# Machine learning of source spectra for large earthquakes

Shang Ma[1], Zefeng Li[1], and Wei Wang[2]

[1]Univerisity of Science and Technology of China
[2]University of Southern California

November 30, 2022



**Abstract**

The shape of earthquake source spectra, traditionally fit by physics-based models, contains important parameters to constrain rupture dimension, duration, and geometry. Here we apply machine learning (ML) to derive single-variable and double-variable data-driven models of source spectra from 3675 Mw>5.5 global earthquakes, assuming that the Fourier transform of source time functions well represent earthquake source spectra below 1 Hz. The single-variable ML model, in the same degree of freedom as the Brune model, improves the goodness of fit by 8.5%. Specifically, the ML model fits the data without systematic bias, whereas the Brune model tends to underestimate at intermediate frequencies and overestimate at high frequencies. The latter discrepancy cannot be modelled by increasing the fall-off exponent in the Brune-type or the Boatwright-type models. The double-variable ML model is compared to existing double-corner-frequency models and is found to capture the second-order features such as the subtle curvature differences around the corner. Our results demonstrate that unsupervised machine learning can extract hidden global characteristics of high-dimensional data and provide observational evidence to amend existing physical models.

Geophysical Journal International

# Machine learning of source spectra for large earthquakes

| | |
|---|---|
| Journal: | *Geophysical Journal International* |
| Manuscript ID | GJI-S-21-0968.R2 |
| Manuscript Type: | Research Paper |
| Date Submitted by the Author: | 17-Apr-2022 |
| Complete List of Authors: | Ma, Shang; University of Science and Technology of China, School of Earth and Space Sciences<br>Li, Zefeng; University of Science and Technology of China, School of Earth and Space Sciences<br>Wang, Wei; University of Southern California, Department of Earth Sciences |
| Keywords: | Earthquake source observations < SEISMOLOGY, Machine Learning < GENERAL SUBJECTS, SEISMOLOGY |
| Additional Keywords: | |

# Machine learning of source spectra for large earthquakes

**Shang Ma[1], Zefeng Li[1,2*], and Wei Wang[3]**

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*
3. *Department of Earth Sciences, University of Southern California, Los Angeles, USA*

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

*Manuscript submitted to Geophysical Journal International*

*April 18, 2022*

## Summary

The shape of earthquake source spectra, traditionally fit by physics-based models, contains important parameters to constrain rupture dimension, duration, and geometry. Here we apply machine learning (ML) to derive single-variable and double-variable data-driven models of source spectra from 3675 $M_w$ >5.5 global earthquakes, assuming that the Fourier transform of source time functions well represent earthquake source spectra below 1 Hz. The single-variable ML model, in the same degree of freedom as the Brune model, improves the goodness of fit by 8.5%. Specifically, the ML model fits the data without systematic bias, whereas the Brune model tends to underestimate at intermediate frequencies and overestimate at high frequencies. The latter discrepancy cannot be modelled by increasing the fall-off exponent in the Brune-type or the Boatwright-type models. The double-variable ML model is compared to existing double-corner-frequency models and is found to capture the second-order features such as the subtle curvature differences around the corner. Our results demonstrate that unsupervised machine learning can extract hidden global characteristics of high-dimensional data and provide observational evidence to amend existing physical models.


**Key words:** Seismology; Machine learning; Earthquake source observations.

## 1 Introduction

The shape of earthquake source spectra is an important constraint on rupture processes. Corner frequency, fall-off rate, and curvature at corner are crucial parameters to infer stress drop, rupture dimension, geometry, duration and rupture velocity (Savage 1972; Chounet *et al.* 2018; Shearer *et al.* 2019). Haskell (1964) assumed that a uniform displacement discontinuity propogates at constant rupture velocity on a rectangular fault, which results in the far-field displacement pulse as the convolution of two boxcar pulses. In frequency domain, the Haskell model spectrum is flat at a level proportional to seismic moment at low frequencies, decays as $\omega^{-1}$ at intermediate frequencies, and $\omega^{-2}$ at high frequencies. Following the Haskell model, Aki (1967) proposed the $\omega^{-2}$ theoretical model and derived a scaling law based on the assumption of self-similarity, meaning that large and small earthquakes are geometrically similar and the long-period amplitude is the inverse-cube power of the corner frequency.

Brune (1970) suggested an influential form of source spectra:

$$A(f) = \mathrm{M}_0/[1 + (f/f_c)^2] \qquad (1)$$

Where $f_c$ is the corner frequency, $\mathrm{M}_0$ is the seismic moment. A more general functional form includes the path attenuation effects (Shearer 2019):

$$A(f) = \mathrm{M}_0 \exp(-\pi f t^*)/[1 + (f/f_c)^{\gamma n}]^{1/\gamma} \qquad (2)$$

Where *n* represents the high-frequency fall-off rate and is associated with high-frequency energy radiation, and $\gamma$ controls the curvature of the spectrum near the corner frequency. In this form, $\gamma$=1 and n=2 for the $\omega^{-2}$ model; n=3 for the $\omega^{-3}$ model; n=2 and $\gamma$=2 for the Boatwright model (Boatwright 1978).

The Brune model ($\omega^{-2}$ model) is widely used to infer source parameters and generally agree with global average source spectra across different magnitudes (Abercrombie 2015; Ye *et al.* 2016a, b; Shearer *et al.* 2019). However, Kaneko & Shearer (2014) considered a two degrees of freedom model:

$$A(f) = \mathrm{M}_0/[1 + (f/f_c)^n] \qquad (3)$$

They obtained *n* = 2.2 to 2.9 for P-wave spectra and *n* =1.5 to 2.6 for S-wave spectra in the frequency range of $0.05f_c \leq f \leq 10f_c$ in the cohesive-zone models of circular rupture. Uchide & Imanishi (2016) observed that the source spectra of most $M_w$ 3.2-4.0 shallow earthquakes in Japan deviate from the Brune model and implies a fall-off exponent slightly higher than 2.

Models with double corner frequencies have also been proposed to fit complex source spectra. Denolle & Shearer (2016) suggested a double-$f_c$ model to explain the source spectra of large subduction earthquakes (referred to as DS16 model hereafter) using the functional form:

$$A(f) = M_0/[\sqrt{1+(f/f_{c1})^2}\sqrt{1+(f/f_{c2})^2}] \qquad (4)$$

Ji & Archuleta (2021) introduced two double-$f_c$ models (one is self-similar, another is not), which can reproduce the mean peak ground acceleration (PGA) and mean peak ground velocity (PGV) for magnitudes 3.3–7.3. They reported that the magnitude dependencies of PGA and PGV are well explained by the nonself-similar double-$f_c$ source model (referred to as JA21 model hereafter).

While the aforementioned physical models provide a generally good fit for source spectra, deviations may exist for specific earthquakes owing to various assumptions and calculations of source spectra made on the rupture process. Comparatively, machine learning allows extracting the hidden features of large datasets which could be otherwise invisible through physical models (Bergen *et al.* 2019). In this study, we employ a generative machine learning algorithm, called variational autoencoder (VAE), to learn data-driven models of source spectra for large earthquakes. We aim to answer a question: whether or not the source spectra of large earthquakes are the Brune-type on global average? Alternatively, are any systematic source characteristics not captured by the Brune-type model?

This paper is organized as follows. First, we transform source time functions to source spectra. The spectra are normalized and fit with the Brune model to obtain the corner frequencies. Second, synthetic tests are used to prove that VAE can derive meaningful average spectra from a large dataset. Third, VAE is applied to learn the real source spectra of global earthquakes. We compare the single- and double-variable VAE models to the Brune and other more complex models and discuss the consistency and discrepancy among them. Finally, we explore the possible physical implications of source spectra derived from the data and the potential limitations.

## 2 Data

Source spectra of global earthquakes used in this study are amplitude spectra of source time functions (STFs). We use the SCARDEC (Seismic source ChActeristics Retrieved from DEConvolvolving teleseismic body waves) (Vallée & Douet 2016) dataset, which consists of 3675 large earthquakes with $M_w$ >5.5 from January 1, 1992 to December 31, 2019. Vallée *et al.* (2011) deconvolved teleseismic waves with the Green's functions and take local surface reflections for both source and receiver crusts as well as mantle attenuation (geometrical spreading and attenuation factors) into account. In addition, the SCARDEC method does not assume the same STF at each station and imposes no constraints on the spatial-temporal complexity of the rupture process. Overall, SCARDEC is advantageous in its full automation and thus provides a larger dataset for machine learning in this study. SCAREDEC provides two types of STFs, the optimal one among all stations and the average one of all stations. In this study we use the average STFs because they are relatively insensitive to possible outliers and generally more robust than the optimal ones. Moreover, it is less impacted by directivity effects that may exist on specific stations (Vallée & Douet 2016).

The STFs are first filled with zeros at the end of the STFs to keep the same time length (and hence the same frequency range and interval after Fourier transform). Each spectrum is normalized by seismic moment to only keep the shape. It is resampled evenly on the logarithmic scale to 128 points, which is the input size of the machine learning model. In this study, we reserve the frequency range up to 1 Hz to avoid the poorly constrained high-frequency contents in STFs, which follows the presumption that the global models of attenuation are better constrained by seismic frequencies lower than 1 Hz (Danré *et al.* 2019; Denolle 2019; Yin *et al.* 2021). However, whether or not the frequency content below 1 Hz is absolutely reliable remains arguable, which may have negative impact on our analyses and will be discussed in the last section.

## 3 Method

Generative modeling is a category of machine learning approaches to model the data implicitly (i.e. without clear physical meaning). It learns a probabilistic model that describes how a dataset is generated from input. With generative modeling, we could generate new plausible examples like a physical model by sampling probabilistic model. Generative adversarial networks (GANs) are a typical generative model and have been applied to earthquake early warning and seismic data augmentation (Li *et al.* 2018; Wang *et al.* 2021). However, GANs often suffers from unstable training (Salimans et al., 2016), as they do not have explicit constraints on the probability distribution. Comparatively, VAE, which is also a generative model, uses explicit constraints on the probability distribution and is able to reconstruct high-dimensional data from a compact latent representation (Kingma & Welling 2013). Its functionality acts similarly to physical modeling (case-dependent parameters embedded in a shared functional form), and could be used to learn a model for earthquake source spectra directly from data.

A typical VAE model is composed of two parts, i.e. an encoder and a decoder. The encoder compresses the input to a compact latent representation, whereas the decoder reconstructs the input from the latent representation. The bottleneck structure forces the model to learn the primary features of the data. The VAE's loss function in this analysis is defined as:

$$loss = ||\frac{S_{model} - S_{raw}}{freq}||_2 + KL[\mathbb{N}(\mu_z, \sigma_z) - \mathbb{N}(0,1\,)]$$

where the first term is the root mean square between the reconstructed and original source spectra, and the second is the Kullback-Leibler divergence which measures the difference between latent variables and normal distribution. After training, the VAE acts in a manner similar to the physical Brune model: for a given source spectrum, the encoder obtains the latent parameter (like the $f_c$ of the Brune model) and the decoder reconstructs the source spectra (like the fit curve of the Brune model). The difference is that VAE model is data-driven (from data directly) and the model parameter is implicit, whereas the Brune model is physics-based (from theoretical assumptions of source process) and the model parameter has explicit physical meaning.

We construct the VAE architecture following Li (2022) (Fig. 1). The encoder and the decoder have two fully connected layers, and each layer has 256 neurons. We first use only one latent variable in VAE to keep the same degree of freedom as the Brune model (one-parameter $f_c$; amplitude is normalized). Hence, the latent variable of VAE acts like $f_c$ of the Brune model, and the trained parameters of the neural network acts like the functional form of the Brune model (Fig. 1). Following a similar procedure, we train another VAE model with two latent variables and compare it with recently proposed double-$f_c$ models (DS16 and JA21). We randomly split 80% of the data as training set and 20% as testing set in both synthetic tests and the SCARDEC dataset.

## 4 Results

### *4.1 Validation of Machine learning with synthetic source spectra*

We follow the procedure in Shearer *et al.* (2006) to fit the SCARDEC spectra with the Brune model by least square of the log spectrum. We estimate the $f_c$ for all earthquakes and use them to generate synthetic data for machine learning tests.

First, to demonstrate the modeling capability of VAE, we apply VAE to learn the Brune model from noise-free synthetic data. We generate 3675 synthetic source spectra using eq. (1) derived from SCARDEC. Fig. 2a shows that an example from test data for which VAE has perfect fit. The misfits ($||\frac{S_{model} - S_{raw}}{freq}||_2$) for most of these spectra are almost negligible (Fig. 2b). This demonstrates that VAE correctly learns the Brune model from noise-free spectra.

Second, to demonstrate that VAE can learn an average model from noisy data, we add random Gaussian perturbation to the fall-off exponent (i.e., $n = 2 \pm \delta,\ \delta \in \mathbb{N}(0, 0.2)$) of the synthetic source spectra in the previous test. Then the average model of all the synthetic data still follows the Brune model, but individual spectra randomly deviate. The training result shows that the VAE model also agrees with the Brune model (Figs. 2c and d). This demonstrates that VAE can derive an average model from a large dataset, even if the data have random deviation individually.

*4.2 Single-variable modeling of real source spectra*

On the basis of the previous tests, we train a single-variable VAE model (referred to as VAE1 model hereafter) with the source spectra from SCARDEC. The results show that, VAE1 generally provides better fit than the Brune model and exhibits some different characteristics (Fig. 2e). Fig. 2f shows the histograms of misfits of two models as well as the differences between them. Overall, the mean misfit of VAE1 is about 8.5% smaller than that of the Brune model. This small difference suggests that two models are largely consistent. Moreover, we observe that the latent variable of VAE1 is strongly correlated with $f_c$ (Fig. 3a). The latent variable also has a similar relationship with $M_w$ (Fig. 3b) because of the inherent scaling relationship between $f_c$ and $M_w$. Since machine learning searches in a much wider parameter space than Brune model, the high correlation with $f_c$ demonstrates that $f_c$ is indeed an effective parameter controlling the spectral shapes. However, with the same degree of freedom, the lower misfit of the data-driven model suggests that the Brune model misses some systematic characteristics of the observed data.

Fig. 4a shows the overall variation of the VAE1 model spectra, the Brune-type source spectra and average real source spectra, with respect to the VAE1 latent parameter. Specifically, as the latent variable value increases, $f_c$ decreases and magnitude increases. It is noteworthy that the reliability of the VAE curves depends on the number of available real data. Therefore, VAE1 provides the most reliable results approximately for $M_w$ 6-7 because of the data abundance in that range. To reveal the differences between VAE1 and the Brune model, we subtract the average real data from their fitting curves (Fig. 4b). The residuals suggest that the VAE1 spectra are more consistent with the observed data across different magnitudes. This suggests that VAE1 has learned an unbiased average model from the data and can serve as a baseline for other physical source models. In comparison, the residuals show that the Brune model systematically underestimates in intermediate frequencies and overestimates in high frequencies for the SCARDEC dataset (Fig. 4b).

*4.3 Double-variable modeling of real source spectra*

For comparison with double-$f_c$ models (DS16, JA21) recently proposed to supplement the Brune model (single-$f_c$) , we train a double-variable VAE model (referred to as VAE2 model hereafter) which also has two degrees of freedom. The VAE2 have two latent variables $Z_1$ and $Z_2$, compared to a single variable Z in the VAE1 model. Fig. 5a shows an example of SCARDEC source spectra fitted by VAE2 and double-$f_c$ models. The median residual of the VAE2 model is near zero, whereas DS16 and JA21 have some deviations (Fig. 5b), similar to the Brune model. We observe that 95-percentile of the VAE2, DS16, JA21 misfits are 0.0065, 0.0088, 0.0278 respectively (Fig. 5c and Table 1). In comparison, VAE1 has 95-percentile misfit at 0.0115 (Table 1). We estimate the statistical relative amount of information loss of these models with Akaike information criterion (AIC), which deals with the trade-off between the goodness of fit and the complexity of the model:

$$\mathrm{AIC} = 2\mathrm{k} + \mathrm{N}ln\left(\frac{\mathrm{RSS}}{\mathrm{N}}\right)$$

Where N is the number of frequency samples, k is the number of estimated parameters, RSS is the residual sum of squares. Note that this definition assumes normally distributed errors. For single-variable models, the parameters are the latent variable or $f_c$ and residual variance so that k = 2. Double-variable models have k = 3. Compared to other models, VAE2 provides improved goodness of fit (Table 1) and captures more detailed features of the source spectra, especially for the curvature at turning corner (Fig. 6).

To investigate the effect of the additional latent variable, we visualize the variations of source spectra in the latent space (Fig. 7a). We observe that the source spectra change more significantly with $Z_2$ than with $Z_1$, indicating a primary effect of $Z_2$ and a secondary effect of $Z_1$. Moreover, the correlation between $Z_1$ and $f_c$ acts more subtle (Fig. 7b), whereas $Z_2$ appears (Fig. 7c) similar to that in VAE1(Fig. 3a), . Therefore, these likely suggests that the role of $Z_2$ in VAE2 is comparable to that in single-variable VAE model, whereas $Z_1$ could catch more secondary details to promote the goodness of fit. Generally, $Z_2$ seems to control the corner frequency like the only variable in the VAE1, whereas $Z_1$ seems to control the abruptness of transition from the low-frequency plateau to the high-frequency fall-off.

## 5 Discussions

### *5.1 Physical implications of the VAE models*

To clarify the causes of the systematic characteristics not captured by the Brune and other physical models, we explore different model parameters and attenuation effect to see if the difference can be reduced to near zero across the frequency range as the VAEs do. First, we experiment different values of high-frequency fall-off rate in eq. (3). Although the fit improves slightly when the high-frequency fall-off rate around 2.3, the mean differential curve cannot be reduced to be flat by simply tuning the fall-off rate (Fig. 8a). Second, we tune the parameters $\gamma$ and high-frequency fall-off rate in Boatwright model, but find it leads to even higher misfit than the Brune model (Fig. 8b). Third, SCARDEC uses an attenuation model to correct the attenuation effect on the spectra. Although there could be attenuation effect not fully corrected, the apparent slopes in high- and low-frequency ranges observed in this study appear too large to be explained by the remaining attenuation effect. Since DS16 and JA21 cannot adequately explain the observed slow fall-off in the intermediate frequencies and the fast fall-off in the high frequencies, we propose a modified double-$f_c$ model to simulate the characteristics of the real data revealed from VAE models (Fig. 8d). This model is similar to DS16 and JA21 but has the $f^{-1}$ fall-off rate in the intermediate frequencies and has a $f^{-2.6}$ slope in the high-frequency region. We find that this combination can generally replicate the major shape. Nonetheless, we can only constrain the first exponent to be <2 and the second exponent to be >2; the actual combination of them can vary.

In the Haskell model, the presence of two corner frequencies results from that the slip risetime is much less than rupture duration time ($\tau_r << \tau_d$). This short risetime phenomenon has been shown by dynamic rupture modeling results (*Beroza & Mikumo 1996; Melgar & Hayes 2017; Wang & Day 2017*) and can be caused by several mechanisms. For example, Das

& Aki (1977) suggested that spatially heterogeneous fault strength (e.g., barriers) may limit slip duration at particular locations on a fault. Heaton (1990) postulated that short risetime can be caused by dynamic fault friction, which decreases with increasing slip velocity. Dynamic changes of normal stress induced by bi-material effects (Andrews & Ben-Zion 1997; Ampuero & Ben-Zion 2008) and low-velocity fault zones (Huang & Ampuero 2011) could also generate short-risetime slip pulses. These explanations are indeed indistinguishable from our observations. However, it needs to be noted that the slow fall-off rate ($<2$) in intermediate frequencies is introduced because of the finite-length long-narrow faults (Haskell 1964; Savage 1972); the higher fall-off rate ($>2$) in high frequencies implies that the rupture processes of large earthquakes may not be strictly self-similar but prefer elongated fault geometries (Shearer 2019). Our hypothesized model is purely data-driven and precise interpretation of its physical meaning remains a subject of future research.

*5.2 Limitations of the STF data and their impact*

Our conclusions largely rely on the presumption that the Fourier transform of the SCARDEC STFs below 1Hz are good representations of source spectra. However, several factors may undermine this presumption and thus have negative impact on our conclusions. Although Vallée & Douet (2016) deconvolved the teleseismic waves by Green's functions which are better constrained at frequencies below 1 Hz, this does not directly support that the frequencies below 1 Hz are perfectly reliable. Besides, the averaging of STFs and time-domain deconvolution tend to further reduce the high frequency content (Vallée & Douet 2016), which may lead to overestimation in the high-frequency fall-off rate. Therefore, one should be aware that the actual appropriate cutoff frequency remains difficult to determine, although the 1-Hz cutoff is used in this study.

Moreover, Danré *et al.* (2019) showed that the STFs in SCARDEC can be fit by a sum of Gaussian pulses, implying the source spectra calculated from this dataset are inherently different from the Brune model. Besides, the calculation of STFs suffers a number of intricate issues such as the assumed Green’s function and the influence of attenuation. The SCARDEC method (Vallée et al. 2011) makes a point source approximation and deconvolves seismic waveforms with assumed Green's function derived from the method of Bouchon (1976) to separate the source term directly. Vallée and Douet (2016) noted that it is difficult to well separate real source effects from spurious moment episodes related to unmodeled seismic phases. In the cases like offshore strike-slip events, long and complex STFs can be questionable. However, Yin et al. (2021) observed that colocated shallow events have distinct degrees of complexity, and therefore inaccuracy in the Green's function should not strongly systematically bias the results.

### *5.3 Potentials of the VAE method*

The VAE approach optimizes global fit and derives the general features directly from real data. It can infer complex, nonlinear and high-dimensional data relationships, and obtain a data-driven model without any prior assumption and human supervision. Although the latent parameters and the functional form to model source spectra are implicit and difficult to interpret, the data-driven model serves as a quasi-optimal baseline for which physical models need to approach. Any systematic shift of physical models from the data-driven model in the same degree of freedom likely indicates some inherent flaws within the physical models (another possibility is that the dataset itself is already biased). In this perspective, our approach has the potential to reveal hidden characteristics of large and high-dimensional seismological data and provide opportunities to amend existing theoretical frameworks.

## Acknowledgements

The authors are grateful to Martin Vallée for providing the source time function databases. Zefeng Li thank Jiuxun Yin and Yifang Cheng for discussion on the preliminary idea of using machine learning to learn source spectra shapes. This research was supported by the National Key R&D Program of China (No. 2021YFC3000700).

## Data Availability

The SCARDEC database is available at http://scardec.projects.sismo.ipgp.fr/ (last accessed on September 3, 2021).

## References

Abercrombie, R.E. (2015) Investigating uncertainties in empirical Green's function analysis of earthquake source parameters. *J. Geophys. Res. Solid Earth*, **120**, 4263–4277. doi:10.1002/ 2015JB011984

Aki, K. (1967) Scaling Law of Seismic Spectrum. *J. Geophys. Res.*, **72**, 1217–1231. doi:10.1029/JZ072i004p01217

Ampuero, J.-P. & Ben-Zion, Y. (2008) Cracks, pulses and macroscopic asymmetry of dynamic rupture on a bimaterial interface with velocity-weakening friction. *Geophys. J. Int.*, **173**, 674–692. doi:10.1111/j.1365-246X.2008.03736.x

Andrews, D.J. & Ben-Zion, Y. (1997) Wrinkle-like slip pulse on a fault between different materials. *J. Geophys. Res. Solid Earth*, **102**, 553–571. doi:10.1029/96JB02856

Bergen, K.J., Johnson, P.A., Hoop, M.V. de & Beroza, G.C. (2019) Machine learning for data-driven discovery in solid Earth geoscience. *Science*, **363**, eaau0323. doi:10.1126/science.aau0323

Beroza, G.C. & Mikumo, T. (1996) Short slip duration in dynamic rupture in the presence of heterogeneous fault properties. *J. Geophys. Res. Solid Earth*, **101**, 22449–22460. doi:10.1029/96JB02291

Boatwright, J. (1978) Detailed spectral analysis of two small New York state earthquakes. *Bull. Seismol. Soc. Am.*, **68**, 1117–1131.

Bouchon, M. (1976) Teleseismic body wave radiation from a seismic source in a layered medium. *Geophys. J. Int.*, **47**, 515–530.

Brune, J.N. (1970) Tectonic Stress and the Spectra of Seismic Shear Waves from Earthquakes. *J. Geophys. Res.*, **75**, 4997–5009. doi:10.1029/JB075i026p04997

Chounet, A., Vallée, M., Causse, M. & Mathieu, F. (2018) Global catalog of earthquake rupture velocities shows anticorrelation between stress drop and rupture velocity. *Tectonophysics*, **733**, 148–158.

Danré, P., Yin, J., Lipovsky, B.P. & Denolle, M.A. (2019) Earthquakes Within Earthquakes: Patterns in Rupture Complexity. *Geophys. Res. Lett.*, **46**, 7352–7360. doi:10.1029/2019GL083093\

Das, S. & Aki, K. (1977) Fault plane with barriers: A versatile earthquake model. *J. Geophys. Res. 1896-1977*, **82**, 5658–5670. doi:10.1029/JB082i036p05658

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

Denolle, M.A. (2019) Energetic Onset of Earthquakes. *Geophys. Res. Lett.*, **46**, 2458–2466. doi:10.1029/2018GL080687

Haskell, N. (1964) Total energy and energy spectra density of elastic waves from propagating faults. *Bull. Seismol. Soc. Am.*, **54**, 1811–1841.

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

Huang, Y. & Ampuero, J.-P. (2011) Pulse-like ruptures induced by low-velocity fault zones. *J. Geophys. Res.*, **116**, B12307. doi:10.1029/2011JB008684

Ji, C. & Archuleta, R.J. (2021) Two Empirical Double-Corner-Frequency Source Spectra and Their Physical Implications. *Bull. Seismol. Soc. Am.*, **111**, 737–761. doi:10.1785/0120200238

Kaneko, Y. & Shearer, P.M. (2014) Seismic source spectra and estimated stress drop derived from cohesivezone models of circular subshear rupture. *Geophys. J. Int.*, **197**, 1002–1015. doi:https://doi.org/10.1093/gji/ggu030

360 Kingma, D.P. & Welling, M. (2013) Auto-Encoding Variational Bayes. *ArXiv E-Prints*, arXiv:1312.6114.

362 Li, Z. (2022). A generic model of global earthquake rupture characteristics revealed by machine learning. *Geophysical Research Letters*, 49, e2021GL096464. https://doi.org/10.1029/2021GL096464

365 Li, Z., Meier, M.-A., Hauksson, E., Zhan, Z., & Andrews, J. (2018). Machine Learning Seismic Wave Discrimination: Application to Earthquake Early Warning. *Geophysical Research Letters*, 45(10), 4773–4779. https://doi.org/10.1029/2018GL077870

368 Melgar, D. & Hayes, G.P. (2017) Systematic Observations of the Slip Pulse Properties of Large Earthquake Ruptures. *Geophys. Res. Lett.*, **44**, 9691–9698. doi:10.1002/2017GL074916

371 Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X. & Chen, X. (2016) Improved Techniques for Training GANs. *Adv. Neural Inf. Process. Syst.* eds. Lee, D., Sugiyama, M., Luxburg, U., Guyon, I. & Garnett, R., Vol. 29, Curran Associates, Inc.

375 Savage, J.C. (1972) Relation of corner frequency to fault dimensions. *J. Geophys. Res. 1896-1977*, **77**, 3788–3795. doi:10.1029/JB077i020p03788

377 Shearer, P.M. (2019) *Introduction To Seismology*, 3rd ed., Cambridge: Cambridge University Press.

379 Shearer, P.M., Abercrombie, R.E., Trugman, D.T. & Wang, W. (2019) Comparing EGF Methods for Estimating Corner Frequency and Stress Drop From P Wave Spectra. *J. Geophys. Res. Solid Earth*, **124**, 3966–3986. doi:https://doi.org/10.1029/2018JB016957

383 Shearer, P.M., Prieto, G.A. & Hauksson, E. (2006) Comprehensive analysis of earthquake source spectra insouthern California. *J. Geophys. Res. Solid Earth*, **111**, B06303.

Uchide, T. & Imanishi, K. (2016) Small earthquakes deviate from the omega-square model as revealed by multiple spectral ratio analysis. *Bull. Seismol. Soc. Am.*, **106**, 1357–1363. doi:https://doi.org/10.1785/0120150322

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. *Geophys. J. Int.*, **184**, 338–358.

Vallée, M. & Douet, V. (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

Wang, T., Trugman, D. & Lin, Y. (2021) SeismoGen: Seismic Waveform Synthesis Using GAN With Application to Seismic Data Augmentation. *J. Geophys. Res. Solid Earth*, **126**, e2020JB020077. doi:10.1029/2020JB020077

Wang, Y. & Day, S.M. (2017) Seismic source spectral properties of crack-like and pulse-like modes of dynamic rupture. *J. Geophys. Res. Solid Earth*, **122**, 6657–6684. doi:10.1002/2017JB014454

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

Ye, L., Lay, T., Kanamori, H. & Rivera, L. (2016b) Rupture characteristics of major and great ($M_w \geq 7.0$) megathrust earthquakes from 1990 to 2015: 2. Depth dependence. *J. Geophys. Res. Solid Earth*, **121**, 845–863.

Yin, J., Li, Z. & Denolle, M.A. (2021) Source Time Function Clustering Reveals Patterns in Earthquake Dynamics. *Seismol. Res. Lett.*, **92**, 2343–2353. doi:10.1785/0220200403

**Tables:**

**Table 1.** Performance of single- and double-variable models

| | Single-variable | | Double-variable | | |
|---|---|---|---|---|---|
| | VAE1 | Brune | VAE2 | DS16 | JA21 |
| Median misfit | 0.0013 | 0.0015 | 0.0008 | 0.0014 | 0.0056 |
| 95% misfit | 0.0115 | 0.0117 | 0.0065 | 0.0088 | 0.0278 |
| Average AIC | -1064.2 | -1029.6 | -1100.2 | -1033.8 | -843.5 |

**Figures:**

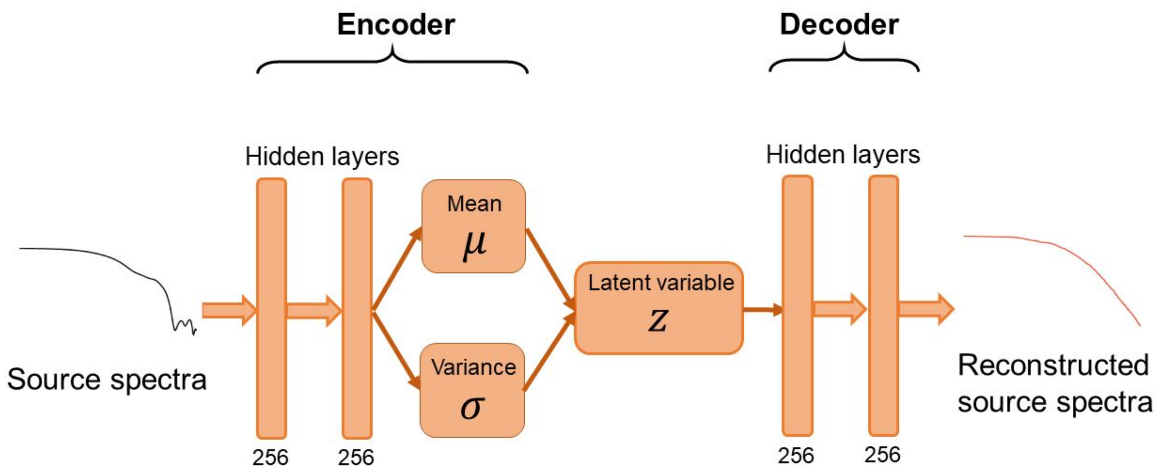



**Figure 1.** The architecture of variational autoencoder (VAE). Both of the decoder and encoder consist of two fully connected layers with 256 neurons each. The lengths of input (real source spectra) and output (reconstructed source spectra) are 128 data points. Note that we limit the latent dimension to 1 (for the single-variable model) and 2 (for the double-variable model) so that each spectrum can be modelled in a low degree of freedom.

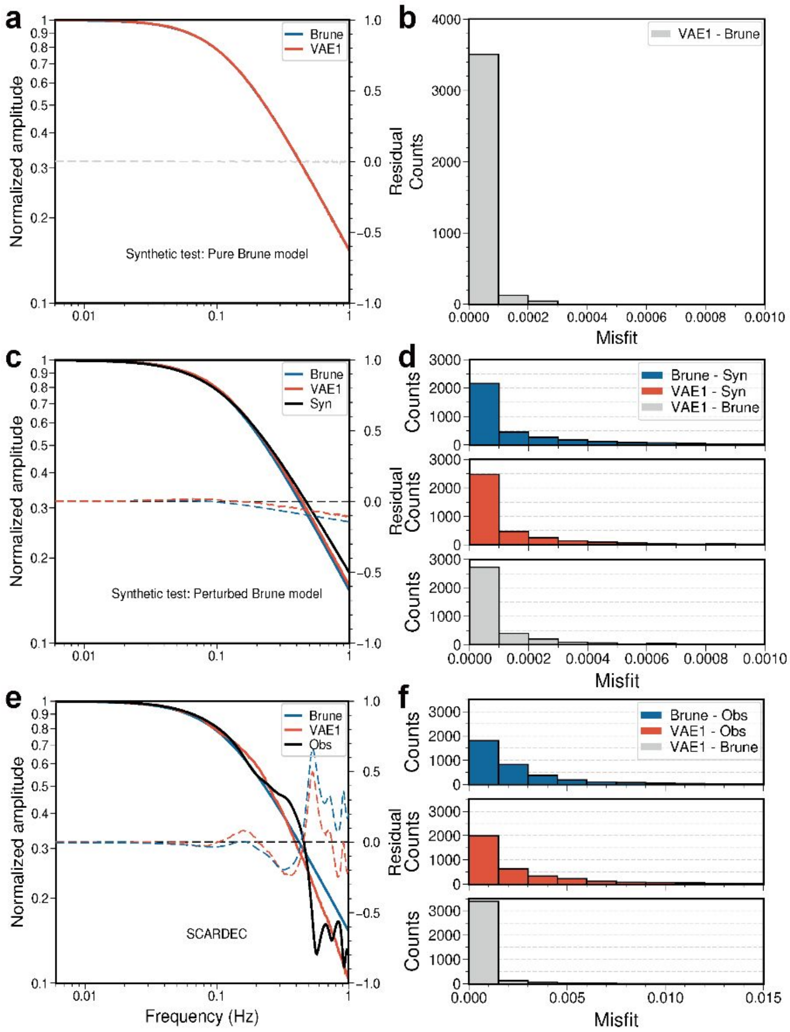



**Figure 2.** Goodness of fit of the single-variable VAE and Brune models for earthquake source spectra. (a) A noise-free synthetic source spectrum example fit by the VAE model (solid red) and the Brune model (solid blue), which overlap in the plot. The grey dashed line denotes the residual (right y-axis). (b) Misfit histogram of the VAE model for the noise-free synthetic spectra. (c) A random-perturbed Brune-type synthetic source spectrum fit by the VAE model (solid red) and the Brune model (solid blue). The blue and red dashed lines denote the residuals of the VAE model and the Brune model (right y-axis). (d) Misfit histograms of the VAE and Brune models and differences between them for the random-perturbed Brune-type synthetic source spectra. (e) A SCARDEC source spectrum example fit by the VAE model (solid red) and the Brune model (solid blue) (f) Misfit histograms of the VAE and Brune models for the SCARDEC source spectra.

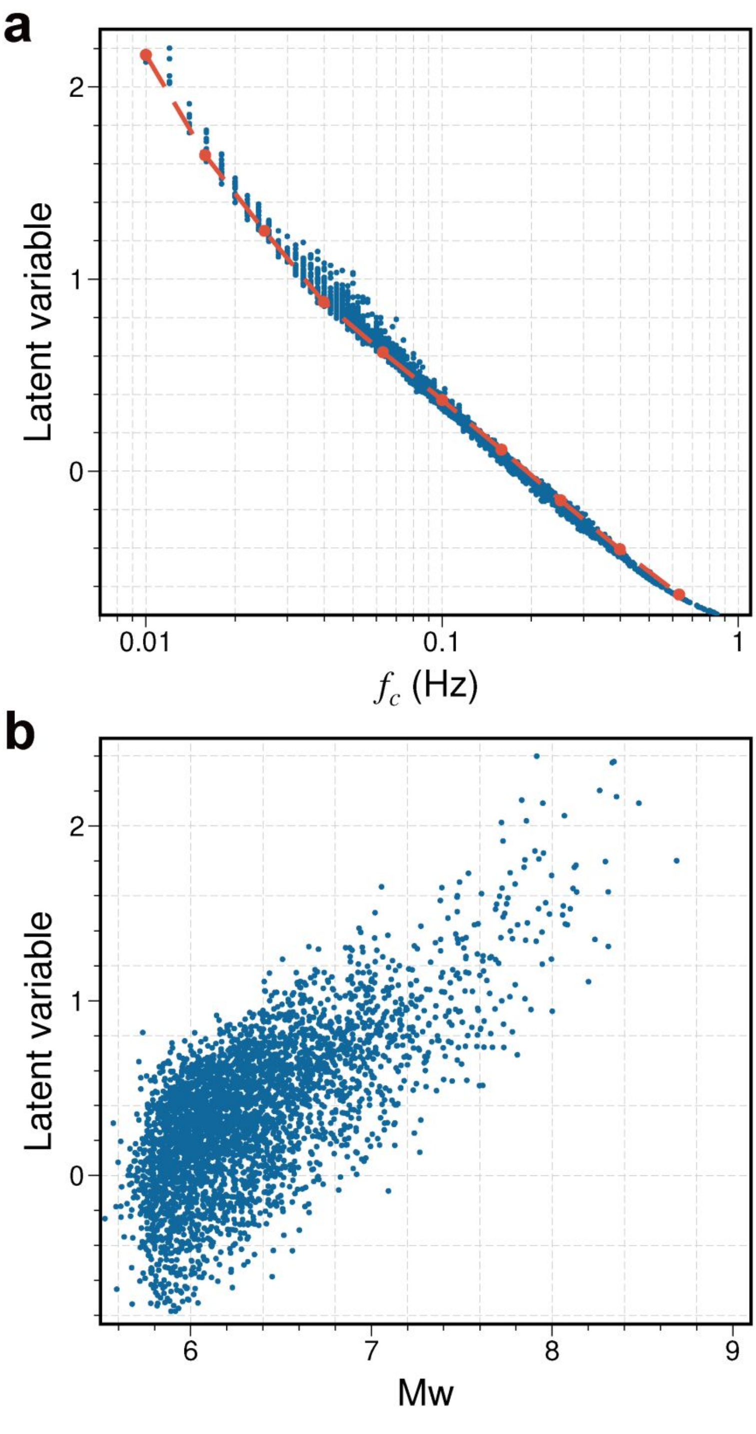



**Figure 3.** Variations of the single VAE latent variable with the Brune model parameter $f_c$ and the earthquake magnitude. (a) Strong correlation between the VAE latent variable and the Brune model parameter $f_c$. The red dotted line is the median latent value for each $f_c$. (b) Correlation between the VAE latent variable and earthquake magnitude.

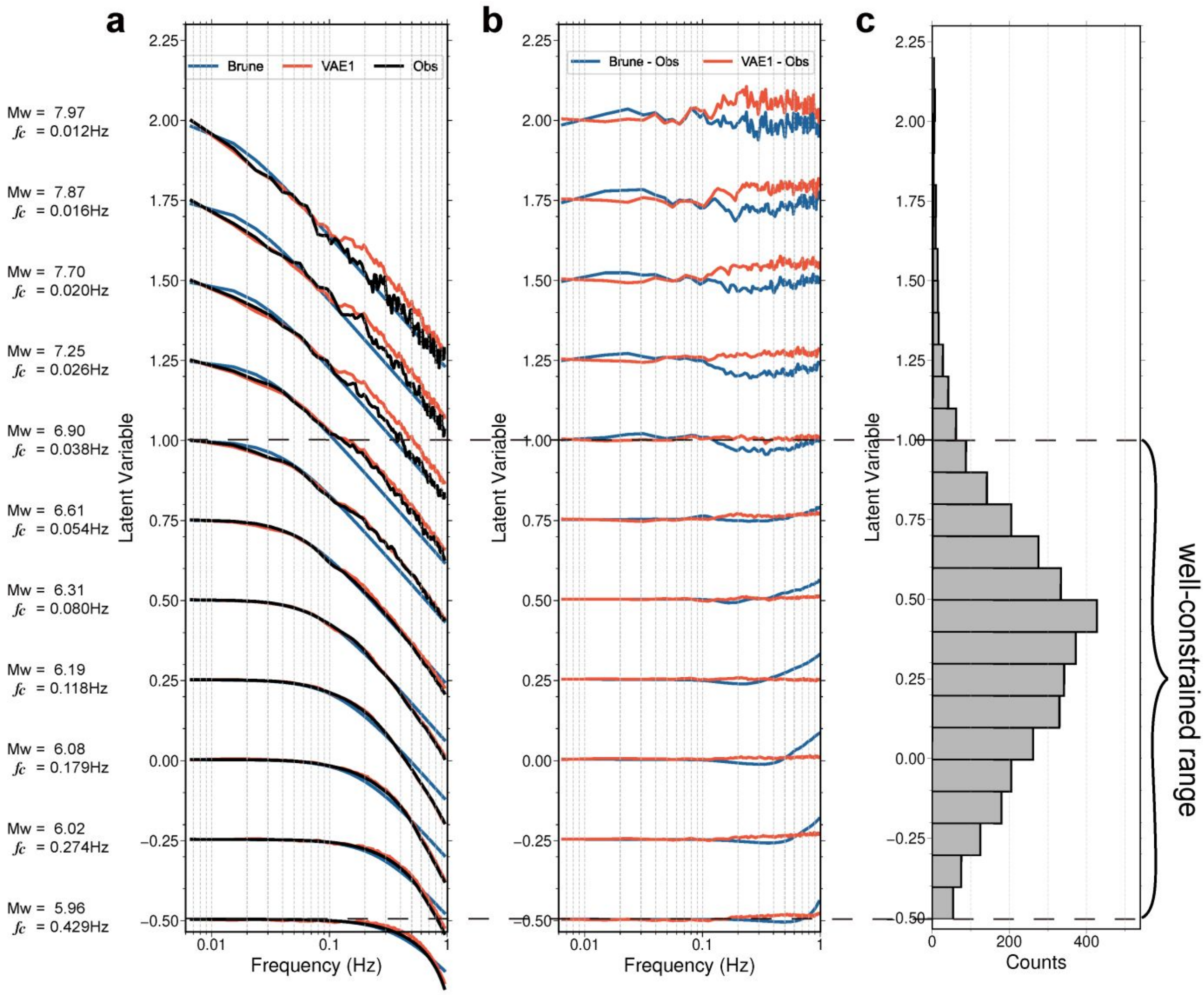



**Figure 4.** Systematic variation of source spectra in the latent space of the single-variable VAE. (a) The VAE spectra, Brune spectra and SCARDEC magnitude-binned average source spectra, with respect to the VAE latent parameter. (b) Residuals of the VAE and Brune models by substracting the SCARDEC data. (c) Histogram of latent variable values. The VAE model is well constrained within M 6-7 due to the abundant observations.

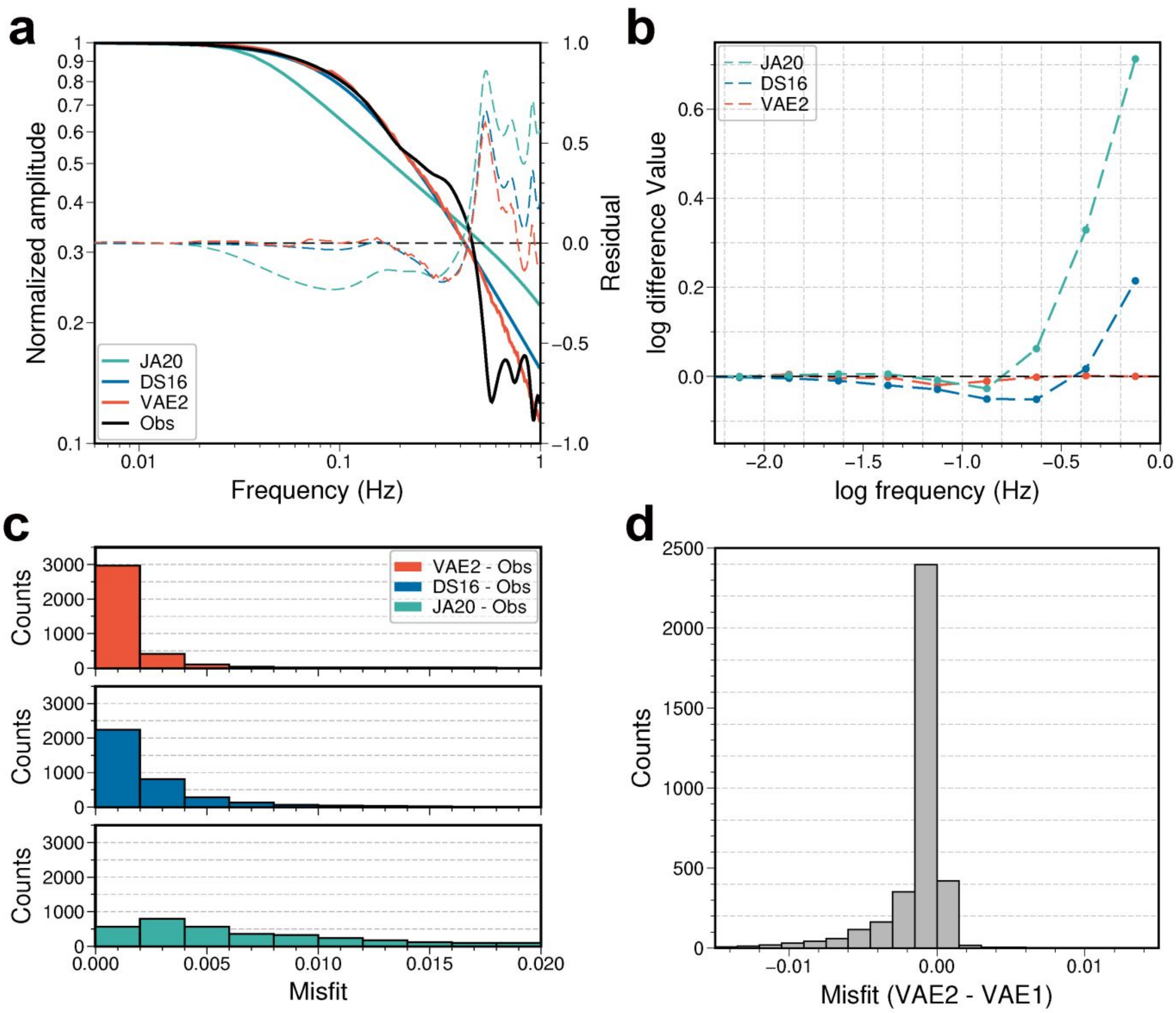



**Figure 5.** Goodness of fit of the double-variable VAE and double-$f_c$ physical models (JA21, DS16). (a) A SCARDEC source spectrum example fit by the VAE model (solid red), DS16 model (solid blue) and JA21 model (solid green). The red, blue and green dashed lines denote the residuals of the VAE model, DS16 model and JA21 model (right y-axis). (b) Median residuals of the double-variable VAE model and double-$f_c$ physical models (JA21, DS16) . (c) Misfit histograms of the VAE, DS16 and JA21 models for the SCARDEC source spectra. (d) Histogram of misfit residuals between the double-variable and single-variable VAE models.

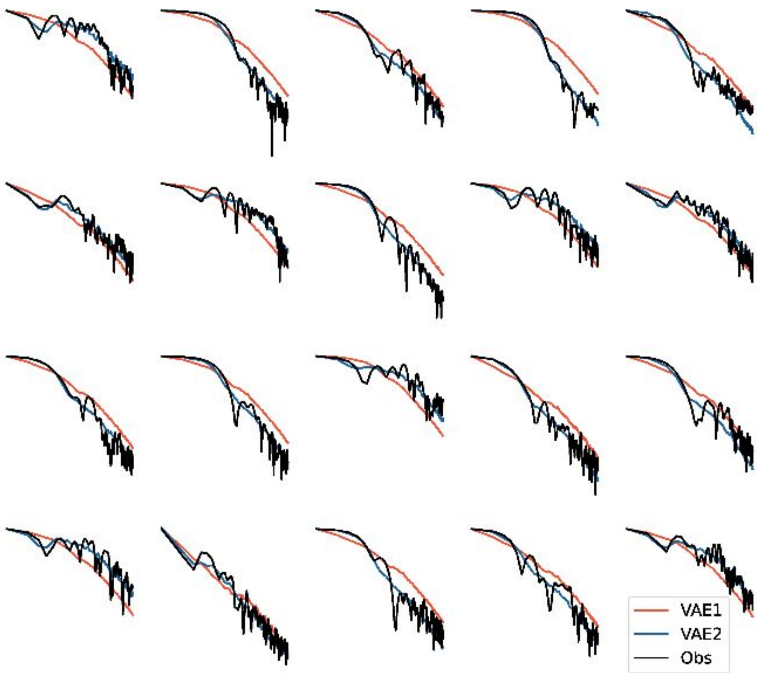



**Figure 6.** Examples of SCARDEC source spectra (black) and their reconstructions from single-variable (red) and double-variable (blue) VAE models.

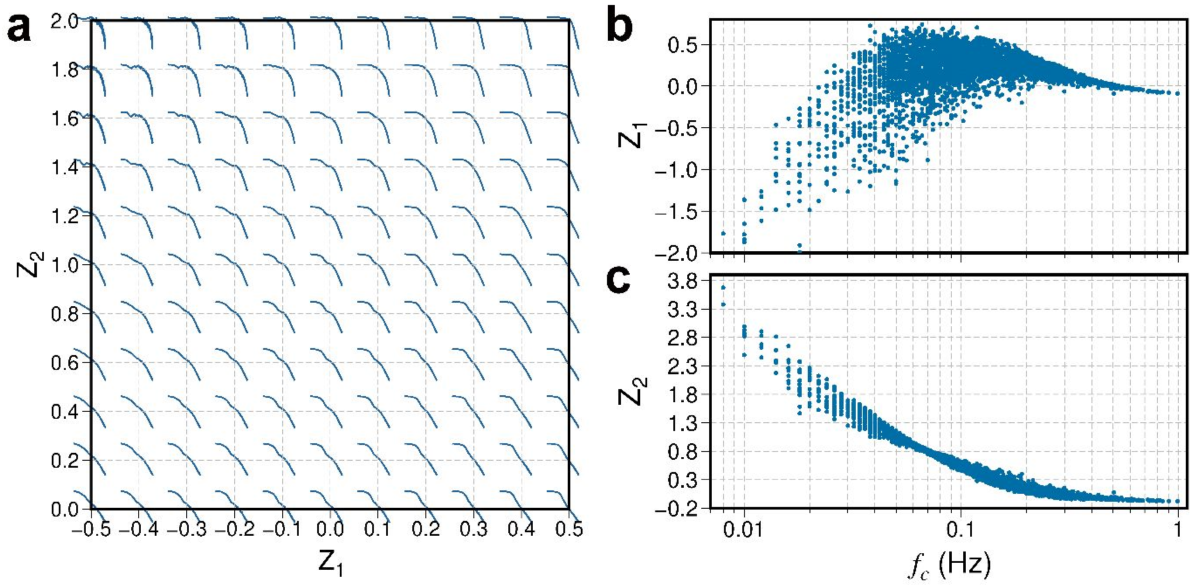



**Figure 7**. (a) Manifold of source spectra reconstructed from evenly sampled latent variables of the double-variable VAE model. (b) Correlation between the latent variable $Z_1$ and the Brune model parameter $f_c$. (c) Correlation between the latent variable $Z_2$ and the Brune model parameter $f_c$.

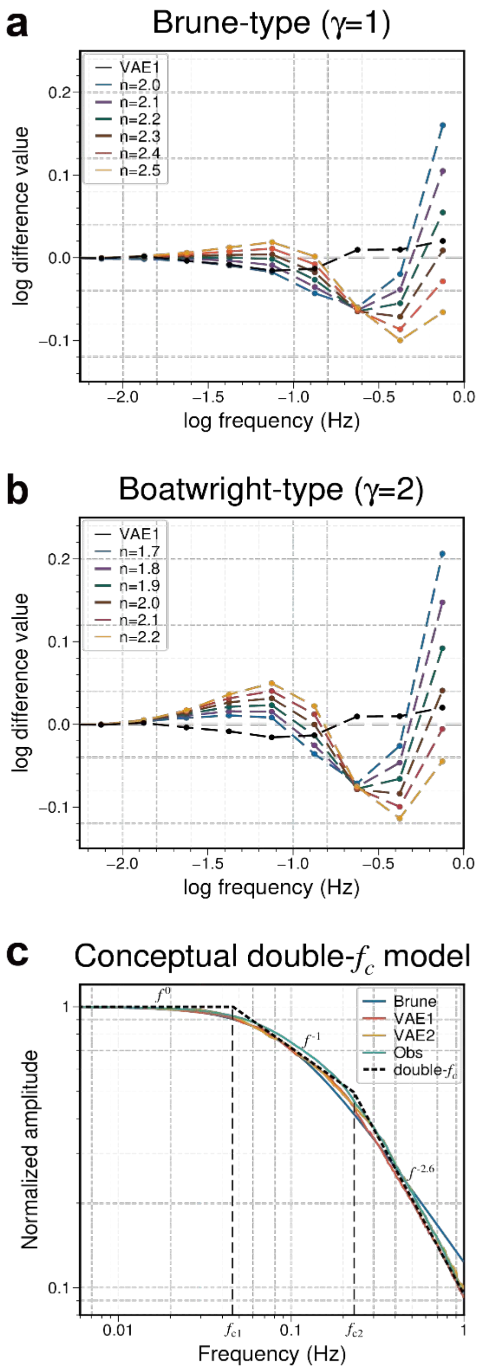



**Figure 8.** The effect of model and attenuation parameters on fitting SCARDEC source spectra. (a) Median residuals between different Brune-type models ($\gamma$=1, n=2.0-2.5) and the observed source spectra. (b) Median residuals between Boatwright-type models ($\gamma$=2, n=1.7-2.2) and the observed source spectra. (c) Average source spectra for Brune model, VAE1, VAE2, and observations for $M_w$ 6-7. The VAE1 and VAE2 curves are obtained by taking the median of all fitting curves for the SCARDEC events within $M_w$ 6-7. Dashed black line represents our proposed double-$f_c$ model.

# Machine learning of source spectra for large earthquakes

**Shang Ma[1], Zefeng Li[1,2*], and Wei Wang[3]**

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*
3. *Department of Earth Sciences, University of Southern California, Los Angeles, USA*

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

*Manuscript submitted to Geophysical Journal International*

*April 18, 2022*

**Style Definition:** Comment Text

**Deleted:** *February 16*

**Summary**

The shape of earthquake source spectra, traditionally fit by physics-based models, contains important parameters to constrain rupture dimension, duration, and geometry. Here we apply machine learning (ML) to derive single-variable and double-variable data-driven models of source spectra from 3675 $M_w$ >5.5 global earthquakes, assuming that the Fourier transform of source time functions well represent earthquake source spectra below 1 Hz. The single-variable ML model, in the same degree of freedom as the Brune model, improves the goodness of fit by 8.5%. Specifically, the ML model fits the data without systematic bias, whereas the Brune model tends to underestimate at intermediate frequencies and overestimate at high frequencies. The latter discrepancy cannot be modelled by increasing the fall-off exponent in the Brune-type or the Boatwright-type models. The double-variable ML model is compared to existing double-corner-frequency models and is found to capture the second-order features such as the subtle curvature differences around the corner. Our results demonstrate that unsupervised machine learning can extract hidden global characteristics of high-dimensional data and provide observational evidence to amend existing physical models.

**Deleted:** a global set of

**Deleted:** intrinsic attenuation effect, nor by

**Deleted:** Also, we compare the

**Deleted:** find that the double-variable ML model captures

**Deleted:** hints

**Key words:** Seismology; Machine learning; Earthquake source observations.

## 1 Introduction

The shape of earthquake source spectra is an important constraint on rupture processes. Corner frequency, fall-off rate, and curvature at corner are crucial parameters to infer stress drop, rupture dimension, geometry, duration and rupture velocity (Savage 1972; Chounet *et al.* 2018; Shearer *et al.* 2019). Haskell (1964) assumed that a uniform displacement discontinuity propogates at constant rupture velocity on a rectangular fault, which results in the far-field displacement pulse as the convolution of two boxcar pulses. In frequency domain, the Haskell model spectrum is flat at a level proportional to seismic moment at low frequencies, decays as $\omega^{-1}$ at intermediate frequencies, and $\omega^{-2}$ at high frequencies. Following the Haskell model, Aki (1967) proposed the $\omega^{-2}$ theoretical model and derived a scaling law based on the assumption of self-similarity, meaning that large and small earthquakes are geometrically similar and the long-period amplitude is the inverse-cube power of the corner frequency.

Brune (1970) suggested an influential form of source spectra:

$$A(f) = \mathrm{M}_0/[1 + (f/f_c)^2] \quad (1)$$

Where $f_c$ is the corner frequency, $\mathrm{M}_0$ is the seismic moment. A more general functional form includes the path attenuation effects (Shearer 2019):

$$A(f) = \mathrm{M}_0 \exp(-\pi f t^*)/[1 + (f/f_c)^{\gamma n}]^{1/\gamma} \quad (2)$$

Where $n$ represents the high-frequency fall-off rate and is associated with high-frequency energy radiation, and $\gamma$ controls the curvature of the spectrum near the corner frequency. In this form, $\gamma$=1 and n=2 for the $\omega^{-2}$ model; n=3 for the $\omega^{-3}$ model; n=2 and $\gamma$=2 for the Boatwright model (Boatwright 1978).

The Brune model ($\omega^{-2}$ model) is widely used to infer source parameters and generally agree with global average source spectra across different magnitudes (Abercrombie 2015; Ye *et al.* 2016a, b; Shearer *et al.* 2019). However, Kaneko & Shearer (2014) considered a two degrees of freedom model:

$$A(f) = \mathrm{M}_0/[1 + (f/f_c)^n] \quad (3)$$

They obtained $n$ = 2.2 to 2.9 for P-wave spectra and $n$ =1.5 to 2.6 for S-wave spectra in the

**Deleted:** falloff

**Deleted:** spreads

**Deleted:** falloff

**Formatted:** Font: Italic

**Deleted:** Numerous studies have shown that the

**Formatted:** Font: Italic

**Formatted:** Font: Italic

frequency range of $0.05f_c \leq f \leq 10f_c$ in the cohesive-zone models of circular rupture. Uchide & Imanishi (2016) observed that the source spectra of most $M_w$ 3.2-4.0 shallow earthquakes in Japan deviate from the Brune model and implies a fall-off exponent slightly higher than 2.

Models with double corner frequencies have also been proposed to fit complex source spectra. Denolle & Shearer (2016) suggested a double-$f_c$ model to explain the source spectra of large subduction earthquakes (referred to as DS16 model hereafter) using the functional form:

$$A(f) = \mathrm{M}_0/[\sqrt{1+(f/f_{c1})^2}\sqrt{1+(f/f_{c2})^2}] \quad (4)$$

Ji & Archuleta (2021) introduced two double-$f_c$ models (one is self-similar, another is not), which can reproduce the mean peak ground acceleration (PGA) and mean peak ground velocity (PGV) for magnitudes 3.3–7.3. They reported that the magnitude dependencies of PGA and PGV are well explained by the nonself-similar double-$f_c$ source model (referred to as JA21 model hereafter).

While the aforementioned physical models provide a generally good fit for source spectra, deviations may exist for specific earthquakes owing to various assumptions and calculations of source spectra made on the rupture process. Comparatively, machine learning allows extracting the hidden features of large datasets which could be otherwise invisible through physical models (Bergen *et al.* 2019). In this study, we employ a generative machine learning algorithm, called variational autoencoder (VAE), to learn data-driven models of source spectra for large earthquakes. We aim to answer a question: whether or not the source spectra of large earthquakes are the Brune-type on global average? Alternatively, are any systematic source characteristics not captured by the Brune-type model?

This paper is organized as follows. First, we transform source time functions to source spectra. The spectra are normalized and fit with the Brune model to obtain the corner frequencies. Second, synthetic tests are used to prove that VAE can derive meaningful average spectra from a large dataset. Third, VAE is applied to learn the real source spectra of global earthquakes. We compare the single- and double-variable VAE models to the Brune

**Deleted:** falloff

**Deleted:** dependence

**Deleted:** is

**Deleted:** Our study

and other more complex models and discuss the consistency and discrepancy among them. Finally, we explore the possible physical implications of source spectra derived from the data and the potential limitations.

## 2 Data

Source spectra of global earthquakes used in this study are amplitude spectra of source time functions (STFs). We use the SCARDEC (Seismic source ChActeristics Retrieved from DEConvolvolving teleseismic body waves) (Vallée & Douet 2016) dataset, which consists of 3675 large earthquakes with $M_w > 5.5$ from January 1, 1992 to December 31, 2019. Vallée *et al.* (2011) deconvolved teleseismic waves with the Green's functions and take local surface reflections for both source and receiver crusts as well as mantle attenuation (geometrical spreading and attenuation factors) into account. In addition, the SCARDEC method does not assume the same STF at each station and imposes no constraints on the spatial-temporal complexity of the rupture process. Overall, SCARDEC is advantageous in its full automation and thus provides a larger dataset for machine learning in this study. SCAREDEC provides two types of STFs, the optimal one among all stations and the average one of all stations. In this study we use the average STFs because they are relatively insensitive to possible outliers and generally more robust than the optimal ones. Moreover, it is less impacted by directivity effects that may exist on specific stations (Vallée & Douet 2016).

The STFs are first filled with zeros at the end of the STFs to keep the same time length (and hence the same frequency range and interval after Fourier transform). Each spectrum is normalized by seismic moment to only keep the shape. It is resampled evenly on the logarithmic scale to 128 points, which is the input size of the machine learning model. In this study, we reserve the frequency range up to 1 Hz to avoid the poorly constrained high-frequency contents in STFs, which follows the presumption that the global models of attenuation are better constrained by seismic frequencies lower than 1 Hz (Danré *et al.* 2019; Denolle 2019; Yin *et al.* 2021). However, whether or not the frequency content below 1 Hz is

**Deleted:** ,

**Deleted:**

**Deleted:** toke

**Deleted:** (t*)

**Deleted:** We reserve the frequency range to 1 Hz maximum to avoid the poorly constrained high-frequency contents in STFs. It is also a common practice as the global models of attenuation are better constrained by seismic frequencies lower than 1 Hz (Danré *et al.* …

**Moved down [1]:** 2019; Denolle 2019; Yin *et al.* 2021).

**Moved (insertion) [1]**

absolutely reliable remains arguable, which may have negative impact on our analyses and will be discussed in the last section.

## 3 Method

Generative modeling is a category of machine learning approaches to model the data implicitly (i.e. without clear physical meaning). It learns a probabilistic model that describes how a dataset is generated from input. With generative modeling, we could generate new plausible examples like a physical model by sampling probabilistic model. Generative adversarial networks (GANs) are a typical generative model and have been applied to earthquake early warning and seismic data augmentation (Li *et al.* 2018; Wang *et al.* 2021). However, GANs often suffers from unstable training (Salimans et al., 2016), as they do not have explicit constraints on the probability distribution. Comparatively, VAE, which is also a generative model, uses explicit constraints on the probability distribution and is able to reconstruct high-dimensional data from a compact latent representation (Kingma & Welling 2013). Its functionality acts similarly to physical modeling (case-dependent parameters embedded in a shared functional form), and could be used to learn a model for earthquake source spectra directly from data.

A typical VAE model is composed of two parts, i.e. an encoder and a decoder. The encoder compresses the input to a compact latent representation, whereas the decoder reconstructs the input from the latent representation. The bottleneck structure forces the model to learn the primary features of the data. The VAE's loss function in this analysis is defined as:

$$loss = ||\frac{S_{model} - S_{raw}}{freq}||_2 + KL[\mathbb{N}(\mu_z, \sigma_z) - \mathbb{N}(0,1)]$$

where the first term is the root mean square between the reconstructed and original source spectra, and the second is the Kullback-Leibler divergence which measures the difference between latent variables and normal distribution. After training, the VAE acts in a manner similar to the physical Brune model: for a given source spectrum, the encoder obtains the latent parameter (like the $f_c$ of the Brune model) and the decoder reconstructs the source spectra (like the fit curve of the Brune model). The difference is that VAE model is data-driven (from data directly) and the model parameter is implicit, whereas the Brune model is physics-based (from theoretical assumptions of source process) and the model parameter has explicit physical meaning.

We construct the VAE architecture following Li (2022) (Fig. 1). The encoder and the decoder have two fully connected layers, and each layer has 256 neurons. We first use only one latent variable in VAE to keep the same degree of freedom as the Brune model (one-parameter $f_c$; amplitude is normalized). Hence, the latent variable of VAE acts like $f_c$ of the Brune model, and the trained parameters of the neural network acts like the functional form of the Brune model (Fig. 1). Following a similar procedure, we train another VAE model with two latent variables and compare it with recently proposed double-$f_c$ models (DS16 and JA21). We randomly split 80% of the data as training set and 20% as testing set in both synthetic tests and the SCARDEC dataset.

**Deleted:** as follows

## 4 Results

### *4.1 Validation of Machine learning with synthetic source spectra*

We follow the procedure in Shearer *et al.* (2006) to fit the SCARDEC spectra with the Brune model by least square of the log spectrum. We estimate the $f_c$ for all earthquakes and use them to generate synthetic data for machine learning tests.

**Deleted:** obtain

**Deleted:** estimates of

**Deleted:** which are then used

First, to demonstrate the modeling capability of VAE, we apply VAE to learn the Brune model from noise-free synthetic data. We generate 3675 synthetic source spectra using eq. (1) derived from SCARDEC. Fig. 2a shows that an example from test data for which VAE has perfect fit. The misfits ($||\frac{S_{model}-S_{raw}}{freq}||_2$) for most of these spectra are almost negligible (Fig. 2b). This demonstrates that VAE correctly learns the Brune model from noise-free spectra.

Second, to demonstrate that VAE can learn an average model from noisy data, we add random Gaussian perturbation to the fall-off exponent (i.e., $n = 2 \pm \delta$, $\delta \in \mathbb{N}(0, 0.2)$) of the synthetic source spectra in the previous test. Then the average model of all the synthetic data still follows the Brune model, but individual spectra randomly deviate. The training result shows that the VAE model also agrees with the Brune model (Figs. 2c and d). This demonstrates that VAE can derive an average model from a large dataset, even if the data have random deviation individually.

*4.2 Single-variable modeling of real source spectra*

On the basis of the previous tests, we train a single-variable VAE model (referred to as VAE1 model hereafter) with the source spectra from SCARDEC. The results show that, VAE1 generally provides better fit than the Brune model and exhibits some different characteristics (Fig. 2e). Fig. 2f shows the histograms of misfits of two models as well as the differences between them. Overall, the mean misfit of VAE1 is about 8.5% smaller than that of the Brune model. This small difference suggests that two models are largely consistent. Moreover, we observe that the latent variable of VAE1 is strongly correlated with $f_c$ (Fig. 3a). The latent variable also has a similar relationship with $M_w$ (Fig. 3b) because of the inherent scaling relationship between $f_c$ and $M_w$. Since machine learning searches in a much wider parameter space than Brune model, the high correlation with $f_c$ demonstrates that $f_c$ is indeed an effective parameter controlling the spectral shapes. However, with the same degree of freedom, the lower misfit of the data-driven model suggests that the Brune model misses some systematic characteristics of the observed data.

Fig. 4a shows the overall variation of the VAE1 model spectra, the Brune-type source spectra and average real source spectra, with respect to the VAE1 latent parameter. Specifically, as the latent variable value increases, $f_c$ decreases and magnitude increases. It is noteworthy that the reliability of the VAE curves depends on the number of available real data. Therefore, VAE1 provides the most reliable results approximately for $M_w$ 6-7 because of the data abundance in that range. To reveal the differences between VAE1 and the Brune model, we subtract the average real data from their fitting curves (Fig. 4b). The residuals suggest that the VAE1 spectra are more consistent with the observed data across different magnitudes. This suggests that VAE1 has learned an unbiased average model from the data and can serve as a baseline for other physical source models. In comparison, the residuals show that the Brune model systematically underestimates in intermediate frequencies and overestimates in high frequencies for the SCARDEC dataset (Fig. 4b).

*4.3 Double-variable modeling of real source spectra*

For comparison with double-$f_c$ models (DS16, JA21) recently proposed to supplement the Brune model (single-$f_c$) , we train a double-variable VAE model (referred to as VAE2 model hereafter) which also has two degrees of freedom. The VAE2 have two latent variables $Z_1$ and $Z_2$, compared to a single variable Z in the VAE1 model. Fig. 5a shows an example of SCARDEC source spectra fitted by VAE2 and double-$f_c$ models. The median residual of the VAE2 model is near zero, whereas DS16 and JA21 have some deviations (Fig. 5b), similar to the Brune model. We observe that 95-percentile of the VAE2, DS16, JA21 misfits are 0.0065, 0.0088, 0.0278 respectively (Fig. 5c and Table 1). In comparison, VAE1 has 95-percentile misfit at 0.0115 (Table 1). We estimate the statistical relative amount of information loss of these models with Akaike information criterion (AIC), which deals with the trade-off between the goodness of fit and the complexity of the model:

$$\mathrm{AIC} = 2\mathrm{k} + \mathrm{N}ln\left(\frac{\mathrm{RSS}}{\mathrm{N}}\right)$$

Where N is the number of frequency samples, k is the number of estimated parameters, RSS is the residual sum of squares. Note that this definition assumes normally distributed errors. For single-variable models, the parameters are the latent variable or $f_c$ and residual variance so that k = 2. Double-variable models have k = 3. Compared to other models, VAE2 provides improved goodness of fit (Table 1) and captures more detailed features of the source spectra, especially for the curvature at turning corner (Fig. 6).

To investigate the effect of the additional latent variable, we visualize the variations of source spectra in the latent space (Fig. 7a). We observe that the source spectra change more significantly with $Z_2$ than with $Z_1$, indicating a primary effect of $Z_2$ and a secondary effect of $Z_1$. Moreover, the correlation between $Z_1$ and $f_c$ acts more subtle (Fig. 7b), whereas $Z_2$ appears (Fig. 7c) similar to that in VAE1(Fig. 3a), . Therefore, these likely suggests that the role of $Z_2$ in VAE2 is comparable to that in single-variable VAE model, whereas $Z_1$ could catch more secondary details to promote the goodness of fit. Generally, $Z_2$ seems to control the corner frequency like the only variable in the VAE1, whereas $Z_1$ seems to control the abruptness of transition from the low-frequency plateau to the high-frequency fall-off.

## 5 Discussions

### *5.1 Physical implications of the VAE models*

To clarify the causes of the systematic characteristics not captured by the Brune and other physical models, we explore different model parameters and attenuation effect to see if the difference can be reduced to near zero across the frequency range as the VAEs do. First, we experiment different values of high-frequency fall-off rate in eq. (3). Although the fit improves slightly when the high-frequency fall-off rate around 2.3, the mean differential curve cannot be reduced to be flat by simply tuning the fall-off rate (Fig. 8a). Second, we tune the parameters $\gamma$ and high-frequency fall-off rate in Boatwright model, but find it leads to even higher misfit than the Brune model (Fig. 8b). Third, SCARDEC uses an attenuation model to correct the attenuation effect on the spectra. Although there could be attenuation effect not fully corrected, the apparent slopes in high- and low-frequency ranges observed in this study appear too large to be explained by the remaining attenuation effect. Since DS16 and JA21 cannot adequately explain the observed slow fall-off in the intermediate frequencies and the fast fall-off in the high frequencies, we propose a modified double-$f_c$ model to simulate the characteristics of the real data revealed from VAE models (Fig. 8d). This model is similar to DS16 and JA21 but has the $f^{-1}$ fall-off rate in the intermediate frequencies and has a $f^{-2.6}$ slope in the high-frequency region. We find that this combination can generally replicate the major shape. Nonetheless, we can only constrain the first exponent to be <2 and the second exponent to be >2; the actual combination of them can vary.

In the Haskell model, the presence of two corner frequencies results from that the slip risetime is much less than rupture duration time ($\tau_r << \tau_d$). This short risetime phenomenon has been shown by dynamic rupture modeling results (*Beroza & Mikumo 1996; Melgar & Hayes 2017; Wang & Day 2017*) and can be caused by several mechanisms. For example,



**Formatted:** Automatically adjust right indent when grid is defined, Pattern: Clear (White)

**Deleted:** does

**Deleted:** falloff

**Deleted:** falloff

**Deleted:** falloff

**Deleted:** falloff

**Deleted:** we add

**Deleted:** intrinsic

**Deleted:** term (eq. 2) to the Brune

**Deleted:** but observe that

**Deleted:** curve deviates more

**Deleted:** even though it fits better

**Deleted:** the high-frequency (Fig. 8c). This is because the attenuation parameter $t^*$ is a monotonic function of frequency, and changing the parameter $t^*$ cannot improve low- and high-frequency simultaneously…

**Deleted:** compared to the regular double-$f_c$ models.

**Deleted:** the

**Deleted:** of exponents -1 and -2.6

**Deleted:** the constraints are

**Deleted:** , but

**Deleted:** may

**Formatted:** Font: Bold, Font color: Black

**Formatted**

**Formatted:** Font: Italic

**Formatted:** Font: Italic

**Formatted:** Font: Italic

**Formatted:** Heading 1 Char, Font: Italic

**Deleted:** ). Two general classes of

**Deleted:** have been proposed to explain it. One postulates that the short risetime is caused by dynamic fault friction, which decreases with increasing slip velocity (Heaton 1990). The other is…

Das & Aki (1977) suggested that spatially heterogeneous fault strength (e.g., barriers) may limit slip duration at particular locations on a fault. Heaton (1990) postulated that short risetime can be caused by dynamic fault friction, which decreases with increasing slip velocity. Dynamic changes of normal stress induced by bi-material effects (Andrews & Ben-Zion 1997; Ampuero & Ben-Zion 2008) and low-velocity fault zones (Huang & Ampuero 2011) could also generate short-risetime slip pulses. These explanations are indeed indistinguishable from our observations. However, it needs to be noted that the slow fall-off rate (<2) in intermediate frequencies is introduced because of the finite-length long-narrow faults (Haskell 1964; Savage 1972); the higher fall-off rate (>2) in high frequencies implies that the rupture processes of large earthquakes may not be strictly self-similar but prefer elongated fault geometries (Shearer 2019). Our hypothesized model is purely data-driven and precise interpretation of its physical meaning remains a subject of future research.

*5.2 Limitations of the STF data and their impact*

Our conclusions largely rely on the presumption that the Fourier transform of the SCARDEC STFs below 1Hz are good representations of source spectra. However, several factors may undermine this presumption and thus have negative impact on our conclusions. Although Vallée & Douet (2016) deconvolved the teleseismic waves by Green's functions which are better constrained at frequencies below 1 Hz, this does not directly support that the frequencies below 1 Hz are perfectly reliable. Besides, the averaging of STFs and time-domain deconvolution tend to further reduce the high frequency content (Vallée & Douet 2016), which may lead to overestimation in the high-frequency fall-off rate. Therefore, one should be aware that the actual appropriate cutoff frequency remains difficult to determine, although the 1-Hz cutoff is used in this study.

Moreover, Danré *et al.* (2019) showed that the STFs in SCARDEC can be fit by a sum of Gaussian pulses, implying the source spectra calculated from this dataset are inherently different from the Brune model. Besides, the calculation of STFs suffers a number of intricate issues such as the assumed Green's function and the influence of attenuation. The SCARDEC

**Deleted:** the
**Deleted:** the fault (Das & Aki 1977).
**Deleted:** two
**Deleted:** need
**Deleted:** falloff
**Deleted:**
**Deleted:** ), and
**Deleted:** falloff
**Deleted:**
**Deleted:** to
**Deleted:** *the VAE method*
**Deleted:** For example
**Formatted:** First line: 1.75 ch

method (Vallée et al. 2011) makes a point source approximation and deconvolves seismic waveforms with assumed Green's function derived from the method of Bouchon (1976) to separate the source term directly. Vallée and Douet (2016) noted that it is difficult to well separate real source effects from spurious moment episodes related to unmodeled seismic phases. In the cases like offshore strike-slip events, long and complex STFs can be questionable. However, Yin et al. (2021) observed that colocated shallow events have distinct degrees of complexity, and therefore inaccuracy in the Green's function should not strongly systematically bias the results.

### *5.3 Potentials of the VAE method*

The VAE approach optimizes global fit and derives the general features directly from real data. It can infer complex, nonlinear and high-dimensional data relationships, and obtain a data-driven model without any prior assumption and human supervision. Although the latent parameters and the functional form to model source spectra are implicit and difficult to interpret, the data-driven model serves as a quasi-optimal baseline for which physical models need to approach. Any systematic shift of physical models from the data-driven model in the same degree of freedom likely indicates some inherent flaws within the physical models (another possibility is that the dataset itself is already biased). In this perspective, our approach has the potential to reveal hidden characteristics of large and high-dimensional seismological data and provide opportunities to amend existing theoretical frameworks.

**Acknowledgements**

The authors are grateful to Martin Vallée for providing the source time function databases. Zefeng Li thank Jiuxun Yin and Yifang Cheng for discussion on the preliminary idea of using machine learning to learn source spectra shapes. This research was supported by the National Key R&D Program of China (No. 2021YFC3000700).

**Data Availability**

The SCARDEC database is available at http://scardec.projects.sismo.ipgp.fr/ (last accessed on September 3, 2021).

**References**

Abercrombie, R.E. (2015) Investigating uncertainties in empirical Green's function analysis of earthquake source parameters. *J. Geophys. Res. Solid Earth*, **120**, 4263–4277. doi:10.1002/ 2015JB011984

Aki, K. (1967) Scaling Law of Seismic Spectrum. *J. Geophys. Res.*, **72**, 1217–1231. doi:10.1029/JZ072i004p01217

Ampuero, J.-P. & Ben-Zion, Y. (2008) Cracks, pulses and macroscopic asymmetry of dynamic rupture on a bimaterial interface with velocity-weakening friction. *Geophys. J. Int.*, **173**, 674–692. doi:10.1111/j.1365-246X.2008.03736.x

Andrews, D.J. & Ben-Zion, Y. (1997) Wrinkle-like slip pulse on a fault between different materials. *J. Geophys. Res. Solid Earth*, **102**, 553–571. doi:10.1029/96JB02856

Moved (insertion) [2]

Bergen, K.J., Johnson, P.A., Hoop, M.V. de & Beroza, G.C. (2019) Machine learning for data-driven discovery in solid Earth geoscience. *Science*, **363**, eaau0323. doi:10.1126/science.aau0323

Beroza, G.C. & Mikumo, T. (1996) Short slip duration in dynamic rupture in the presence of heterogeneous fault properties. *J. Geophys. Res. Solid Earth*, **101**, 22449–22460. doi:10.1029/96JB02291

Boatwright, J. (1978) Detailed spectral analysis of two small New York state earthquakes. *Bull. Seismol. Soc. Am.*, **68**, 1117–1131.

Bouchon, M. (1976) Teleseismic body wave radiation from a seismic source in a layered medium. *Geophys. J. Int.*, **47**, 515–530.

Brune, J.N. (1970) Tectonic Stress and the Spectra of Seismic Shear Waves from Earthquakes. *J. Geophys. Res.*, **75**, 4997–5009. doi:10.1029/JB075i026p04997

Chounet, A., Vallée, M., Causse, M. & Mathieu, F. (2018) Global catalog of earthquake rupture velocities shows anticorrelation between stress drop and rupture velocity. *Tectonophysics*, **733**, 148–158.

Danré, P., Yin, J., Lipovsky, B.P. & Denolle, M.A. (2019) Earthquakes Within Earthquakes: Patterns in Rupture Complexity. *Geophys. Res. Lett.*, **46**, 7352–7360. doi:10.1029/2019GL083093\

Das, S. & Aki, K. (1977) Fault plane with barriers: A versatile earthquake model. *J. Geophys. Res. 1896-1977*, **82**, 5658–5670. doi:10.1029/JB082i036p05658

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

Denolle, M.A. (2019) Energetic Onset of Earthquakes. *Geophys. Res. Lett.*, **46**, 2458–2466. doi:10.1029/2018GL080687

Haskell, N. (1964) Total energy and energy spectra density of elastic waves from propagating faults. *Bull. Seismol. Soc. Am.*, **54**, 1811–1841.

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

Huang, Y. & Ampuero, J.-P. (2011) Pulse-like ruptures induced by low-velocity fault zones. *J. Geophys. Res.*, **116**, B12307. doi:10.1029/2011JB008684

Ji, C. & Archuleta, R.J. (2021) Two Empirical Double-Corner-Frequency Source Spectra and Their Physical Implications. *Bull. Seismol. Soc. Am.*, **111**, 737–761. doi:10.1785/0120200238

Kaneko, Y. & Shearer, P.M. (2014) Seismic source spectra and estimated stress drop derived from cohesivezone models of circular subshear rupture. *Geophys. J. Int.*, **197**, 1002–1015. doi:https://doi.org/10.1093/gji/ggu030

Kingma, D.P. & Welling, M. (2013) Auto-Encoding Variational Bayes. *ArXiv E-Prints*, arXiv:1312.6114.

Li, Z. (2022). A generic model of global earthquake rupture characteristics revealed by machine learning. *Geophysical Research Letters*, 49, e2021GL096464. https://doi.org/10.1029/2021GL096464

Li, Z., Meier, M.-A., Hauksson, E., Zhan, Z., & Andrews, J. (2018). Machine Learning Seismic Wave Discrimination: Application to Earthquake Early Warning. *Geophysical Research Letters*, 45(10), 4773–4779. https://doi.org/10.1029/2018GL077870

Melgar, D. & Hayes, G.P. (2017) Systematic Observations of the Slip Pulse Properties of Large Earthquake Ruptures. *Geophys. Res. Lett.*, **44**, 9691–9698. doi:10.1002/2017GL074916

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X. & Chen, X. (2016) Improved Techniques for Training GANs. *Adv. Neural Inf. Process. Syst.* eds. Lee, D., Sugiyama, M., Luxburg, U., Guyon, I. & Garnett, R., Vol. 29, Curran Associates, Inc.

Savage, J.C. (1972) Relation of corner frequency to fault dimensions. *J. Geophys. Res. 1896-1977*, **77**, 3788–3795. doi:10.1029/JB077i020p03788

Shearer, P.M. (2019) *Introduction To Seismology*, 3rd ed., Cambridge: Cambridge University Press.

Shearer, P.M., Abercrombie, R.E., Trugman, D.T. & Wang, W. (2019) Comparing EGF Methods for Estimating Corner Frequency and Stress Drop From P Wave Spectra. *J. Geophys. Res. Solid Earth*, **124**, 3966–3986. doi:https://doi.org/10.1029/2018JB016957

Shearer, P.M., Prieto, G.A. & Hauksson, E. (2006) Comprehensive analysis of earthquake source spectra insouthern California. *J. Geophys. Res. Solid Earth*, **111**, B06303.

**Deleted:** Z rly

**Formatted:** Bibliography, Justified, Indent: Left: 0 cm, First line: 0 ch, Line spacing: single

**Moved up [2]:** *Geophys. Res.*

**Deleted:** *Lett.*,

**Formatted:** Font: Not Bold

**Deleted:** ,

**Deleted:** :

Uchide, T. & Imanishi, K. (2016) Small earthquakes deviate from the omega-square model as revealed by multiple spectral ratio analysis. *Bull. Seismol. Soc. Am.*, **106**, 1357–1363. doi:https://doi.org/10.1785/0120150322

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. *Geophys. J. Int.*, **184**, 338–358.

Vallée, M. & Douet, V. (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

Wang, T., Trugman, D. & Lin, Y. (2021) SeismoGen: Seismic Waveform Synthesis Using GAN With Application to Seismic Data Augmentation. *J. Geophys. Res. Solid Earth*, **126**, e2020JB020077. doi:10.1029/2020JB020077

Wang, Y. & Day, S.M. (2017) Seismic source spectral properties of crack-like and pulse-like modes of dynamic rupture. *J. Geophys. Res. Solid Earth*, **122**, 6657–6684. doi:10.1002/2017JB014454

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

Ye, L., Lay, T., Kanamori, H. & Rivera, L. (2016b) Rupture characteristics of major and great ($M_w \geq 7.0$) megathrust earthquakes from 1990 to 2015: 2. Depth dependence. *J. Geophys. Res. Solid Earth*, **121**, 845–863.

Yin, J., Li, Z. & Denolle, M.A. (2021) Source Time Function Clustering Reveals Patterns in Earthquake Dynamics. *Seismol. Res. Lett.*, **92**, 2343–2353. doi:10.1785/0220200403

485 **Tables:**

486

487 **Table 1.** Performance of single- and double-variable models

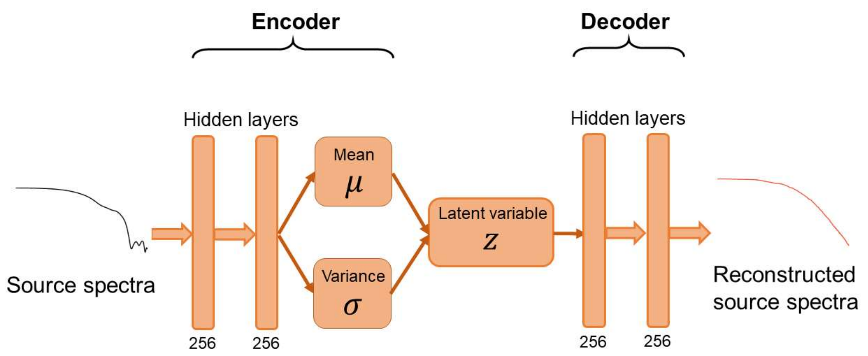
| | Single-variable | | Double-variable | | |
|---|---|---|---|---|---|
| | VAE1 | Brune | VAE2 | DS16 | JA21 |
| Median misfit | 0.0013 | 0.0015 | 0.0008 | 0.0014 | 0.0056 |
| 95% misfit | 0.0115 | 0.0117 | 0.0065 | 0.0088 | 0.0278 |
| Average AIC | -1064.2 | -1029.6 | -1100.2 | -1033.8 | -843.5 |

Formatted Table

488

489

490 **Figures:**




491 **Figure 1.** The architecture of variational autoencoder (VAE). Both of the decoder and
492 encoder consist of two fully connected layers with 256 neurons each. The lengths of input
493 (real source spectra) and output (reconstructed source spectra) are 128 data points. Note that
494 we limit the latent dimension to 1 (for the single-variable model) and 2 (for the double-
495 variable model) so that each spectrum can be modelled in a low degree of freedom.

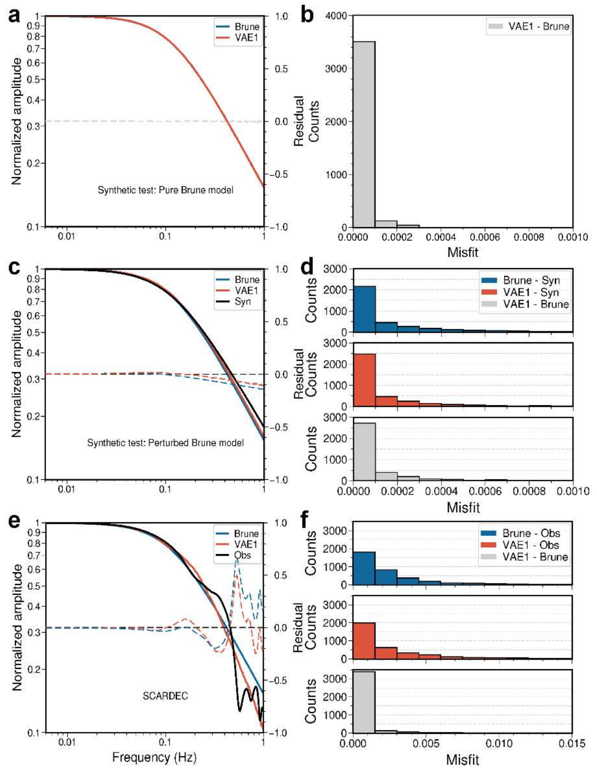



**Figure 2.** Goodness of fit of the single-variable VAE and Brune models for earthquake source spectra. (a) A noise-free synthetic source spectrum example fit by the VAE model (solid red) and the Brune model (solid blue), which overlap in the plot. The grey dashed line denotes the residual (right y-axis). (b) Misfit histogram of the VAE model for the noise-free synthetic spectra. (c) A random-perturbed Brune-type synthetic source spectrum fit by the VAE model (solid red) and the Brune model (solid blue). The blue and red dashed lines denote the residuals of the VAE model and the Brune model (right y-axis). (d) Misfit histograms of the VAE and Brune models and differences between them for the random-perturbed Brune-type synthetic source spectra. (e) A SCARDEC source spectrum example fit by the VAE model (solid red) and the Brune model (solid blue) (f) Misfit histograms of the VAE and Brune models for the SCARDEC source spectra.

**Deleted:** [redacted]

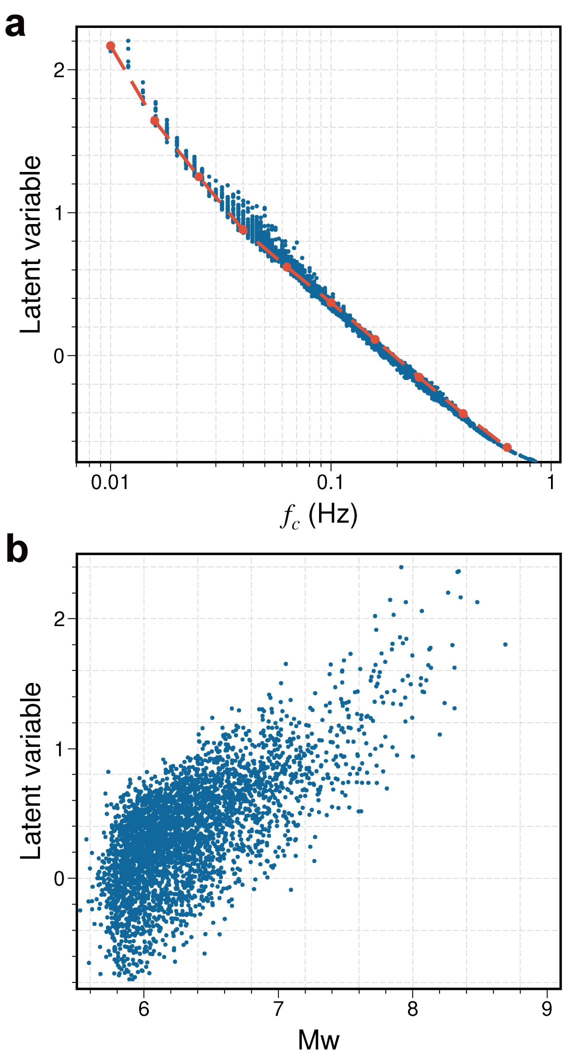



**Figure 3.** Variations of the single VAE latent variable with the Brune model parameter $f_c$ and the earthquake magnitude. (a) Strong correlation between the VAE latent variable and the Brune model parameter $f_c$. The red dotted line is the median latent value for each $f_c$. (b) Correlation between the VAE latent variable and earthquake magnitude.

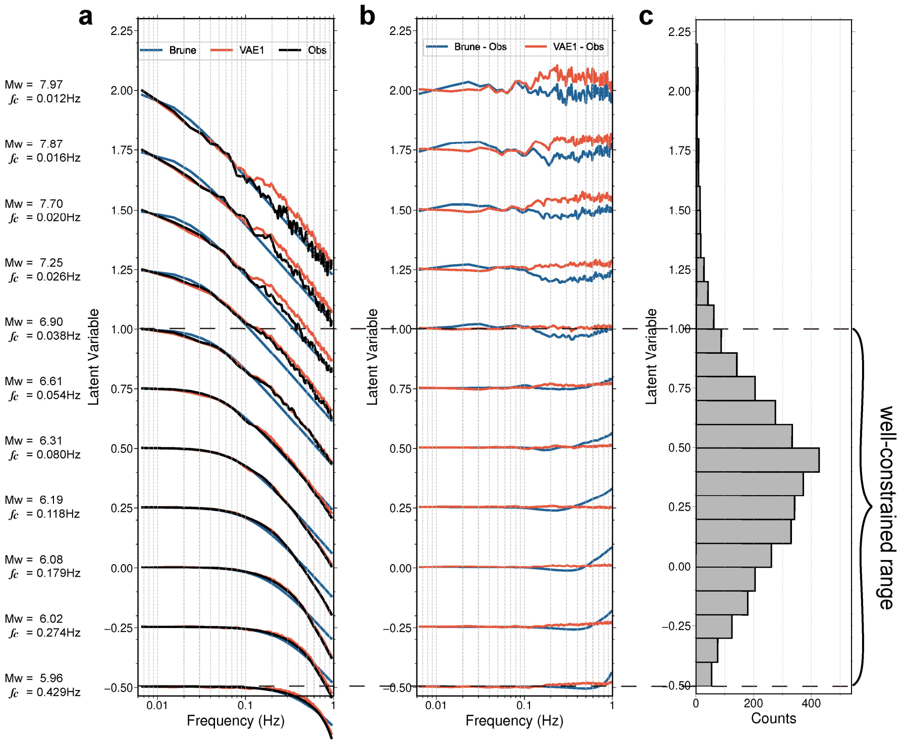



**Figure 4.** Systematic variation of source spectra in the latent space of the single-variable VAE. (a) The VAE spectra, Brune spectra and SCARDEC magnitude-binned average source spectra, with respect to the VAE latent parameter. (b) Residuals of the VAE and Brune models by substracting the SCARDEC data. (c) Histogram of latent variable values. The VAE model is well constrained within M 6-7 due to the abundant observations.

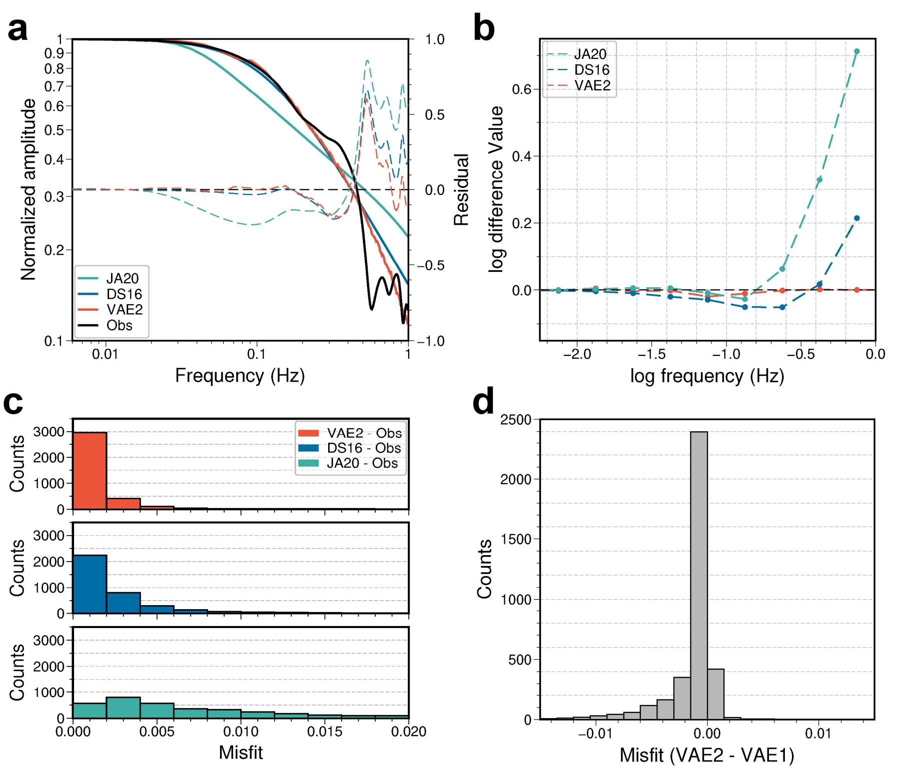



**Figure 5.** Goodness of fit of the double-variable VAE and double-$f_c$ physical models (JA21, DS16). (a) A SCARDEC source spectrum example fit by the VAE model (solid red), DS16 model (solid blue) and JA21 model (solid green). The red, blue and green dashed lines denote the residuals of the VAE model, DS16 model and JA21 model (right y-axis). (b) Median residuals of the double-variable VAE model and double-$f_c$ physical models (JA21, DS16) . (c) Misfit histograms of the VAE, DS16 and JA21 models for the SCARDEC source spectra. (d) Histogram of misfit residuals between the double-variable and single-variable VAE models.

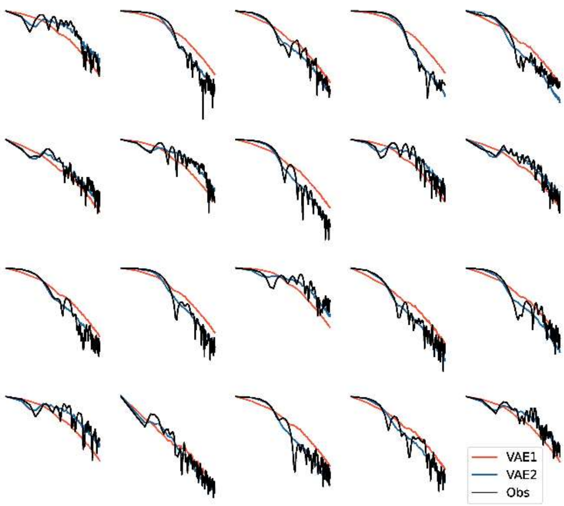



**Figure 6.** Examples of SCARDEC source spectra (black) and their reconstructions from single-variable (red) and double-variable (blue) VAE models.

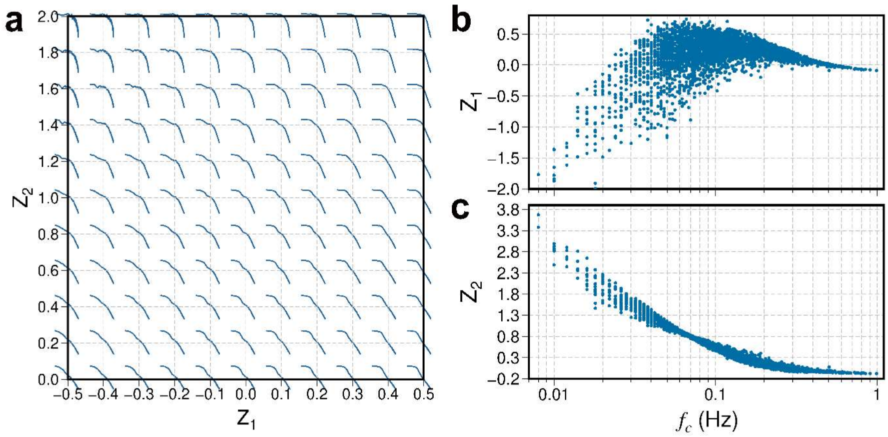



**Figure 7**. (a) Manifold of source spectra reconstructed from evenly sampled latent variables of the double-variable VAE model. (b) Correlation between the latent variable $Z_1$ and the Brune model parameter $f_c$. (c) Correlation between the latent variable $Z_2$ and the Brune model parameter $f_c$.

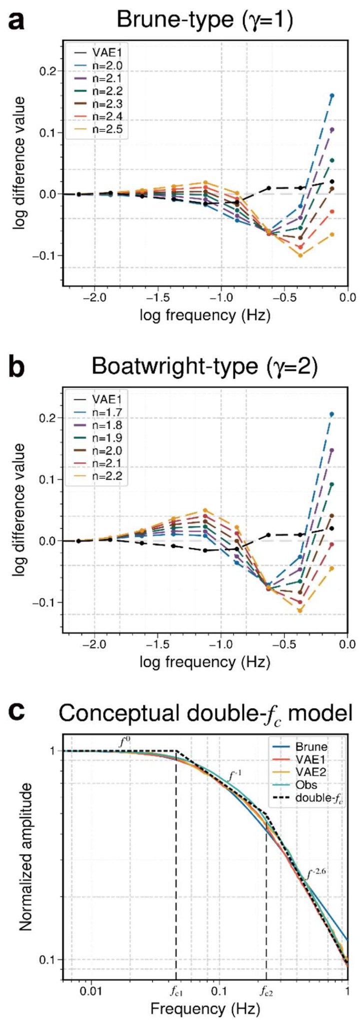



**Figure 8.** The effect of model and attenuation parameters on fitting SCARDEC source spectra. (a) Median residuals between different Brune-type models ($\gamma$=1, n=2.0-2.5) and the observed source spectra. (b) Median residuals between Boatwright-type models ($\gamma$=2, n=1.7-2.2) and the observed source spectra. (c) Average source spectra for Brune model, VAE1, VAE2, and observations for $M_w$ 6-7. The VAE1 and VAE2 curves are obtained by taking the median of all fitting curves for the SCARDEC events within $M_w$ 6-7. Dashed black line represents our proposed double-$f_c$ model.

**Deleted:** [redacted]

**Formatted:** Left

**Deleted:** Median residuals between

**Deleted:** with intrinsic attenuation

**Deleted:** the observed source spectra. (d) Modeling

**Deleted:** and

**Deleted:** observed source spectra

**Formatted:** Subscript

**Formatted:** Font: Bold