

Source Time Function Clustering Reveals Patterns in Earthquake Dynamics

Jiuxun Yin¹, Zefeng Li^{*2,3}, and Marine A. Denolle¹

Abstract

We cluster a global database of 3529 $M_w > 5.5$ earthquakes in 1995–2018 based on a dynamic time warping distance between earthquake source time functions (STFs). The clustering exhibits different degrees of complexity of the STF shapes and suggests an association between STF complexity and earthquake source parameters. Most of the thrust events have simple STF shapes across all depths. In contrast, earthquakes with complex STF shapes tend to be located at shallow depths in complicated tectonic regions, exhibit long source duration compared with others of similar magnitude, and tend to have strike-slip mechanisms. With 2D dynamic modeling of dynamic ruptures on heterogeneous fault properties, 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 frictional properties. In particular, the characteristic slip-weakening distance could be constrained to be short (< 0.1 m) and depth dependent if stress drop is in general constant.

Cite this article as Yin, J., Z. Li, and M. A. Denolle (2021). Source Time Function Clustering Reveals Patterns in Earthquake Dynamics, *Seismol. Res. Lett.* **XX**, 1–11, doi: [10.1785/0220200403](https://doi.org/10.1785/0220200403).

[Supplemental Material](#)

Introduction

Earthquakes are known to break in diverse manners: some events rupture on a geometrically simple fault with a relatively smooth slip distribution (Yagi and Fukahata, 2011), whereas 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 interevent variability among earthquakes (Allmann and Shearer, 2009; Atik *et al.*, 2010; Chounet and Vallée, 2018; Ye *et al.*, 2018; Denolle, 2019). Such variability may partly come from differences in data processing strategy (Ide and Beroza, 2001). Therefore, large catalogs of STFs (or their spectra) obtained from a uniform approach is preferable to analyze relative differences among earthquakes (Allmann and Shearer, 2009; Convers and Newman, 2011; Denolle and Shearer, 2016; Vallée and Douet, 2016; Chounet and Vallée, 2018; Ye *et al.*, 2018).

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 and Newman, 2011; Baltay *et al.*, 2014; Denolle and 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

1. Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, U.S.A., <https://orcid.org/0000-0003-0641-5680> (JY); 2. Laboratory of Seismology and Physics of Earth's Interior, School of Earth and Space Sciences, University of Science and Technology of China, Hefei, China, <https://orcid.org/0000-0003-4405-8872> (ZL); 3. Mengcheng National Geophysical Observatory, University of Science and Technology of China, Mengcheng, China

*Corresponding author: zefengli@ustc.edu.cn

© Seismological Society of America

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 STF shapes differ as well. For instance, Meier *et al.* (2017) find that average STF shapes have rather a triangle shape, whereas Denolle (2019) suggests a rather skewed-Gaussian functional form.

Here, we propose to weaken the assumption of a choice in source duration metrics and instead use dynamic time warping (DTW) to compare the shapes of the STF shapes. DTW has been widely used in speech recognition (Berndt and Clifford, 1994; Müller, 2007) to handle varying speech speeds. The DTW algorithm performs a nonuniform stretching of time and amplitude to match the shape of two time series via the optimal warping path with minimum distance (Fig. S1, available in the supplemental material to this article). We measure the similarity between STF shapes with the DTW distance and cluster the STF shapes according to the DTW distance. We apply this to the global Seismic source ChAracteristics Retrieved from DEConvolving teleseismic body waves (SCARDEC) catalog of STF shapes (Vallée and Douet, 2016, see Data and Resources), which contains 3529 earthquakes of magnitude > 5.5 from 1 January 1992 and until 31 December 2018. Using the developed DTW methodology, we finally regroup the clusters based on the complexity of their centroid and form four groups. We find that the association with a particular group or degree of complexity 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 STF shapes. 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 proportion of groups in the SCARDEC database is most similar to those simulated STF shapes with small values of D_c . Thus, the group proportion of a large number of STF shapes can potentially provide observational constraints to earthquake dynamics.

DTW 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) nonuniformly in the time dimensions to optimally match two series (Fig. S1). This algorithm is widely used in automated speech recognition in which different audio sequences may have different speaking speeds (Berndt and Clifford, 1994; Müller, 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 preconditioning, (2) DTW distance calculation, (3) clustering, and (4) regrouping around a centroid event.

We first perform minimal preconditioning of the STF shapes. The STF shapes are built from the deconvolution of teleseismic P waves that are relatively well constrained at frequencies <1 Hz (Vallée and Douet, 2016). Given that the maximum duration of the STF in the catalog is ~100 s, we resample the data to 100 points giving a minimum sampling rate of 1 point per second. We then normalize the amplitude STF shapes 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 STF shapes, which did not affect the conclusions of this analysis.

Second, we apply the DTW to each pair of STF shapes. The DTW distance is the Euclidean distance between two STF shapes warped along the optimal warping path and is chosen here as the measure of similarity between two STF shapes (see Fig. S1a,b). We apply a global search of the shortest warping distance and thus allow for the maximum level of distortion. The warping follows causality such that the order of the peaks is kept. Knowing that the distortion induced by nonlinear stretching is unphysical, we will mainly focus on the STF general shape, which is well preserved when choosing an appropriate maximum distortion (Supplementary File 2). Unlike in Danré *et al.* (2019), we do not seek to interpret individual stretched peaks as physical subevents.

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 and Fig. S1c). Here, we choose the threshold of clustering cutoff distance to be 0.45, which corresponds to 20 clusters. The DTW distance threshold, or choice in number of clusters, controls the degree of distortion. Thus, by choosing a proper number of clusters, we keep the diversity of clusters and limit the unphysical distortion. Figure 1 shows the STF shapes within each cluster share consistently similar features, which demonstrates our cluster method is effective to group STF shapes alike.

For each of these clusters, we choose the representative STF (defined as the centroid event) that has the minimum median distance to all of the other members of the cluster. It is similar to the stack of all stretched STF shapes within each cluster (Fig. 2), which in turn exhibits the common features of all cluster members.

Next, 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. To be counted as a peak, we choose a threshold of prominent

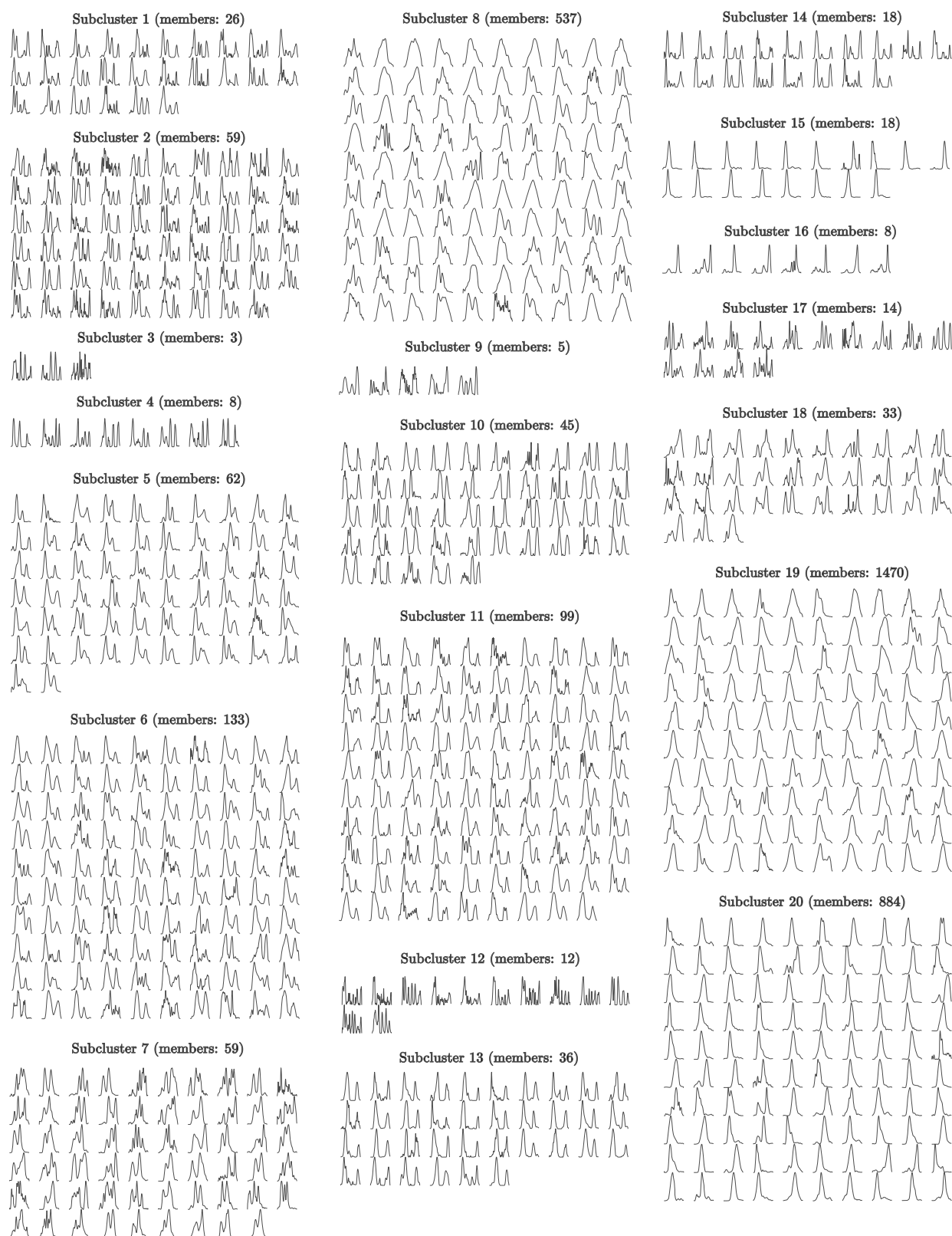


Figure 1. Examples of source time functions (STFs) for 20 clusters resulted from dynamic time warping (DTW). If a cluster has >100

members, only 100 randomly selected ones are shown.

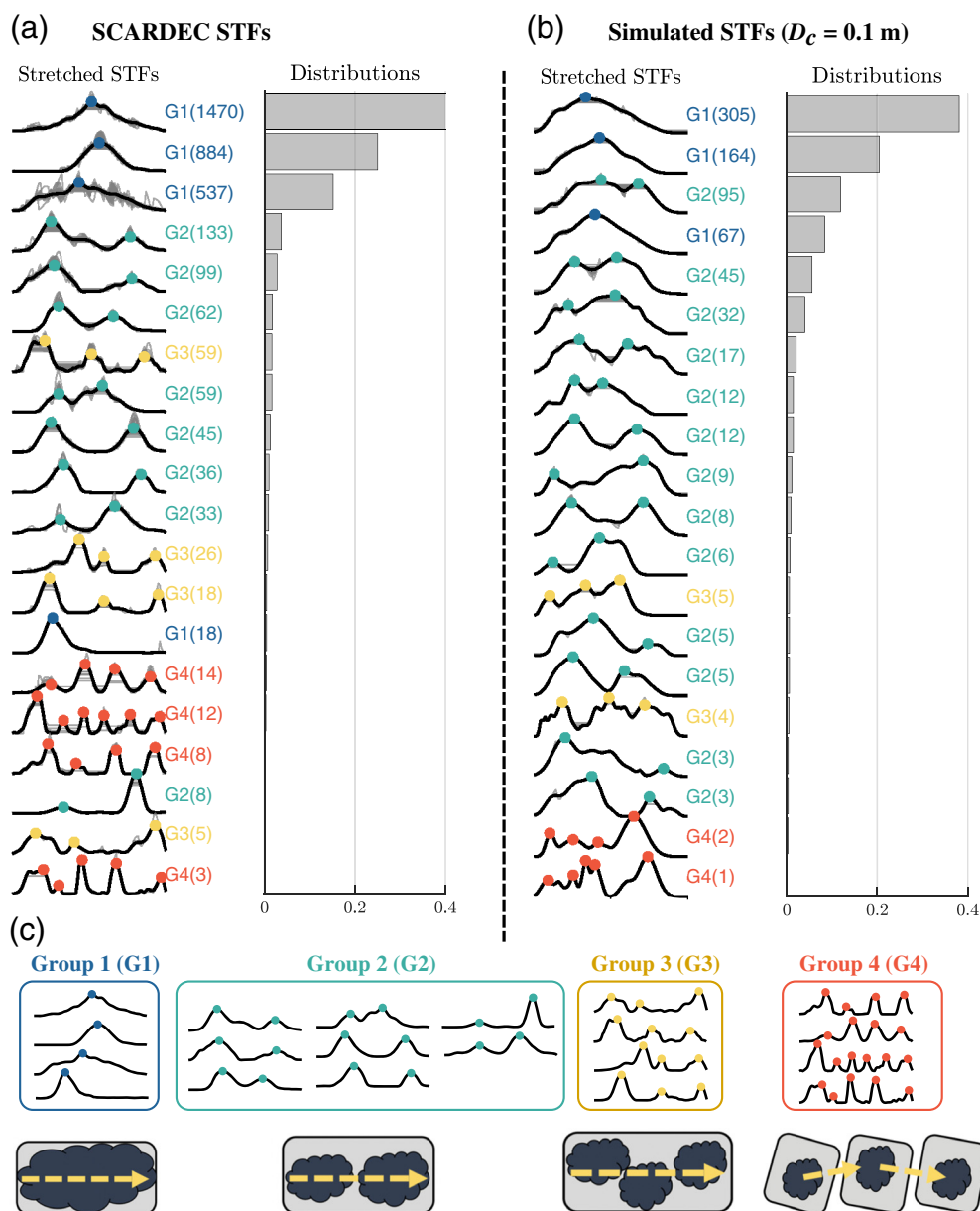


Figure 2. STF clustering, grouping, and conceptual interpretation. (a) Individual STFs after DTW and clustering are shown by gray thin lines. Black thick lines are the STFs of the centroid event of each cluster. Colored dots indicate the prominent peaks of the centroid STF as well as the associated group. 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. The color version of this figure is available only in the electronic edition.

peak amplitude to be 10% of the global maximum of the STF. The purpose of this step is to avoid counting the small-amplitude spurious peaks, which are usually caused by imperfect Green's function removal, as the false prominent peaks. Previous studies have used zero crossing of the time derivative of the STF (Persh and Houston, 2004), which is also sensitive to spurious peaks because it does not account for the signal

amplitudes. We also test 1% and 5% threshold, and the conclusions remain unchanged. We notice that the stretched STFs have a lot fewer prominent peaks than individual peaks from the Gaussian decomposition by Danré *et al.* (2019) (Fig. S2). Moreover, the stretched STFs have fewer prominent peaks than the raw STFs, but in general, the same number of prominent peaks as the centroid event (Fig. S3). For instance, an STF may have multiple separated amplitude peaks but only one single prominent peak (Fig. 2a,b). These differences are because our new metric is defined for the complexity of general shape extracted from large number of STFs instead of the detailed features of each individual STF.

Finally, we group the clusters based on the number of prominent peaks of the centroid event: G1 is the group in which the centroid event has one prominent peak, G2 is the group in which the centroid event has two prominent peaks, and so on (Fig. 2c). G4 is the group in which the centroid event has at least four prominent peaks. Examples of detected prominent peaks are found in Figure 2a,b (see Supplementary File 2 for each individual STF). In this study, we define the STFs to be “complex” if their DTW stretched STFs have multiple prominent peaks. The first-order result from the grouping is that most events have a single

prominent peak, whereas ~20% events are more complex.

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.

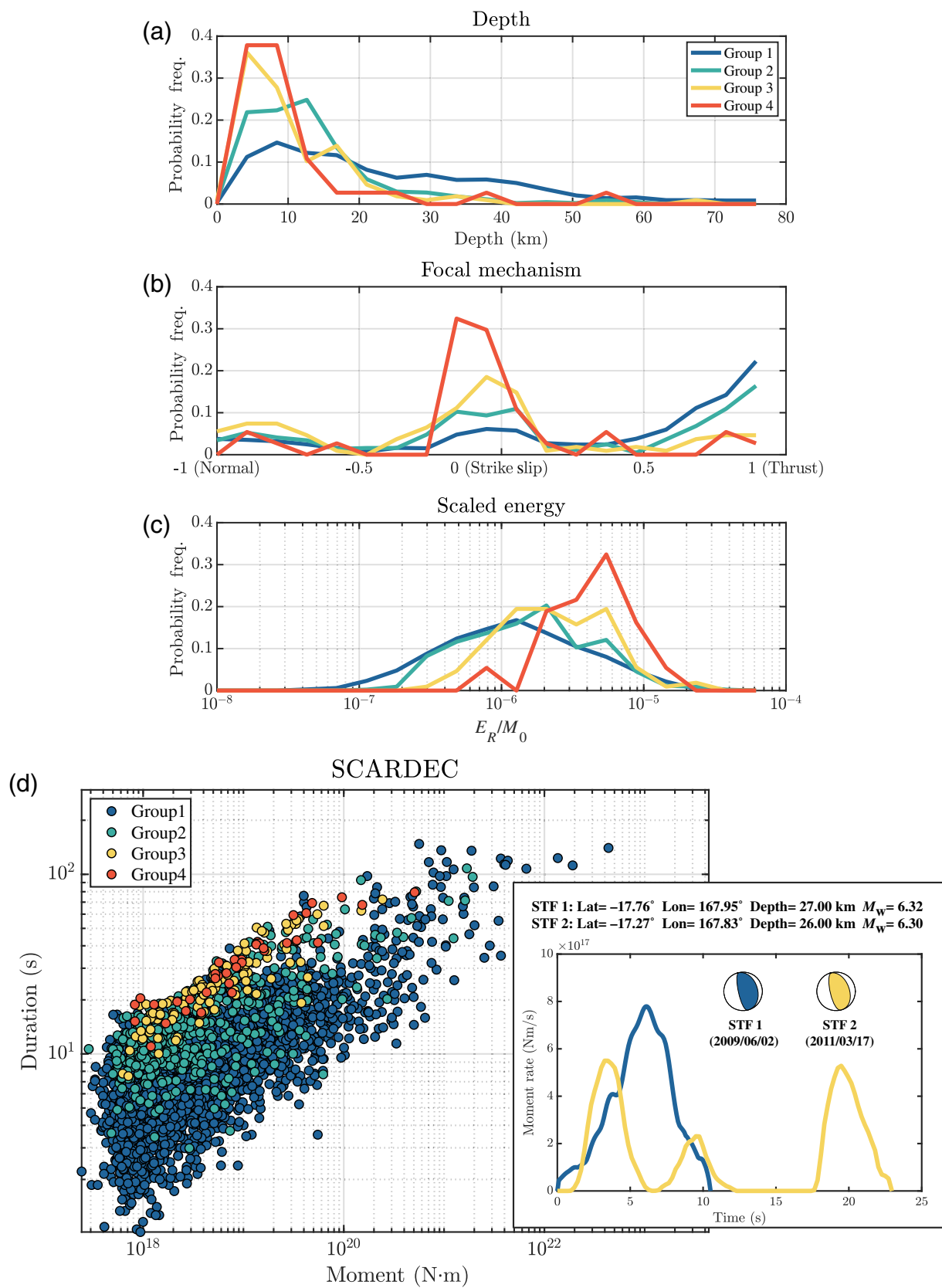
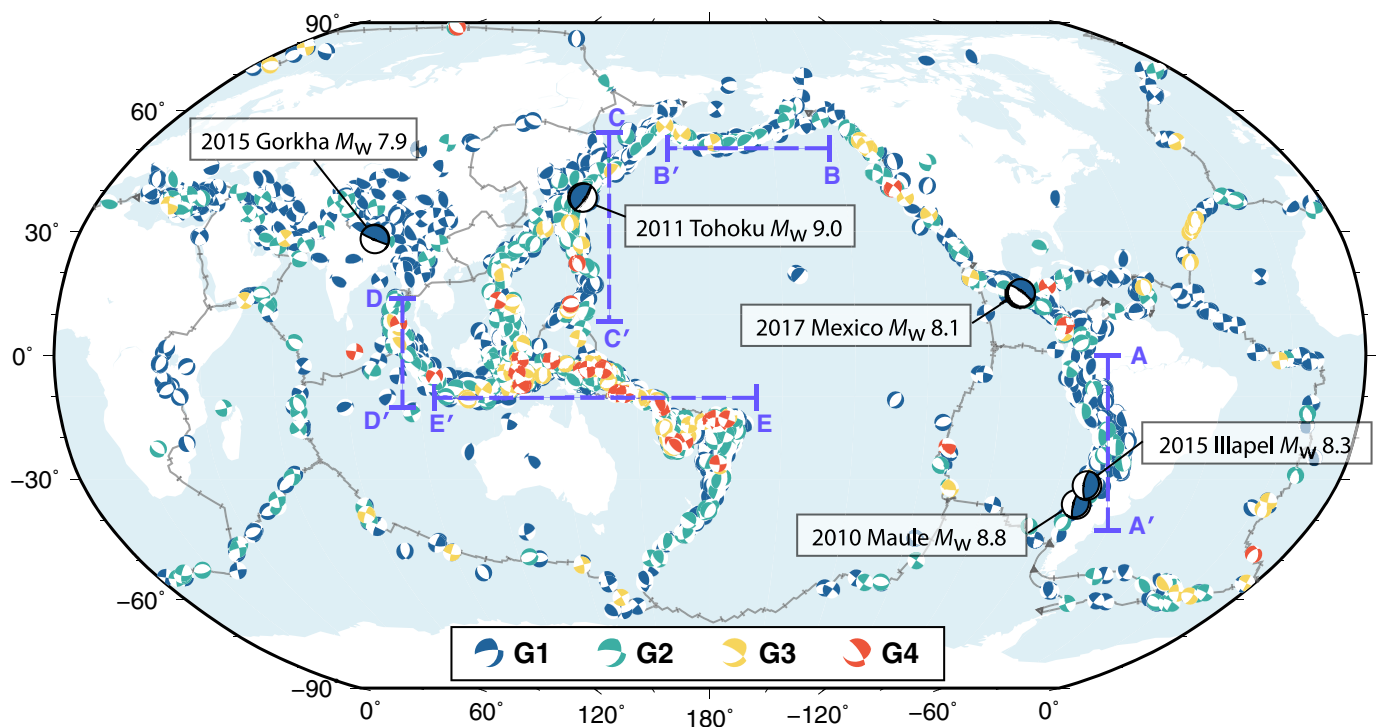


Figure 3. Population distribution of four complexity groups and correlation with different source parameters: (a) centroid depth, (b) focal mechanism (scalar defined by Shearer et al., 2006, that varies from -1 [normal], 0 [strike slip] to 1 [reverse]), (c) and scaled radiated energy

$e = E_R/M_0$. (d) The earthquake duration against earthquake moment, colored with the respective group labels. One pair of collocated events with different complexity are also shown in the inset. The color version of this figure is available only in the electronic edition.



The first property we investigate is the source depth. Complex STFs (groups G2–G4) are mostly shallow crustal events, whereas the simple STFs (group G1) can be found at all depths (Fig. 3a): G1 35.37%; G2 73.68%; G3 86.11%; and G4 89.19% with depth ≤ 20 km. Because collocated events have various degrees of complexity (Figs. 3d and 4), inaccuracy in the Green's function does not strongly bias our results.

The second property we investigate is the focal mechanism (Fig. 3b). 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 Fig. 3d and Fig. S4). For example in Figure 3d, we see that the largest events in SCARDEC database may only have one prominent peak in their stretched STF, whereas 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 (Fig. 3d). It is illustrated in Figure 3d by visualization of two STFs of collocated 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 with the G1 STFs that have a single high amplitude and short duration peak. Simple models of crack ruptures yield a relation between moment,

Figure 4. 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 5. The color version of this figure is available only in the electronic edition.

source duration, and stress drop that could indicate low stress drops for the G4 events (Fig. S5a–c; Eshelby, 1957; Brune, 1971).

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 $E_R = (\frac{1}{15\pi\rho V_p^3} + \frac{1}{10\pi\rho V_s^3}) \int_0^\infty (\ddot{M}_0(t))^2 dt$. We select depth-dependent bulk properties (V_p P -wave velocity, V_s shear-wave velocity, and ρ density) from the preliminary reference Earth model (Dziewonski and Anderson, 1981). Radiated energy scales almost linearly with seismic moment and we calculate the scaled energy as the ratio of both radiated energy and seismic moment, which is about invariant with earthquake size (Convers and Newman, 2011; Denolle and Shearer, 2016). Figure 3c 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

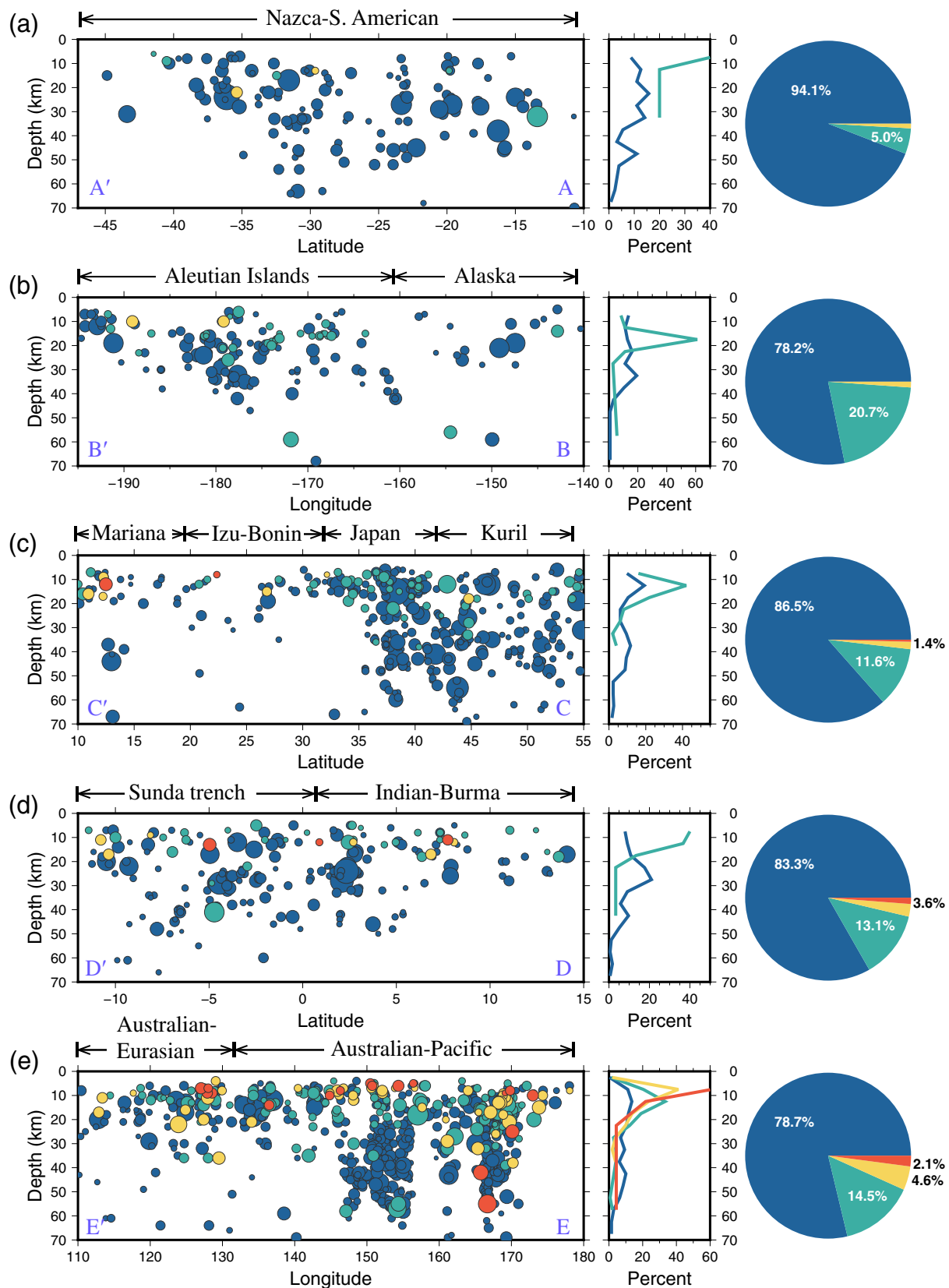


Figure 5. Earthquake distributions of different complexity groups on the vertical profiles (from 0 to 70 km) AA' (a), BB' (b), CC' (c), DD' (d), and EE' (e), whose locations are indicated by blue dashed lines in Fig. 4. The regional along-depth and total group distributions are

also shown to the right. The colors blue, green, yellow and red represent G1, G2, G3, and G4, respectively. The circle size in depth profiles is scaled to earthquake magnitude. The color version of this figure is available only in the electronic edition.

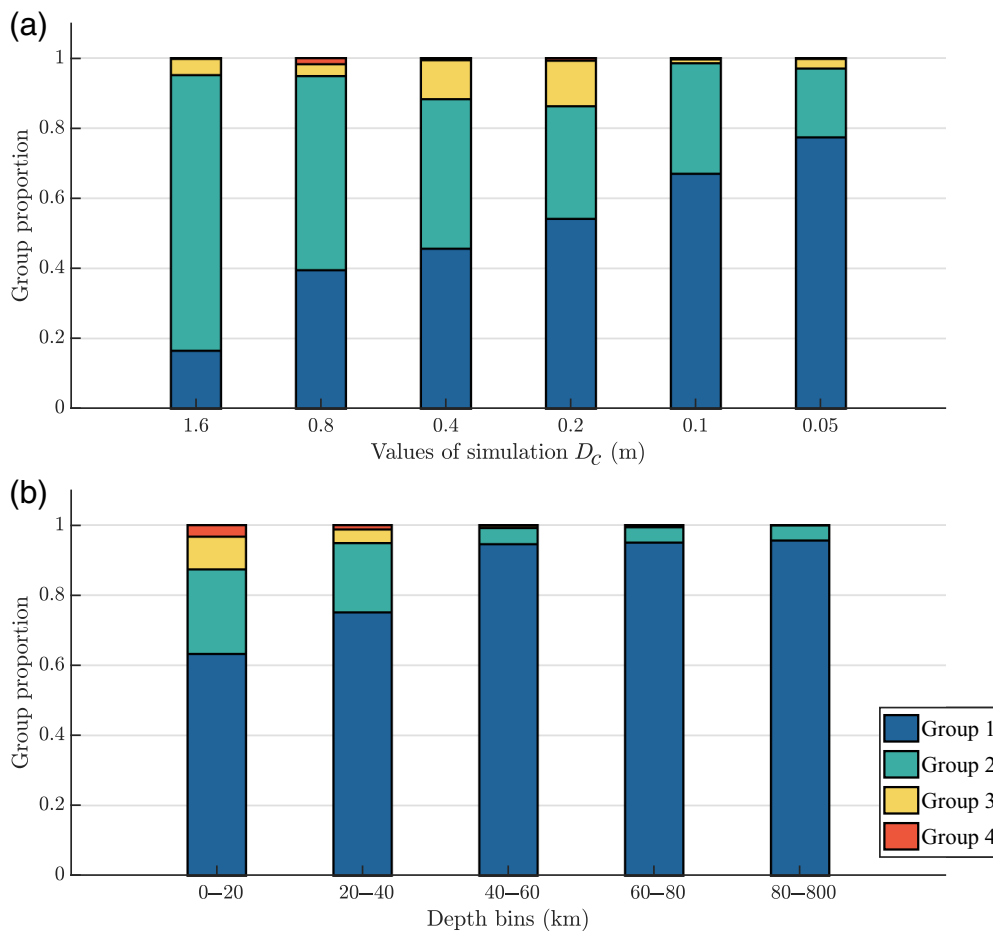


Figure 6. Group proportion distributions: (a) simulated STFs clustering with different values of D_c . (b) Group proportions of real STFs (Seismic source ChAracteristics Retrieved from DEConvolving teleseismic body waves [SCARDEC]) within different depth bins. The color version of this figure is available only in the electronic edition.

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 and Lapusta, 2010) with repeated rupture acceleration and deceleration as they travel across segments (Kanamori *et al.*, 1992; Peyrat *et al.*, 2001; Cesca *et al.*, 2017).

Because 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 with the marked variation of tectonic environments (Fig. 4). 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 or within the subducting slabs at various depths. For example, since 1992, there were only two events ($M_w > 5.5$) belonging to the G3 group along the Southern American and Aleutian subduction zones, respectively (Fig. 5a,b). Similarly, in other subduction zone regions such as in Japan and Sumatra, the Indian–Eurasian collision zone

are also dominated by simple-type earthquakes (Fig. 5c,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 (Figs. 4 and 5e). Bird (2003) explored and documented the kinematics at plate boundaries and found that this region is characterized by a particularly extensive number of microplates, whose boundaries exhibit varied relative motions and kinematics (their fig. 6). Therefore, we propose that the complexity in the STF may reflect the complexity in the regional stress field.

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 and Beroza, 2002; Ripperger *et al.*, 2007; Trugman and Dunham, 2014; Graves and Pitarka, 2016; Mai *et al.*, 2017). To investigate possible factors that control the STF complexity patterns, we perform a large number of 2D dynamic rupture simulations with stochastic distributions of prestress 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 2D medium in an antiplane setting. Prestress 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 ~ 1 MPa (Fig. S6). Danré *et al.* (2019) find that heterogeneity is necessary to reproduce realistically rough STFs. There is a

trade-off between strength excess and D_c in controlling rupture velocity and the resulting ground motions (Guatteri and Spudich, 2000). Here, we choose to focus on D_c . The use of dynamic simulation in this study is not intended to compare the shape of individual subevents or to relate the time and shape of these subevents to physical properties on the field. Instead, we use these simulations to provide an ensemble of realistic STF and analyze the statistical properties of their shapes. Although 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 prestress distributions that we use in each simulations. The dynamic rupture is solved with the 2D boundary integral method Spectral Boundary Integral Equation Method in MATLAB (see Data and Resources). We discard the rupture models that unsuccessfully nucleated with a source dimension <20 km or rupture beyond the zone of heterogeneous prestress and obtain 800 qualified simulations for each D_c value. Finally, the STF is calculated from the integral of the moment-density-rate functions over the fault surface (more details in Text S2).

Our results indicate that the small values of $D_c < 0.1$ m are probably necessary to produce the general level of complexity of the SCARDEC STF (Fig. 6a). Furthermore, we notice that the distributions among group numbers vary systematically with depth in the SCARDEC database (Fig. 6b). The STF of shallow crustal earthquakes present a diversity in complexity similar to that obtained in the simulations when using $D_c \sim 0.1$ m. The STF of deep mantle earthquakes present a diversity in complexity that can be obtained with much smaller values $D_c (\leq 0.05$ m). 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 should 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 and Takeo, 1997). Our results may provide a supporting evidence that the characteristic slip-weakening distance, or more generally, the fracture energy that is proportional to the product of D_c and stress drop (Guatteri and Spudich, 2000), varies at depth over crustal scales.

Except for the on-fault frictional properties simulated in this study, different faulting mechanisms for shallow (0–80 km) and deep (80–800 km) earthquakes can potentially be another factor for the variation of STF complexity along depth. For instance, the deep earthquakes are inferred to have different faulting mechanisms such as transformational faulting, dehydration embrittlement, and thermal runaway (Zhan, 2020). However, whether these different faulting mechanisms leads to various STF shapes is beyond the scope of our simulations in this study.

Discussion and Conclusion

In this study, we develop a DTW methodology to cluster a large number of earthquake STF into different complexity groups based on similarity of their general shapes. We find the patterns of STF shape complexity correlate with different source parameters such as depth, duration, focal mechanism, and scaled energy. Through dynamic rupture simulation, we find that the frictional parameter D_c affects the proportions of different complexity groups. Comparing the diversity in the STF complexity from SCARDEC database with that from the simulated STF, we suggest that small values of D_c are a viable explanation to the distribution between simple and complex events. Moreover, the results of SCARDEC STF show that the shallow crust has relatively larger proportion of complex groups than the deep region, as does the STF simulated with a large D_c . Based on this, the observed variations of complexity with depth could be explained with a depth variation of D_c , or more generally, the fracture energy.

Furthermore, we compare our results with previous work in the same database (Danré *et al.*, 2019). The definition of complexity in Danré *et al.* (2019) is the total number of individual peaks, which they referred to as subevents. Danré *et al.* (2019) also found that strike-slip events had more complex STF than thrust earthquakes. This study adds to the Danré *et al.* (2019) study in three ways. First, there is no obvious correlation between earthquake magnitude and the number of prominent peaks. This suggests that earthquakes have a limited number of large subevents (prominent peak). Second, this study analyzes 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 STF exhibit large radiation ratio (proportion of radiated energy over available energy). Finally, the modeled STF exhibit different degrees of complexity depending on the frictional properties.

We apply our method to two additional established data sets of STF: 187 STF from U.S. Geological Survey (see Data and Resources) and 114 STF of megathrust earthquakes from Ye *et al.* (2018) (Fig. S13). Both were created from finite-fault inversion of teleseismic body and surface waves and use two conventional inversion approaches (Kikuchi and Kanamori, 1991; Hartzell *et al.*, 2007). We observe that the STF from these two data sets present less complexity: most of the events are clustered into the simple groups (G1 and G2). We also get a similar correlation between the shape complexity and the source parameters (Fig. S14). Despite these similarities, statistical robustness is limited by the low number of events in the two data sets.

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 3D dynamic rupture simulation framework, which is impractical to implement because of high computational expense and the statistical approaches used. Nevertheless, because fault geometry and fault properties seem to play a dominant role in shaping the source and the resulting strong ground motions, further 3D modeling and observations are necessary.

Data and Resources

All the source time functions (STFs) are downloaded from Seismic source Characteristics Retrieved from DEconvolving teleseismic body waves (SCARDEC) STF database (<http://scardec.projects.sismo.ipgp.fr>, last accessed January 2020). The dynamic rupture simulation code Spectral Boundary Integral Equation Method in MATLAB (SBIEMLAB) is developed by Jean-Paul Ampuero (available at <https://github.com/jpampuero>, last accessed February 2021). The MATLAB scripts to reproduce the results and figures can be obtained from the Github (https://github.com/yinjiuxun/STF_DTW, last accessed January 2020). Global maps are made by Generic Mapping Tools (GMT; Wessel *et al.*, 2013; available at <https://www.generic-mapping-tools.org>, last accessed January 2020). Other relevant data are from U.S. Geological Survey (<https://earthquake.usgs.gov/earthquakes/search>, last accessed January 2020). Supplemental material for this article includes descriptions of cluster analysis Text S1 and dynamic rupture simulation Text S2, Figures S1–S14, and a PDF file of all individual STFs before and after stretching to the cluster mean shape.

Declaration of Competing Interests

The authors declare no competing interests.

Acknowledgments

The authors sincerely thank Martin Vallée, the assistant professor William B. Frank, and the two anonymous reviewers for their constructive suggestions that greatly helped improve this work. The authors are also grateful to Gavin Hayes at U.S. Geological Survey (USGS) and Lingling Ye at Sun Yat-sen University for providing their source time functions (STFs) for cross validation.

References

- Allmann, B. P., and P. M. Shearer (2009). Global variations of stress drop for moderate to large earthquakes, *J. Geophys. Res.* **114**, no. B1, doi: [10.1029/2008JB005821](https://doi.org/10.1029/2008JB005821).
- Ammon, C. J., C. Ji, H.-K. Thio, D. Robinson, S. Ni, V. Hjorleifsdottir, H. Kanamori, T. Lay, S. Das, D. Helmberger, *et al.* (2005). Rupture process of the 2004 Sumatra-Andaman earthquake, *Science* **308**, no. 5725, 1133–1139.
- Andrews, D. J. (1976). Rupture propagation with finite stress in antiplane strain, *J. Geophys. Res.* **81**, no. 20, 3575–3582.
- Atik, L. A., N. Abrahamson, J. J. Bommer, F. Scherbaum, F. Cotton, and N. Kuehn (2010). The variability of ground-motion prediction models and its components, *Seismol. Res. Lett.* **81**, no. 5, 794–801.
- Baltay, A. S., G. C. Beroza, and S. Ide (2014). Radiated energy of great earthquakes from teleseismic empirical Green's function deconvolution, *Pure Appl. Geophys.* **171**, no. 10, 2841–2862.
- Berndt, D. J., and J. Clifford (1994). Using dynamic time warping to find patterns in time series, *KDD Workshop*, Vol. 10, Seattle, Washington, 359–370.
- Bird, P. (2003). An updated digital model of plate boundaries, *Geochem. Geophys. Geosys.* **4**, no. 3, doi: [10.1029/2001GC000252](https://doi.org/10.1029/2001GC000252).
- Brune, J. N. (1971). Correction to Brune, 1970, *J. Geophys. Res.* **76**, 5002.
- Candela, T., F. Renard, Y. Klinger, K. Mair, J. Schmittbuhl, and E. E. Brodsky (2012). Roughness of fault surfaces over nine decades of length scales, *J. Geophys. Res.* **117**, no. B8, doi: [10.1029/2011JB009041](https://doi.org/10.1029/2011JB009041).
- Cesca, S., Y. Zhang, V. Mouslopoulou, R. Wang, J. Saul, M. Savage, S. Heimann, S. K. Kufner, O. Oncken, and T. Dahm (2017). Complex rupture process of the M_w 7.8, 2016, Kaikoura earthquake, New Zealand, and its aftershock sequence, *Earth Planet. Sci. Lett.* **478**, 110–120.
- Chouneet, A., and M. Vallée (2018). Global and interregion characterization of subduction interface earthquakes derived from source time functions properties, *J. Geophys. Res.* **123**, no. 7, 5831–5852, doi: [10.1029/2018JB015932](https://doi.org/10.1029/2018JB015932).
- Convers, J. A., and A. V. Newman (2011). Global evaluation of large earthquake energy from 1997 through mid-2010, *J. Geophys. Res.* **116**, doi: [10.1029/2010JB007928](https://doi.org/10.1029/2010JB007928).
- Danré, P., J. Yin, B. P. Lipovsky, and M. A. Denolle (2019). Earthquakes within earthquakes: Patterns in rupture complexity, *Geophys. Res. Lett.* **46**, no. 13, 7352–7360.
- Denolle, M. A. (2019). Energetic onset of earthquakes, *Geophys. Res. Lett.* **46**, no. 5, 2458–2466.
- Denolle, M. A., and P. M. Shearer (2016). New perspectives on self-similarity for shallow thrust earthquakes, *J. Geophys. Res.* **121**, no. 9, 6533–6565.
- Dziewonski, A. M., and D. L. Anderson (1981). Preliminary reference Earth model, *Phys. Earth Planet. In.* **25**, no. 4, 297–356.
- Eshelby, J. D. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems, *Proc. Math. Phys. Sci.* **241**, 376–396.
- Graves, R., and A. Pitarka (2016). Kinematic ground-motion simulations on rough faults including effects of 3D stochastic velocity perturbations kinematic ground-motion simulations on rough faults, *Bull. Seismol. Soc. Am.* **106**, no. 5, 2136–2153.
- Guatteri, M., and P. Spudich (2000). What can strong-motion data tell us about slip-weakening fault-friction laws? *Bull. Seismol. Soc. Am.* **90**, no. 1, 98–116.
- Hartzell, S., P. Liu, C. Mendoza, C. Ji, and K. M. Larson (2007). Stability and uncertainty of finite-fault slip inversions: Application to the 2004 Parkfield, California, earthquake, *Bull. Seismol. Soc. Am.* **97**, no. 6, 1911–1934.
- Houston, H. (2001). Influence of depth, focal mechanism, and tectonic setting on the shape and duration of earthquake source time functions, *J. Geophys. Res.* **106**, no. B6, 11,137–11,150.
- Ide, S., and G. C. Beroza (2001). Does apparent stress vary with earthquake size, *Geophys. Res. Lett.* **28**, no. 17, 3349–3352.

- Ide, S., and M. Takeo (1997). Determination of constitutive relations of fault slip based on seismic wave analysis, *J. Geophys. Res.* **102**, no. B12, 27,379–27,391.
- Kanamori, H., T. Hong-Kie, D. Doug, H. Egill, and T. Heaton (1992). Initial investigation of the Landers, California, earthquake of 28 June 1992 using TERRAscope, *Geophys. Res. Lett.* **19**, no. 22, 2267–2270.
- Kaneko, Y., and N. Lapusta (2010). Supershear transition due to a free surface in 3-D simulations of spontaneous dynamic rupture on vertical strike-slip faults, *Tectonophysics* **493**, no. 3, 272–284.
- Kaneko, Y., E. Fukuyama, and I. J. Hamling (2017). Slip-weakening distance and energy budget inferred from near-fault ground deformation during the 2016 M_w 7.8 Kaikōura earthquake, *Geophys. Res. Lett.* **44**, no. 10, 4765–4773.
- Kikuchi, M., and H. Kanamori (1991). Inversion of complex body waves—III, *Bull. Seismol. Soc. Am.* **81**, no. 6, 2335–2350.
- Klinger, Y. (2010). Relation between continental strike-slip earthquake segmentation and thickness of the crust, *J. Geophys. Res.* **115**, no. B7, doi: [10.1029/2009JB006550](https://doi.org/10.1029/2009JB006550).
- Li, Y.-G., K. Aki, J. E. Vidale, W. H. K. Lee, and C. J. Marone (1994). Fine structure of the Landers fault zone: Segmentation and the rupture process, *Science* **265**, no. 5170, 367–370.
- Mai, P. M., and G. C. Beroza (2002). A spatial random field model to characterize complexity in earthquake slip, *J. Geophys. Res.* **107**, no. B11, ESE 10-1–ESE 10-21.
- Mai, P. M., M. Galis, K. K. S. Thingbaijam, J. C. Vyas, and E. M. Dunham (2017). Accounting for fault roughness in pseudo-dynamic ground-motion simulations, *Pure Appl. Geophys.* **174**, no. 9, 3419–3450.
- Massonnet, D., M. Rossi, C. Carmona, F. Adragna, G. Peltzer, K. Feigl, and T. Rabaute (1993). The displacement field of the Landers earthquake mapped by radar interferometry, *Nature* **364**, no. 6433, 138–142.
- Meier, M.-A., J. P. Ampuero, and T. H. Heaton (2017). The hidden simplicity of subduction megathrust earthquakes, *Science* **357**, no. 6357, 1277–1281.
- Meng, L., J.-P. Ampuero, J. Stock, Z. Duputel, Y. Luo, and V. C. Tsai (2012). Earthquake in a maze: Compressional rupture branching during the 2012 M_w 8.6 Sumatra earthquake, *Science* **337**, no. 6095, 724–726.
- Müller, M. (2007). Dynamic time warping, in *Information Retrieval for Music and Motion*, Springer, Berlin/Heidelberg, Germany, 69–84.
- Persh, S. E., and H. Houston (2004). Strongly depth-dependent after-shock production in deep earthquakes, *Bull. Seismol. Soc. Am.* **94**, no. 5, 1808–1816.
- Peyrat, S., K. Olsen, and R. Madariaga (2001). Dynamic modeling of the 1992 Landers earthquake, *J. Geophys. Res.* **106**, no. B11, 26,467–26,482.
- Ripperger, J., J.-P. Ampuero, P. M. Mai, and D. Giardini (2007). Earthquake source characteristics from dynamic rupture with constrained stochastic fault stress, *J. Geophys. Res.* **112**, no. B4, doi: [10.1029/2006JB004515](https://doi.org/10.1029/2006JB004515).
- Shearer, P. M., G. A. Prieto, and E. Hauksson (2006). Comprehensive analysis of earthquake source spectra in southern California, *J. Geophys. Res.* **111**, no. B6, doi: [10.1029/2005JB003979](https://doi.org/10.1029/2005JB003979).
- Trugman, D. T., and E. M. Dunham (2014). A 2D pseudodynamic rupture model generator for earthquakes on geometrically complex faults, *Bull. Seismol. Soc. Am.* **104**, no. 1, 95–112.
- Vallée, M. (2013). Source time function properties indicate a strain drop independent of earthquake depth and magnitude, *Nat. Comm.* **4**, 2606.
- Vallée, M., and V. Douet (2016). A new database of source time functions (STFs) extracted from the SCARDEC method, *Phys. Earth Planet. In.* **257**, 149–157.
- Vallée, M., J. Charléty, A. M. G. Ferreira, B. Delouis, and J. Vergoz (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, *Geophys. J. Int.* **184**, no. 1, 338–358.
- Wessel, P., W. H. Smith, R. Scharroo, J. Luis, and F. Wobbe (2013). Generic Mapping Tools: Improved version released, *Eos Trans. AGU* **94**, no. 45, 409–410.
- Wibberley, C. A. J., and T. Shimamoto (2005). Earthquake slip weakening and asperities explained by thermal pressurization, *Nature* **436**, no. 7051, 689–692.
- Yagi, Y., and Y. Fukahata (2011). Rupture process of the 2011 Tohoku-Oki earthquake and absolute elastic strain release, *Geophys. Res. Lett.* **38**, no. 19, doi: [10.1029/2011GL048701](https://doi.org/10.1029/2011GL048701).
- Ye, L., H. Kanamori, and T. Lay (2018). Global variations of large megathrust earthquake rupture characteristics, *Sci. Adv.* **4**, no. 3, eaao4915.
- Zhan, Z. (2020). Mechanisms and implications of deep earthquakes, *Ann. Rev. Earth Planet. Sci.* **48**, no. 1, 147–174, doi: [10.1146/annurev-earth-053018-060314](https://doi.org/10.1146/annurev-earth-053018-060314).

Manuscript received 28 October 2020

Published online 31 March 2021