New perspectives on self-similarity for shallow thrust earthquakes

Citation

Published Version
doi:10.1002/2016JB013105

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:34712309

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story.

Accessibility
New perspectives on self-similarity for shallow thrust earthquakes

Marine A. Denolle and Peter M. Shearer

Abstract Scaling of dynamic rupture processes from small to large earthquakes is critical to seismic hazard assessment. Large subduction earthquakes are typically remote, and we mostly rely on teleseismic body waves to extract information on their slip rate functions. We estimate the $P$ wave source spectra of 942 thrust earthquakes of magnitude $M_w$ 5.5 and above by carefully removing wave propagation effects (geometrical spreading, attenuation, and free surface effects). The conventional spectral model of a single-corner frequency and high-frequency falloff rate does not explain our data, and we instead introduce a double-corner-frequency model, modified from the Haskell propagating source model, with an intermediate falloff of $f^{-1}$. The first corner frequency $f_1$ relates closely to the source duration $T_1$, its scaling follows $M_0 \propto T_1^3$ for $M_w < 7.5$, and changes to $M_0 \propto T_2^5$ for larger earthquakes. An elliptical rupture geometry better explains the observed scaling than circular crack models. The second time scale $T_2$ varies more weakly with moment, $M_0 \propto T_2^5$, varies weakly with depth, and can be interpreted either as expressions of starting and stopping phases, as a pulse-like rupture, or a dynamic weakening process. Estimated stress drops and scaled energy (ratio of radiated energy over seismic moment) are both invariant with seismic moment. However, the observed earthquakes are not self-similar because their source geometry and spectral shapes vary with earthquake size. We find and map global variations of these source parameters.

1. Introduction

How earthquake dynamics scale with event size is a fundamental issue in seismology; it is a key question in seismic hazard assessment yet has not achieved scientific consensus. Are rupture processes of small earthquakes similar to those of large earthquakes? Aki [1967] introduced the concept of self-similarity among earthquakes: if the source spectrum follows a simple $\omega^{-2}$ decay at high frequencies and a single-corner frequency describes the source duration $T$, then scaling of the earthquake size, or moment $M_0$, with duration follows $M_0 \propto T_1^3$, which implies that stress drop is independent of earthquake size. If the shape of the spectrum (in a log(frequency) versus log(amplitude) plot) is also preserved [Aki, 1967; Prieto et al., 2004], then the ratio of earthquake radiated energy $E_r$ over moment, or scaled energy $E_r/M_0$, is also independent of earthquake size. While self-similarity implies that stress drop and scaled energy are independent of moment, we will show that the converse is not necessarily true.

Static measures of earthquakes provide the source dimension and total slip distribution on the fault and can be measured most directly using field or geodetic observations in the near field. Far-field observations of seismic waves at long periods generally provide good constraints on the seismic moment, $M_0 = \mu DA$, where $\mu$ is the shear modulus, $D$ is the average slip on the fault, and $A$ is the fault area. When the fault area $A$ is known, we can estimate the average slip $D$ on the fault from $M_0$, and approximate knowledge of $\mu$. The average static stress drop $\Delta\sigma$ can also be estimated from the slip and source dimension.

However, the source dimension can be difficult to determine directly using teleseismic observations alone, and it is often inferred from the total rupture duration and an assumed rupture velocity. Estimates of source duration are hampered by signal-to-noise limitations and the often emergent nucleation and termination phases of the rupture. We can, however, roughly constrain the moment-rate release using seismic waves at teleseismic distances. Long-period fundamental-mode surface waves are used to construct smoothed source-time functions [Velasco et al., 1994]. Finite-fault inversions of short-period body waves target the evolution of the large slip pulses on the fault plane [Trifunac, 1974]. Because the onset of slip and changes in rupture
velocity generate high frequencies, backprojection of high-frequency teleseismic $P$ waves likely tracks the rupture propagation [Ishii et al., 2005], thereby also providing clues on the fault extent. For smaller earthquakes, we often rely on scaling relations between source duration and seismic moment [Aki, 1967; Houston, 2001]. A measure of source duration can be approximated from the source amplitude spectrum by estimating the source corner frequency where the low-frequency asymptote intersects with the high-frequency asymptote [Brune, 1970; Silver, 1983]. When the source duration is the only known static parameter, the fault is typically assumed circular. When finite source effects can be mapped (e.g., using backprojection or finite slip inversions or aftershock distributions), one can estimate the fault aspect ratio. Leonard [2010] provides a comprehensive overview on the scaling of the fault dimensions with earthquake size for canonical faulting styles.

Dynamic measures of earthquakes capture the amount of seismic energy that was excited, or spent, during the rupture itself. Whether earthquakes radiate more or less high-frequency seismic waves is critical to characterizing strong ground motion. Dynamic rupture processes, such as abrupt changes in rupture velocity [Madariaga, 1977; Boatwright, 1982], strength-stress evolution as the fault slips [Ida, 1972; Andrews, 1976], and rupture through barriers [Das and Aki, 1977], excite high-frequency seismic waves. For the same moment, a high-stress drop and/or high rupture velocity event radiates more high frequencies than a low-stress drop and/or a slow earthquake. At macroscopic scales, part of the energy released during an earthquake partitions into frictional energy (heat), fracture energy (energy spent to propagate the rupture), and energy spent radiating seismic waves (see review by Kanamori and Brodksy [2004]). Since most of the kinetic seismic energy is carried by high frequencies, teleseismic $P$ waves carry vital information on the earthquake energy budget.

There are two main approaches to estimate radiated energy. Boatwright [1980] directly measured the radiated energy flux by preserving the amplitude information in the waveforms, which, when integrated over enough stations (azimuth and takeoff angles), is a direct measure of radiated energy. In all studies that measure energy with this approach [Convers and Newman, 2011; Pérez-Campos and Beroza, 2001], the path effects are removed using a simplified Earth Green’s function with attenuation and geometrical spreading. However, because absolute amplitude calibration can be difficult to perform accurately, a second approach normalizes the seismogram (i.e., to the zero-frequency asymptote of its spectrum), assumes a consistent spectral shape over all azimuths and takeoff angles, and often uses an independent measure of seismic moment to calibrate the long-period spectral level. This second approach, referred to as the single-station approach in Denolle et al. [2015], has been widely used in studies that remove path effects using empirical Green’s functions [Mayeda and Walter, 1996; Baltay et al., 2010, 2014]. However, when the data are missing or unreliable at long periods below the corner frequency where the spectrum should become flat, normalization of the spectrum has to be done carefully. Extrapolation to low and high frequencies is common [Baltay et al., 2014], but here we propose instead an interpolation scheme that takes advantage of the best fitting source model.

Accurately estimating stress drops and radiated energy for a wide range of earthquake sizes remains a challenge. With fewer large events than small earthquakes, evaluation of these parameters is more robust for small earthquakes because greater numbers of events can be averaged. The data processing techniques are also different for $M_w < 2$ events with corner frequencies above ∼40 Hz [Abercrombie, 2014] and for recordings of $M_w > 8$ events where the direct $P$ waves, which carry the source pulse, become contaminated by depth phases and corner frequencies become lower than ∼0.005 Hz [Baltay et al., 2014]. Estimates of source duration (and hence stress drop) from corner frequencies are more common for small earthquakes, while those from finite-slip inversions are more common for large earthquakes. Comparing estimates of the same source parameters obtained from different data sets and methodologies may be misleading because they might reflect differences in underlying model assumptions rather than actual differences in source properties [Ide and Beroza, 2001; Allmann and Shearer, 2009].

Subduction zones are particularly active plate boundaries as they host the largest recorded earthquakes of $M_w > 9$ and their aftershocks. In this study, we focus on a single data set of $P$ waves from shallow (∼50 km) thrust earthquakes of magnitudes ranging from $M_w$ 5.5 to 9.2 from 1990 to mid-2015. We find that 942 of these events had high quality $P$ waves. For large and shallow earthquakes, estimates of source duration and spectral shape from teleseismic body waves are complicated by free-surface reflections. The upgoing $P$ to downgoing $P$ reflection ($pP$) and the upgoing $S$ to downgoing $P$ reflection ($sP$), commonly called depth phases, arrive immediately after direct $P$ and mask the end of the direct $P$ pulse. For large earthquakes, the difference in traveltine is less than the width of the source pulse, such that the direct and depth phases interfere. Boatwright and Choy [1986] developed a correction for the energy of the depth phases, but they ignore the
frequency dependence of the interference between the direct and depth phases. Other studies [Houston and Kanamori, 1986; Boatwright and Choy, 1989, 1992] used the amplitude spectrum of the P wave train to construct the P source spectrum, without discussing the impact on radiated energy. For most thrust earthquakes, the interference is destructive, with the energy of the $P + pP + sP$ phase smaller than the sum of the individual energies.

This study aims to provide a self-consistent approach to measuring static and dynamic properties of earthquakes for a wide range of magnitudes. We adopt a two-step approach to removing path effects: (1) a simplified Earth model that we combine with a depth-phase transfer function for the smaller earthquakes and (2) the use of nearby small earthquakes to calibrate 3-D paths for the larger earthquakes. Most commonly used source spectral shape models adopt a single time scale, manifested in a single-corner frequency $f_c$ with a flat asymptote for frequencies lower that $f_c$ and a high-frequency falloff varying from $f^{-1.6}$ to $f^{-3}$ [Aki, 1967; Brune, 1970; Boatwright, 1980; Ye et al., 2016a], $f^{-1.6}$ being chosen to bound the seismic energy. In contrast, the presence of a low falloff rate at intermediate frequencies in the spectra reflects two dominant time scales. Madariaga [1979] suggested that the source P pulse of thrust earthquakes recorded at steep take-off angles could exhibit a clear distinction between the starting/stopping phase pulse time scale and the total source duration time scale. Boatwright and Choy [1989] and Boatwright and Choy [1992] reconstructed source acceleration spectra from moderate and large earthquakes and found that source spectra of the larger ones exhibited an intermediate falloff slope of about $f^{-1.25}$. Our study confirms that a double-corner-frequency model (i.e., two time scales) better explains our observations than a single-corner frequency. Inspired by the Haskell source model of unilateral ruptures [Haskell, 1964], we construct a two-corner-frequency model with an intermediate falloff of $f^{-1}$ and a high-frequency falloff of $f^{-2}$. The implication of this finding is that there are two dominant time scales, $T_1$ and $T_2$, in the slip-rate functions of thrust earthquakes. We find that duration follows well the conventional scaling $M_0 \sim T_1^2$ for earthquakes of $M_0 < 7.5$ and that the scaling changes to $M_0 \sim T_2^2$ for larger events. The second time scale has a much weaker moment dependence, $M_0 \sim T_2^2$, and we discuss its possible physical interpretations. We review these results in the context of scaling relations, the dependence of static and dynamic measures of earthquakes on earthquake size, and spatial variations in earthquake properties with depth and along-strike location in the major subduction zones.

2. Data and Methods
2.1. Data
We select all thrust earthquakes that have rakes between 45° and 75° and dips from 0° to 45°, with magnitudes $M_w \geq 5.5$ and greater from 1 January 1990 to 30 June 2015 with GCMT (Global Centroid Moment Tensor) [Ekström et al., 2012] depth less than 50 km from the catalog made available by Incorporated Research Institutions for Seismology (IRIS) (http://ds.iris.edu/spud/momenttensor, last accessed 23 January 2016). We find 4417 events that are recorded by the Global Seismic Network (GSN) and select the stations of the networks G, GE, II, IU, IC, and MN that are within 30° and 90° angular distance from the source to retrieve only P waves that are not affected by triplications in the mantle and diffraction along the core mantle boundary. We remove the instrument response of the raw seismic data in Seismic Analysis Code (SAC) using the transfer function and a band-pass filter between 0.001 Hz and 10 Hz and integrate to displacement. We find the P arrival time using the IASP91 velocity model [Kennett and Engdahl, 1991] and estimate attenuation from Warren and Shearer [2000], given preliminary estimates of the earthquake depth and the angular distance between source and receiver.

When selecting P wave pulses for a wide range of magnitudes, a fixed window length would both increase the amount of noise for the small (short duration) events and cut the main pulse for the large (long duration) events. Instead, we window the P wave pulse according to its magnitude using a conservative assumption based on scaling relations between corner frequency and earthquake size [Eshelby, 1957; Brune, 1970, 1971]. We increase the window duration $T_d$ with magnitude as $T_d(M_w) = 10^{M_w/-3.3}/2$. $T_d$ has to be at least 25 s long to capture the depth phases and less than 150 s to avoid the PP arrivals. We remove the trend and the mean of the window, add 15 s of a cos2 taper (Tukey window) on either side of the window, and pad with zeros the trace until 250 s. To increase the weight of the long periods in the spectral fit, we interpolate the Fourier amplitude spectra of the time window onto a uniform logspace vector of frequency between 0.01 Hz and 1.5 Hz. We find that the effects of attenuation are too great above 1.5 Hz to use the higher frequencies.

To select only high-quality data, we measure the ratio of the mean spectral amplitude of the data to the noise in three frequency bands (0.01 Hz–0.1 Hz, 0.1 Hz–0.5 Hz, and 0.5 Hz–1.5 Hz) and require the signal-to-noise
Figure 1. Effects of depth phases in $P$ wave amplitude spectra: (a) times series from synthetic stick seismograms (depth-phase Green’s function) and observed $P$ wave train from a $M_w$ 6.83 2011 earthquake (latitude = −37.62° and longitude = −73.45° recorded at II.COCO) and (b) their displacement spectrum, used to construct the synthetic spectrum (orange) from the convolution of the Green’s function (red) with a source model (green) to match the observed data (blue).

ratio (SNR) to be at least 2.5 in all bands. We also require that an earthquake is recorded by at least five stations meeting the SNR criterion, which further reduces our data set from 4417 events down to 942 thrust earthquakes. To remove wave propagation effects from the observed displacement spectra, we choose a Green’s function from a simplified Earth model where the Green’s function is simply expressed by attenuation

\[ \hat{G}(f) = \exp(-\pi f t^*(f)) \]

and where the geometrical spreading is accounted for in our spectral normalization. Attenuation is expressed using a $t^*$ approach [Warren and Shearer, 2000] such that $t^*(f)$ varies as a function of angular distance between the source and the receiver and we provide a MATLAB function to evaluate $t^*(f)$ in the supporting information. We estimate the source amplitude displacement spectrum \( \hat{U}(f) \) from the vertical component of the observed $P$ wave train displacement amplitude spectrum \( \hat{Z}(f) \) such that \( \hat{U}(f) = \hat{Z}(f) \exp(\pi f t^*(f)) \).

2.2. Depth-Phase Effects in Spectra

Our preliminary path removal does not account for three-dimensional wave propagation effects. In particular, the Earth’s free surface is a primary reflector of body waves, generating the depth phases $pP$ and $sP$, which follow direct $P$ at times that increase with the source depth. The time separation between the depth phases and the direct phases is comparable to the source duration for moderate and large earthquakes shallower than 50 km, resulting in overlapping pulses that produce interference effects in the amplitude spectrum. However, these effects can be modeled assuming that the radiation pattern is known from the GCMT solution.

In the far field, all the phases of the $P$ wave train, $P$, $pP$, and $sP$, arrive with a similar ray parameter (related to the incidence angle and the seismic wave speed at the receiver), such that we can constrain the takeoff
angles for $P$, $pP$, and $sP$ given the source depth. The $P$ waves recorded at angular distances between 30° and 90° have a steep takeoff angle, between 16° and 37°, leaving the focal sphere. With this geometry, for shallow-dipping thrust events, $sP$ typically arrives with comparable amplitude but opposite polarity to $P$. We construct a depth-phase transfer function, similar to that introduced in Houston and Kanamori [1986], that contains the amplitude and phase information of the $P$ wavetrain, given an impulse point source and estimate its spectrum, which we refer to as the depth-phase spectrum.

We show in Figure 1 the effects of depth phases in altering the shape of the $P$ wave train spectrum [Warren and Shearer, 2005; Denolle et al., 2015]. With $sP$ of opposite polarity, the amplitude decreases at very low frequencies because of the destructive interference, and thus, the low-frequency asymptote of the amplitude spectrum is not flat and cannot be used to infer moment. The characteristic frequencies of the troughs in the spectra (Figure 3) and the low-frequency asymptote are controlled by the source depth for a given source-receiver geometry. The troughs in the depth-phase spectrum are sources of instability in attempts to perform deconvolution of the observed spectrum. For large earthquakes, where the point-source approximation fails to represent the moment release over a range of depths, the depth-phase spectrum will provide an incomplete measure of the depth-phase effect. In Appendix A1, we show that we can reestimate the source focal depths, similar to Warren and Shearer [2005] and Denolle et al. [2015], of the earthquakes to better account for the depth phases when removing path effects. We provide the new depths in our catalog (see Table S1 in the supporting information and Figure A1). For earthquakes of $M_w \geq 8$, where a single “point-source depth” loses meaning due to the finite extent of the fault, we use the GCMT depths as the best representative depth as it should represent the depth of highest moment release. For the $M_w 9$ 2011 Tohoku event, we choose a focal depth of 16 km, which was found by Minson et al. [2014] to be that of highest potency.

Another bias to interpreting the depth-phase-contaminated spectrum as a source spectrum is the apparent corner frequency. The depth-phase corner frequency is controlled by the source radiation pattern and depth and is generally equal or greater to the corner frequency of the source at $M_w 6$ and above. Denolle et al. [2015] showed that for the 2015 Nepal earthquake sequence, the depth-phase effect on estimating corner frequency was significant for estimates of stress drop. Further analysis of source depth and earthquake size (Appendix A2) shows that earthquakes with source dimensions equal or greater than their centroid depth need careful handling due to depth-phase effects. However, the high-frequency asymptote remains relatively unaffected, indicating that the shape of the high-frequency spectrum is robust with respect to depth phase effects.

Figure 2a shows the stacked $P$ spectra within magnitude bins for $M_w 7.5$ and smaller events; results at larger magnitudes are less reliable owing to the much smaller number of events to stack. For each event, we normalize the $P$ wavetrain spectrum at high frequencies (0.2–1.5 Hz), where we know that the depth phases do not have large effects. We geometrically average the spectra over stations, requiring that each event has...
Step 1: find source model for small eGf

Step 2: find source model for large earthquake

Figure 3. Strategies to remove path effects and estimate the true source spectrum: (a) Method 1: use the depth-phase spectrum to remove the path effects of the small events and recover their source spectrum (and best fitting source depth), then (b) Method 2: use the same small events and their best fitting source model to create empirical 3-D Green’s functions, remove the path effects for the large events, and estimate their source spectrum.

at least five good records. Within moment magnitude bins of 0.5, we then stack all individual event stacks (in log amplitude) to construct an average $P$ wavetrain spectrum for thrust earthquakes of magnitudes $M_w \geq 5.5$. The first obvious feature is the downgoing low-frequency asymptote.

Figure 2b shows the stacks of computed depth-phase spectra for each event, which we call $\tilde{G}_d(f)$, constructed exactly as in Figure 3b. The synthetic depth-phase stacks have a downgoing low-frequency asymptote, an apparent corner frequency of about 0.04 Hz (25 s), and a nearly flat spectrum at high frequencies. This confirms that depth phases have little effect on the high frequencies but large effects on the low frequencies. *Boatwright and Choy* [1989, 1992] constructed similar spectra that they named the “free-surface spectrum.”

2.3. Removing Path Effects

Removing path effects entails correcting for geometrical spreading and attenuation due to wave propagation in a realistic 3-D Earth. Calibration of the path effects in the records of a large earthquake using nearby earthquakes of smaller magnitudes, often called the empirical Green’s function (eGf) approach, accounts for realistic 3-D Earth wave propagation. The eGf is simply the deconvolution of the observed displacement of a small event by its inferred source model. To retrieve the source models of the small earthquakes (Method 1), we construct an approximate Green’s function using the attenuation from a simplified Earth model convolved with a depth-phase transfer function. We correct for geometrical spreading by normalizing the spectrum (using its flat low-frequency asymptote) to its independently known seismic moment (from GCMT solutions). For large earthquakes (Method 2), for which our frequency bandwidth does not capture the flat asymptote, we estimate the geometrical spreading using the eGf constructed with nearby earthquakes using the Method 1 approach.
2.3.1. Method 1 (Small Earthquakes)
For earthquakes of $M_w \leq 7$, we use the depth-phase spectrum to correct for the free-surface effect. It allows us to better recover the flat part of the source spectrum at low frequencies, providing confidence that we can normalize the spectra to absolute seismic moment. We illustrate this approach in Figure 3a.

For each small earthquake, we construct the source spectrum $\hat{S}_s(f)$ by stacking over $N$ stations the observed displacement spectra, corrected for attenuation, normalized to one at 1 Hz (averaged over 20 points of the log-space frequency vector between 0.89 Hz and 1.12 Hz), $\hat{U}_i(f)$ recorded at station $i$, and the source-receiver specific depth-phase spectrum $\hat{G}_d(f)$ such that

$$\log_{10} \left( \hat{S}_s(f) \right) = \frac{1}{N} \sum_{i=1}^{N} \left[ \log_{10} \left( \hat{U}_i(f) \right) - \log_{10} \left( \hat{G}_d(f) \right) \right].$$

(1)

At this stage, we do not have a true estimate of the geometrical spreading, but we have confidence in the average level of the low-frequency asymptote that can be normalized to the independently known seismic moment. We will fit $\hat{S}_s(f)$ to a source model $\hat{M}_s(f)$.

2.3.2. Method 2 (Large Earthquakes)
For earthquakes larger than $M_w > 7$, our limited frequency bandwidth does not fully capture the source corner frequency. We cannot normalize the spectra to their flat, low-frequency asymptote and thus need to adopt another approach to removing path effects. Any event of smaller magnitude within a source-dimension distance from the target event is a candidate to construct empirical Green’s functions. Both events share nearly the same 3-D wave-propagation effects from source to receiver. At a given seismic station, we use the displacement records for a large event and a nearby small event and calibrate the relative amplitude to true moment using the known seismic moment of the small event. We illustrate this approach in Figure 3b.

Our first and approximate measure of source dimension uses the window length duration $T_d$, with an assumed rupture velocity of $0.9v_p$ to estimate rupture length as $L = 2 \times 0.9\beta T_d$ for bilateral ruptures. If there is no event of smaller magnitude within that distance range that has high enough SNR, we simply keep the results from depth phases (Method 1). The data only contain events that have similar focal mechanisms and similar depths (hence similar relative amplitudes of $P, sP$, and $sS$). Let us assume that two events, a small one $s$ and a big one $b$, share the same 3-D Green’s function $G(f)$ and that the observed displacement at a given station of the small event is $\hat{U}_s(f) = G(f) \cdot \hat{M}_s(f)$, $\hat{M}_s(f)$ being the best fit source model of the small event and that of the big event is $\hat{U}_b(f) = G(f) \cdot \hat{M}_b(f)$. Then we can estimate the source spectrum of the big event $\hat{S}_b(f)$ at each station using

$$\log_{10} \left( \hat{S}_b(f) \right) = \log_{10} \left( \hat{U}_b(f) \right) - \log_{10} \left( \hat{U}_s(f) \right) + \log_{10} \left( \hat{M}_s(f) \right).$$

(2)

and we then stack this result over the number of stations (averaging the azimuthal variations in spectral shapes). Figure 3b illustrates this deconvolution. We obtain an ensemble of source spectra $\hat{S}_b(f)$ from which we bootstrap over the contributing stations to estimate uncertainties in the resulting stack. The shape of the source model $\hat{M}_s(f)$ affects the high-frequency asymptote of $\hat{S}_b(f)$ (Figure 2c), such that using an eGf as small as possible with corner frequencies as high as possible reduces bias in estimating the high frequencies in $\hat{S}_b(f)$. Our approach attempts to address this issue by best fitting for the eGf source model and by incorporating as many eGf events as possible to estimate $\hat{S}_b(f)$. Note that we did not use $\hat{S}_s(f)$ in equation (2) because the deconvolution with the troughs in $\hat{S}_b(f)$ is unstable. In general, the use of the eGf approach approximates the path effects between the target event and the receiver as a (point source) Green’s function. Large faults expand over a range of depth, and the depth-phase spectrum will be smoother than that of a point source. Supporting information Figure S13 shows the likely uncertainties that arise from this approximation.

In this procedure, we do not have the flat low-frequency asymptote for the large earthquakes that allowed us the normalization to the true seismic moment in the first step, but we are able to calibrate the spectra to their true amplitudes with the eGfs from the smaller events (for which accurate moment calibration is possible).

The two combined approaches to removing path effects allow us to use the same data sets in a unified way while maximizing the range of earthquake magnitudes, as illustrated in Figure 2c. The choice of the threshold magnitude (here $M_w 7$) was determined as a trade-off between (i) having larger uncertainties in the spectral fit if the threshold is lower ($M_w 6.5$) and (ii) avoiding the bias in corner frequency estimates when the duration corner frequency is lower than the low-frequency cutoff for a higher threshold ($M_w 7.5$, see Figure S12).
In this study, we ignore the azimuthal distribution of the stations even if some azimuths may contribute more to the stacks. Directivity and other finite source effects are not within the scope of this study, but future work using this approach could incorporate more details of the ruptures.

3. Modeling P Wave Source Spectra

To look at the overall average shape of the source spectra $\hat{S}(f)$ and extract features common to all events as a function of moment, we stack by taking the median of the $P$ wave spectra $\hat{S}(f)$ within magnitude bins of 0.4, as shown in Figure 4. Note that we normalized the amplitudes of the estimated $P$ spectra for the smaller events using their known moments and calibrated the amplitudes for the larger events ($M_w \geq 7.3$ in Figure 4) using eGfs from nearby smaller events.

A common approach is to then fit a simple $f^{-2}$ model with a single-corner frequency $f_c$ that reflects the source duration $T$, assuming $T = 1/\pi f_c$ [Aki, 1967; Brune, 1970; Madariaga, 1976]:

$$\hat{M}_b(f) = \frac{M_0}{1 + (f/f_c)^2},$$

(3)

where $M_0$ is the seismic moment. To fit equation (3) to the stacks, we focus on the difference in spectral shape between $\hat{M}_b(f)$ and $\hat{S}(f)$, and not on the absolute amplitude levels. We perform a simple grid search to find the best $f_c$ that minimizes the misfit function $R$:

$$R(f_c) = \| \log_{10} (\hat{M}_b (f_c)) - \log_{10} (\hat{S}) - S_0 \|_2,$$

(4)
where $S_0(f_c)$ is a calibration factor, $S_0(f_c) = \text{mean}(\log_{10}(\hat{M}_d(f_c)) - \log_{10}(\hat{S}))$ (mean of the residuals) for the source spectra estimated from Method 1, and $S_0(f_c) = 0$ for the source spectra estimated from Method 2 (i.e., where we know the absolute amplitude level). For Method 2 events, the spectral fit is thus not an extrapolation to the flat low-frequency asymptote [Baltay et al., 2011, 2014] but rather an informed interpolation between the amplitude spectrum at its lowest reliable frequency and the zero-frequency limit defining the true seismic moment. Note that the misfit in equation (4) is calculated within the data frequency band (0.01 – 1.5 Hz), and we grid search through corner frequencies between $10^{-3.5}$ Hz and 10$^1$ Hz. No grid search is needed on the amplitudes because this can be estimated from the mean residual. Ignoring the amplitude information (Figure 4a), we normalize our stacks to the best fitting source model $\hat{M}_d(f)$ to highlight the fit in the spectral shapes (rather than the amplitudes). We best fit equation (3) to our data and find an apparent scaling of moment with corner frequency of about $M_0 \propto f^{-4}$, which is unreasonably steep. Moreover, the corresponding source durations seem unrealistic (e.g., 15 s for $M_w$ 8+), and the fit to the spectra is poor for the large events near the modeled corner frequency. A much better fit can be achieved by adding an intermediate falloff slope in the model for spectral shapes, introducing two corner frequencies, $f_1$ and $f_2$:

$$\hat{M}_d(f) = \frac{M_0}{\sqrt{(1 + (f/f_1)^2)^2}} \sqrt{(1 + (f/f_2)^2)}.$$  

(5)

This spectral shape was proposed by Haskell [1964] for a moving line source and by Savage [1972] for a rectangular fault with two source dimensions. Note that when $f_1 = f_2$, $\hat{M}_d(f) = \hat{M}_d(f)$. Figure 4b shows the fit of the estimated source spectra to the best fitting double-corner frequency model $\hat{M}_d(f)$. The estimates of source duration from $f_1$ are more reasonable, 1.5 s for $M_w$ 5.7, 4 s for $M_w$ 6.5, and 40 s for $M_w$ 8+. The model for the intermediate falloff region with a slope of $f^{-1}$ fits the observed P spectra. Nearly self-similar scaling of $M_0 \propto f^{-3}$ is observed for magnitudes less than 7.3 but appears to fail for large magnitudes.

The double-corner frequency model fits the data better than the single-corner frequency model, both in the moment-binned spectral stacks and in fits to the individual events. The root-mean-square (RMS) misfit residuals between the data and the best fit synthetic spectra in logspace for all earthquakes $i$, $\text{res}_i^k$ and $\text{res}_{i\mu}^k$ are visually nearly normally distributed (see Figure S1), but the total RMS residual over all events is 1.8 times higher for the single-corner frequency model. However, some improvement in fit can always be achieved by adding additional model parameters. To evaluate the statistical significance of the fit improvement, we apply the Akaike information criterion (AIC), which penalizes model complexity in assessing misfit significance to favor parsimony,

$$\text{AIC} = 2k + N \ln \left( \frac{1}{N} \sum_{i=1}^{N} \left( \text{res}_i^k \right)^2 \right)$$  

(6)

where $N$ is the number of fitted frequencies and $k$ is the number of free parameters. For small earthquakes (Method 1 events), the parameters are the corner frequency(s), the long-period spectral level, and the residual variance, thus $k = 3$ (equation (3)) or $k = 4$ (equation (5)). For larger earthquakes (Method 2 events), the long-period level is fixed, and the number of free parameters is reduced to $k = 2$ (equation (3)) and $k = 3$ (equation (5)). To compare whether the double-corner frequency model is statistically preferable to the simpler single-corner frequency model, we compare both AIC values, $\Delta \text{AIC} = \text{AIC}_1 - \text{AIC}_0$. We find that $\Delta \text{AIC}$ is mostly negative for events of small magnitude ($M_w \leq 7$) for which the single- and double-corner frequency model share similar significance, but for larger earthquakes, $\Delta \text{AIC}$ is always significantly negative, thus favoring the double-corner-frequency model. $\Delta \text{AIC}$ shows significant dependence on earthquake size when considering the spectral stacks (see Figures 7 and S2).

A measure of uncertainties in the corner frequency estimates can be obtained from the behavior of the misfit functions. For the small earthquakes (i.e., the ones with the depth-phase correction only), we estimate the uncertainties using the skewness of the misfit function. The misfit function $R(f_1, f_2)$ now has two dimensions and is symmetric ($R(f_1, f_2) = R(f_2, f_1)$) because of the symmetry in the source model (equation (5)). We find its minimum, imposing $f_1$ lower than $f_2$. We estimate the uncertainty in $f_1$ by fixing $f_2$ to its lower value, $R = R|_{f_2}$, and measuring the error bar within with $R|_{f_1} < 0.1$ max ($R|_{f_2}$). While $f_1$ and $f_2$ are anticorrelated, our measure of
uncertainty remains reasonable. We show in Figure 5a an example of the misfit for the \( M_w 6.5 \) stack of Figure 4b. This measure of uncertainty solely relies on the spectral shape, and not its absolute amplitude.

For the large earthquakes (i.e., those obtained from the empirical Green’s function approach), the measure of uncertainty presented above is not appropriate. We know the amplitude level, and our fitting acts as an interpolation between the lowest reliable frequency point of the estimated source spectra and the known static moment level. We can however evaluate uncertainties due to averaging of the source spectra that we estimate from each eGf. At each frequency, we estimate the source amplitude spectrum by measuring the mean of all individual estimates (i.e., from each eGf). At each frequency, we calculate the uncertainties from \( \pm 1 \) standard deviation (std) of the amplitude spectrum estimates. This provides a lower energy spectrum and a higher-energy spectrum as viable bounds for the uncertainty. We then carry out the same model fitting for the lower and upper bound spectra and use their best fitting \( f_1 \) and \( f_2 \) as corner frequency errors (shown in Figure 5b).

The four largest events in our data set are the 2004 \( M_w 9.2 \) Andaman-Sumatra, the 2011 \( M_w 9.0 \) Tohoku, the 2005 \( M_w 8.7 \) Nias-Sumatra, and the 2010 \( M_w 8.8 \) Maule earthquakes. Figure 6 presents the estimated source \( P \) spectra of all four events from each of their eGf and their average (with our bootstrapping method). Note that our lowest corner frequency \( f_1 \) is lower than the data bandwidth, such that we can estimate source durations longer than our window length. The choice of seismic moment directly affects \( f_1 \), that is, underestimation of the moment will yield a high value of \( f_1 \), and vice versa. We chose seismic moments that were used in Ye et al. [2016a, 2016b] from the GCMT for the Tohoku, Maule, and Nias earthquakes from systematic comparison with results from finite fault inversions.
Figure 6. Source spectra estimated from individual empirical Green's functions (colored lines), spectra averaged in log space (thick black line), best fit source model (solid dash line), and best estimated corner frequencies $f_1$, $f_2$ (white dots).

For the 2011 $M_w 9$ Tohoku earthquake, we choose $M_0 = 5.31 \times 10^{22}$ Nm, which is slightly larger than the estimates from Koketsu et al. [2011] but quite close to that of Minson et al. [2014]. We find a source duration of approximately $T_1 = 125$ s, which is close to Minson et al. [2014] estimates and slightly lower than other duration values [Ide et al., 2011; Meng et al., 2011]. The second time scale for the Tohoku earthquake is $T_2 = 11$ s, but likely poorly constrained given that it is close to the cutoff frequency of our data bandwidth.

The $M_w 8.8$ 2010 Maule ($M_0 = 1.86 \times 10^{22}$ Nm) and the $M_w 8.7$ 2005 Nias ($M_0 = 1.05 \times 10^{22}$ Nm) earthquakes have more distinct spectral shapes that follow the double-corner-frequency model (Figures 6b and 6c). The duration of the Nias event is $T_1 = 153$ s, again comparable with other studies [Konca et al., 2007], that of the Maule earthquake is $T_1 = 176$ s (150 s, http://www.tectonics.caltech.edu/slip_history/2010_chile/index.html, last accessed 31 January 2016). The intermediate falloff slope is very clear in the two spectra and the second time scale is quite short, $T_2 = 2.2$ s and $T_2 = 1.9$ s for Nias and Maule, respectively.

For the $M_w 9$ 2004 Sumatra earthquake, we choose $M_0 = 5 \times 10^{22}$ Nm to be closer to the values provided by Banerjee et al. [2005], which yields a reasonable duration for the teleseismic $P$ wave pulse of $T_1 = 577$ s [Ishii et al., 2005; Ammon et al., 2005]. This earthquake was unusually large, and the estimates of moment release from normal-mode estimates differ from that of body and surface waves [Ammon et al., 2005; Park et al., 2005]. Our study focuses on accurate measures of source duration, which influenced our choice of moment. We find a short second time scale, $T_2 \sim 2$ s, though the fit of the intermediate falloff region (Figure 6a) is not as good as for the other events, owing to complexity in the source spectrum.

4. Scaling of Time Scales and of Source Dimension

To identify the main patterns in the spectral shapes, we stack the source models found by fitting all source spectra to equation (5) and by binning the stacks every 0.4 in $M_w$. Figure 7 shows the stacked models for $M_w < 8.6$. The frequency bandwidth of the intermediate falloff region increases with magnitude. We fit again the stacked source model to a double-corner frequency model to find the best $f_1$ and $f_2$ and verify that the
best fitting corner frequencies of the stacks are similar to the average (in log space) of the individual best fitting corner frequencies.

4.1. Source Duration—$T_1$

The first corner frequency follows relatively well the expected duration scaling ($M_0 \propto T^3$) for sources of $M_w < 7.4$. Durations are a few seconds for $M_w = 5–6$, tens of seconds for $M_w = 7–8$, and up to hundreds for $M_w = 8+$. There is a clear change of scaling in the source duration at magnitudes $M_w = 7–7.5$. The exact magnitude above which the scaling breaks is likely blurred by the global averaging. A change of duration scaling due to the finite width of the fault has been proposed for strike-slip events [Hanks and Bakun, 2008; Leonard, 2010]. When the fault is bounded by a finite width, the fault growth only occurs in the long-dimension $L$, such as $L = V_T T$, and the width $W$ stays constant. With this constraint, the seismic moment becomes

$$M_0 = \mu s W L \propto \mu s W V_T T.$$  \hspace{1cm} (7)
Figure 8. Individual measurements (solid circles, solid triangles for tsunami earthquakes $M_w 7.8$ 2006 Java, $M_w 7.9$ 2010 Mentawai [Kanamori, 1972; Ammon et al., 2006; Lay et al., 2011], and their uncertainties (thin gray for Method 1 processing and thin red for Method 2 processing), their bootstrapped mean (solid squares) and standard deviation (solid dark error bars) with seismic moment. White solid squares are the most uncertain bootstrapped mean with only two $M_w 9$. Most of the error bars are smaller than the square symbols. (a) Corner frequency ($f_1$) with constant stress drop lines when assuming a circular crack ($f_c = 0.42 \beta M^{-1/3} \sigma_1^{1/3}$, $\beta = 3900$ km/s, $\sigma_1 = 0.1$ MPa, 1 MPa, 10 MPa, and 100 MPa). (b) Source durations $T_1 = 1/\pi f_1$ in seconds with indications of likely moment-duration scaling. (c) Second corner frequency ($f_2$) with moment and indication of a likely scaling (Figure 7). (d) Short time scale $T_2 = 1/\pi f_2$ based on the second corner frequency.

The seismic moment should only scale with $T$, and not $T^3$. This scaling has been extensively discussed for slow slip earthquakes [Ide et al., 2007; Peng and Gomberg, 2010]. The same phenomena may occur for both slow and fast earthquakes [Gomberg et al., 2016], and we propose that our data reflect the finiteness of the megathrust width.

The change of scaling is also visible from measurements of corner frequencies of individual events. Figures 8a and 8c show measurements for both corner frequency estimates and their interpretation in terms of pulse duration. We find that $T_1 = 1/\pi f_1$ represents quite reasonably the time scale of the source duration. Using $f_1$ as a “duration” corner frequency, we can compare the individual estimates to that predicted from a circular crack with constant rupture velocity ($V_r = 0.9/\ell$) and uniform static stress drop of 0.1 MPa, 1 MPa, 10 MPa, and 100 MPa [Madariaga, 1976; Kaneko and Shearer, 2014]. We clearly see that the scaling of $f_1$ changes for $M_w 7+$ and that the use of a circular crack to estimate stress drop would predict a decreasing scaling of static stress drop with moment, down to unreasonably low stress drops.

We highlight the corner frequencies and duration estimates for specific tsunami earthquakes in Figure 8. The $M_w 7.8$ 2006 Java and the $M_w 7.9$ 2010 Mentawai tsunami earthquakes had an abnormally large tsunami for their radiation [Kanamori, 1972], and we find durations that are shorter than that obtained by other studies: 125 s for Java (against 165 s from Ammon et al. [2006]) and 110 s for Mentawai (against 185 s from Lay et al. [2011]). Our source duration estimates for the tsunami earthquakes are shorter than other studies likely because our simple model does not adequately capture the complex spectral shape (see Figure S3). Nevertheless, our tsunami earthquake duration estimates are longer than those of other earthquakes of similar moment.
The uncertainties for the small earthquakes (from Method 1) show that we may overpredict $f_1$ or underpredict the source duration. The circular crack model predicts a 1 MPa stress drop for those events. The uncertainties for the larger events (from Method 2), which we derived from bootstrapping the source spectrum from each empirical Green’s function, are larger for the $M_w 7$ than for the $M_w 8–9$ events. This difference may arise from our methods to estimate uncertainty but gives us some confidence in the eGf approach to constrain corner frequencies lower than that bounded by the data. Our corner frequency estimates are, in general, slightly lower than that of the Allmann and Shearer [2009] study that we refer to as AS09. In AS09 study, the spectral fit ignores depth phases and the source model uses a single-corner frequency estimate with a high-frequency falloff rate of $n = 1.6$. The difference in the source model is likely the reason for this discrepancy.

To study the behavior of $T_1$ with source depth, we should correct for the moment dependence of $T_1$. Houston [2001] constructed a scaled duration to investigate source-time-function duration and shape with respect to depth without magnitude-dependent effects. We employ a similar approach and correct for the effect of the moment on duration by removing the trend, performing $T \times M_w^{-1/3}$ and adjusting the level to the median duration between magnitude 6 and 6.5, $T_6 = 10^{\text{median}(\log_{10}(T \times M_w^{-1/3}))}$, such as $T' = T \times M_w^{-1/3} \times T_6$. Figure 9a shows the scaled durations with depth. There is no obvious variation of the source duration with source depth from the surface down to 35–40 km depth. A slight decrease of the source duration occurs at 40 km, a jump that was also highlighted by Houston [2001] who proposed that it was the downdip end of most megathrusts, again an indication of a finite fault width effect.

Figure 10a shows spatial distributions of $T'_1$, which we will interpret later regarding spatial variations of stress drop.

### 4.2. Short Time Scale — $T_2$

The second time scale $T_2$ has a weaker moment dependence (Figures 7 and 8) and varies between 0.5 and 5 s. It is clearly not as well constrained as the duration $T_1$. Our simple model fits the stacks, which likely averages the complexities of the individual spectra. However, our uncertainty bounds show that we are more likely...
Figure 10. (a) Global variations of $T'_1$: low average values of $T'_1$ (short duration earthquakes) are seen for intraplate events (Tibet) and Chile events, while high values (long duration earthquakes) are found in the Mariana Trench and in isolated cases. (b) Global variations of $T'_2$ with lower values in Sumatra and in the Kuril subduction zones, in Papua New Guinea, with high values having a less distinguishable pattern.

Figure 9b has more scatter in the duration values but shows a weak depth dependence with a possible reduction of the time scale around 20 km. The median $T'_2$ is 1.5 s for the shallow events and 0.8 s for the deep events.

Figure 10b shows few clear spatial patterns in $T'_2$. The Kuril subduction zones seems to have lower values, intraplate earthquakes seem to have less variability as an earthquake ensemble than regions along plate boundaries. The Central America subduction zone has events with higher values of $T'_2$, probably indicating...
Table 1. Rupture Velocities and References for Particular Earthquakes

<table>
<thead>
<tr>
<th>Earthquake</th>
<th>Rupture Velocity</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mw 8.4 Peru 2001</td>
<td>2.25 km/s</td>
<td>Lay et al. [2010a]</td>
</tr>
<tr>
<td>Mw 9.0 Sumatra 2005</td>
<td>2.5 km/s</td>
<td>Ishii et al. [2005] and de Groot-Hedlin [2005]</td>
</tr>
<tr>
<td>Mw 8.7 Nias 2005</td>
<td>2.5 km/s</td>
<td>Ishii et al. [2007]</td>
</tr>
<tr>
<td>Mw 7.8 Java 2006</td>
<td>1.25 km/s</td>
<td>Ammon et al. [2006]</td>
</tr>
<tr>
<td>Mw 8.3 Kuril 2006</td>
<td>2 km/s</td>
<td>Lay et al. [2009]</td>
</tr>
<tr>
<td>Mw 8.3 Peru 2007</td>
<td>1 km/s</td>
<td>Lay et al. [2010a]</td>
</tr>
<tr>
<td>Mw 8.5 Sumatra 2007</td>
<td>2.5 km/s</td>
<td>Konca et al. [2008]</td>
</tr>
<tr>
<td>Mw 8.8 Maule 2010</td>
<td>2.25 km/s</td>
<td>Lay et al. [2010b]</td>
</tr>
<tr>
<td>Mw 9.0 Tohoku 2011</td>
<td>2.5 km/s</td>
<td>Meng et al. [2011] and Yagi and Fukahata [2011]</td>
</tr>
<tr>
<td>Mw 8.1 Iquique 2014</td>
<td>1 km/s</td>
<td>Meng et al. [2015]</td>
</tr>
<tr>
<td>Mw 7.9 Nepal 2015</td>
<td>2.9 km/s</td>
<td>Fan and Shearer [2015]</td>
</tr>
</tbody>
</table>

relatively high energy events. A planned follow up study will investigate, along with the mechanisms for $T'_2$, the correlation, or lack of correction, of $T'_2$ with geodynamic parameters.

4.3. Rupture Area

We now attempt to estimate the dimensions of the source given our measures of source duration. For small earthquakes, it is common to assume that the rupture is 90% of the shear wave speed, $V_s = 0.9\beta$ and that the rupture propagates like an expanding crack in a circular fault [Brune, 1970; Madariaga, 1976; Kaneko and Shearer, 2014]. We use CRUST 1.0, [Laske et al., 2013] for the local estimate of shear wave speed $\beta$. For the largest earthquakes, we use rupture velocities found by finite source inversion studies and by backprojection of teleseismic $P$ waves (see Table 1).

For most earthquakes, we assume a bilateral rupture, such that the total length is $L = 2V_s T_1$. For those that have clear unilateral propagation, such as the $M_w$ 9 2004 Sumatra [Ishii et al., 2005] and $M_w$ 7.9 2015 Nepal [Galetzka et al., 2015; Fan and Shearer, 2015] earthquakes, we assume $L = V_s T_1$. Figure 11a shows the length estimates and their averages over magnitude bins. We propagate the uncertainties in our measurements of $f_1$ to evaluate source length uncertainties and color code them similar to Figure 8. We also highlight what a 250 km source dimension would be.

The transition from brittle to ductile rheology occurs fairly deep in subduction zones, which allows the seismogenic width of the plate interface to be as wide as $\sim$250 km. The largest subduction zone events will exceed this dimension in length, such that these earthquakes become more elongated, mode II type ruptures. It is obvious from Figure 8a that $L$ becomes too large to represent the diameter of a circular crack and that the circular crack model fails for large earthquakes, as the fault area would be significantly overestimated.

The limit of fault width has been discussed, in particular, for large strike-slip faults that are bounded by a narrow width of $\sim$15 km [Shaw and Scholz, 2001; Romanowicz and Ruff, 2002]. Hanks and Bakun [2008] and Leonard [2010] discuss how we can expect a break of scaling in measuring fault area and/or dimension. Leonard [2010] proposes a simple smooth relation between fault width and fault length $W \approx 20 L^{2/3}$. Though we understand that this general relation may not capture the complexity of individual sources, it allows us to smoothly increase the aspect ratio of the fault and not introduce a spurious break of scaling. Our estimates of fault width always fall under 250 km, except for the $M_w$ 9.2 Sumatra earthquake where we keep the width fixed to 250 km (Figure 11b).

Using the Leonard [2010] relation between fault length and width and our estimates of source duration, the fault aspect ratios vary between 1 and 4 for most earthquakes. Finite-fault inversions almost always assume a wide planar and rectangular fault area to capture all slip within the bounds of the fault. It is likely that our spectral measurement applies to the fault regions with the greatest moment release and thus may not include the full dimension of slip, i.e., low slip regions near the initiation or termination of rupture. However, the aspect ratios of the two largest events have reasonable values, the $M_w$ 9.2 2011 Tohoku event was compact [Ide et al., 2011], while the $M_w$ 9.2 2004 Sumatra event was particularly elongated Ishii et al. [2005].
Figure 11. Individual estimates of length and width with seismic moment, same conventions as in Figure 8. (a) Rupture length $L$ estimated from $V_r T_1$ ($V_r = 0.9 \beta$ except when $V_r$ is known (Table 1). (b) Width $W$ estimated from Leonard [2010]. Maximum fault width in subduction zone highlighted at 250 km. Asymptotes of scaling between length and moment shown in dark lines. Individual measurements are in solid orange circles, and magnitude-bin (bootstrapped) means in large solid blue squares and their standard deviation in solid dark error bars. Solid white circles represents the most uncertain bootstrapped means. Uncertainties propagated from $f_1$ uncertainties are in gray (Method 1 events) and red (Method 2 events).

5. Static Stress Drop

Static stress drop is the total change of shear stress on the fault from the earthquake, integrated over the fault area. As the moment is generally well constrained, difficulties in reliably estimating stress drop are mainly due to uncertainties in the fault dimension and the elastic properties, i.e., shear modulus (rigidity), near the fault.

With an incomplete knowledge of the elastic structure at the scale of the fault, we often approximate the shear modulus to be constant ($\mu = 40$ GPa) and ignore material contrasts across the fault [Rubin and Ampuero, 2007]. While density is likely the least well constrained parameter, we estimate the shear modulus given the location and depth of the earthquakes such that $\mu = \rho \beta^2$ from global crustal velocity models (e.g., CRUST1.0) [Laske et al., 2013]. This allows us to capture at least the depth dependence of the elastic properties, though spatial variability might prevail over the depth range of shallow (< 50 km) earthquakes. Uncertainties in shear modulus are linearly related to that of stress drop, indicating that strain drop is likely better constrained than stress drop.

Fault area is a nontrivial, yet fundamental, static measure of earthquakes. Scaling relations provide approximate relations between earthquake size and fault dimensions [Mai and Beroza, 2000; Hanks and Bakun, 2008; Leonard, 2010]. For large earthquakes, finite-fault inversions provide estimates of the spatial variation of slip history. Classical slip inversions are performed on large fault planes, often rectangular and planar, to allow the slip to be contained within the parameter space, such that the area of nonzero slip is smaller than the area of...
the fault chosen for the inversion. To better estimate the area where slip occurred, a threshold of slip is defined to constrain the minimum of slip required to count each pixel in the area [Noda et al., 2013], which can yield trimming factors between the area of most slip and the total fault surface [Somerville et al., 1999; Noda et al., 2013; Ye et al., 2016a, 2016b].

Smoothing in the slip inversions is anticorrelated with stress drop because the stronger the slip heterogeneity, the stronger the strain [Ihmlé, 1998]. The $M_{w}$ 9 2011 Tohoku earthquake serves as a good example of this nonuniqueness as many fault models were produced from the extensive recordings of this earthquake [Brown et al., 2015]. Given that Tohoku was a surface-rupturing event, the use of teleseismic $P$ waves for slip inversion is likely strongly affected by destructive interference with the depth phases. For small earthquakes, we cannot properly image the finite source effects, and it is common practice to assume a simple model of uniform stress drop, for which the slip solution follows elliptical functionals [Eshelby, 1957].

Taken together, the choice of rigidity, of smoothing in the finite slip inversion, of the fault area, and of averaging scheme is critical to estimating stress drop. Noda et al. [2013] differentiates three types of averaging, moment-based, area-based, and energy-based averaging, which agree if the stress drop is uniform on the fault. Note that the Noda et al. [2013] study finds that moment-based stress drops likely underestimate the estimate of the average stress drop if the stress change is heterogeneous by up to a factor of 8. Brown et al. [2015] confirms that stress drop varies strongly locally, and the choice of the threshold (trimming factor) especially affects total measures of stress drop.

In this study, we assume an elliptical fault, with major axis $a = L/2$ and minor axis $b = W/2$, with area $\pi ab$, and with uniform stress drop as a consistent reference model for our entire data set. We estimate average slip from the seismic moment,

$$s = \frac{M_0}{\mu \pi ab},$$

(8)

since we do not solve for variable slip on the fault. The static stress drop for a buried fault embedded in a homogeneous whole space is

$$\Delta \sigma = \mu C(a, b, \nu) \frac{s}{b} = C(a, b, \nu) \frac{M_0}{\pi ab^2},$$

(9)

where $C(a, b, \nu)$ is a factor that depends on the aspect ratio $a/b$ and the Poisson ratio $\nu$ and $\mu C(a, b, \nu)$ acts as an effective fault stiffness. Changes in $\nu$ obtained from the 3-D velocity model only vary $C(a, b, \nu)$ by 5%, which is a small error given the uncertainty and variability of stress drop estimates. $C$ varies between 1.22 and 1.5 for aspect ratios between 1 and 10 (see Appendix B for more details on our derivation of $C$, also discussed in Madariaga [1977]). There are other expressions for moment-based stress drops that are used for rectangular faults but that are not fully compatible with those of Eshelby [1957], in particular, for circular faults. For consistency, we use Eshelby [1957] for the circular and elliptical cases. Appendix B discusses those differences in terms of stress drop. A key factor in estimating stress drop is the presence of the free surface. Knopoff [1958] and Parsons et al. [1988] addressed the issue for strike-slip faults, Minson et al. [2014] suggests that for dip-slip faults the presence of the free surface reduces the stress drop by a factor of 2, although we have not found in the literature rigorous proof for that factor. Here we do not evaluate equation (9) for strike-slip events, but other authors define $C$ using the fault area, not the smallest dimension, and find that effective stiffness can vary by a factor of 2 for large aspect ratios [Parsons et al., 1988; Noda et al., 2013].

### 5.1. Scaling of Stress Drop With Earthquake Size

Figure 12 shows a simple distribution of the $\log_{10}$ stress drops that appears roughly Gaussian, with a mean (or median) of 2 MPa. This average value was also found by Allmann and Shearer [2009]. We estimate stress drop from equation (9) for all of our events. Figure 13a shows all individual measurements versus seismic moment. The large scatter of the stress drops of small events may reflect both large uncertainties (due to SNR levels) and high natural variability (due to the large number of events). Our estimated uncertainties in stress drop are reduced for the larger earthquakes, which nonetheless show stress drop variability that spans 3 orders of magnitude.

Our results suggest no systematic scaling of average stress drop with seismic moment, if one accounts for the scaling of aspect ratio with earthquake size. In contrast, if a circular fault is assumed for all events, then we would obtain a decrease of stress drop with seismic moment that is directly explained by an overestimation.
Figure 12. (a) Variations of stress drop with seismic moment (same conventions as in Figure 8). (b) Variations of stress drop with source depth. Solid circles are individual measurements, solid triangles are those of tsunami earthquakes \cite{Kanamori, 1972; Ammon et al., 2006; Lay et al., 2011}, and solid squares are mean (bootstrapped) over magnitude bins. Gray lines are uncertainty measurements for the Method 1 events, and red lines are those of the Method 2 events.

Our results suggest that stress drop is relatively invariant with earthquake size, but that self-similarity is nonetheless violated by the changing fault geometry of the largest earthquakes.

Our estimates of stress drop for the large earthquakes are comparable to other studies. We obtain a stress drop of 9 MPa for the $M_w 9.0$ 2011 Tohoku earthquake, a value in reasonable agreement with Minson et al. [2014] who obtained 5–7 MPa for their static stress drop estimates of a buried fault using a rupture area constrained by slip contours of 20\% of maximum slip. Brown et al. [2015] chose a fixed shear modulus of 40 GPa and analyzed 40 slip-inversion models to find a static stress drop of about 2 MPa for slip contours of 5 m, though higher slip contours provides higher stress drop. We use the 16 km depth of highest potency found by Minson et al. [2014] and the CRUST1.0 value of 45 GPa, which yields 8.6 MPa with $\mu = 40$ GPa. We estimate an average stress drop of 1.8 MPa for the $M_w 9$ 2004 Sumatra earthquake, though we did not find estimates from the literature from finite source inversions. We obtain 1.6 MPa for the 2010 Maule event (3.6 MPa from Ye et al. [2016b]) and 0.9 MPa for the 2005 Nias event (2.6 MPa from Ye et al. [2016b]). We compare most $M_w 7+$ estimates with those found by Ye et al. [2016a, 2016b] and with those of AS09 in supporting information (Figures S4 and S5).

5.2. Stress Drop With Depth
We find no systematic variation of stress drop with earthquake depth on a global scale, after accounting for the likely increase in rupture velocity with depth using CRUST1.0 to extract elastic properties at the source locations. Global variations of strain drop also do not show great depth dependence (not presented in our study). However, there is a suggestion in Figure 9a for a weak jump in stress drop around 40 km \cite{Houston, 2001; Bilek and Lay, 1999} that is naturally explained by shorter source pulses. Locally, an increase of stress...
Figure 13. (a) Histogram of stress drop estimates in log space and best fitting normal distribution. (b) Histogram of scaled energy estimates in log space and the best fitting (though poorly fitting) normal distribution.

drop with depth is clearly observed in the Tohoku region (see Figure 16a); such an increase with depth was also observed by Uchide et al. [2014] in smaller earthquakes before the Tohoku main shock.

5.3. Global Variations of Stress Drops From Thrust Earthquakes

There are interesting global patterns of stress drop (Figure 14). Average stress drops in the western Pacific are lower than those of the eastern Pacific, about 1.5 MPa compared to 3 MPa. Average intraplate stress drops (from the intra-Asia events) are larger than interplate stress drops, 1.5 MPa compared to 7.5 MPa, as also noted by many other studies [Kanamori and Anderson, 1975; Scholz et al., 1986; Allmann and Shearer, 2009]. The Aleutian subduction zones have lower average stress drops, 1 MPa. Supporting information figures show close-up maps of stress drop estimates for different subduction regions (Central America Figure S6, Aleutians Figure S7, Papua New Guinea and Marianas trenches Figure S8, South America Figure S9, Sumatra Figure S10, and Solomon Islands Figure S11). AS09 used a different technique to remove path effects from teleseismic P spectra: a $t^*$ operator and regional $\Delta t^*$ to account for large-scale Q variations, a station-event stacking scheme to reduce site effects at each station, and computation of an empirical correction spectrum based on a self-similar source spectral model with a single-corner frequency. They found that a high-frequency falloff rate of 1.6 better fit their spectra and imposed that falloff to find the corner frequency for which the match between model and data is the best. Note that this falloff rate is nearly midway between the $f^{-1}$ falloff we find at intermediate frequencies and the $f^{-2}$ falloff we find at high frequencies in our double-corner-frequency modeling approach. Figure S4 compares the stress drop values from our study and AS09 for the common events.
Aside from the intraplate/interplate difference in stress drops, our results on spatial variations of stress drop differ from AS09 for individual events. Figure S4 shows that we have relatively common low and high stress drop events, except that our study finds some events with lower stress drop than AS09. Those events are smaller ($M_w \leq 6$) such that we attribute this discrepancy to a signal-to-noise level issue. One possible contributor to this difference is the laterally varying $t^*$ correction used in AS09. In general, a lack of high frequencies in a source spectral estimate could be due to a true source effect (i.e., a low corner frequency) or unmodeled strong near-source attenuation. We investigated the approach of AS09 using Warren and Shearer [2002] $t^*$ perturbations and corrected our spectra for variable $Q$. However, we found that our estimates of stress drop did not vary as anticipated and decided to ignore laterally varying attenuation. We believe that the differences between our results and those of AS09 are most likely due to the assumed high-frequency falloff rate of $f^{-1.6}$ and single-corner frequency model used in AS09.

6. Radiated Energy

Our measure of radiated energy is based on the single-station approach, in which absolute amplitudes are not required and we use the spectral stack over all stations as a representation of the source spectral shape (applying depth-phase corrections and the eGf corrections for the larger events, see below). Assuming the success of our approach (i.e., $\tilde{S}(f)$ is an unbiased moment-normalized estimate of the shape of the source spectrum), then the radiated energy can be obtained from

$$E_p = \frac{2\pi M_w^2 (R_p^2)}{\rho a^2} \int_0^\infty \left| \tilde{S}(f) \right|^2 df,$$  

(10)

$\langle R_p^2 \rangle$ being the squared $P$ wave radiation pattern averaged over the focal sphere and equal to $4\pi/15$. Because we measure $P$ waves only, to estimate the total radiated energy, we must make assumptions about the $S$ spectra. If the $S$ spectra have the same shape and corner frequency as the $P$ spectra, then $E_s = \frac{3}{2 \beta} E_p$, where $a$ and $\beta$ are the local $P$ and $S$ velocities at the source. For a Poisson solid, $E_s/E_p = \frac{3\nu}{2\beta} = 23.4$, a factor often referred to as $q$. We use the CRUST1.0 velocity model to account for the ratio $a/\beta$ to be consistent throughout our study. However, it should be noted that some observations have found that $P$ corner frequencies are higher than $S$ corner frequencies and estimates of $q$ have ranged from 9 to 25 [Boatwright and Fletcher, 1984; Prieto et al., 2004], although often assumed constant in large data set studies [Convers and Newman, 2011].

6.1. Radiated Energy With Depth Phases (Method 1)

Depth phases alter the spectral shape such that the normalization to zero frequency may be biased by the downgoing low-frequency asymptote of the depth-phase Green’s function. Appendix A2 discusses the...
impact of an uninformed spectral normalization, biased interpretation of the source and depth-phase corner
frequency, and possible impacts on radiated energy estimates.

For small earthquakes, which are contaminated with depth phases, we construct a correction as in Denolle
et al. [2015]. From the best fitting source model and the station-averaged depth-phase Green’s functions, we
estimate the stack of the synthetics for all frequencies (from $10^{-4}$ Hz to 10 Hz). We then calculate the radiated
energy from the synthetics (equation (10)) within the finite frequency bandwidth (from $10^{-2}$ Hz to 1.5 Hz) and
that from the best fitting source model (from $10^{-4}$ Hz to 10 Hz). The bias of the depth phases in radiated energy
can be quantified and corrected for using the ratio of the radiated energy estimated from the synthetics with
that from the true model. We use this ratio to correct the radiated energy estimated from the observed stacked
spectra, similar to Denolle et al. [2015]. With this approach, the true absolute amplitude of the stacked spectra is
not necessary to evaluate the radiated energy, provided an independent estimate of the moment is available.

### 6.2. Radiated Energy From the eGf (Method 2)

For the large earthquakes, the corner frequency is not necessarily sampled by the data, but most of the radiated
energy is concentrated around the corner frequency. The eGf approach provides us the true amplitude
level for the spectra but is bandwidth limited in data processing. To remedy this, for instance, Baltay et al.
[2014] extrapolate the spectrum between the amplitude sampled at the lowest reliable frequency to their
independently known seismic moments and apply a similar technique to the high frequencies. Our best fitting
source model serves as a guide to approximately know the distribution of the energy within various frequency
bands. For the frequencies below the low-frequency cutoff (<0.01 Hz) and for the frequencies above the
high-frequency cutoff (>1.5 Hz), we use the area under the source squared-velocity model to construct the total $P$ energy:

$$E_{P\text{tot}} = \frac{2\pi M_0^2 \langle R_0^2 \rangle}{\rho a^5} \left( \int_0^{0.01 \text{Hz}} \| f\hat{\mathcal{M}}_f(f) \|^2 \text{d}f + \int_{0.01 \text{Hz}}^{1.5 \text{Hz}} \| f\hat{\mathcal{M}}_f(f) \|^2 \text{d}f \right) + E_{P\text{.}} .$$

Although our methods address depth-phase interference and data bandwidth issues, our results are still limited and potentially biased by our incomplete sampling of the focal sphere, i.e., the teleseismic $P$ waves we use are restricted to a relatively narrow range of takeoff angles. Other parts of the focal sphere potentially carry more information on finite source effects, which we hope to address in future work. Radiated energy is more commonly calculated for large earthquakes than stress drop because in principle it is less model-dependent and thus a more robust measure of source parameters. However, energy estimates for individual earthquakes often vary considerably among different studies owing to different data processing methods and modeling assumptions. Compared to Baltay et al. [2014], our radiated energy estimates of the 2011 Tohoku, 2004 Sumatra, 2005 Nias, and 2010 Maule earthquakes differ by factors of 0.9, 0.24, 0.41, and 2, respectively.

It is convenient to estimate the apparent stress [Wyss and Brune, 1968]:

$$\tau_a = \mu \frac{E_{P\text{.}}}{M_0},$$

which is often used to compare radiated energy (dynamic) estimates to stress drop (static) estimates as $\tau_a$ has the dimensions of stress.

Figure 12 shows that the distribution of scaled energy is not quite log Gaussian but has a tail for higher values. Nonetheless, we measure a median value of scaled energy of $1.7 \times 10^5 \text{ J/N/m}$, which is very consistent with many studies [Convers and Newman, 2011; Baltay et al., 2014] and corresponds to $\tau_a = 0.7 \text{ MPa}$. There seems to be less variability in scaled energy than in stress drop, but both coefficients of variation, ratio of standard deviation with mean in logspace, are equal.

### 6.3. Scaling of Radiated Energy With Earthquake Size

Figure 15 shows the lack of systematic variation of scaled energy, $E_{P\text{.}}/M_0$, with seismic moment, i.e., that radiated energy scales linearly with moment. However, our data suggest that average spectral shape varies with seismic moment, with the presence of a second time scale or frequency and an intermediate falloff slope of one. If $f_1$ were to follow the scaling $f_1 \propto M_0^{-3}$, we would expect an increase of scaled energy with seismic moment because the area under the spectrum between $f_1$ and $f_2$ grows with moment. Instead, $f_1$ decreases
more rapidly with seismic moment and limits the increase of radiated energy. Both effects roughly cancel out and yield an invariance with moment of the scaled energy. When the spectral shape is the same across earthquake sizes (i.e., when earthquakes are self-similar), the scaled energy is invariant with earthquake size. Here we show that the average spectral shapes vary systematically with earthquake size (i.e., earthquakes are not truly self-similar with increasing moment), but the average scaled energy remains approximately invariant with moment.

6.4. Radiated Energy With Depth
Our global scaled energy estimates suggest a weak depth dependence with an increase occurring at 30–40 km of source depth. It likely correlates with the small stress drop increase observed at similar depths. Locally, we see a depth dependence of scaled energy in the Sumatra and Tohoku regions (Figures 16b and S10b), where there was a high seismicity rate following the giant megathrust events.

6.5. Global Variations of Scaled Energy
Figure 17 shows the global variations of scaled energy, which have clear regional patterns. High radiated energy is explained by a relatively high $f_1$ and/or a relatively high $f_2$ given the earthquake size. Intraplate earthquakes have systematically higher scaled energy than interplate earthquakes, i.e., a log average of approximately $1.0 \times 10^{-4}$ (J/N/m) and apparent stress of $\tau_a = 4$ MPa, compared to $1.5 \times 10^{-5}$ (J/N/m) and apparent stress of $\tau_a = 0.6$ MPa for interplate events. This was first seen by Houston [1990]. This is consistent with the higher stress drop average noted above for intraplate events (note that stress drop is related to $f_1$; the $f_2$ values for the intraplate events are within the normal range, see $T'_2$ in Figure 10b).

Figure 16 shows a closeup of the stress drop and radiated energy variations in the Western Pacific region. Regions of low scaled energy are the Aleutians (average $E_r/M_0 \sim 10^{-5}$ (J/N/m), $\tau_a = 0.4$ MPa), the Solomon and Tonga subduction zones ($E_r/M_0 \sim 10^{-4.9}$ (J/N/m), $\tau_a = 0.5$ MPa). In contrast, earthquakes in Chile/Peru and in Papua New Guinea have relatively high average values of scaled energy, $E_r/M_0 \sim 10^{-4.6}$ (J/N/m), $\tau_a = 1$ MPa. Mexico subduction zone events seems to have relatively high stress drops and high radiated energy, whereas Costa Rica seems to have low stress drops and low scaled energy events. We could not explain the observed along-strike variations as artifacts of uncorrected near-source $Q$ variations (resulting from lateral variation in mantle attenuation); thus, we are confident that these variations are source related. Possible correlation to geodynamical parameters (plate convergence rate, oceanic plate age, etc.) will be the topic of a future study.

7. Discussion
Our study offers a comprehensive analysis and catalog of static and dynamic source parameters for 942 shallow $M_w \geq 5.5$ thrust earthquakes globally recorded between 1990 and mid-2015. We find that teleseismic

![Figure 17. Global variations of scaled energy estimates of thrust earthquakes.](image)
$P$ wave records of these events suffer destructive interference from near-source free-surface reflections, the depth-phases $pP$, and $sP$. This interference alters the shape of $P$ wave spectra by introducing an apparent corner frequency and troughs that are controlled by the source focal mechanism, the source depth, and the source-receiver geometry. We use these properties to constrain focal depth by fitting depth-phase synthetic amplitude spectra to the observed $P$ wave train amplitude spectra. For small earthquakes ($M_w < 7$), where the depth-phase apparent corner frequency is lower than that of the source, we find the best fitting source model by fitting synthetics (depth-phase Green's function convolved with a source model) to the observed $P$ wave train spectra. For larger earthquakes ($M_w > 7$), we use smaller nearby events as empirical Green's functions to remove path effects (true 3-D velocity and attenuation structure and depth-phase effects). Our study does not incorporate the body waves that reverberate in the water column about the source region, often called $pwp$, that are potentially present in our waveforms for the very shallow events [Engdahl and Billington, 1986]. Furthermore, it does not account for the depth extent of large earthquakes when correcting for the point source empirical Green's function, and we show the uncertainties that arise from this approximation in supporting information Figure S13.

After applying both approaches to remove path effects, we find that a double-corner frequency model best fits our source spectral estimates, with an intermediate falloff of $f^{-1}$ between the lower corner frequency $f_1$ and the upper corner frequency $f_2$, and a $f^{-2}$ falloff at higher frequencies. The first corner frequency $f_1$ predicts well the source duration $T_1$, from which we can infer a source dimension. For smaller earthquakes, $f_1$ obeys the commonly assumed scaling of $M_0 \propto T_1^2$, but this changes for earthquakes above $M_{w} \sim 7$ to $M_0 \propto T_1^3$, which we explain by a change in the aspect ratio of the rupture area. Moment-corrected source duration $T_1$ varies weakly with depth, with a small but resolvable decrease near $35–40$ km, similar to what was suggested by Houston [2001]. We use $T_1$ as a measure of source length assuming bilateral rupture (except for Sumatra 2004 and Nepal 2015, which have clear unilateral ruptures). We assume a constant rupture velocity proportional to the $S$ velocity found at the source location, except for the large earthquakes where backprojection analysis provides constraints on absolute rupture velocity. We use empirical scaling relations between fault length and width to find the second source dimension of the rupture area [Leonard, 2010].

$f_2$ should not be confused with $f_{\text{max}}$ [Hanks, 1982; Papageorgiou and Aki, 1983] as our model uses an $f^{-2}$ decay above $f_2$. However, a double-corner frequency source spectrum will have a critical influence on strong ground motion. For single-time-scale source models, most of the seismic energy concentrates around the (single) corner frequency. It is common to assume that a high stress drop event (high single-corner frequency) yields strong ground motions at frequencies that are relevant to urban structures. Our results show that even a low-stress drop event (low first corner frequency) with a high value of $f_2$ can yield strong ground motion at high frequencies.

The second corner frequency $f_2$ is generally too high to be related to a second fault dimension (width). The meaning of this second time scale remains to be investigated. Our main results are that $f_2$ scales with moment as $M_0 \propto f_2^2$ and is weakly depth dependent. One possibility is that we are sampling the source time function such that we capture the short time scale of the starting and stopping phases that alter the shape of the far-field pulse at steep takeoff angles [Madariaga, 1979]. Since we constructed the double-corner frequency model by modifying the Haskell propagating source model, a second possibility is that the second time scale $T_2$ represents the rise-time (pulse duration) of pulse-like ruptures in some cases, although $T_2$ is often short compared to the values reported in finite source inversion studies. A third possibility comes from slip-weakening mechanisms, [e.g., Ohnaka and Yamashita, 1989; Tinti et al., 2005]. The slip-rate function increases and decreases at different rates, and the first phase often is faster than the second phase. Weakening mechanisms with onset of slip can control the acceleration of slip through a short interval. It is possible that $T_2$ relates to this acceleration phase, time from zero-to-peak slip rate. Tinti et al. [2005] proposed a relation between average slip, source duration, acceleration time scale, and critical slip weakening distances, which resembles that of Ohnaka and Yamashita [1989]. The asymmetry in the slip-rate function has also been parameterized from dynamic source models [Guatteri et al., 2004; Liu et al., 2006], correlated to source duration [Schmedes et al., 2010], and is now used in ground motion prediction [Graves and Pitarka, 2010]. Galetzka et al.'s [2015] inversion of near-field measurements of the $M_{w}7.8$ 2015 Nepal earthquake to infer slip history on the fault also finds asymmetry in the slip-rate function and measured the time-to-peak slip rate to be short, on the order of 2 s.
To explain the break of scaling in source duration, we propose an increasing fault aspect ratio for larger earthquakes. We use an elliptical source model, for which we have a closed-form solution between the average slip, the short source dimension, and the static uniform stress drop, for a mode II crack embedded in a homogeneous whole space [Eshelby, 1957]. This method allows us to have a consistent relation between slip and stress for both the popular circular crack model and the elongated source models. Unlike strike-slip faults, the relation does not vary much with aspect ratio for dip-slip faults. It allows us to estimate static stress drop, with all the assumptions mentioned above. We find an average value of 2 MPa for the entire data set, which is very consistent with other studies [Allmann and Shearer, 2009]. Despite our increased fault aspect ratio with seismic moment (i.e., earthquakes are not perfectly self-similar), we also find no systematic scaling of stress drop with seismic moment. Thus, invariant stress drop does not necessarily imply self-similarity because it can, as in our study, result from systematic variations in fault geometry.

The presence of an intermediate falloff rate that spans a wider frequency band for larger earthquakes represents a change in the spectral shape, which also violates true earthquake self-similarity. However, the break of scaling of $f_1$ is compensated by the scaling of $f_2$ such that the average scaled energy remains approximately invariant with seismic moment. Assuming a near-source shear modulus of $\mu = 40$ GPa, we find an average apparent stress $\tau = \mu E_\text{r}/M_0$ of 0.7 MPa, which is comparable to values found in other studies [Ye et al., 2016a] and from other earthquakes [Baltay et al., 2010, 2011; Convers and Newman, 2011; Baltay et al., 2014]. For simple models of slip weakening, the radiation efficiency, which is the ratio of radiated energy over available energy, can be expressed as twice the ratio of apparent stress to static stress drop [Husseini, 1977; Venkataraman and Kanamori, 2004]. Our high values of radiation efficiency, often greater than one, either suggest that we underestimate the static stress drop (overpredict the fault area) or that our findings suggest undershoot [Venkataraman and Kanamori, 2004; Beeler, 2006; Noda et al., 2013]. The conventional approach to estimating fracture energy [Abercrombie and Rice, 2005] from estimates of stress drop, radiated energy, and slip (from moment and fault area) often yields negative values of fracture energy (rarely discussed in the literature). Nonetheless, high values of apparent stress compared to stress drop may indicate strong dynamic weakening mechanisms [Viesca and Garagash, 2015]. Although some negative fracture energy results might be expected due to scatter from measurement uncertainties, this issue deserves more attention. Methodologies to estimate radiated energy and stress drop should be revisited to account for the full energy budget of shallow earthquakes near the free surface. Stress drop must decrease at the surface or strain drop would become infinite.

8. Conclusions

In summary, we analyze teleseismic P spectra for over 942 earthquakes using a systematic and self-consistent approach that accounts for depth phases and limited data bandwidth. Our main findings are as follows: (1) a double-corner frequency model for large global thrust earthquakes, (2) a lower corner frequency related to source duration, which exhibits a break in scaling above $M_0$ 7–7.5 and that we explain by likely changes in fault geometry, (3) an upper corner frequency suggesting a shorter time scale unrelated to source duration, which exhibits its own scaling relation, (4) the invariance of average stress drop with respect to moment and changes in rupture aspect ratio, and (5) the invariance of average scaled energy with respect to moment despite systematic changes in spectral shapes. Although we find that average stress drop and scaled energy do not vary with moment, our results do not imply complete earthquake self-similarity for these shallow thrust earthquakes, as we find that the spectral shapes and inferred fault aspect ratios change with increasing moment. We find only a weak dependence of source parameters on depth, consistent with the findings of Ye et al. [2016a].

There are many opportunities for future work to test and build on these findings. We limited our study to reverse-faulting events because they include most of the very largest earthquakes but will expand to other source mechanisms in the future. We also only examined P waves because attenuation prevents reliable observations at high frequencies for S waves, but for large earthquakes with relatively low corner frequencies it should be possible to apply our methods to resolve S wave source spectra. We limited our epicentral distance range to avoid complicated wave propagation effects, but in principle the eGf approach can correct for these effects, meaning that one need not restrict the data selection to specific takeoff angles. There is thus potential to improve not only the azimuthal coverage but also the takeoff (vertical) angle coverage, providing additional information on source directivity and other finite source effects.
Figure A1. Comparison of our estimates of focal depths with the PDE (blue circles) and CMT (red circles) catalogs. A line of unit slope is also indicated. There is a slight systematic bias toward overpredicting the depth of the CMT and PDE that could be explained by our approximate homogeneous half space above deep sources, the use of the PREM Earth model for the CMT solution [Ekström et al., 2012], and the bias of the PDE solutions due to the body-wave reflections in the water column [Engdahl and Billington, 1986].

Appendix A: Miscellaneous Investigations of Depth Phases

A1. Constraining Focal Depth With Depth Phases

The presence of the notches in the depth-phase spectrum provides an interesting tool to constrain focal depth [e.g., Warren and Shearer, 2005]. Denolle et al. [2015] used such an approach to constrain the source depth of the Nepal earthquake and its largest aftershocks. The source is embedded in a homogeneous purely elastic (attenuation free) half space with elastic properties of a Poisson medium \( \alpha = \sqrt{3} \beta \), P wave speed \( \alpha \), and shear wave speed \( \beta \). In this example, the elastic shear wave speed is taken as \( \beta = 3900 \text{ km/s} \), and the density is \( \rho = 3000 \text{ kg m}^{-3} \). In this study, we estimate the source depth with the following steps: construct a toy model source spectrum with a corner frequency that follows \( f_c = 2.3 \times 10^4 \text{ M}_0^{1/3} \) (similar to Allmann and Shearer [2009]); construct the depth-phase spectrum at each receiver for various source depths \( H \); estimate the residual at each receiver between the observed and predicted spectra using our spectral fitting algorithm for all \( H \); stack the misfit functions for each station; and find \( H \) that minimizes the stacked misfit function using a simple grid search.

This approach works best for small and moderate earthquakes that can be well approximated as point sources for short- and long-period body waves. It assumes a homogeneous half space in the vicinity of the source, which is, of course, a crude approximation given the complex velocity structure in some subduction zones. The stacked misfit function is generally sensitive enough to find a well-constrained source depth. For the larger events, where the fault dimension is large compared to the frequencies, the misfit function is not as well behaved, and we restrict our grid search to \(-15 \text{ km} \) and \(+15 \text{ km} \) from the GCMT depth. We show in Figure A1 the distributions of our focal depth estimates compared to those of the preliminary determination of epicenters (PDE) and GCMT catalogs.

Once the depth is constrained, we use CRUST1.0 [Laske et al., 2013] to extract the elastic parameters \( (\rho, \alpha, \beta) \) for an attenuation-free medium. CRUST1.0 provides a high-resolution (1°) velocity model of the crust for a layered structure. For each source location, we extract a velocity and density profile from CRUST1.0 and linearly interpolate the elastic properties to the source depth.

A2. Effects of Depth Phases on Source Spectra and Energy Estimates

We have explored the effects of the destructive interference between direct P and the depth phases pP and sP. Our canonical example is a thrust event of magnitude 7 at 10 km depth with a pure thrust motion on a 15° dipping fault. We set up the problem using the media properties from the previous section. We assume that P and S have the same source spectrum \( \tilde{S}(f) \).

Figure A2a shows the stick seismograms for such a source and a receiver located at 46° azimuth and a takeoff angle of 29.4°. The stick seismograms show that P and sP have opposite polarity and that pP is weak. The resulting amplitude spectrum \( |\tilde{S}_{pP}(f)| \) is shown in Figure A2b, along with the \( |\tilde{S}(f)| \) and the best fit model
Figure A2. (a) Stick seismogram normalized to the absolute amplitude of direct $P$ for a pure thrust moment tensor with 15° dip, a receiver at 46°, and a takeoff angle of 29.4°. (b) True source model (equation (3)) and its corner frequency $f_P = 0.01$ Hz, blue synthetic (depth-phase stick seismogram convolved with source model), red the single-corner frequency model best-fitting the synthetic and the biased corner frequency $f_{gP}$ within two standard deviations of the estimates (shaded red area).

This example illustrates that corner frequency estimated from a direct fit of $|\tilde{S}_{gP}(f)|$ is biased by the apparent corner frequency $f_{gP}$ of the depth-phase Green’s function and overpredicts the true source corner frequency $f_P$.

The degree of overlap between the depth phases is controlled by the difference in arrival time of the phases, their relative amplitudes, and the source pulse width. Source depth $H$ controls the arrival time of the different phases in the $P$ wave train. Radiation pattern and source depth control the relative amplitudes between the phases. Earthquake source size controls the source pulse width and can be approximated for the circular crack case as $a = k\beta/f_c$ [Brune, 1970; Madariaga, 1976; Kaneko and Shearer, 2014], where $f_c$ is inversely proportional to the pulse width $T$. We quantify the biases in corner frequency estimates and radiated energy with respect to the ratio of the earthquake depth over its size with the parameter $\eta = H f_p / k\beta$.

We continue this analysis by randomly selecting takeoff angles between 16° and 37° and azimuths between 0° and 360°. We quantify the bias in confusing $f_{gP}$ for $f_p$ by estimating a distribution of ratios $f_{gP}/f_p$. Figure A2c

Figure A3. Bias in estimating radiated energy from the single-station approach (equation (10)) from the depth-phase contaminated spectra ($\mathcal{E}_{gP}$) with the $P$ wave train amplitude spectra compared to the true radiated energy $\mathcal{E}_P$ against $\eta$. The estimates of energy for shallow and/or low stress drop events are largely overpredicting the true energy and need to be accounted for.
shows the ratio for various $\eta$ and highlights the bias for large shallow earthquakes. The altered spectrum greatly affects our measure of radiated energy.

With the single-station approach to estimating radiated energy, the normalization of the spectra to its low-frequency asymptote is clearly affected by the destructive interference of the depth phases with the direct phase. The bias in normalizing the spectrum to its apparent low-frequency asymptote biases the estimates of radiated energy. Figure A3 shows that for a thrust earthquake, the bias can overpredict the energy up to a factor of 10 for shallow and/or large events and underpredict the energy by up to 70% for deep and/or small events.

**Appendix B: The “C” Factor in Stress Drop Estimates**

In these notes, we describe the factor that relates a uniform stress drop from the fault-averaged slip for a buried crack embedded in a homogeneous whole space. We define this relation following Eshelby’s [1957] work:

$$\Delta \sigma = C \mu \frac{s}{b}$$

(B1)

$$= C \frac{M_0}{\pi ab^2}.$$  

(B2)

For a circular crack problem ($a = b$), $C = 7\pi/16$, a result widely applied in the community to estimate stress drop when the geometry of the fault is not known and simply assumed circular. Published studies do not always make clear what model they are using; here we reaffirm that we use a uniform stress drop and not a uniform slip model.

We start with equation (5.7) from Eshelby [1957], which provides the relation between a uniform stress drop and an elliptical profile of slip bounded by the maximum slip offset $s$ on the crack for an ellipse of axis $a$ and $b$ ($a > b$) along the $x$ and $y$ axis of the crack plane:

$$\frac{b \Delta \sigma}{\mu \eta} \sqrt{1 - x^2/a^2 - y^2/b^2} = \frac{1}{2} s, \quad (B3)$$

where $\eta$ is a function of Poisson’s modulus $\nu$ and aspect ratio $b/a$ given by equation (5.3) [Eshelby, 1957]. The slip distribution on the fault is elliptical. We perform an surface integral over an ellipse $A$ using the ellipse equation ($x'^2/a^2 + y'^2/b^2 = 1$), a change of variables ($x' = x/a$, $y' = y/b$) and finally a change of variable from Cartesian to polar ($dx'dy' = r dr d\theta$):

$$\int \int_A \sqrt{1 - x^2/a^2 - y^2/b^2} dA = \frac{2\pi ab}{3}. \quad (B4)$$

We average the slip over an elliptical surface (right-hand side of equation (B3)):

$$\int \int_A \frac{s}{2} dA = \frac{\pi ab \bar{s}}{2}. \quad (B5)$$

Integrating both sides of equation (B3) on an elliptical surface and using $M_0 = \mu b \pi ab$,

$$\frac{b \Delta \sigma \pi ab}{\mu \eta} \frac{2\pi ab}{3} = \frac{\pi ab \bar{s}}{2} \quad (B6)$$

$$\Delta \sigma = \frac{3}{4} \frac{M_0}{\pi ab^2} \eta \quad (B7)$$

$\eta$ is a function of aspect ratio $a/b$ and of the complete elliptic integrals $E(k), K(k)$ [Eshelby, 1957, equation (5.3)]:

$$\eta = E(k^2) + \frac{\nu}{1 - \nu} \frac{E \left( k^2 \right) - k^2 K \left( k^2 \right)}{k^2}; \quad k = \sqrt{1 - b^2/a^2}, \quad (B8)$$

$$\eta = \pi \frac{2 - \nu}{4(1 - \nu)}; \quad b = a. \quad (B9)$$
The data used can be found at IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to waveforms, related metadata, and/or derived products used in this study. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience and EarthScope (SAGE) Proposal of the National Science Foundation under cooperative agreement EAR-1261681. The data used can be found in doi:10.1002/2014GL062079.


