A generic spectrum of global earthquake rupture characteristics revealed by machine learning

Zefeng Li[1]

[1]Univerisity of Science and Technology of China

November 30, 2022

Abstract

Rupture processes of global large earthquakes have been observed to exhibit great variability, whereas recent studies suggest that the average rupture behavior could be unexpectedly simple. To what extent do large earthquakes share common rupture characteristics? Here we use a machine learning algorithm to derive a generic spectrum of global earthquake source time functions. The spectrum indicates that simple and homogeneous ruptures are pervasive whereas complex and irregular ruptures are relatively rare. Despite the standard long-tail and near-symmetric moment release processes, the spectrum reveals two special rupture types: runaway earthquakes with weak growing phases and relatively abrupt termination, and complex earthquakes with all faulting mechanisms but mostly shallow origins (<40 km). The diversity of temporal moment release patterns imposes a limit on magnitude predictability in earthquake early warning. Our results present a panoptic view on the collective similarity and diversity in the rupture processes of global large earthquakes.

A generic spectrum of global earthquake rupture characteristics revealed by machine learning

Zefeng Li[1,2*]

1. *Laboratory of Seismology and Physics of Earth's Interior, School of Earth and Space Sciences, University of Science and Technology of China, Hefei, China*
2. *Mengcheng National Geophysical Observatory, University of Science and Technology of China, Mengcheng, China*

Manuscript submitted to GRL

*Corresponding author: Zefeng Li (zefengli@ustc.edu.cn)

Highlights:

1. A generic spectrum of characteristic source time functions is derived from global earthquake observations using machine learning.
2. The spectrum presents a panoptic view of the similarity and the diversity in the rupture processes of large earthquakes.
3. The diversity of temporal moment release patterns imposes a limit on magnitude predictability in earthquake early warning.

Abstract

Rupture processes of global large earthquakes have been observed to exhibit great variability, whereas recent studies suggest that the average rupture behavior could be unexpectedly simple. To what extent do large earthquakes share common rupture characteristics? Here we use a machine learning algorithm to derive a generic spectrum of global earthquake source time functions. The spectrum indicates that simple and homogeneous ruptures are pervasive whereas complex and irregular ruptures are relatively rare. Despite the standard long-tail and near-symmetric moment release processes, the spectrum reveals two special rupture types: runaway earthquakes with weak growing phases and relatively abrupt termination, and complex earthquakes with all faulting mechanisms but mostly shallow origins (<40 km). The diversity of temporal moment release patterns imposes a limit on magnitude predictability in earthquake early warning. Our results present a panoptic view on the collective similarity and diversity in the rupture processes of global large earthquakes.

Plain language summary

In past decades, the rupture processes of many large earthquakes have observed to exhibit great variability. However, some recent studies suggest that the average rupture behavior could be unexpectedly simple. Can the average behavior be representative of most earthquakes? To what extent do large earthquakes share common rupture characteristics? Here we use machine learning to derive a panoptic picture, i.e. a generic spectrum of source time functions, for global earthquake. The spectrum show that simple and homogeneous ruptures are pervasive whereas complex and irregular ruptures are relatively rare. Besides, it reveals two special rupture types: runaway earthquakes with weak initial phases, and complex earthquakes with all faulting mechanisms but mostly shallow origins (<40 km). Our results present a panoptic view on the collective similarity and diversity in the rupture processes of global large earthquakes, which affects how well we can predict earthquake magnitude in earthquake early warning.

Introduction

Large earthquakes start, propagate, and terminate in diverse manners owing to complex interplay between rupture dynamics and fault properties. In the past decades, observations of large earthquake rupture processes have shown various degrees of peculiarity (Ammon, 2005; Ammon *et al.*, 2006; Meng *et al.*, 2012; Ross *et al.*, 2019), suggesting that each large earthquake probably has its own unique characteristics. However, understanding the general physical laws that govern earthquake phenomena requires derivation of the underlying patterns from the seemingly diverse behaviors (Houston and Vidale, 1994; Vallée, 2013; Meier *et al.*, 2017; Denolle, 2019). The average behaviors, often obtained by stacking a large set of seismological data, tend to show relatively simple characteristics, implying common similarity hidden behind the diverse ruptures (Houston and Vidale, 1994; Meier *et al.*, 2017). The distinct emphases on collective rupture peculiarity and similarity raises a critical question: to what extent do large earthquakes share common rupture characteristics? To answer this question calls for a panoptic view of the variability in the rupture processes of global earthquakes.

Earthquake source time functions (STFs) describes the history of seismic moment release during rupture. As an important observation constraining on the source processes, STFs have been routinely extracted from seismograms for large earthquakes. However, because of the high-dimensionality and great variations of amplitude and duration (Tanioka and Ruff, 1997; Duputel *et al.*, 2013; Vallée, 2013), STFs cannot be compared directly. Hence, comparison is often performed on individual STF properties such as duration, peak amplitude, peaks and skewness (Houston, 2001; Persh and Houston, 2004), as well as other derived parameters, such as scale energy (Denolle, 2019) and relative radiated energy efficiency (Ye *et al.*, 2018). Although these individual properties constrain on specific aspects of earthquake ruptures, it remains challenging to examine the variability of overall moment release processes.

Here we employ a machine learning algorithm, called variational autoencoder (VAE), to illuminate the systematic variability of STF shapes among global earthquakes (Figure 1). We train the VAE model with normalized STFs of 3675 M>5.5 global earthquakes from 1992 to 2019 (SCARDEC database, Vallée & Douet, 2016). This trained model is applied to another independent database of 112 STF of M>7.0 megathrust earthquakes (Ye *et al.*, 2016), to validate the model's generalization capability. With the model, we derive a standard STF spectrum that contains a systematic set of characteristic shapes along with corresponding earthquake population density. The spectrum exhibits a broad range of rupture characteristics for global earthquakes and sheds light on special classes of earthquakes that are not well attended before. Moreover, the deviation of individual earthquakes from the standard spectrum is measured by the reconstruction misfit. Hence, large reconstruction misfits naturally detect earthquakes outside the norm, i.e. earthquakes with unusual rupture processes.

VAE for STFs

VAE is widely used in signal and image processing to uncover the intrinsic structure of a large data set (Kingma and Welling, 2014). It consists of an encoder to compress the high-dimensional data into a low-dimensional latent representation and a decoder to reconstruct the high-dimensional data from the latent representation (Figure 1). The bottleneck architecture forces the model to learn the key characteristics of STFs and discards the noise in individual samples. After training, VAE can take input of virtual latent values to generate synthetic data constrained by real observations, and therefore belongs to generative learning methods.

Following Yin et al. (2021), the STFs of SCARDEC and YE2016 are resampled to 128 points, given that the maximum duration is approximately 100 s and that a minimum sampling rate would be greater than 1 Hz. The amplitude of STFs are then normalized to the event seismic moment to retain the shapes only. The VAE model is constructed as follows: an input layer of 128 neurons, an encoder of two 512 neurons, a bottleneck of two neurons (latent representation), a decoder of two 512 neurons, and an output layer of 128 neurons (Figure 1).

The loss function to train the VAE is defined as:

$$loss = \|STF' - STF\| + KL[\mathbb{N}(\mu_x, \sigma_x) - \mathbb{N}(0,1\ )]$$

where the first term is the root mean square between the reconstructed and original STFs, and the second is the Kullback-Leibler divergence which measures the difference in probability density between latent variables and normal distribution. The Kullback-Leibler divergence essentially acts a regularizer of the latent space (Kingma and Welling, 2014). We use the Adam solver for network parameter optimization (Kingma and Ba, 2017).

The SCARDEC STFs are split in 80% for training and 20% for validation. The convergence of train and validation loss (Figure S1) ensures the model generality and warrants analysis of the entire data set altogether. An important sign for a well-trained VAE is the successful reconstruction of STFs for both SCARDEC and YE2016 (Figures 1b and 1c). Except for some complex events (e.g. 2000 Mw 8.1 New Ireland earthquake, Papua New Guinea earthquake, 2006 Mw 8.3 Kuril earthquake), the reconstructed STFs capture the primary characteristics of the most observed STFs, such as the variations in skewness and peakedness variations, demonstrating the learned low-dimensional latent variables are good representations of the high-dimensional STFs.

The standard STF spectrum

The VAE model allows us to visualize the STFs orderly in both low- and high- dimensional spaces, which is important for evaluating their systematic variability. The encoder projects the STFs into a 2D latent space (Figure 2), whose affinity property ensures that similar STFs are located closely in the latent space (Figure S2). Because of the imposed regularizer on the latent variables, the earthquake population in the latent space generally follows a normal distribution, i.e. approximately ~67% of earthquakes within the radius of 1 and 95% of earthquakes within the radius of 2. Based on these two properties, the most common STF shapes are mapped near the center, whereas the rare ones are mapped outwards.

To visualize the overall STF variations, we input virtual latent parameters at every 0.5 interval from -3.0 to 3.0 into the decoder to construct a set of synthetic STFs (Figure 2b). Because each of these STFs is constrained by real STFs near its locality in the latent space, they represent the generic variations of global earthquakes. Therefore, we call this synthetic collection as the standard STF spectrum. It is noteworthy that the synthetic STFs are constrained with a different number of STFs, given the population distribution of real STFs. Overall, the standard spectrum exhibits three outstanding characteristics that vary continuously:

number of peaks, skewness, and peakedness (Figures 2b and S3). For convenience of discussion, the spectrum is divided into four quadrants based on the changes of characteristics:

$$Q_1: Z_1>-0.5, Z_2>0.5, \quad Q_2: Z_1<-0.5, Z_2>0$$

$$Q_3: Z_1<-0.5, Z_2<0, \quad Q_4: Z_1<-0.5, Z_2<0.5$$

Where Z_1 and Z_2 are the first and second latent variable, respectively. Q_1 represents the complex type with two or more subevents with varying relative size and timings. In comparison, events in Q_2, Q_3, and Q_4 are single peaked but negative-skewed (or runaway), symmetric, and positive-skewed (or long-tail) types, respectively. Overall, the one-peak types Q_2-Q_4 account for 83% of global events (16% runaway, 15% symmetric, 52% long-tail), whereas the complex type in Q_1 account for 17%.

Skewness measures the relative duration of moment acceleration and deceleration phases. The long-tailed shapes suggest the rupture breaks away energetically but die away slowly. They have the largest population among all the types, consistent with generally energetic onsets of large earthquakes (Denolle, 2019). However, part of them, especially those with very long tails, could be due to artifacts from imperfect modeling of P wave coda (Vallée and Douet, 2016). The symmetric type suggests near equivalent acceleration and deceleration duration, which is often considered a generic STF shape in standard models (e.g. Tanioka and Ruff, 1997). Finally, the runaway type has a relatively weak onset, representing ruptures that culminate in the late stage of rupture. This type of events is of particular interest, because they are misguided as small events at the beginning but their final magnitude is largely determined by the later phase. The runaway type is comparably populated as the symmetric type.

Another characteristic revealed in the spectrum peakedness measures the temporal homogeneity of moment release. From the center to periphery, peak width changes from being broad to narrow. The rounded peaks near the center suggests relatively homogenous moment release during the rupture, whereas the spiky peaks in the periphery suggest that a predominant amount of energy is released within a compact prime time. The population distribution suggests that earthquakes with homogenous moment release is much more populated than the highly concentrated ones, reflecting the prevalence of homogenous faulting in nature.

Earthquakes outside the norm

The spectrum describes a comprehensive set of standard shapes that represent a majority of global earthquakes. However, some earthquakes cannot be adequately described by the spectrum, which is quantified by the misfit between the reconstructed and original STFs $\frac{\|STF_{rec}-STF_{raw}\|}{\|STF_{raw}\|}$. While the majority of the SCARDEC events have predominantly small misfits (Figure 3a), some have unusually high misfits, such as the 2006 M 7.7 Java tsunami earthquake, 2006 M 8.3 Kuril earthquake and 2007 M 8.1 Solomon tsunami earthquake (Figure S4). These events have complex STFs and many are documented with unusual rupture characteristics (Ammon *et al.*, 2006, 2008; Furlong *et al.*, 2009). Hence, the events with high misfits represent a special class of unusual earthquakes outside the norm.

The complex events can be categorized according to their latent locality, i.e. the systematic ones in Q_1 and the scatters in other quadrants. The high misfits observed in Q_1 indicate the actual shapes there are even more complicated than depicted in the spectrum. These events have temporally separated subevents, representing ruptures of relatively distant asperities/faults or inter-event triggering. In comparison, the high-misfit events outside Q_1 are fit with simple one-peak shapes, yet exhibiting complex characteristics. They appear to be temporally more compact and generally “rougher”. This could be interpreted as ruptures of spatially concentrated asperities and/or heterogeneous frictional properties along faults (Ye *et al.*, 2018).

Figure 3 shows that the complex events exhibits three intriguing characteristics: 1) they are shallower than 40 km; 2) they exist in all faulting environments; 3) many are located along the northern boundaries of the Australian plate and in Southeast Asia. Houston (2001) examined the number of peaks of subduction zone events and found a group of shallow complex ones. She attributed this phenomenon to heterogeneous and rapidly deformation in interplate boundary region. However, the complex events found here exist in all different mechanisms, suggesting more universal depth-dependent rupture complexity. Applying cluster analysis to the SCARDEC data, Yin et al. (2021) also identified depth dependence of earthquake complexity. They proposed that variations of frictional properties with depth, such as slip weakening distance or other equivalent rupture parameters could play an important role in controlling rupture complexity. Alternatively, we speculate that there could be more geometric and stress irregularities (e.g. faults and asperities) at shallow depths owing to low confining pressure and temperature. In statistical perspective, the populated irregularities allow higher chance of triggering and accidental activation during rupture. This hypothesis is partly supported by the pervasiveness of complex events in the apparently complex fault systems, such as in the northern boundaries of the Australian plate and in Southeast Asia, although the causes for specific complex events could be case dependent.

Implications for earthquake rupture diversity

The machine learning model condenses a large number of global earthquake observations into the standard STF spectrum along with the information of population density, which unravels more diverse rupture characteristics than the mean shape from commonly used stacking methods. Complementary to the spectrum, the reconstruction misfit quantifies the deviation of individual earthquakes from it and naturally measures the rupture uniqueness. These components taken together provide a panoramic view of earthquake rupture variability and showcase the extent that large earthquakes share common rupture characteristics.

Our observations, first of all, confirm that earthquakes with apparent simple rupture processes are predominant and that the extremely skewed or complex ones are rare. Around the spectrum center with highest population density, the shapes are weakly skewed and gently peaked, representing relatively homogenous and apparent one-patch-like rupture processes. This is generally consistent with the simple mean shape reported by Meier et al. (2017). However, the spectrum offers a much richer collection of standard shapes other than the simple mean. Beyond one standard deviation, for example, the shapes start to exhibit significant skewness and more complexity, reflecting increasingly irregular rupture processes. These irregular types comprise a non-negligible proportion of the entire earthquake population (e.g.

32% if counting those outside radius 1), which should be interpreted by more sophisticated rupture models.

The diversity of temporal moment release patterns imposes a limit on magnitude determinism, i.e. the predictability of earthquake size before the rupture is complete. A near isosceles triangular shape averaged from all STFs suggests that approximately half the duration is required to predict final magnitude (Meier *et al.*, 2017). For the runaway events, the early amplitude is small and the peak moment rate arrives much later. The break of strong asperities in the later rupture stage implies the existence of drastic dynamic weakening (Denolle *et al.*, 2015), such as flash heating and thermal pressurization. This type of event poses a greater challenge for early warning, because using the early P waves tend to underestimate the final size. In addition, recent discoveries of similar initial waveforms between large and small earthquakes imply that whether or not a small event can develop into a big one could be in part a stochastic result (Okuda and Ide, 2018; Ide, 2019). Although STFs predicted by the Brune-Haskell model (Haskell, 1964; Brune, 1970) as well as those empirically derived tend to emphasize the importance of symmetric STFs (Tanioka and Ruff, 1997), our results show that the population of the runaway type is actually comparable to that of the symmetric type, examples including the 1996 M 7.7 and 2001 M 7.6 Peru, 2010 M 8.8 Chile, 2011 M 7.3 Honshu earthquakes (Figure S4). In addition to the runaway type, complex events represented in Q_1 have subevents with unknown relative size and timing, which can further confuse early warning systems. It is more challenging that there seems no effective way to diagnose the event type when the rupture is developing. Therefore, early estimates of earthquake magnitude would be expectedly often offshoot, even though it could be partly inferred (Melgar and Hayes, 2019).

An important questions in earthquake science is whether or not rupture processes are magnitude-dependent (Meier *et al.*, 2017; Ye *et al.*, 2018; Melgar and Hayes, 2019; Renou *et al.*, 2019). Figure 2a shows a disproportion of M>8 events in the Z_1>0 and Z_2<0 quadrants, implying a plausible preferential rupture mode for M>8 events. In fact, the largest earthquakes seem to shift systematically to the negative direction of the first latent variable (Figure 4). This trend, if true, implies that the largest earthquakes prefer to begin with small events (symmetric or run-away types) rather than release most of the energy in the early stage (long-tail type). Yet, this hypothesis remains to be tested because of the obvious concern on the scarcity of large events. Despite this, the encoder-decoder system provides a general tool to investigate potential variations with source parameters and thus could offer useful insights into the collective patterns of physical rupture processes. Moreover, our study illustrates that generative machine learning could be helpful in uncovering underlying patterns of high-dimensional seismic data.

Acknowledgements

The author is grateful for Jiuxun Yin, Hiroo Kanamori, Men-Andrin Meier, Han Yue, Lingling Ye for helpful discussion. I thank Martin Vallee and Lingling Ye for providing their source time function databases.

Data Availability

The SCARDEC database is available at http://scardec.projects.sismo.ipgp.fr/ (last accessed on September 3, 2021). The STFs of megathrust earthquakes can be accessed from the supplementary material in Ye et al. (2016) (doi.org/10.1002/2015JB012426, last accessed on September 3, 2021).

Author contributions

Z. L. conceive the idea, analyze the data and write the manuscript.

Competing interests

The author declares no competing interests.

References

Ammon, C. J. (2005). Rupture Process of the 2004 Sumatra-Andaman Earthquake, *Science* **308**, no. 5725, 1133–1139, doi: 10.1126/science.1112260.

Ammon, C. J., H. Kanamori, and T. Lay (2008). A great earthquake doublet and seismic stress transfer cycle in the central Kuril islands, *Nature* **451**, no. 7178, 561–565, doi: 10.1038/nature06521.

Ammon, C. J., H. Kanamori, T. Lay, and A. A. Velasco (2006). The 17 July 2006 Java tsunami earthquake, *Geophys. Res. Lett.* **33**, no. 24, doi: 10.1029/2006GL028005.

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res. 1896-1977* **75**, no. 26, 4997–5009, doi: 10.1029/JB075i026p04997.

Denolle, M. A. (2019). Energetic Onset of Earthquakes, *Geophys. Res. Lett.* **46**, no. 5, 2458–2466, doi: 10.1029/2018GL080687.

Denolle, M. A., W. Fan, and P. M. Shearer (2015). Dynamics of the 2015 M7.8 Nepal earthquake, *Geophys. Res. Lett.* **42**, no. 18, 7467–7475, doi: 10.1002/2015GL065336.

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

Furlong, K. P., T. Lay, and C. J. Ammon (2009). A Great Earthquake Rupture Across a Rapidly Evolving Three-Plate Boundary, *Science* **324**, no. 5924, 226–229, doi: 10.1126/science.1167476.

Haskell, N. A. (1964). Total energy and energy spectral density of elastic wave radiation from propagating faults, *Bull. Seismol. Soc. Am.* **54**, no. 6A, 1811–1841.

Houston, H. (2001). Influence of depth, focal mechanism, and tectonic setting on the shape and duration of earthquake source time functions, *J. Geophys. Res. Solid Earth* **106**, no. B6, 11137–11150, doi: 10.1029/2000JB900468.

Houston, H., and J. E. Vidale (1994). The Temporal Distribution of Seismic Radiation During Deep Earthquake Rupture, *Science* **265**, no. 5173, 771–774, doi: 10.1126/science.265.5173.771.

Ide, S. (2019). Frequent observations of identical onsets of large and small earthquakes, 7772, *Nature* **573**, no. 7772, 112–116, doi: 10.1038/s41586-019-1508-5.

Kingma, D. P., and J. Ba (2017). Adam: A Method for Stochastic Optimization, *ArXiv14126980 Cs*.

Kingma, D. P., and M. Welling (2014). Auto-Encoding Variational Bayes, in *Proceedings of the International Conference on Learning Representations (ICLR)*.

Meier, M.-A., J. P. Ampuero, and T. H. Heaton (2017). The hidden simplicity of subduction megathrust earthquakes, *Science* **357**, no. 6357, 1277–1281, doi: 10.1126/science.aan5643.

Melgar, D., and G. P. Hayes (2019). Characterizing large earthquakes before rupture is complete, *Sci. Adv.* **5**, no. 5, eaav2032, doi: 10.1126/sciadv.aav2032.

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 Mw 8.6 Sumatra Earthquake, *Science* **337**, no. 6095, 724–726, doi: 10.1126/science.1224030.

Okuda, T., and S. Ide (2018). Hierarchical rupture growth evidenced by the initial seismic waveforms, 1, *Nat. Commun.* **9**, no. 1, 3714, doi: 10.1038/s41467-018-06168-3.

Persh, S. E., and H. Houston (2004). Deep earthquake rupture histories determined by global stacking of broadband P waveforms, *J. Geophys. Res. Solid Earth* **109**, no. B4, doi: https://doi.org/10.1029/2003JB002762.

Renou, J., M. Vallée, and P. Dublanchet (2019). How Does Seismic Rupture Accelerate? Observational Insights From Earthquake Source Time Functions, *J. Geophys. Res. Solid Earth* **124**, no. 8, 8942–8952, doi: 10.1029/2019JB018045.

Ross, Z. E., B. Idini, Z. Jia, O. L. Stephenson, M. Zhong, X. Wang, Z. Zhan, M. Simons, E. J. Fielding, S.-H. Yun, *et al.* (2019). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence, *Science* **366**, no. 6463, 346–351, doi: 10.1126/science.aaz0109.

Tanioka, Y., and L. J. Ruff (1997). Source Time Functions, *Seismol. Res. Lett.* **68**, no. 3, 386–400, doi: 10.1785/gssrl.68.3.386.

Vallée, M. (2013). Source time function properties indicate a strain drop independent of earthquake depth and magnitude, *Nat. Commun.* **4**, no. 1, 2606, doi: 10.1038/ncomms3606.

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

Ye, L., H. Kanamori, and T. Lay (2018). Global variations of large megathrust earthquake rupture characteristics, *Sci. Adv.* **4**, no. 3, eaao4915, doi: 10.1126/sciadv.aao4915.

Ye, L., T. Lay, H. Kanamori, and L. Rivera (2016). Rupture characteristics of major and great (M_w $\geq$ 7.0) megathrust earthquakes from 1990 to 2015: 1. Source parameter scaling relationships, *J. Geophys. Res. Solid Earth* **121**, no. 2, 826–844, doi: 10.1002/2015JB012426.

Yin, J., Z. Li, and M. A. Denolle (2021). Source Time Function Clustering Reveals Patterns in Earthquake Dynamics, *Seismol. Res. Lett.* **92**, no. 4, 2343–2353.

Figures

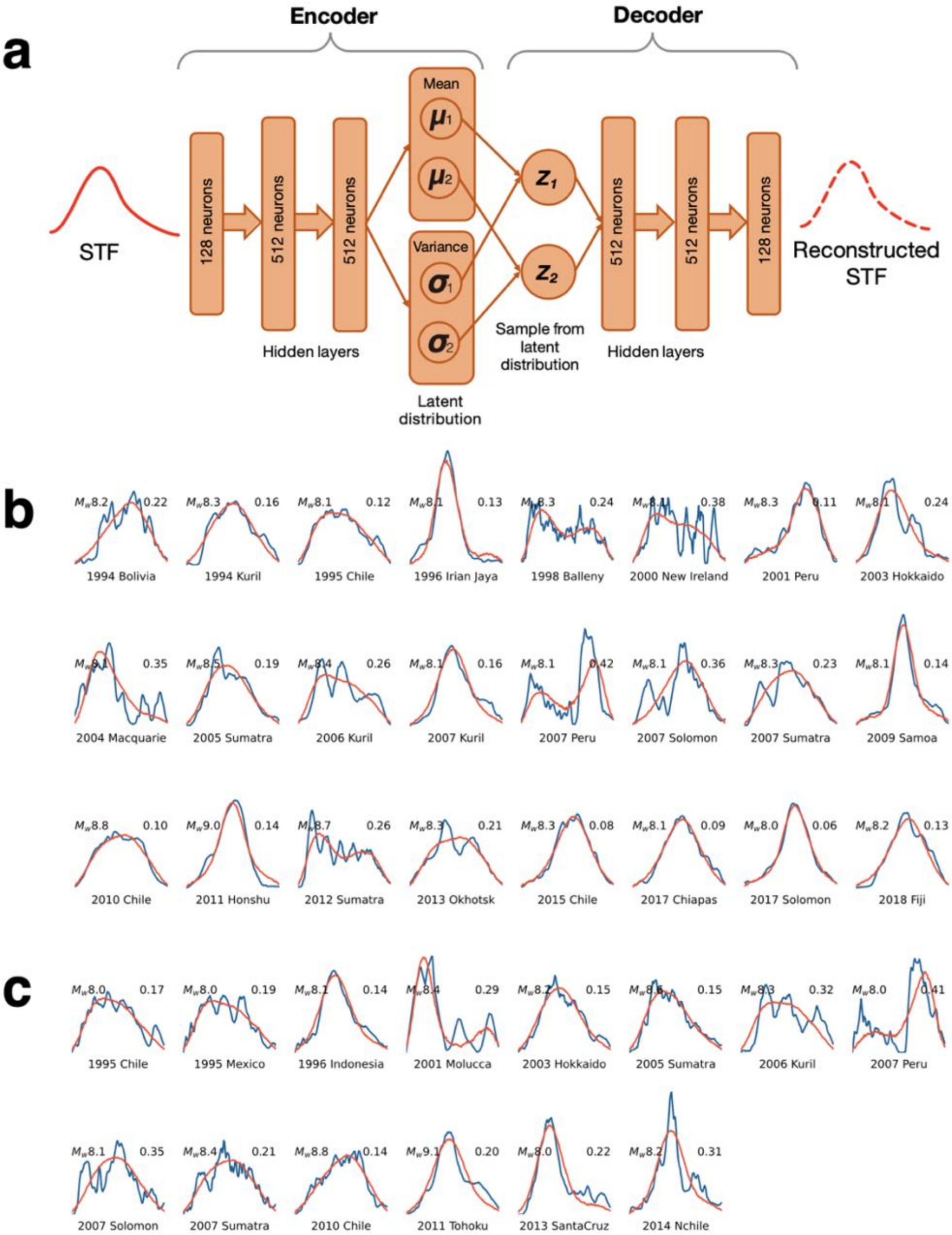

Figure 1. Variational autoencoder (VAE) for earthquake source time functions (STFs). a. The VAE architecture. Both of the decoder and encoder consist of two fully connected layers with 512 neurons. The bottleneck comprises of two latent variables constrained to follow normal distribution. b. Original STFs (blue) and VAE reconstructed STFs (red) from SCARDEC. The numbers in the top right mark the misfits. c. Same as b but for the STF database of Ye et al. (2016) (refer to as YE 2016 thereinafter).

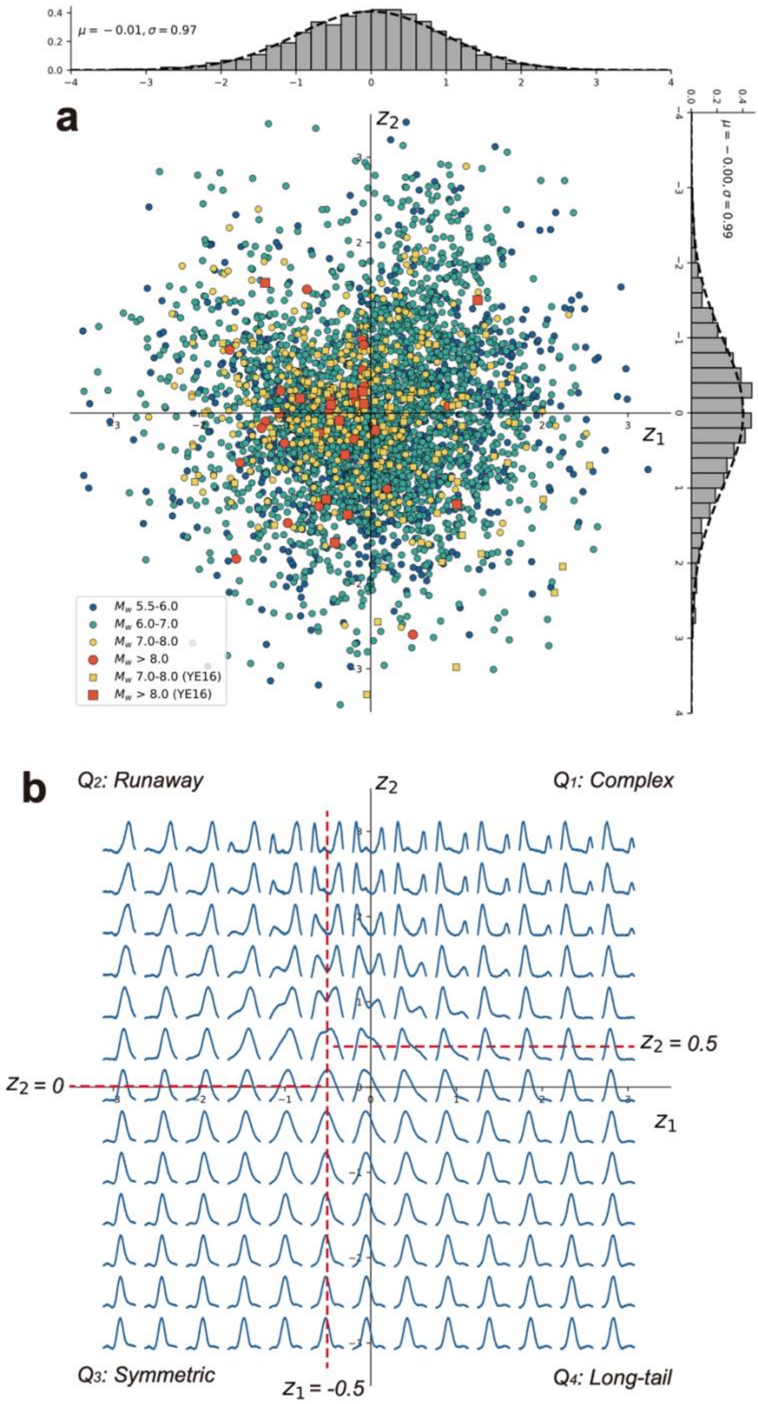

Figure 2. The latent representation and the standard spectrum of STFs. a. Low-dimensional latent representations of SCARDEC (dots) and YE2016 (squares) STFs. The population density in both Z_1 and Z_2 follows normal distribution. b. The standard spectrum of synthetic STFs as reconstructed from virtual latent values. The spectrum shows systematic variations of STFs: complex type in Q_1, runaway type in Q_2, symmetric type in Q_3, long-tail type in Q_4.

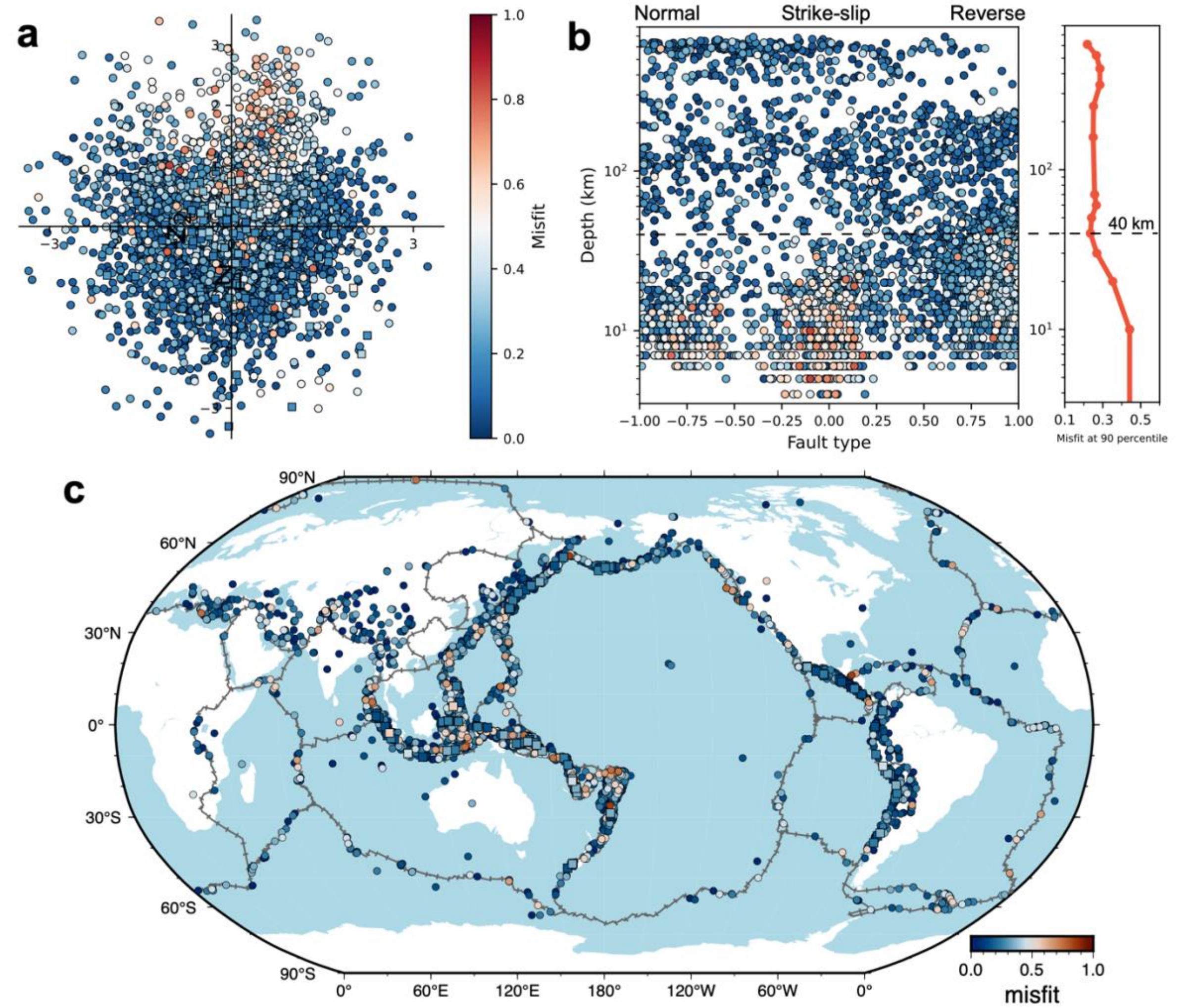

Figure 3. Reconstruction misfit as a proxy of unusual earthquakes. a. Reconstruction misfits of all the SCARDEC (dots) and YE2016 (squares) events in latent space. Note that most high-misfits are located in Q_1 and some are in other quadrants. b. Reconstruction misfit as a function of earthquake depth and focal mechanisms. On the right is the median misfit across different depths. The depth at 40 km marks the transition of earthquake complexity. c. Geographic distribution of reconstruction misfit of global earthquakes. Note that the high-misfits are predominantly located in the northern boundaries of the Australian plate and Southeast Asia.

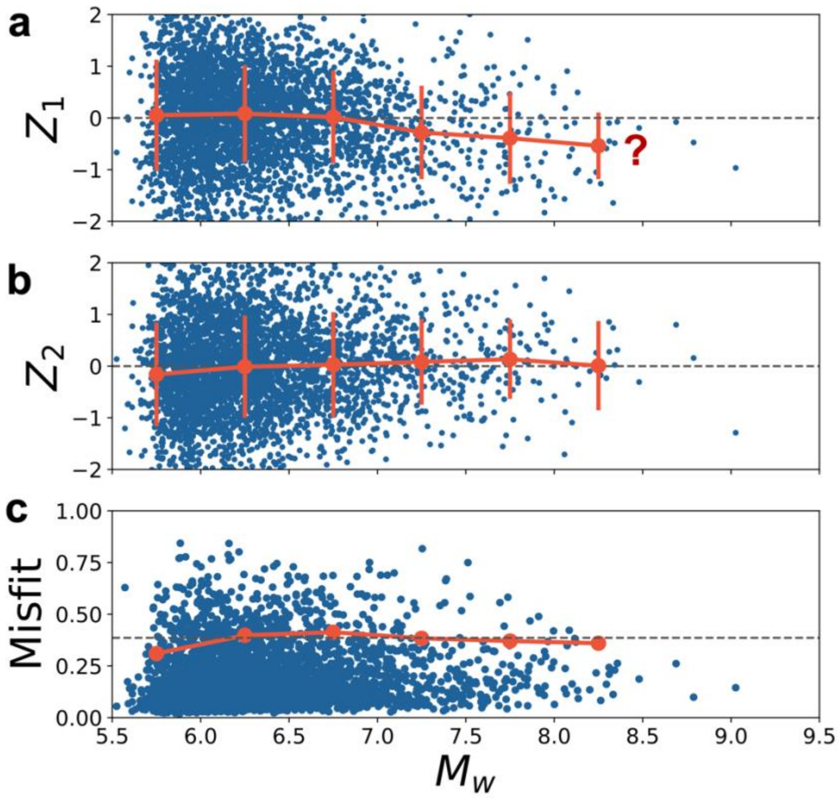

Figure 4. Latent variables *Z_1, Z_2* and reconstruction misfits as a function of Mw. Note that Z_2 and misfit remain nearly constant across different magnitudes, whereas Z_1 decreases with magnitude, indicating a plausible magnitude dependence of moment release processes.

Supplementary Information for

A generic spectrum of global earthquake rupture characteristics revealed by machine learning

Zefeng Li[1,2*]

1. *Laboratory of Seismology and Physics of Earth's Interior, School of Earth and Space Sciences, University of Science and Technology of China, Hefei, China*
2. *Mengcheng National Geophysical Observatory, University of Science and Technology of China, Mengcheng, China*

*Corresponding author: Zefeng Li (zefengli@ustc.edu.cn)

This file includes:

Figure S1: History of training and validation loss.

Figure S2: Relationship between the STF dissimilarity and distance in the latent space.

Figure S3: Variations of skewness and kurtosis across the standard STF spectrum.

Figure S4: Comparison of original and VAE reconstructed STFs for all the 112 megathrust earthquakes in Ye et al. (2016).

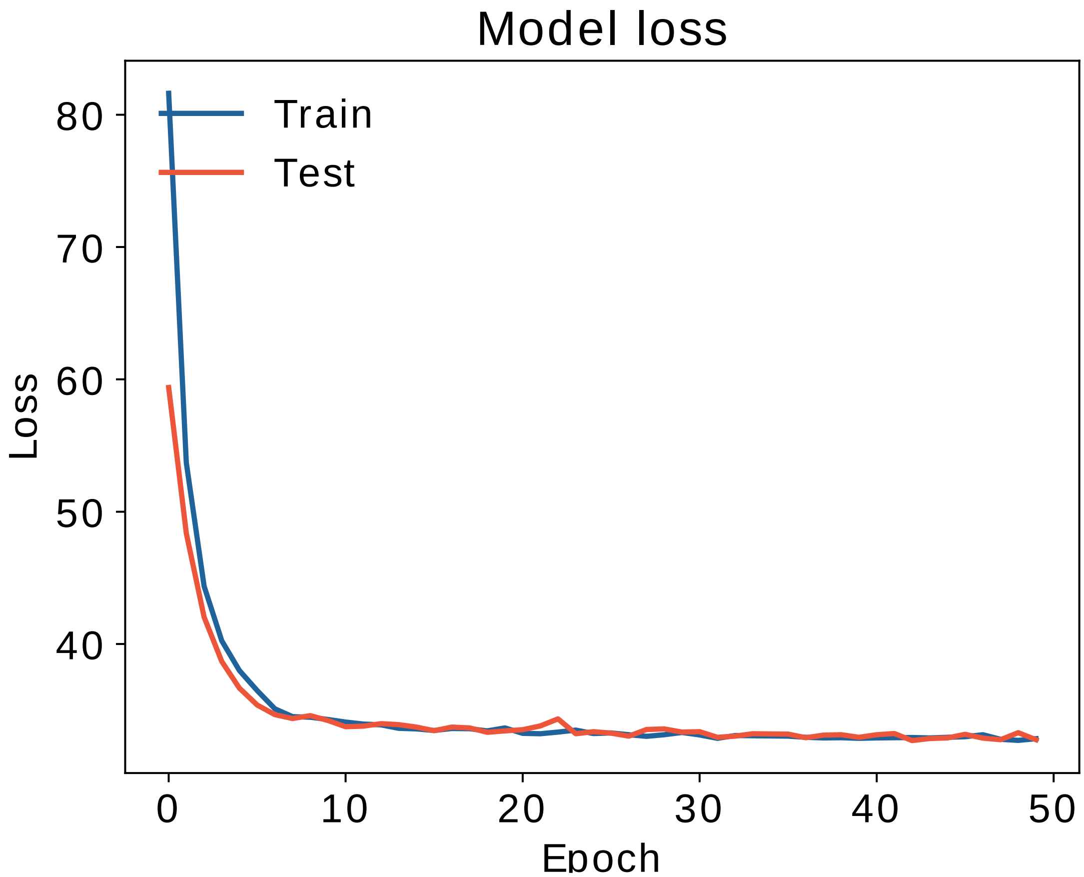

Figure S1. History of train and test loss for the VAE model.

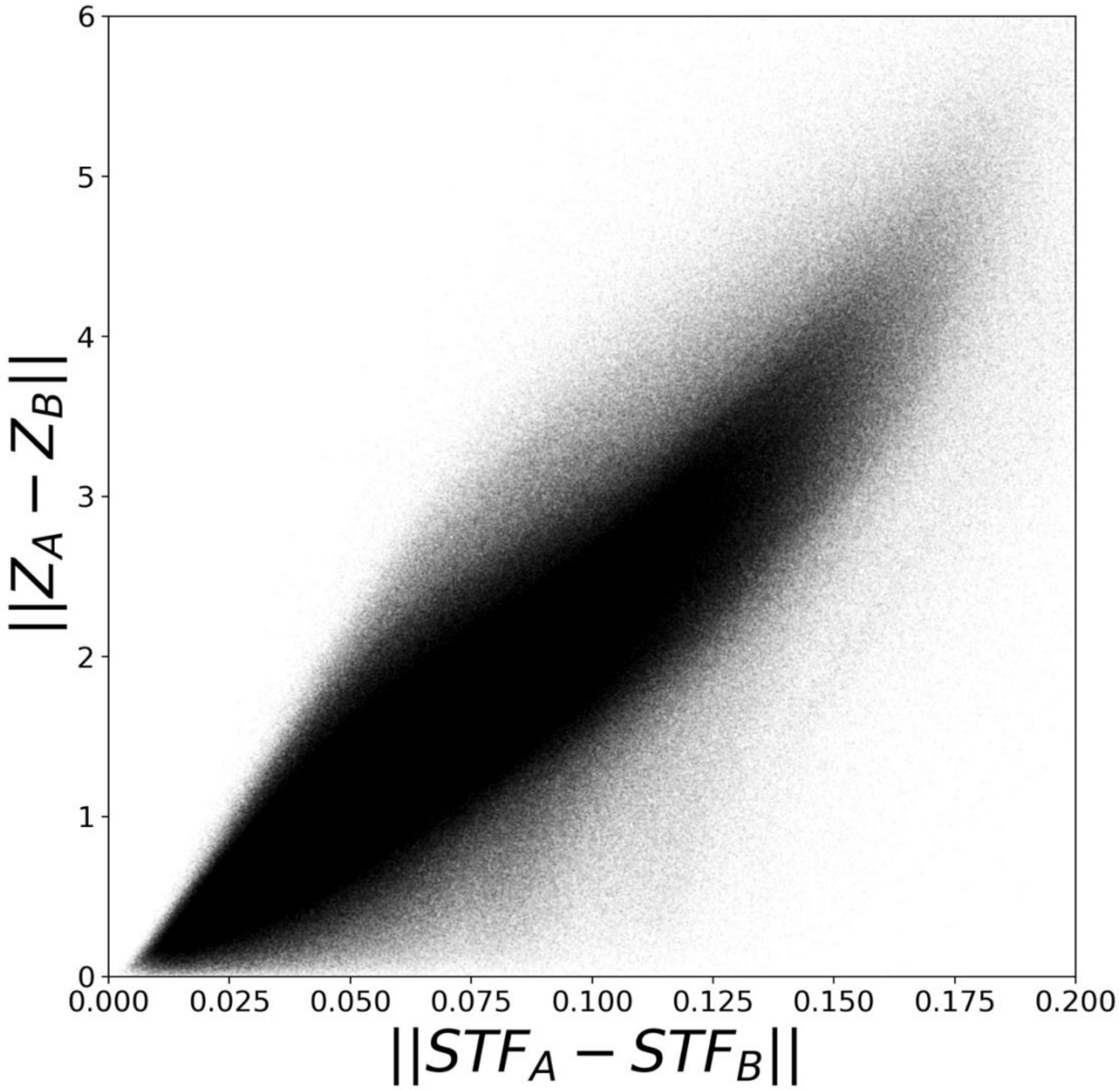

Figure S2. Affinity property of the latent space: the distance between two STFs in latent space is proportional to the dissimilarity between them.

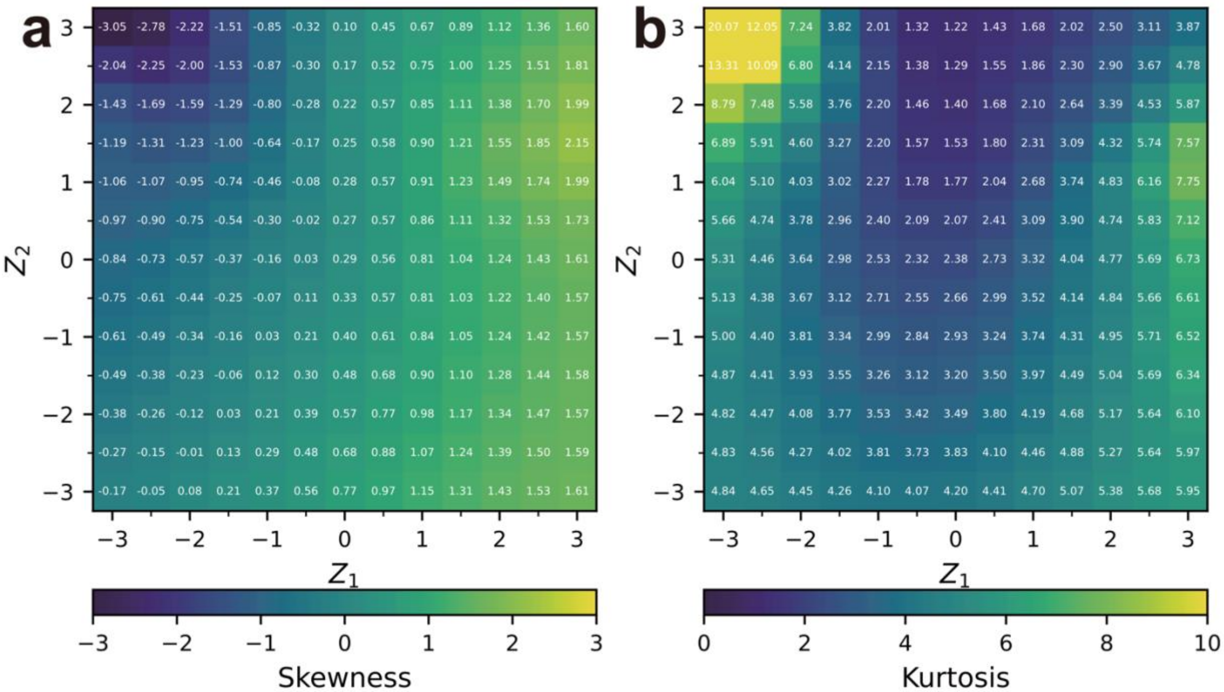

Figure S3. The variations of skewness and kurtosis across the standard STF spectrum.

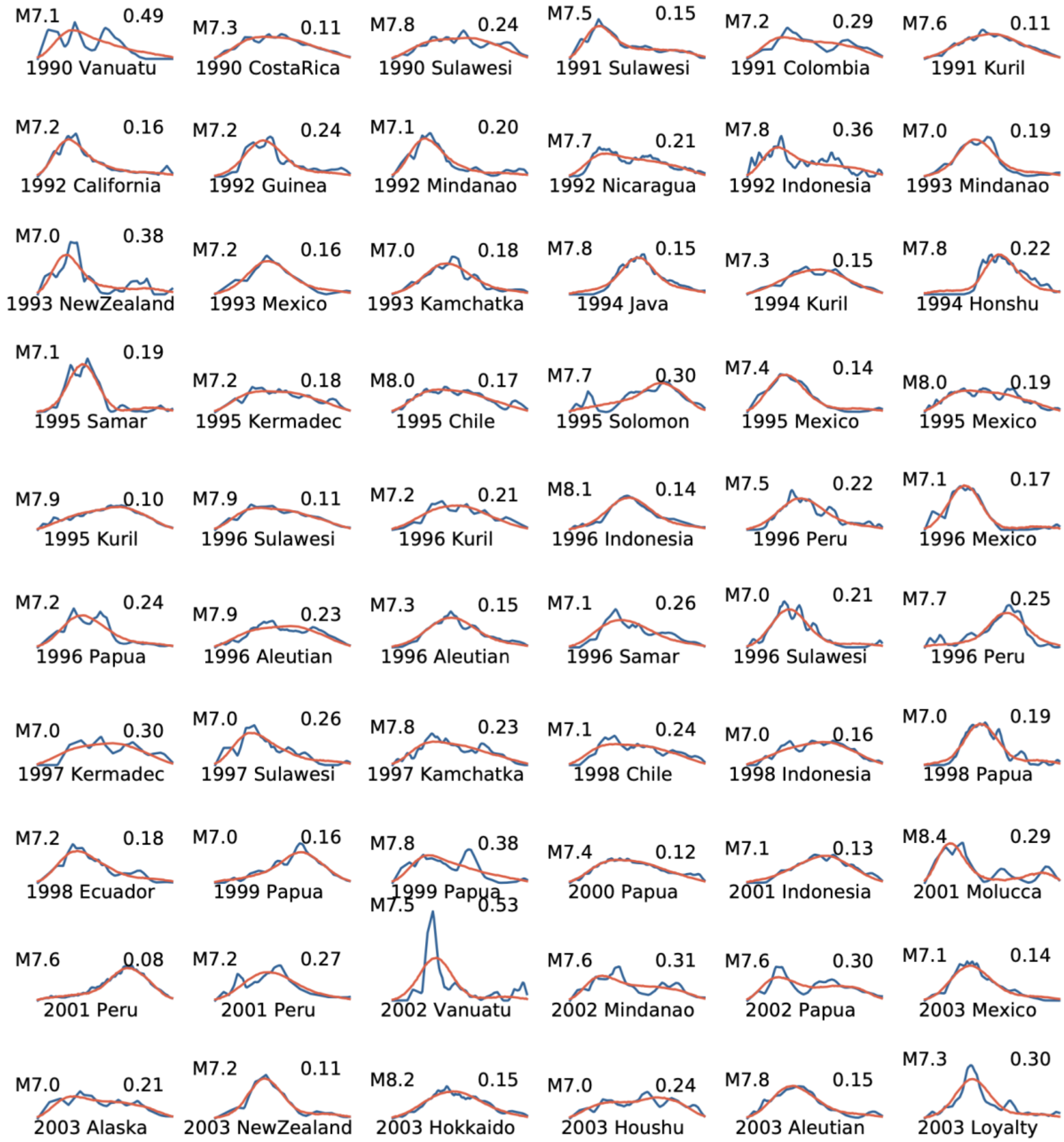

Figure S4. Original STFs (blue) and VAE reconstructed STFs (red) for all the 112 megathrust earthquakes in YE2016 (b). The numbers in the top right mark the misfits.

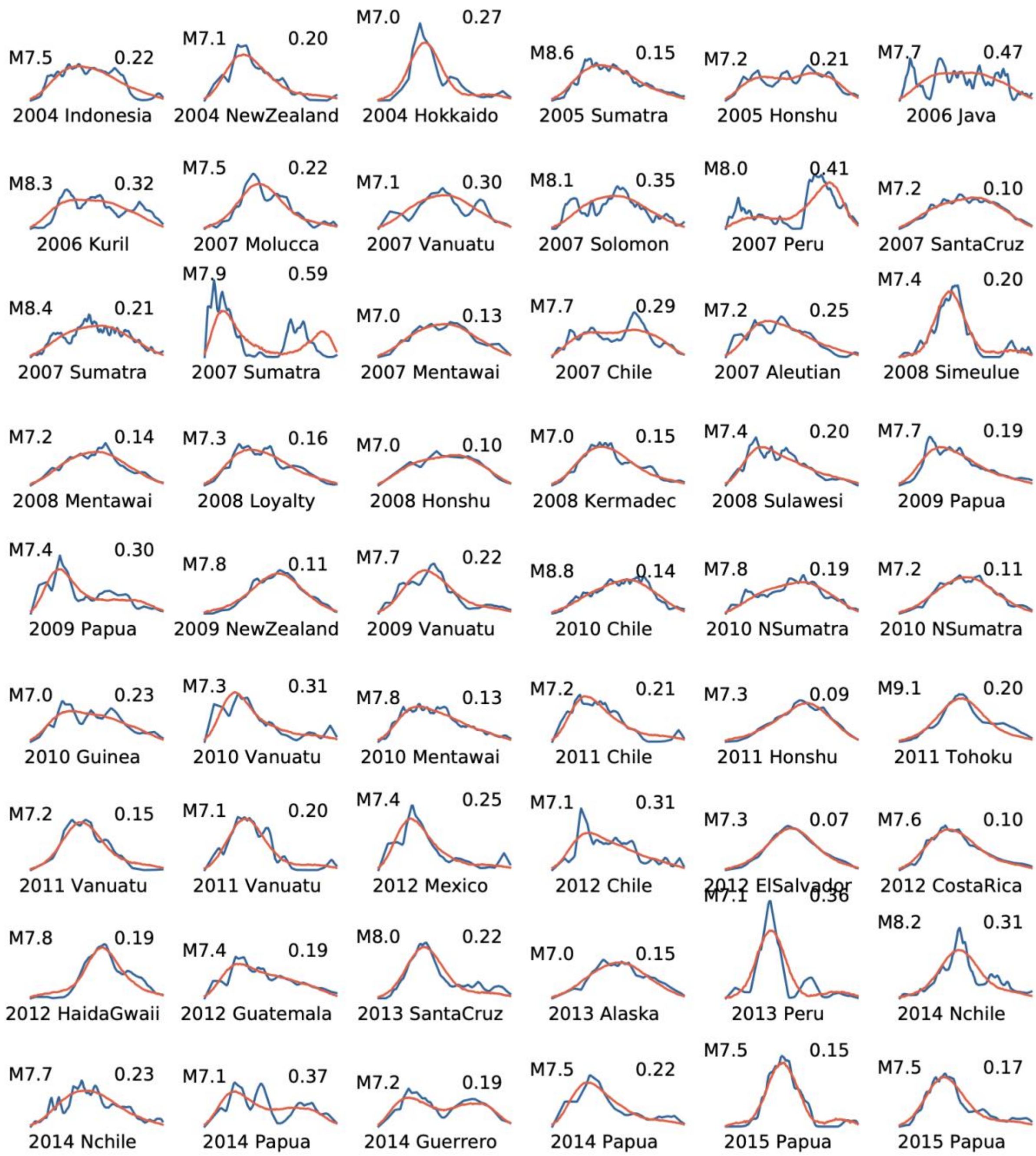

Figure S4 (continued). Original STFs (blue) and VAE reconstructed STFs (red) for all the 112 megathrust earthquakes in YE2016 (b). The numbers in the top right mark the misfits.