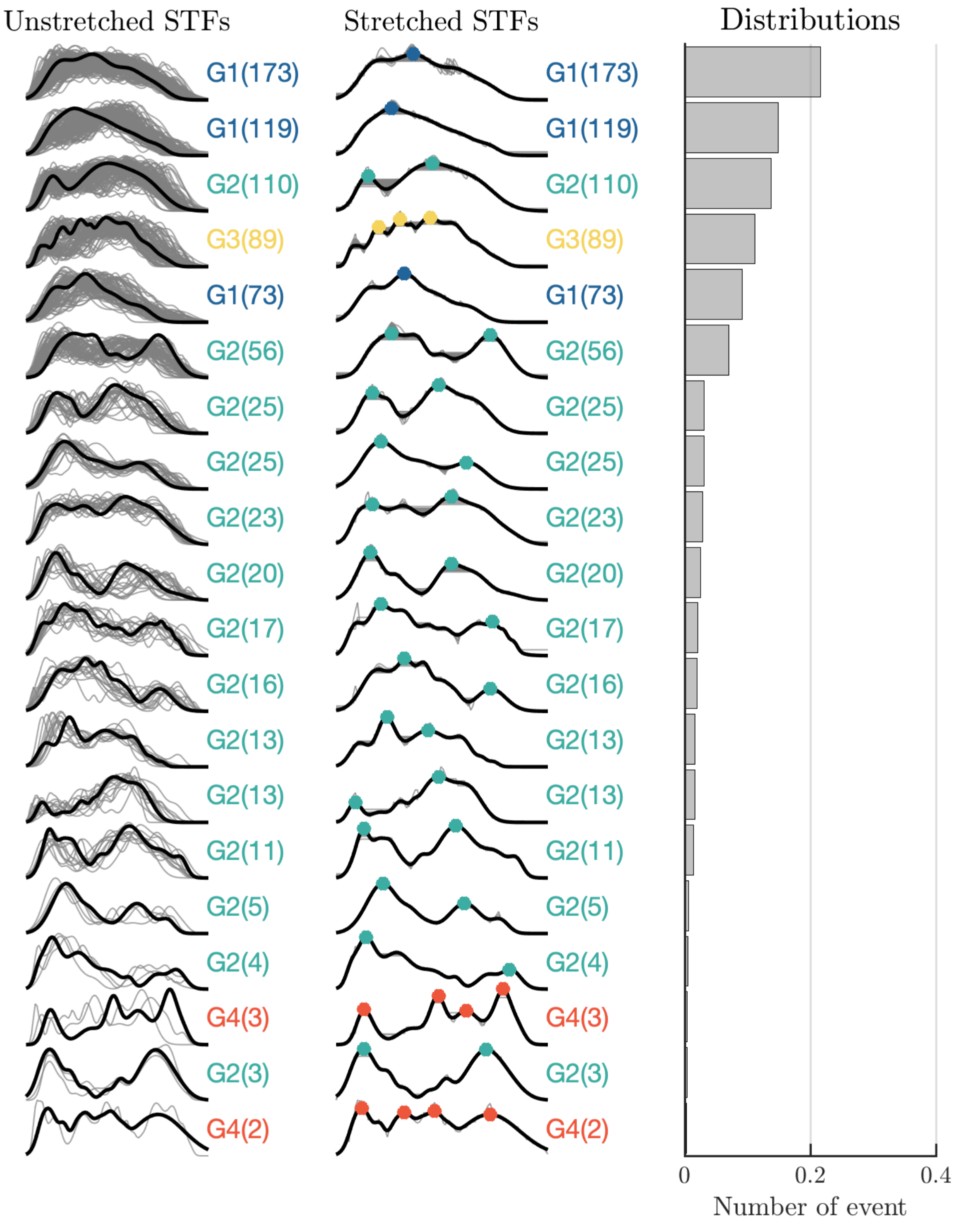

Figure S10. DTW clustering results for the simulated STFs with $D_c = 0.4$ m.

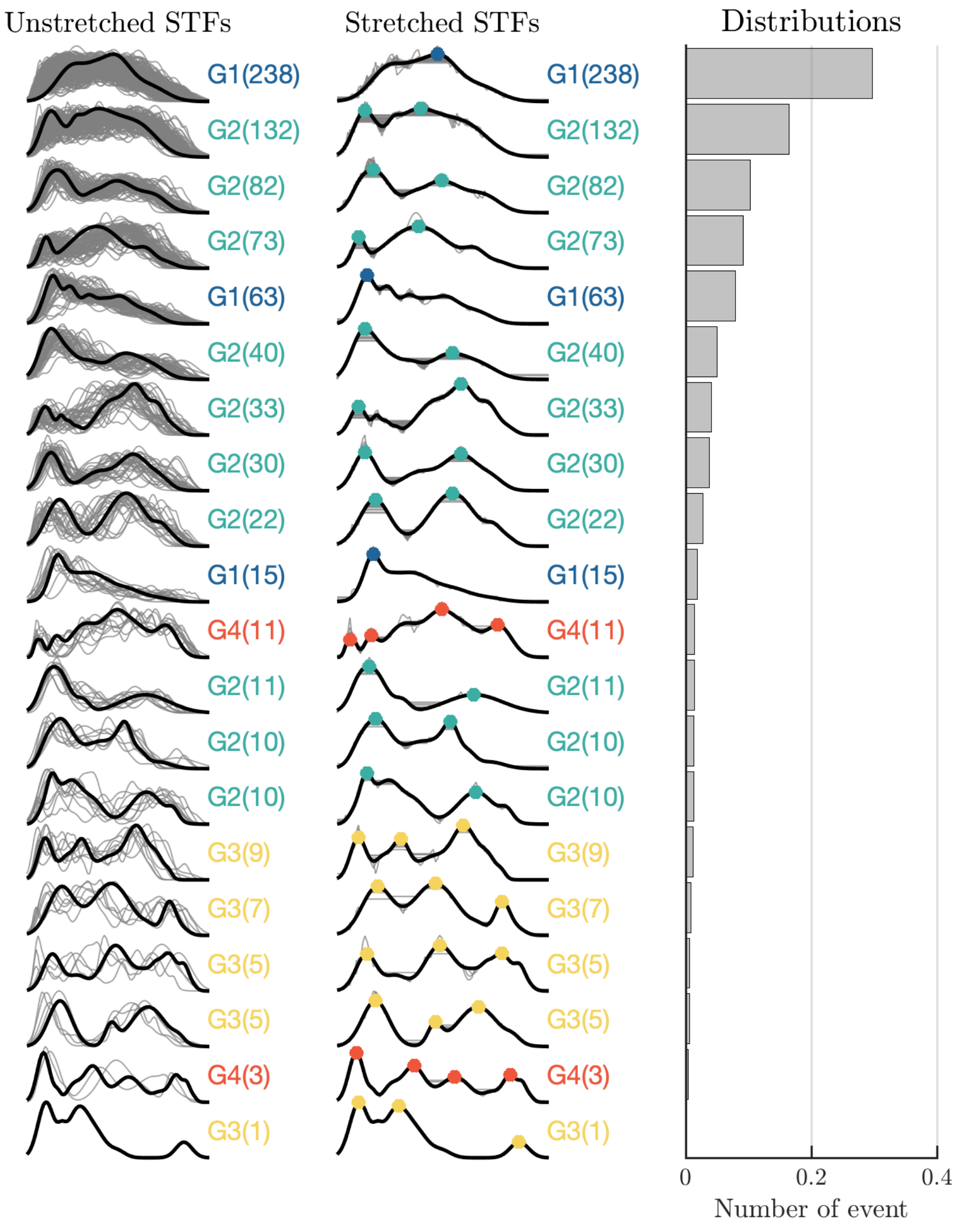

Figure S11. DTW clustering results for the simulated STFs with $D_c = 0.8$ m.

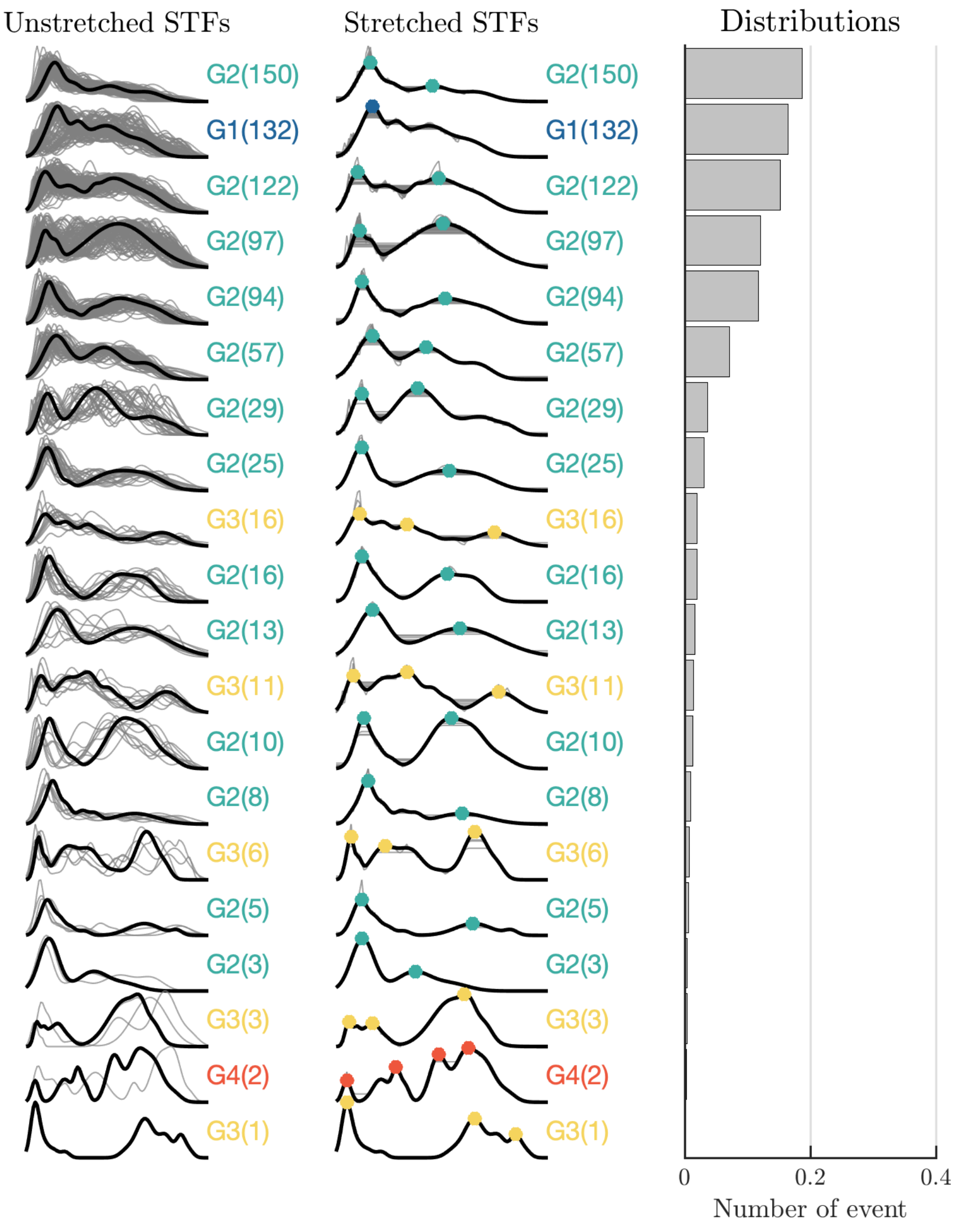

Figure S12. DTW clustering results for the simulated STFs with $D_c = 1.6$ m.