Geophysical Research Letters

RESEARCH LETTER
10.1029/2020GL089342

Key Points:
- We used a Large-*n* Seismic Array to investigate the robustness of direct wave spectral estimates
- *P* wave single spectra reveal azimuthal dependency of corner frequency and long- period spectral amplitudes
- *P* wave single spectra and spectral ratio corner frequency estimates are highly biased by using a small number of stations

Supporting Information:
- Supporting Information S1

Correspondence to:
K. B. Kemna,
kilian.kemna@rub.de

Citation:
Kemna, K. B., Peña Castro, A. F., Harrington, R. M., & Cochran, E. S. (2020). Using a Large-*n* Seismic Array to explore the robustness of spectral estimations. *Geophysical Research Letters*, *47*, e2020GL089342. https://doi.org/10.1029/2020GL089342

Received 15 JUN 2020
Accepted 15 OCT 2020
Accepted article online 22 OCT 2020

Using a Large-*n* Seismic Array to Explore the Robustness of Spectral Estimations

K. B. Kemna[1], A. F. Peña Castro[1,2], R. M. Harrington[1], and E. S. Cochran[3]

[1]Institute of Geology, Mineralogy and Geophysics, Ruhr University Bochum, Bochum, Germany, [2]Department of Earth and Planetary Sciences, McGill University, Montréal, Quebec, Canada, [3]U.S. Geological Survey, Earthquake Science Center, Pasadena, CA, USA

Abstract Spectral analysis is widely used to estimate and refine earthquake source parameters such as source radius, seismic moment, and stress drop. This study aims to quantify the precision of the single spectra and empirical Green's function spectral ratio approach using the Large-*n* Seismic Survey in Oklahoma (LASSO) array. The dense station coverage in an area of local saltwater disposal offers a unique opportunity to observe and quantify radiation pattern effects and subsequent precision of spectral estimates of small earthquakes ($M < 3$). The results suggest that the precision of source properties estimated from direct phase arrivals for arrays with less than 20 stations should be assumed to be not less than 30% and could be as high as 150% if less than five stations are used. Furthermore, we do not see clear evidence for, or against, a scaling of stress drop with magnitude of small earthquakes ($M < 3$) as observed by other studies.

Plain Language Summary Seismologists use ground motion recordings of seismic waves (seismograms) to infer details about the earthquake rupture process. While large earthquakes often generate a physical imprint on the earth's surface through surface rupture, small earthquakes can often only be studied from seismograms. Nevertheless, small earthquakes are of particular interest to learning about the rupture process for many reasons. For example, they are much more numerous than larger magnitude earthquakes and might rupture by the same physical process(es). While seismic arrays are usually restricted to a few to tens of stations, here we use a very large seismic array with >1,800 temporary stations to study small earthquakes and how station resolution may bias source property estimates. Source properties include the physical size of the rupture surface and the corresponding slip on that surface, which relate to the amount of stress released by the earthquake. The large number of stations allows us to estimate the source properties in unique detail and test the variability in measurements using different numbers of stations to estimate the precision. We find that the estimation of source properties is highly biased when using a small number of stations (<20), which should be taken under consideration in future studies.

1. Introduction

The shear stress released during an earthquake is a fundamental parameter that describes the rupture process and controls resulting ground motions. Despite the relatively straightforward approach of estimating the average decrease in shear stress (stress drop) across a fault before and after an earthquake from the spectral corner frequency (e.g., Boatwright, 1980; Brune, 1970; Madariaga, 1976), the natural high variability of stress drop estimates (around 0.1–100 MPa; e.g., Hanks, 1977), and the significant discrepancies between different methods for the same data sets (e.g., Abercrombie, 2014, 2015; Neely et al., 2020; Shearer et al., 2019) contribute to the ongoing controversy regarding the self-similarity of earthquakes. Deconvolving the trade-off between the degree to which earthquakes vary in their stress-drop scaling (e.g., Abercrombie, 1995, 2013; Aki, 1967; Kanamori et al., 1993; Harrington & Brodsky, 2009) versus observational artifacts of spectral estimation methods is critical to understanding earthquake rupture processes and rests on constraining the resolution of station coverage needed to constrain source parameters robustly.

Most spectral estimation methods assume that a seismogram $s(t)$ can be written as a convolution of the earthquake source $e(t)$, path propagation effects $G(t)$, and the instrument response $I(t)$. The individual phase spectrum can be expressed as

$$s(t) = e(t) * G(t) * I(t), \quad (1)$$

where removing the $I(t)$ term permits the estimation of the true long-period spectral amplitude, which is proportional to magnitude. Spectral corner frequency (f_c) estimates, from which we estimate the static stress drop, depend strongly on the separation between the source $e(t)$ and path $G(t)$ terms. Both direct (P/S wave) phases and coda (typically S wave) waves can be used to estimate f_c (e.g., Abercrombie, 2013; Mayeda & Walter, 1996; Mayeda et al., 2003), and a number of prior studies show wide variation in f_c estimates depending on the method used to constrain them (Ide et al., 2003; Neely et al., 2020; Shearer et al., 2019). Several methods use multiple stations to estimate individual event f_c values using either stacked spectra or station-averaged estimates in efforts to remove statistical scatter (e.g., Abercrombie, 1995; Shearer et al., 2006; Viegas et al., 2010), where the use of coda waves typically shows higher precision (e.g., Mayeda et al., 2007). However, multiple station estimates will not reduce uncertainties related to limited azimuthal coverage (e.g., Abercrombie, 2015; Kaneko & Shearer, 2015; Kane et al., 2011; Prieto et al., 2006). In addition, the number of viable station estimates is naturally coupled to the size of the earthquake because of signal-to-noise-ratio (SNR) limitations, which causes more substantial uncertainties. The small sample size problem resulting from limited station coverage is inherent to most seismic deployments and pervasive in several academic fields (e.g., Spall, 2002; Steele et al., 1993). Nevertheless, the study of small, more numerous, earthquakes ($M < 3$) and their source parameters is paramount in understanding the rupture process, particularly where their scaling properties could be transferable to larger-magnitude, less-frequent earthquakes (Brodsky, 2019).

Here we use the Large-n Seismic Survey in Oklahoma (LASSO, Dougherty et al., 2019) to investigate the robustness of source property estimates, particularly the effect of number of stations and azimuthal coverage, of earthquakes with $M < 3$, including potential radiation pattern effects, where the corner frequency f_c is predicted to vary over the focal sphere (Madariaga, 1976). The array operated 1,829 vertical-component 500-Hz sampling nodal seismometers with a nominal station spacing of around 400 m for a period of 1 month in spring 2016 in Grant County, an area known for extensive operation of wastewater-disposal wells related to injection-induced seismicity (e.g., Cochran et al., 2020; Dougherty et al., 2019). We use the superior resolution afforded by the LASSO array to investigate the robustness and variability of P wave f_c estimates from two independent spectral methods, the single spectra and empirical Green's function spectral ratio, and to quantify the precision as a function of varying station coverage for small earthquakes ($M < 3$). We do not include S wave and/or coda wave estimates due to the difficulty of determining S wave arrivals on the vertical component seismic stations.

2. Method

We consider the earthquakes listed in the Cochran et al. (2020) catalog, which contains the 1,104 events within 5 km of the LASSO array footprint that generated automatic P wave arrivals on the vertical-component stations with a magnitude range of 0.01–3.0 (brown dots in Figure 1). Numbered events represent earthquakes with well resolved focal mechanisms and source parameter estimates and are listed in Table S1. Supporting information also contains full details of the picking procedure.

We use two independent methods in the frequency domain to estimate f_c and long-period spectral amplitude (Ω_0), which is then converted to seismic moment M_0 (see Equation 3; cf. Figure S3). The first method (referred to as "single-spectral estimates") uses the theoretical displacement spectra $\Omega_t(f)$:

$$\Omega_t(f) = \frac{\Omega_0 e^{-(\pi f t/Q)}}{[1 + (f/f_c)^{\gamma n}]^{1/\gamma}}, \tag{2}$$

described by the long-period spectral amplitude Ω_0, corner frequency f_c, frequency fall-off rate n, frequency dependent quality factor Q, and γ, which controls the shape of the corner (e.g., Abercrombie, 1995; Boatwright, 1980; Brune, 1970). We estimate displacement spectral amplitude using the multitaper approach implemented by Prieto et al. (2009) from 0.8-s time windows for each available P wave arrival (starting 0.2 s before the arrival). We resample the spectra in the log domain to 100 constant $\log(f)$ increments and use a least squares algorithm (Newville et al., 2014) to fit Equation 2 on individual nodes and averaging estimates over all nodes that meet the quality control criteria. Using a Boatwright-type model ($\gamma = 2$) with $Q = 600$ and $n = 3.5$ produces the lowest overall residuals after extensive testing of different models (Figure S4; see supporting information for parameter fitting details, including details of the instrument response and filtering; Dougherty et al., 2019).

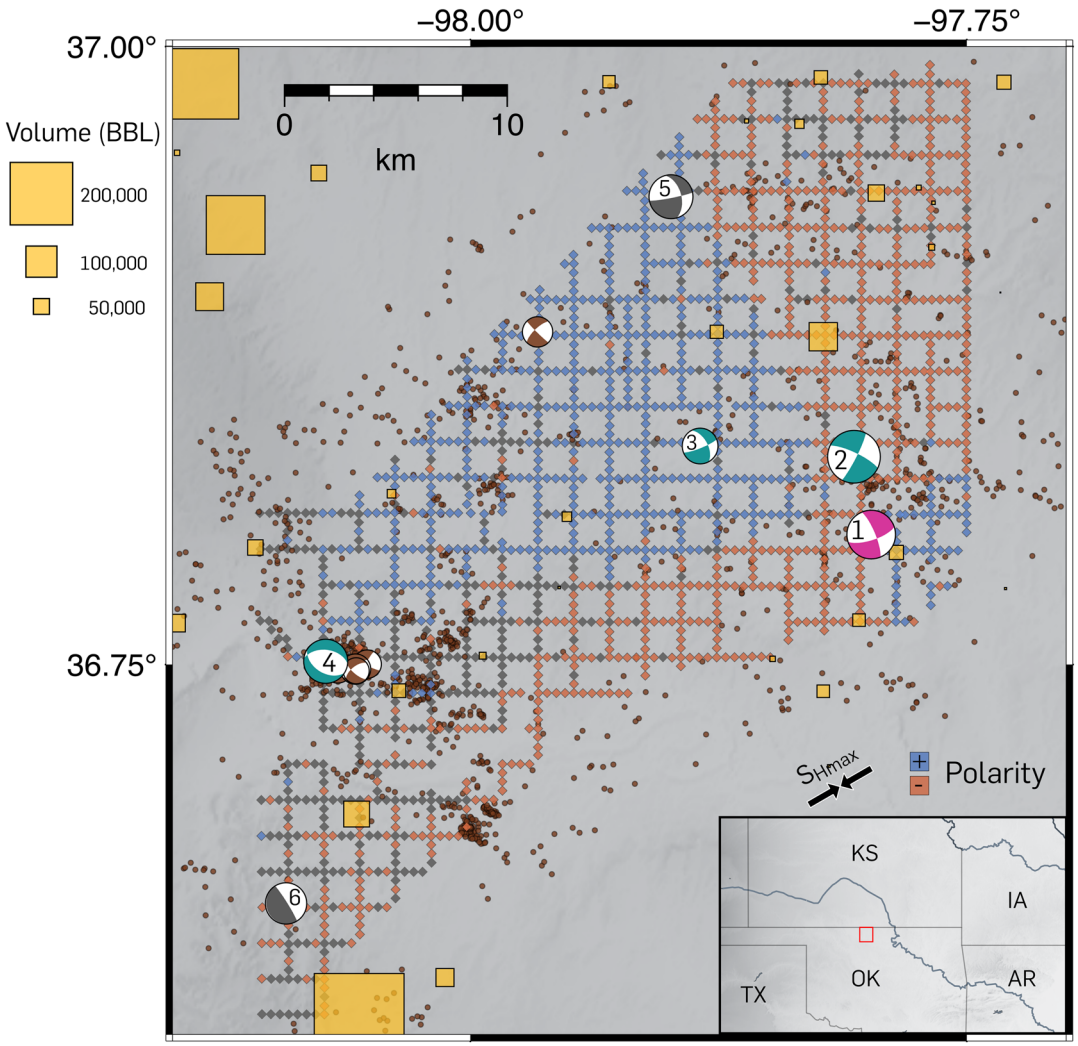

Figure 1. Overview of the Large-*n* LASSO array in northern Oklahoma, with inset showing the regional location. Diamonds denote deployed stations, brown circles show recorded seismicity (Cochran et al., 2020), and gold rectangles show injection wells scaled by cumulative injection volume during the deployment (OCC, 2020). Focal mechanisms for larger events were computed using HASH (see supporting information). Numbered earthquakes have robust focal mechanism and source parameter estimates and are listed in Table S1. Cyan and magenta focal mechanisms denote earthquakes for which single-spectra estimates are discussed in the text, and gray focal mechanisms denote main events of spectral ratio pairs. Stations (nodes) indicate the *P* wave polarity of the magenta focal mechanism (blue = positive, red = negative, gray = undefined).

We estimate the seismic moment M_0 from the single-spectrum Ω_0 estimates (Equation 2) with the following equation:

$$M_0 = \frac{4\pi\rho c^3 R\Omega_0}{U_{\phi\theta}}, \tag{3}$$

where ρ and c are the density and *P* wave velocity at the earthquake hypocenter, R is the hypocentral distance, and $U_{\phi\theta}$ is the mean radiation pattern coefficient (0.52 for *P* waves Madariaga, 1976). We assume a value of $\rho = 2,700$ kg/m^3 and use the *P* wave velocity at the hypocenter indicated by the local 1-D velocity model used for location estimates (Rubinstein et al., 2018).

The second method for estimating spectral corner frequency is based on a spectral ratio approach (Hartzell, 1978; Hough, 1997), in which we stack normalized event spectral ratios at all nodes for which the quality control criteria were met (see supporting information for details). In general, spectral ratio estimates allow better constraints on f_c values by minimizing the imprint of nonsource related terms, such as the $G(t)$ and $I(t)$ terms in Equation 1, on the spectral shape. The following analytical expression, derived

from Equation 1 for two different events (1 and 2), demonstrates how the nonsource related terms are the same between colocated pairs of earthquakes and therefore cancel in the spectral ratio:

$$\Omega_r(f) = \Omega_{0r}\left[\frac{1+(f/f_{c2})^{\gamma n}}{1+(f/f_{c1})^{\gamma n}}\right]^{1/\gamma}, \quad (4)$$

where f_{c1} and f_{c2} are the corner frequencies of the larger (target) event and the smaller (empirical Green's function, eGf) event, respectively (e.g., Abercrombie, 2014; Viegas et al., 2010). We use the same fitting procedure to estimate the f_c values in Equation 4 but retain the Ω_0 values from single spectrum fits, as the long-period spectral amplitude is less sensitive to attenuation effects in single-spectrum fitting. The parameter γ controls the shape of the corner as in Equation 2. We identify 90 candidate event pairs for the spectral ratio fitting using a three step approach. Supporting information provide details regarding the spectral approach, including event pair identification, and parameter estimation.

Following spectral fitting, we then estimate static stress drop $\Delta\sigma$ using the single-spectrum M_0 estimates with the f_c estimates from both single spectra and spectral ratios, where available. We assume a circular crack model (Eshelby, 1957; Madariaga, 1976), where r denotes the crack radius, and β the depth-dependent seismic shear wave velocity:

$$\Delta\sigma = \frac{7}{16}\frac{M_0}{r^3}, \quad (5)$$

$$r = \frac{k\beta}{f_c}. \quad (6)$$

We use $k = 0.38$ following Trugman et al. (2017) to facilitate in comparing our stress drop values to estimates in southern Kansas.

For each event with single-spectra estimates on >40 stations, we calculate the mean value and confidence intervals for f_c, M_0, and $\Delta\sigma$ using the delete-one jackknife mean from all station estimates (Prieto et al., 2007). For f_c and $\Delta\sigma$ values estimated from spectral ratios, we stack all available station ratios (using a minimum of 5) and calculate stress drop values using the f_c estimates from target events (and eGf events, in cases where both f_c estimates fall within the bandwidth of high SNR). Tables S1 and S2 provide spectral parameter estimates for the numbered earthquakes in Figure 1 for which focal mechanism solutions exist, and additional details of precision estimation are included in the supporting information.

3. Results

We were able to recover single spectra source property estimates for 336 events with a total of 34,813 individual station fits, as well as corner frequency estimates from eGf spectral ratios for 126 events (70 target events and 56 eGf events). The single spectrum station estimates in Figures 2a and 2b (as well as Figure S5) show deviations of f_c and M_0 values estimated on individual nodes for two representative events with well-constrained focal mechanisms and hypocentral depths of 6.1 and 4.5 km (Events 1 and 2 in Figure 1 and Table S1, respectively). Corner frequency estimates (Figure 2a) are higher than the mean in the nodal plane directions (yellow dashed lines in a), whereas M_0 values are generally lower (Figure 2b). The inset plot a1 shows individual spectra scaled to a uniform Ω_0 value that also demonstrate the same trend of higher relative f_c and lower M_0 values in the nodal plane direction. The lower relative M_0 values are apparent in the unscaled spectra (Figure 2, inset b1).

The eGf spectral ratio approach shows a similar variation in spectral shape depending on the number of stations used. For example, Figure 2c (as well as Figure S6) shows that the f_c estimates vary significantly for small sample sizes (<10 stations). Increasing the number of stations (>20) significantly reduces the scatter in the stacked ratios and resulting corner frequency estimates. For both earthquakes (Event 5, Figure 2c; Event 6, Figure S6), comparison of the red and black dashed lines qualitatively shows whether the Boatwright (Event 5) or the Brune (Event 6) spectral ratio model provides a better fit to the data.

Uncertainties from f_c estimates are cubed in the static stress drop calculation (Equation 5), meaning that scatter in f_c values (Figure 2) propagates to the resulting stress drop estimates. Figure 3 shows the deviation in static stress drop $\Delta\sigma$ for all events with single-spectrum estimates using at least 10 stations in each quadrant and for all spectral ratio pairs with at least five station ratios in each quadrant. Each shaded rectangular pixel shows $\Delta\sigma$ as the relative difference between the delete-one jackknife mean (stacked ratio) using all stations

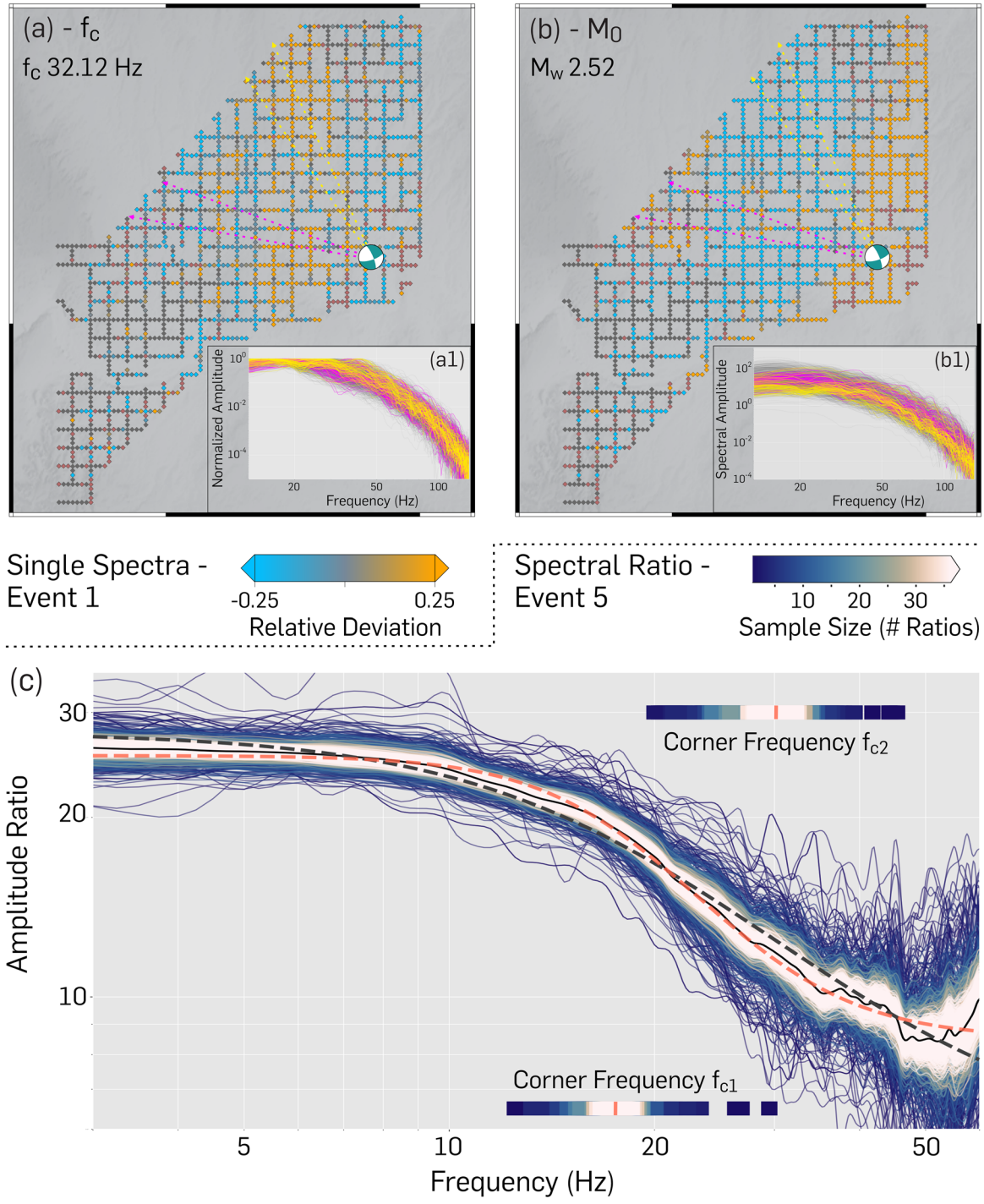

Figure 2. (a, b) Single-spectrum estimates of f_c (a) and M_0 (b) from individual stations fits for Event 1 (Figure 1). Symbols and map are the same as in Figure 1. Colored nodes show the relative deviation in (a) f_c and (b) M_0 (blue = negative, orange = positive) of individual station measurements from the jackknife mean using all available stations (red nodes indicate SNR criteria not satisfied; nodes indicate no *P* wave pick). Insets show multitaper spectral estimates (normalized = a1; original = b1) at individual stations satisfying SNR criteria. Most spectra are shown in gray and subsets of stations inside the dashed colored lines (10 aperture angle) are shown in yellow and magenta, respectively. Spectra shown in the insets of a1 are normalized to the largest spectral amplitude to demonstrate differences in corner frequencies. Spectra shown in the insets of b1 are not scaled and demonstrate differences in the long-period spectral amplitude. (c) Stacked spectral ratios and corresponding spectral fits with corner frequency estimates of Event 5 (Figure 1) using random subsets of station ratios (50 random subsets per sample size). The solid black line denotes the stacked spectral ratio using all available stations. The red and black dashed lines are the Boatwright, and Brune model fits, respectively, with $n = 2$ for the former (where Brune fit is not used). Stacked subsets of available spectra are colored by sample size (number of stations). Colorbars at the upper and lower right show the corner frequency estimates (vertical red bars) and the corresponding range governed by the range of sample sizes (f_{c1} = main event, f_{c2} = eGf event). Colorbar color scheme is the same as the stacked ratios.

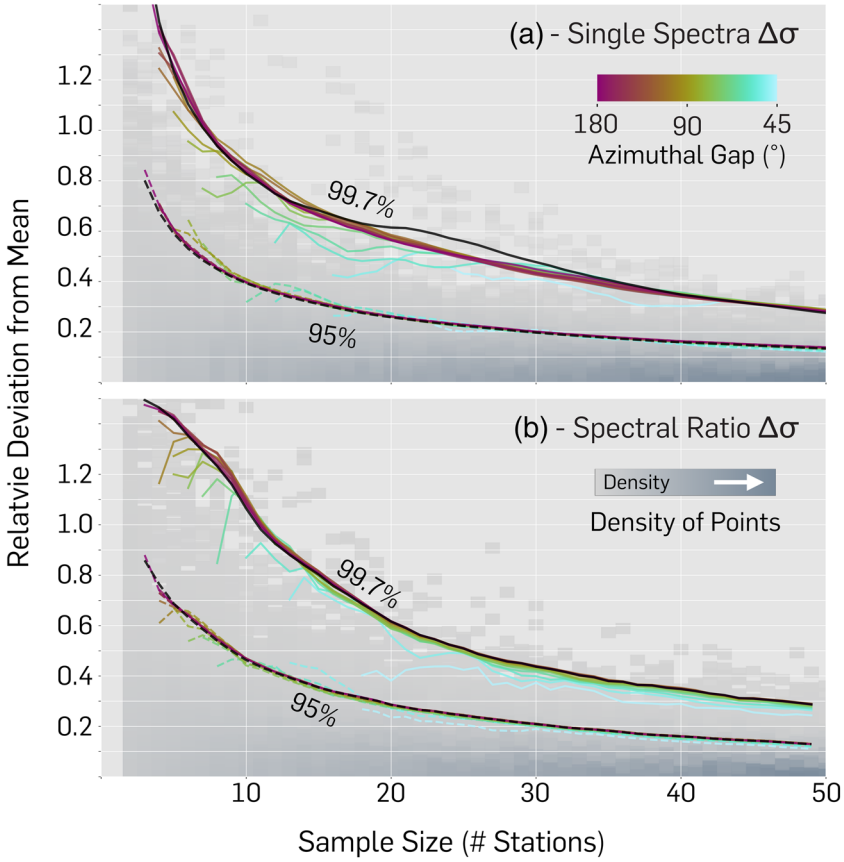

Figure 3. (a) Deviation of single-spectrum stress drop estimates for all events with 10 single station estimates in each quadrant (76 events) using different random sample sizes (1,000 iterations per sample size). The deviation is written as a fraction of the mean using all stations $(\Delta\sigma_{subset} - \Delta\sigma_{all})/\Delta\sigma_{all}$). The grayscale colorbar represents the density of points around a given estimate. The solid, dashed black line shows the 95%, 99.7% confidence bound for each sample size as a rolling mean of three samples, respectively. Colored lines corresponding to the colorbar (top right) illustrate confidence boundaries based on a maximum azimuthal gap of the subset of stations as a rolling mean over four samples. (b) Deviation of the spectral ratio stress drop estimate (f_{c1}) for all spectral ratio pairs with at least five station ratios in each quadrant (21 pairs) using different random sample sizes of stacked station ratios (1,000 iterations per sample size). Same colors and symbols apply as in (a), where the stacked ratio fit using all station ratios for a single event is used as the baseline estimate. The seismic moment is fixed in both plots to the jackknife mean estimate from the single spectra.

and the mean (stacked ratio) of a randomly selected subset of stations as a function of sample size for single spectra and spectral ratios, respectively. Figure 3 demonstrates that the deviation from the mean estimate steadily decreases with an increasing number of stations, which could also be related to the decrease of the maximum azimuthal gap of a given subset of stations (indicated by the colored lines). The distribution of lines suggests that if the number of stations is between 5 and 20, setting a maximum allowable azimuthal gap (e.g., 45–90) can significantly improve the precision in the stress drop estimate. The distribution also suggests that restricting the azimuthal gap with increasing sample size (above 30 stations) brings no added benefit to reducing the range of estimates. The higher density rectangular pixels (darker gray shades) also suggest that individual estimates of random subsets become more tightly clustered around the mean as the station count surpasses 20. The deviation decreases to 30% (2 σ, 95% confidence interval) or 60% (3 σ, 99.7% confidence interval) with 20 to 30 stations, and to 15% (2 σ) or 30% (3 σ) with 50 stations. Further increasing the number of stations beyond 50 does not significantly improve the precision. Figure 3b suggests that for events with <20 station ratios, requiring a maximum azimuthal gap <90 significantly decreases the range in $\Delta\sigma$ estimates to <40%.

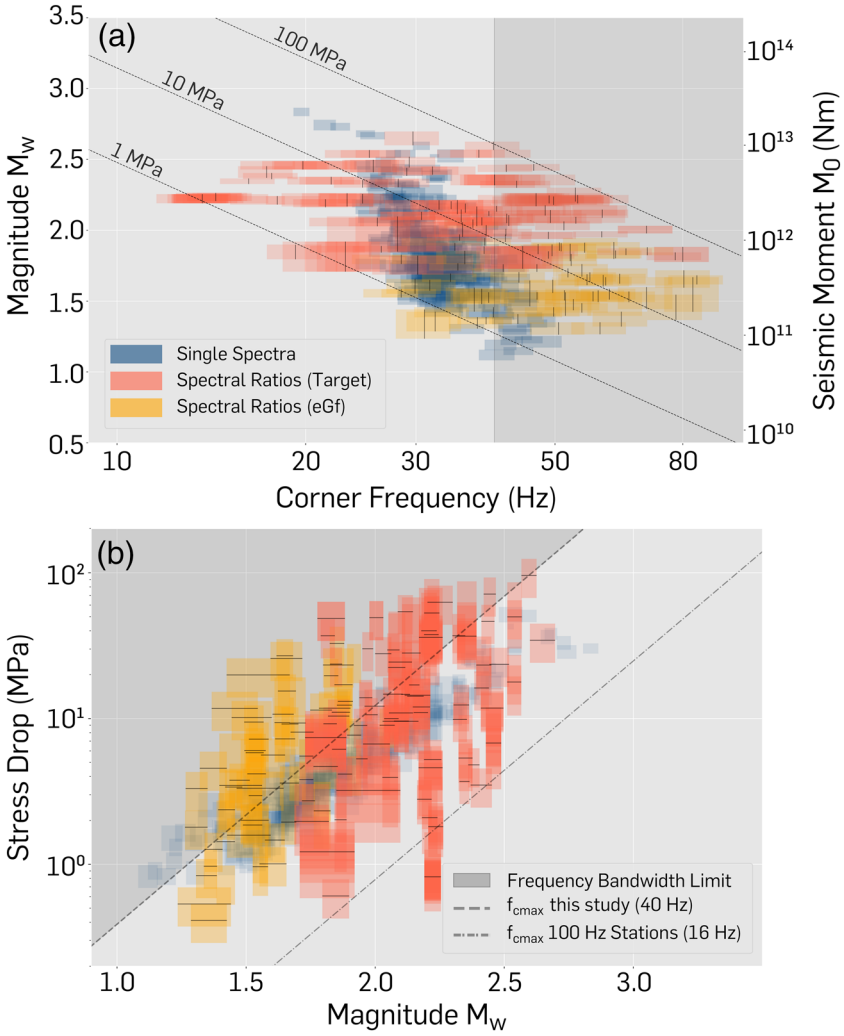

Figure 4. (a) Corner frequency versus moment magnitude for all events (336 single spectra [blue] and 126 spectral ratio estimates [orange and yellow]). The dimension of each rectangle is defined by the 95% jackknife confidence interval of f_c and M_0. Spectral ratio estimates (f_{c1} = orange; f_{c2} = yellow) are manually reviewed and have an assumed f_c error of 10% (based on a sample size of 20 or higher and the 95% confidence interval in Figure S9), and the M_0 error is taken from the single spectra estimate. As target events could have multiple eGfs, rectangles could overlap and increase the opacity, implying a more robust estimate. Fine vertical lines represent the mean estimate, and dashed lines denote stress drop isolines for a shear-wave velocity of 3,600 km/s corresponding to a depth range between 2 and 8 km, where majority of the events are located (Rubinstein et al., 2018). (b) Stress drop calculated from the parameters shown in panel a and Equation 5, where same colors and symbols apply. Each rectangle is defined by the 95% jackknife confidence interval of $\Delta\sigma$ and M_0 and the spectral ratio estimates (f_{c1} = red; f_{c2} = yellow) have an assumed $\Delta\sigma$ error of 30% (based on a sample size of 20 or higher and the 95% confidence interval in Figure 3). The dashed and dashed-dotted lines show the approximate upper frequency bandwidth limit for this study and a network of 100 Hz stations, respectively. The fine black horizontal lines show the stress drop estimate from the stacked spectral ratios for individual event pairs or the median for an event with eGfs. The upper frequency bandwidth limit is defined by one third of the highest resolvable frequency either bounded by the Nyquist frequency or the SNR criteria.

4. Discussion

The LASSO array offers a unique opportunity to explore the effects of station number and azimuthal coverage on earthquake source parameter estimates. In particular, how should the azimuthal dependency of f_c and M_0 exhibited in Figures 2a and 2b and the resulting stress drop estimates be interpreted? The circular fault model of Madariaga (1976) predicts the highest corner frequency estimates at takeoff angles of approximately 25–40 from the fault plane due to a narrowing of the *P* wave pulse in each quadrant. The corner frequency should then decrease as the pulse widens with increasing azimuth. We observe a good agreement

with the circular fault model, particularly for Events 1 and 2 (Figures 2, S5, and S7). The observed pattern in the seismic moment estimation could be explained by the use of a constant for the radiation pattern in Equation 3 and may be reduced by using a take-off angle dependent value (Kwiatek & Ben-Zion, 2013). Deviations could be explained by the small number of estimates in the azimuthal range, uncertainties in the (fully automated) focal mechanism calculation, and/or more complex source processes, which would violate the assumption of the instantaneous rupture initiation. Furthermore, the single spectra of surface stations are likely to be influenced by site effects, as suggested in the south-western part of the array where amplitudes are relatively higher, suggesting lower seismic impedence (Dougherty et al., 2019). More pronounced site effects would also be expected to cause higher frequency fall-off rates above the corner frequency compared to other station locations (Figure S8). Other studies have also shown M ~2 earthquakes can exhibit nonnegligible source complexity (Dreger et al., 2007; Fan & McGuire, 2018; Kim et al., 2016). Additional directivity effects are difficult to observe with these data, as most of the larger magnitude (larger SNR) earthquakes occurred at the edge of the array (Gallovič & Burjanek, 2007; Latour et al., 2013; Pacor et al., 2016). However, they suggest that more detailed directivity studies using Large-n array data sets are warranted.

One of the main objectives of this study is to demonstrate how the LASSO array enables robust constraints of the precision of stress drop estimates. We observe deviations of up to 50% ($3\,\sigma$) when using a large number of independent station estimates (>20–30), using a similar jackknife approach combined with a t-test as seen in Prieto et al. (2007). Deviations increase to values of >150% as the number of stations decreases (<5) (Figures 3 and S9). The significant increase in the range of deviation could be explained by problems with the t-test and resampling techniques in the presence of small (<20) or very small sample sizes (<5), as well as an unknown standard deviation (Huang, 2017). An alternative approach to estimating small sample size uncertainty could be to use well-constrained uncertainties from independent estimates of larger sample sizes and applying their constraints (Steele et al., 1993). The results above would imply that the precision of maximum stress drop estimates from single spectra is roughly 30% ($3\,\sigma$) for >20 independent station estimates and as high as 150% for <5 station estimates. The uncertainty could be reduced for an intermediate number of stations (e.g., 5–20) by requiring a maximum azimuthal gap of 45 or 90 (Figure 3). The data in Figure 3 also suggest that the maximum expected error from a random subset of 40+ stations could be as high as 50% when using the jackknife mean and that an error of 25% is likely (see also Figure S9), regardless of the number of stations.

To consider the precision of spectral ratio stress drop estimates, Kane et al. (2011) looked at different station groupings, source regions, and earthquake pairs and concluded that uncertainties of at least 30% should be expected. Abercrombie (2015) suggests that uncertainties could be reduced by using >5 station measurements and by using multiple high-quality eGfs that are restricted to approximately one source region, where possible. Our results suggest that uncertainties of ~30% likely represent a lower limit in the range afforded by nodal-scale station coverage. A small sample size of station ratios (<10) for a single event pair may mean deviations of stress drop estimates as high as 150% ($3\,\sigma$), which could be reduced to <30% by increasing the number of station ratios. Restricting the azimuthal gap for small sample sizes also decreases the precision to <30% variation. We could not constrain the effect of different focal mechanism (FM) orientations between target and eGf events, as we were only able to constrain FM solutions for select target events (Figure 1). We instead based the selection of event pairs on a commonly applied cross-correlation criteria requirement (e.g., Abercrombie, 2015; Viegas et al., 2010). In addition, Figures 2c and S6 show that the Brune- and Boatwright-type source models may be equally valid in spatially restricted study regions, which suggests variable surface geological conditions (Dougherty et al., 2019). Unfortunately, the observation of two source models using the spectral ratio approach in a small study region reduces the precision of stress drop estimates even further, since the Brune-type spectral model usually returns higher corner frequency estimations (e.g., Huang et al., 2016).

We find stress drops ranging between 0.41 and 95.89 MPa for the single spectra and spectral ratio fits. Single spectra and spectral ratio results generally overlap, with spectral ratios having somewhat higher values (range: 0.41–95.89; median: 9.21 MPa compared to range: 0.83–54.02; median: 4.06 MPa). Wu et al. (2018) also observed a wide range of stress drop values in Oklahoma, and Fan and McGuire (2018) found no difference in the source properties of a M2 earthquake in Oklahoma and typical large earthquakes. In contrast, other studies of induced seismicity in the Central United States observed low stress drops and a scaling dependence with magnitude (Boyd et al., 2017; Sumy et al., 2017; Trugman et al., 2017). Here we find a range of values consistent with both high and low stress drops relative to tectonic events and note that the

resolution and close source-station distances may explain why we were able to recover high values for small ($M < 2$) events compared to regional studies using *P* waves. When considering that *P* wave corner frequency values are expected to be a factor of 1.5 to 2, and possibly up to 3, larger than for *S* waves (Madariaga, 1976; Molnar et al., 1973; Prieto et al., 2004), it reduces the differences between our results compared with studies using *S* wave estimates (Sumy et al., 2017).

At first glance, the earthquakes in this study appear to exhibit scale dependence with magnitude (Figure 4). But the apparent scaling could also be related to the use of surface stations with limited maximum resolvable corner frequency through higher attenuation (Abercrombie, 1995; Viegas et al., 2010) (shaded areas in Figure 4). If we take the upper frequency bandwidth limit ($\sim 3 \cdot f_c$) into account (Abercrombie, 2014), it suggests that the stress drop scaling is mainly governed by observational limits (Figure 4b). Spectral ratio stress drop estimates, especially for events with $M > 2$, show a wide range of values (over two orders of magnitude). As corner frequency estimates are conceivably beyond the frequency bandwidth limit for many small events ($M < 2$), the same wide range of stress drops may be even higher due to an underestimation of the corner frequencies beyond 40 Hz (Aron & Hardebeck, 2009; Ide & Beroza, 2001). The observation of a wide range of stress drop values for different magnitudes is a strong argument for fault complexity at a wide range of length scales and would be consistent with self-similar stress drop scaling in regions with both tectonic and induced events (Kirkpatrick et al., 2020). As induced events are usually of small magnitude ($M < 3$), one might expect to commonly observe low stress drop values, especially where frequency bandwidth is limited (Figure 4b shows an upper frequency limit of a 100 Hz station). However, the coverage and sample rate afforded by the LASSO array provides a wider glimpse into the precision and range of stress drop values of small, induced earthquakes, and provides evidence that their stress drop values are not all on the lower end of the range observed for tectonic events.

5. Conclusion

We use a Large-*n* Seismic Array (>1,800 stations) to evaluate the robustness of source property calculations from spectral methods of small earthquakes (M <3) and constrain the range of precision as a function of station number and azimuthal coverage. We used both single event and event spectral ratios to estimate corner frequency, seismic moment, and stress drop values from *P* waves. We find that

1. single spectra show significant azimuthal dependency of corner frequency and long-period spectral amplitude estimates;
2. the 3σ precision of single spectra stress drop estimates may be no less than approximately 30% for >20 stations or up to >150% for <5 stations;
3. the 3σ precision stress drop estimates from spectral ratio methods should be expected to be at least 30% for >20 stations and >100% for <10 stacked station ratios;
4. requiring a maximum azimuthal gap of <90 and <180 could significantly reduce stress drop uncertainties (~25%) for estimates using <10 stations or stacked station ratios for single spectra and spectral ratio estimates, respectively;
5. single spectra corner frequency estimates show higher precision by using a small (<10) number of stations relative to spectral ratios but may be less accurate due to theoretical constraints;
6. stress drop estimates from wastewater-induced seismicity in Oklahoma show no obvious scaling with magnitude (M 1–3) in the limited resolvable frequency bandwidth and exhibit a wide range of values, suggesting fault complexity at rupture length scales. Earthquakes also exhibit values within the typical range of tectonic earthquakes (10–100 MPa), as well as low stress drop values often associated with induced seismicity (0.1–10 MPa);
7. the discrepancy with previous studies in this region (e.g., Sumy et al., 2017; Trugman et al., 2017) could be explained by the expanded frequency bandwidth and coverage afforded by the LASSO array and the expected differences between *P* and *S* wave corner frequency estimates.

Data Availability Statement

Data collected by the Large-*n* Seismic Survey in Oklahoma (LASSO) experiment are archived at the Incorporated Research Institutions for Seismology (IRIS) Data Management Center under network code 2A in PH5 format and can be accessed at https://ds.iris.edu/mda/2A/?starttime=2016-04-11T13:20:06&endtime=2016-05-24T15:41:19 (last accessed March 2020). The open-source Python package Obspy 1.0.1

was used for data processing (Krischer et al., 2015) available at https://github.com/obspy/obspy (last accessed March 2020). The mtspec Python wrapper was used for the spectral estimations, available at https://zenodo.org/record/321789#.XqKyiZMzZhE (last accessed March 2020). Several figures were created using GMT 6 (Wessel et al., 2019), available at https://docs.generic-mapping-tools.org/latest/ (last accessed March 2020). The other figures were created using Matplotlib 3.1.3 (Hunter, 2007) available at https://matplotlib.org/ (last accessed March 2020). Many colormaps from Scientific Color Maps 6 were used during plotting, available at https://doi.org/10.5281/zenodo.1243862 (last accessed March 2020). Polarities were determined using PhasePApy (Chen & Holland, 2016), available at https://github.com/austinholland/PhasePApy (last accessed March 2020). Focal mechanisms were determined using HASH1.2 (Hardebeck & Shearer, 2002), available at https://www.usgs.gov/software/hash-12 (last accessed September 2020).

Acknowledgments
We thank Editor G. Prieto, K. Mayeda, A. Baltay, O. Boyd, and one anonymous reviewer for their comments. This project is partially funded by Ruhr University Bochum (RUB) new faculty startup funds and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project number 428868223. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Abercrombie, R. E. (1995). Earthquake source scaling relationships from −1 to 5 *M* L using seismograms recorded at 2.5-km depth. *Journal of Geophysical Research*, *100*(B12), 24,015–24,036. https://doi.org/10.1029/95JB02397

Abercrombie, R. E. (2013). Comparison of direct and coda wave stress drop measurements for the Wells, Nevada, earthquake sequence. *Journal of Geophysical Research: Solid Earth*, *118*, 1458–1470. https://doi.org/10.1029/2012JB009638

Abercrombie, R. E. (2014). Stress drops of repeating earthquakes on the San Andreas Fault at Parkfield. *Geophysical Research Letters*, *41*, 8784–8791. https://doi.org/10.1002/2014GL062079

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

Aki, K. (1967). Scaling law of seismic spectrum. *Journal of Geophysical Research*, *72*(4), 1217–1231. https://doi.org/10.1029/JZ072i004p01217

Aron, A., & Hardebeck, J. L. (2009). Seismicity rate changes along the central California coast due to stress changes from the 2003 M 6.5 San Simeon and 2004 M 6.0 Parkfield earthquakes. *Bulletin of the Seismological Society of America*, *99*(4), 13. https://doi.org/10.1785/0120080239

Boatwright, J. (1980). A spectral theory for circular seismic sources; simple estimates of source dimension, dynamic stress drop, and radiated seismic energy. *Bulletin of the Seismological Society of America*, *70*(1), 1–27.

Boyd, O. S., McNamara, D. E., Hartzell, S., & Choy, G. (2017). Influence of lithostatic stress on earthquake stress drops in North America. *Bulletin of the Seismological Society of America*, *107*(2), 856–868. https://doi.org/10.1785/0120160219

Brodsky, E. E. (2019). The importance of studying small earthquakes. *Science*, *364*(6442), 736–737. https://doi.org/10.1126/science.aax2490

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research*, *75*(26), 4997–5009. https://doi.org/10.1029/JB075i026p04997

Chen, C., & Holland, A. A. (2016). PhasePApy: A robust pure Python package for automatic identification of seismic phases. *Seismological Research Letters*, *87*(6), 1384–1396. https://doi.org/10.1785/0220160019

Cochran, E. S., Wickham-Piotrowski, A., Kemna, K. B., Harrington, R. M., Dougherty, S. L., & Peña Castro, A. F. (2020). Minimal clustering of injection-induced earthquakes observed with a Large-*n* Seismic Array. *Bulletin of the Seismological Society of America*, *110*, 2005–2017. https://doi.org/10.1785/0120200101

Dougherty, S. L., Cochran, E. S., & Harrington, R. M. (2019). The Large-*n* Seismic Survey in Oklahoma (LASSO) experiment. *Seismological Research Letters*, *90*(5), 2051–2057. https://doi.org/10.1785/0220190094

Dreger, D., Nadeau, R. M., & Chung, A. (2007). Repeating earthquake finite source models: Strong asperities revealed on the San Andreas Fault. *Geophysical Research Letters*, *34*, L23302. https://doi.org/10.1029/2007GL031353

Eshelby, J. D. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proceedings of the Royal Society of London Series A*, *241*, 376–396. https://doi.org/10.1098/rspa.1957.0133

Fan, W., & McGuire, J. J. (2018). Investigating microearthquake finite source attributes with IRIS Community Wavefield Demonstration Experiment in Oklahoma. *Geophysical Journal International*, *214*(2), 1072–1087. https://doi.org/10.1093/gji/ggy203

Gallovič, F., & Burjanek, J. (2007). High-frequency directivity in strong ground motion modeling methods. *Annales Geophysicae*, *50*(2), 203–211.

Hanks, T. C. (1977). Earthquake stress drops, ambient tectonic stresses and stresses that drive plate motions. *Pure and Applied Geophysics*, *115*(1), 441–458. https://doi.org/10.1007/BF01637120

Hardebeck, J. L., & Shearer, P. M. (2002). A new method for determining first-motion focal mechanisms. *Bulletin of the Seismological Society of America*, *92*(6), 2264–2276. https://doi.org/10.1785/0120010200

Harrington, R. M., & Brodsky, E. E. (2009). Source duration scales with magnitude differently for earthquakes on the San Andreas Fault and on secondary faults in Parkfield, California. https://doi.org/10.1785/0120080216

Hartzell, S. H. (1978). Earthquake aftershocks as Green's functions. *Geophysical Research Letters*, *5*(1), 1–4. https://doi.org/10.1029/GL005i001p00001

Hough, S. E. (1997). Empirical Green's function analysis: Taking the next step. *Journal of Geophysical Research*, *102*(B3), 5369–5384. https://doi.org/10.1029/96JB03488

Huang, H. (2017). Uncertainty estimation with a small number of measurements, part I: New insights on the t-interval method and its limitations. *Measurement Science and Technology*, *29*(1), 15004. https://doi.org/10.1088/1361-6501/aa96c7

Huang, Y., Beroza, G. C., & Ellsworth, W. L. (2016). Stress drop estimates of potentially induced earthquakes in the Guy-Greenbrier sequence. *Journal of Geophysical Research: Solid Earth*, *121*, 6597–6607. https://doi.org/10.1002/2016JB013067

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Ide, S., & Beroza, G. C. (2001). Does apparent stress vary with earthquake size? *Geophysical Research Letters*, *28*(17), 3349–3352. https://doi.org/10.1029/2001GL013106

Ide, S., Beroza, G. C., Prejean, S. G., & Ellsworth, W. L. (2003). Apparent break in earthquake scaling due to path and site effects on deep borehole recordings. *Journal of Geophysical Research*, *108*(B5), 2271. https://doi.org/10.1029/2001JB001617

Kanamori, H., Mori, J., Hauksson, E., Heaton, T. H., Hutton, L. K., & Jones, L. M. (1993). Determination of earthquake energy release and ML using TERRAscope. *Bulletin of the Seismological Society of America*, *83*(2), 330–346.

Kane, D. L., Prieto, G. A., Vernon, F. L., & Shearer, P. M. (2011). Quantifying seismic source parameter uncertainties. *Bulletin of the Seismological Society of America*, *101*(2), 535–543. https://doi.org/10.1785/0120100166
Kaneko, Y., & Shearer, P. M. (2015). Variability of seismic source spectra, estimated stress drop, and radiated energy, derived from cohesive-zone models of symmetrical and asymmetrical circular and elliptical ruptures. *Journal of Geophysical Research: Solid Earth*, *120*, 1053–1079. https://doi.org/10.1002/2014JB011642
Kim, A., Dreger, D. S., Taira, T., & Nadeau, R. M. (2016). Changes in repeating earthquake slip behavior following the 2004 Parkfield main shock from waveform empirical Green's functions finite-source inversion. *Journal of Geophysical Research: Solid Earth*, *121*, 1910–1926. https://doi.org/10.1002/2015JB012562
Kirkpatrick, J. D., Edwards, J. H., Verdecchia, A., Kluesner, J. W., Harrington, R. M., & Silver, E. A. (2020). Subduction megathrust heterogeneity characterized from 3D seismic data. *Nature Geoscience*, *13*(5), 369–374. https://doi.org/10.1038/s41561-020-0562-9
Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., & Wassermann, J. (2015). ObsPy: A bridge for seismology into the scientific Python ecosystem. *Computational Science & Discovery*, *8*(1), 14003. https://doi.org/10.1088/1749-4699/8/1/014003
Kwiatek, G., & Ben-Zion, Y. (2013). Assessment of P and S wave energy radiated from very small shear-tensile seismic events in a deep South African mine. *Journal of Geophysical Research: Solid Earth*, *118*, 3630–3641. https://doi.org/10.1002/jgrb.50274
Latour, S., Schubnel, A., Nielsen, S., Madariaga, R., & Vinciguerra, S. (2013). Characterization of nucleation during laboratory earthquakes. *Geophysical Research Letters*, *40*, 5064–5069. https://doi.org/10.1002/grl.50974
Madariaga, R. (1976). Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America*, *66*(3), 639–666.
Mayeda, K., Hofstetter, A., O'Boyle, J., & Walter, W. (2003). Stable and transportable regional magnitudes based on coda-derived moment-rate spectra. *Bulletin of the Seismological Society of America*, *93*, 224–239. https://doi.org/10.1785/0120020020
Mayeda, K., Malagnini, L., & Walter, W. R. (2007). A new spectral ratio method using narrow band coda envelopes: Evidence for non-self-similarity in the Hector Mine sequence. *Geophysical Research Letters*, *34*, L11303. https://doi.org/10.1029/2007GL030041
Mayeda, K., & Walter, W. R. (1996). Moment, energy, stress drop, and source spectra of western United States earthquakes from regional coda envelopes. *Journal of Geophysical Research*, *101*(B5), 11,195–11,208. https://doi.org/10.1029/96JB00112
Molnar, P., Tucker, B. E., & Brune, J. N. (1973). Corner frequencies of P and S waves and models of earthquake sources. *Bulletin of the Seismological Society of America*, *63*(6–1), 2091–2104.
Neely, J. S., Stein, S., & Spencer, B. D. (2020). Large uncertainties in earthquake stress-drop estimates and their tectonic consequences. *Seismological Research Letters*, *91*, 2320–2329. https://doi.org/10.1785/0220200004
Newville, M., Stensitzki, T., Allen, D. B., & Ingargiola, A. (2014). LMFIT: Non-linear least-square minimization and curve-fitting for Python. Zenodo. https://doi.org/10.5281/zenodo.11813
OCC (2020). OCC oil and gas data files. http://www.occeweb.com/og/ogdatafiles2.htm
Pacor, F., Gallovič, F., Puglia, R., Luzi, L., & D'Amico, M. (2016). Diminishing high-frequency directivity due to a source effect: Empirical evidence from small earthquakes in the Abruzzo region, Italy. *Geophysical Research Letters*, *43*, 5000–5008. https://doi.org/10.1002/2016GL068546
Prieto, G. A., Parker, R. L., & Vernon III, F. L. (2009). A Fortran 90 Library for multitaper spectrum analysis. *Computers & Geosciences*, *35*(8), 1701–1710. https://doi.org/10.1016/j.cageo.2008.06.007
Prieto, G. A., Parker, R. L., Vernon, F. L., Shearer, P. M., & Thomson, D. J. (2006). Uncertainties in earthquake source spectrum estimation using empirical Green functions. In R. Abercrombie, A. McGarr, H. Kanamori, G. Di Toro (Eds.), *Geophysical monograph series* (Vol. 170, pp. 69–74). Washington, D. C.: American Geophysical Union. https://doi.org/10.1029/170GM08
Prieto, G. A., Shearer, P. M., Vernon, F. L., & Kilb, D. (2004). Earthquake source scaling and self-similarity estimation from stacking P and S spectra. *Journal of Geophysical Research*, *109*, B08310. https://doi.org/10.1029/2004JB003084
Prieto, G. A., Thomson, D. J., Vernon, F. L., Shearer, P. M., & Parker, R. L. (2007). Confidence intervals for earthquake source parameters. *Geophysical Journal International*, *168*(3), 1227–1234. https://doi.org/10.1111/j.1365-246X.2006.03257.x
Rubinstein, J. L., Ellsworth, W. L., & Dougherty, S. L. (2018). The 2013–2016 induced earthquakes in Harper and Sumner Counties, Southern Kansas. *Bulletin of the Seismological Society of America*, *108*(2), 674–689. https://doi.org/10.1785/0120170209
Shearer, P. M., Abercrombie, R. E., Trugman, D. T., & Wang, W. (2019). Comparing EGF methods for estimating corner frequency and stress drop from *P* wave spectra. *Journal of Geophysical Research: Solid Earth*, *124*, 3966–3986. https://doi.org/10.1029/2018JB016957
Shearer, P. M., Prieto, G. A., & Hauksson, E. (2006). Comprehensive analysis of earthquake source spectra in southern California. *Journal of Geophysical Research*, *111*, B06303. https://doi.org/10.1029/2005JB003979
Spall, J. C. (2002). Uncertainty bounds in parameter estimation with limited data. In M. Dror, P. L'Ecuyer, F. Szidarovszky (Eds.), *Modeling uncertainty: An examination of stochastic theory, methods, and applications* (pp. 685–709), International Series in Operations Research & Management Science. New York, NY: Springer US. https://doi.org/10.1007/0-306-48102-2_27
Steele, W. G., Taylor, R. P., Burrell, R. E., & Coleman, H. W. (1993). Use of previous experience to estimate precision uncertainty of small sample experiments. *AIAA Journal*, *31*(10), 1891–1896. https://doi.org/10.2514/3.11864
Sumy, D. F., Neighbors, C. J., Cochran, E. S., & Keranen, K. M. (2017). Low stress drops observed for aftershocks of the 2011 Mw 5.7 Prague, Oklahoma, earthquake. *Journal of Geophysical Research: Solid Earth*, *122*, 3813–3834. https://doi.org/10.1002/2016JB013153
Trugman, D. T., Dougherty, S. L., Cochran, E. S., & Shearer, P. M. (2017). Source spectral properties of small to moderate earthquakes in Southern Kansas. *Journal of Geophysical Research: Solid Earth*, *122*, 8021–8034. https://doi.org/10.1002/2017JB014649
Viegas, G., Abercrombie, R. E., & Kim, W.-Y. (2010). The 2002 M5 Au Sable Forks, NY, earthquake sequence: Source scaling relationships and energy budget. *Journal of Geophysical Research*, *115*, B07310. https://doi.org/10.1029/2009JB006799
Wessel, P., Luis, J. F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H. F., & Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, *20*, 5556–5564. https://doi.org/10.1029/2019GC008515
Wu, Q., Chapman, M., & Chen, X. (2018). Stress-drop variations of induced earthquakes in Oklahoma. *Bulletin of the Seismological Society of America*, *108*(3A), 1107–1123. https://doi.org/10.1785/0120170335