Source Spectral Properties of Small to Moderate Earthquakes in Southern Kansas

Daniel T. Trugman, Sara L. Dougherty, Elizabeth S. Cochran, and Peter M. Shearer

Abstract The source spectral properties of injection-induced earthquakes give insight into their nucleation, rupture processes, and influence on ground motion. Here we apply a spectral decomposition approach to analyze P wave spectra and estimate Brune-type stress drop for more than 2,000 ML 1.5–5.2 earthquakes occurring in southern Kansas from 2014 to 2016. We find that these earthquakes are characterized by low stress drop values (median $\sim$0.4 MPa) compared to natural seismicity in California. We observe a significant increase in stress drop as a function of depth, but the shallow depth distribution of these events is not by itself sufficient to explain their lower stress drop. Stress drop increases with magnitude from $M_{L}$1.5 to $M_{L}$3.5, but this scaling trend may weaken above $M_{L}$4 and also depends on the assumed source model. Although we observe a nonstationary, sequence-specific temporal evolution in stress drop, we find no clear systematic relation with the activity of nearby injection wells.

Plain Language Summary The rate of earthquake occurrence in regions of oil and gas production in the central and eastern United States has increased sharply over the last 8 years. In this study, we analyze the source spectra, or frequency content, of earthquakes occurring in one such prominent region of active oil and gas production: southern Kansas. This study is one of the first and the largest to date that provides a quantitative comparison between the spectral properties of these earthquakes, which are potentially induced by human activity, and those of earthquakes that occur in California due to natural tectonic processes. We find that earthquakes in southern Kansas are depleted in high-frequency energy compared to natural earthquakes in California but that their relative frequency content increases significantly with depth and with magnitude. We also observe significant spatial and temporal variations in source spectral properties that may in part be driven by widespread wastewater disposal during oil and gas production. Characterizing the source spectral properties of these earthquakes is important because it lends insight into the physical processes causing these events and because the frequency content of the source has a strong influence on the intensity of shaking felt by the local population.

1. Introduction

Over the past 10 years, seismicity rates have risen to historically unprecedented levels within the oil-producing regions of Oklahoma and southern Kansas. Scientific consensus has attributed much of the elevated seismicity rate to anthropogenic activity and in particular to the injection of wastewater from the oil production process into the Arbuckle Group that is stratigraphically above the granitic basement (e.g., Buchanan, 2015; Ellsworth, 2013; Ellsworth et al., 2015; Rubinstein & Mahani, 2015; Walsh & Zoback, 2015; Yeck et al., 2017). The abrupt increase in seismic hazard within this region (Petersen et al., 2016, 2017) has spurred numerous observational studies focused on connections between fluid injection and seismicity rates (Barbour et al., 2017; Choy et al., 2016; Goebel, 2015; Keranen et al., 2013, 2014; Weingarten et al., 2015), the source properties of a subset of the larger events (Boyd et al., 2017; Choy et al., 2016; Cramer, 2017; Sumy et al., 2017; Walter et al., 2017), and observed ground motion amplitudes (Atkinson & Assatourians, 2017; Atkinson et al., 2016; Hough, 2014; Yenier et al., 2017). Although these studies have rapidly advanced scientific understanding of these earthquakes, there is still much that remains unanswered, and the nonstationary nature of the seismicity warrants continued monitoring.

One of the key unresolved questions is whether the dynamic source properties of these likely induced events, such as corner frequency and stress drop, differ from events within active tectonic regions like California.
(Huang et al., 2016). This question is of particular importance for ground motion estimation due to the positive correlation of ground motion amplitude and stress drop (e.g., Atkinson and Morrison, 2009; Baltay et al., 2017; Boore, 2003). Several recent studies have used various parametric spectral fitting methods (Cramer, 2017; Sumy et al., 2017) or empirical Green’s function spectral ratio approaches (Boyd et al., 2017; Huang et al., 2017; Walter et al., 2017) to analyze source parameters of a subset of the largest of these events, including the 2011 Prague, Oklahoma, sequence and the 2014 Milan, Kansas, earthquakes. In this study, we apply a spectral decomposition technique (Trugman & Shearer, 2017a) to perform a comprehensive analysis of the source spectra and source parameters of seismicity occurring in southern Kansas from 2014 to 2016. We use $P$ wave spectra to derive source parameter estimates—seismic moment, corner frequency, and stress drop—for more than 2,000 events with local magnitudes ranging from 1.5 to 5.2 that we have relocated using waveform cross correlation-based techniques. These source parameter estimates present an opportunity for a quantitative comparison to the source properties of naturally occurring events in California that were analyzed using identical methodology.

We focus our analysis on the distribution and variability in source parameter estimates of the southern Kansas data set as whole, rather than on individual target events or earthquake sequences. We begin by providing an overview of the southern Kansas study region and the associated waveform data. We then briefly describe the methodology we use to derive relocated event positions, source parameter estimates, and parameter uncertainties for each event in our data set. We next examine the most robust statistical features of our source parameter estimates, including the depth dependence of corner frequency and stress drop, the scaling of stress drop and seismic moment, and the nonstationary temporal evolution of stress drop during our study period (2014–2016). We compare our source parameter observations to those of natural (tectonic) earthquakes in Southern California and discuss the physical and practical implications of our results for the scientific understanding of earthquake rupture processes and occurrence in southern Kansas and for the probabilistic assessment of ground motion amplitudes and seismic hazard in oil-producing regions of the central United States.

2. Data and Study Region

Seismicity rates in southern Kansas began to sharply increase in 2013 compared to historical norms (Buchanan, 2015; Choy et al., 2016; Hildebrand et al., 1988), with this increase attributed in part to the proliferation of wastewater injection related to oil and gas production within the Mississippian limestone play that underlies southern Kansas and northern Oklahoma (Buchanan et al., 2014; Ellsworth, 2013; Ellsworth et al., 2015; Weingarten et al., 2015). The Precambrian basement formation begins at a depth of 1.7 to 2.0 km and is crosscut by numerous in situ fault systems, the most notable of which is the northeast striking Nemaha fault (Baars & Watney, 1991; McBee, 2003; Niemi, 2004; Steeple et al., 1979). Oil production wells tap into the shallower sedimentary strata, with wastewater disposal typically occurring in the permeable Arbuckle Group that directly overlies the granitic basement (Buchanan et al., 2014; Kroll et al., 2017). We use in this study the wastewater injection and enhanced oil recovery well locations publicly archived by the Kansas Corporation Commission (http://kcc.ks.gov, last accessed April 2017).

Here we analyze seismicity occurring within southern Kansas from 21 March 2014 to 31 December 2016 (Figure 1). The start date for our study period was chosen based on the installation date (19–21 March 2014) of the U.S. Geological Survey (USGS) Induced Seismicity Menlo Park Project (ISMP) network that was established to monitor seismicity within this region (Rubinstein et al., 2014). Azimuthal station coverage in this area is generally good once the ISMP network was fully installed (late summer 2014), and as such we have few source parameter estimates prior to September 2014. We take initial locations and magnitudes for earthquakes in our data set from the ISMP catalog, which is a subsidiary of the Advanced National Seismic Systems (ANSS) Comprehensive Earthquake Catalog (ComCat, https://earthquake.usgs.gov/earthquakes/search/, last accessed May 2017) that lists events by local magnitude. For our study, we convert waveform data with sampling rates of 100 Hz and 200 Hz from the USGS, Central and Eastern U.S., NetQuakes, Oklahoma Seismic, and U.S. National Seismic networks (network codes GS, N4, NQ, OK, and US) into multiplexed event-based files for later analysis and processing. In total, we consider 5,269 events occurring within our study region during this time period, though only a well-recorded subset of 2,069 of these events met the quality control criteria for our source parameter estimates (section 3).
Figure 1. Map view of 2014–2016 southern Kansas seismicity, with the relocated epicenters of events considered in this study shown in gray, and the subset of these events for which we obtain source parameter estimates shown in black. Local station coverage (red triangles) and locations of active wastewater disposal and enhanced oil recovery wells (green inverted triangles) obtained from the Kansas Corporation Commission (http://kcc.ks.gov, last accessed April 2017), and ANSS ComCat focal mechanism estimates (https://earthquake.usgs.gov/earthquakes/search/, last accessed May 2017) for $M_3$ and greater events are shown for reference. Thin and thick brown lines correspond to county and state boundaries, respectively.

3. Methods: Relocations and Source Parameter Estimates

The primary focus of this article is the analysis and interpretation of earthquake source spectra and source parameters. Although highly accurate earthquake locations are not essential for the spectral analysis, they are useful in the interpretation of spatial variations in source parameters such as systematic trends with depth or with distance from injection wells. Because of this, as a preliminary step in our analysis of source properties, we apply the GrowClust algorithm (Trugman & Shearer, 2017b) to obtain relocated event positions and location uncertainties for the southern Kansas earthquakes in our data set. For these relocations, we perform time domain waveform cross correlation of pairs of events within 4 km catalog distance of each other, saving cross-correlation results for all event pairs with a minimum of eight differential times with cross-correlation values greater than 0.7. We input the resulting approximately one million differential times and cross-correlation values into the GrowClust program, which uses a hybrid hierarchical clustering and relocation algorithm that provides stable relocation results for large-scale catalogs with multiple discrete clusters. The relocated seismicity has median horizontal and vertical location errors of 131 m and 281 m, respectively.

We next obtain source parameter estimates for a subset of the relocated seismicity using the spectral decomposition method. We follow closely the algorithm described in detail by Trugman and Shearer (2017a), and summarize only the main points here. The central idea underlying this approach is that for large and well-recorded seismicity data sets, each earthquake is recorded by many stations, each station records many earthquakes, and each approximate source station travel path is traversed many times. If this assumption is valid, then it is possible to decompose the waveform data spectra from event $i$ recorded at station $j$ into relative source, station, and path-dependent terms, plus a residual error term ($r_{ij}$) for each trace. Working in the log frequency domain, the relative contributions add linearly, and the spectral decomposition at each frequency can be written in the form

$$d_{ij} = e_i + st_j + \tau_{t(i,j)} + r_{ij},$$

where $d_{ij}$ is the recorded waveform data spectra, $e_i$ is the relative source term, $st_j$ is the relative station term, and $\tau_{t(i,j)}$ is a relative path-dependent term that is typically assumed to be isotropic and discretized by source-receiver travel time such that there is a single relative path term for all traces in the $k$th travel time bin.
The spectral decomposition method consists of four basic steps:

1. Compute the amplitude spectra $d_i(f)$ from the waveform data of each trace.
2. Decompose the amplitude data spectrum of all traces into relative source, station, path, and residual terms by solving equation (1) at each frequency point using an iterative, robust least squares inversion algorithm with outlier suppression.
3. Infer the empirical Green’s function correction spectrum (EGF) that best captures path effects common to all sources, such as average near-source and near-receiver attenuation.
4. Subtract the EGF from each source spectra: $s_i(f) = e_i(f) - \text{EGF}(f)$, and use the corrected source spectra to obtain source parameter estimates and uncertainties for seismic moment $M_0$, corner frequency $f_c$, and stress drop $\Delta \sigma$.

For steps (1) and (2), we consider $P$ wave spectra of earthquakes with local magnitude $M_L \geq 5$ and greater, recorded on vertical-component, high-broadband, and short-period channels (HHZ, HNZ, and EHZ) at stations within 150 km distance. For the spectral estimates, we use a magnitude-dependent window length ranging from a minimum length of 1.5 s to a maximum length of 4.5 s, where longer windows correspond to larger events in order to permit adequate corner frequency resolution (Abercrombie et al., 2017; Ross & Ben-Zion, 2016; Trugman & Shearer, 2017a). We define the signal window to begin 0.05 s before the catalog-listed $P$ phase arrival time, truncate each window before the catalog-listed $S$ phase arrival when necessary, and define a noise window that immediately precedes the signal window and is of equal length. We discard clipped waveforms using an automated detection algorithm (Trugman & Shearer, 2017a) and resample the spectra obtained from longer window lengths to the frequency points corresponding to the minimum window length (1.5 s). We compute the average signal-to-noise amplitude in each of five frequency bands (2.5–6, 6–10, 10–15, 15–20, and 20–25 Hz), and only further consider events that are recorded at a minimum of six stations with a signal-to-noise ratio greater than 3 in each frequency band.

To estimate the EGF correction term (step 3), we use the technique described by Trugman and Shearer (2017a) that fits stacked relative source spectra, averaged in bins of spectral moment $\Omega_s$ to a Brune-type theoretical spectrum of the form

$$\hat{s}(f | \Omega_0, f_c, n) = \frac{\Omega_0}{1 + (f/f_c)^n},$$

where $f_c$ and $\Omega_0$ are the corner frequency and spectral moment of each stacked spectra, and the high-frequency falloff rate $n$ is fixed to 2 per the widely used $\omega^{-2}$ model (Aki, 1967; Brune, 1970). In contrast to previous implementations of the spectral decomposition method (e.g., Shearer et al., 2006), our technique does not require an assumption of self-similar or constant stress drop scaling and instead infers the optimal scaling directly from the shape of the stacked spectra. Here we find that the optimal fit requires an EGF with non-self-similar scaling such that stress drop increases with spectral moment (Figure 2), a result in agreement with a recent analysis of earthquakes in California (Trugman & Shearer, 2017a). Inference of the EGF is an essential part of the spectral decomposition technique because the source terms $e_i$ produced by the solution to equation (1) are relative (median amplitude 0) and thus must be corrected for propagation effects that are common to all sources. This includes the spatially averaged near-source attenuation that is not removed by the nearest travel time terms $t_{tt,k}$, as well as average near-station attenuation, since the station terms $s_t$ isolate only relative differences in near-station and instrument effects on the observed spectra.

We account for the possibility of lateral variations in attenuation by applying a modified spectral stacking technique that uses a distance weighting to allow for spatial variations in the EGF. This technique is similar to the nearest-neighbors EGF approach first used by Shearer et al. (2006) but in practice tends to be more stable. In brief, the modified technique uses cluster analysis to define a nonuniform set of grid points that conform to the contours of the observed seismicity. For each grid point (we use six in this study), an EGF is inferred from stacks of relative source spectra that are weighted by inverse distance to the event locations in the study region, such that events that are closer to the grid point assume more weight. The EGF correction to the source spectrum of each event is then computed as a linear combination of the set of EGFs, again weighted by inverse distance. We note that applying this distance-weighted stacking algorithm to account for lateral variations in attenuation does not significantly influence the results presented in this study but may be an important consideration for study regions that extend over larger length scales.
Figure 2. Example of stacked source spectra from earthquakes in southern Kansas, before and after the EGF correction for common path effects (see text for details). (a) Stack-averaged relative source spectra, binned by spectral moment $\Omega_0$ (logarithmic scale), prior to the EGF correction. (b and c) EGF-corrected stacked source spectra (solid black lines) and a comparison to theoretical source spectra (dashed blue lines), with the implied corner frequency of each stack (blue dots) and the EGF spectrum (red line) marked for reference. In Figure 2b, the EGF is estimated with a self-similarity (constant stress drop) constraint and provides a significantly worse fit than the EGF in Figure 2c with no such constraint. It is also important to recognize that the increase in median stress drop with moment for individual events that we present in section 4 is a direct consequence of applying an EGF based on non-self-similar scaling. The evidence for such scaling is the markedly superior fit that we obtain to the stacked spectra (Figure 2c) compared to the fit for a self-similar model (Figure 2b: requiring self-similarity increases the overall misfit by more than a factor of 3). Note that this result is based only on observations within the 2.5 to 25 Hz band where we have good signal-to-noise ratio and does not require resolving corner frequencies outside of this band. However, as discussed in Trugman and Shearer (2017a), the case for an increase in average stress drop with moment does depend upon the assumption of the Brune spectral model and its $f^{-2}$ high-frequency falloff rate, as reasonable fits to the stacked spectra are possible for self-similar models with high-frequency falloff rates less than 2.

Lastly (step 4), we use the EGF-corrected spectra $s_i(f)$ to estimate source parameters and associated uncertainties. To estimate seismic moment $M_0$, we assume that on average the observed spectral moment $\Omega_0$ is proportional to the seismic moment $M_0$ and perform a regression analysis between $\Omega_0$ and $M_0$ to calibrate the appropriate scale factor (Shearer et al., 2006). This analysis can be used to obtain a linear relationship between $M_w$ and $M_0$ that is valid for the smaller earthquakes for which $M_w$ is not routinely estimated. The smallest events in our data set ($M_L 1.5$) correspond to $M_w 1.9$, and the inferred $M_w-M_0$ slope of 0.72 (Figure 3) is comparable to the slope of 0.75 obtained for Southern California by Ross et al. (2016). Although the relationship between $M_w$ and $M_0$ may be slightly nonlinear for larger events (Ben-Zion & Zhu, 2002; Edwards & Douglas, 2014; Goertz-Allmann et al., 2011; Munafo et al., 2016), we do not observe a significant bias between our $M_w$ estimates and those obtained through moment tensor analysis and listed by ANSS ComCat (https://earthquake.usgs.gov/earthquakes/search/, last accessed May 2017). We then estimate the corner frequency $f_c$ using a bounded optimization algorithm that minimizes the root-mean-square residual between the observed, EGF-corrected source spectrum $s_i(f)$ and the Brune the-
theoretical spectrum $\hat{S}(f|f_c)$ in the 2.5–25 Hz band in which we have measured adequate signal-to-noise ratio, and that is insensitive to site resonances and spectral contamination from leaky-mode surface waves that are prevalent for earthquakes in this region (Cramer, 2017). Given $f_c$ and $M_0$, we compute the Brune-type stress drop

$$\Delta \sigma = \frac{7}{16} M_0 \left( \frac{f_c}{k\beta} \right)^3,$$

appropriate for a simplified circular crack model of the earthquake source with constant stress drop and elliptical distribution of slip (Brune, 1970; Madariaga, 1976). We obtain a depth-dependent estimate of the shear wave speed $\beta$ from the velocity model used by the ISMP network to locate the events (Rubinstein et al., 2015) and set the numerical constant $k = 0.38$ following the recent numerical analysis of Kaneko and Shearer (2014). We derive uncertainty estimates for $M_0$, $f_c$, and $\Delta \sigma$ using the statistical resampling techniques detailed in Trugman and Shearer (2017a) that are based upon the variability in the apparent source spectra recorded at each station. Normalized uncertainties in corner frequency ($\Delta f_c/f_c$) tend to be higher for the lowest and highest magnitude events, which have $f_c$ that approach the 2.5–25 Hz limits of the spectral bandwidth. However, the spectral decomposition results for the data set in aggregate are insensitive to resolution of source parameters of these individual events (which are few in number) but are instead controlled primarily by the relative shape of the stacked spectra (see Trugman & Shearer, 2017a for a complete discussion).

4. Results

We apply the spectral decomposition method (section 3) to analyze the earthquake source parameters of 2,069 well-recorded earthquakes in our study region. The southern Kansas earthquakes in our data set have relatively low stress drop values, with an overall distribution that is approximately lognormal with a median stress drop of 0.41 MPa and log$_{10}$ standard deviation of 0.35 (Figure 4a). The relative simplicity of this total marginal histogram of $\Delta \sigma$ does, however, obscure several notable trends. As a concrete example, in Figure 4 we also show conditional histograms of $\Delta \sigma$ for four distinct magnitude ranges (1.5–2.0, 2.0–2.5, 2.5–3.0, and 3.0–3.5) and plot the distribution of corner frequency $f_c$ and stress drop $\Delta \sigma$ as a function of seismic moment $M_0$. From this perspective it is apparent that median stress drop tends to increase as a function of moment, a result that is consistent with the inferred scaling of the stacked spectra (Figure 2) but is in direct violation of the classical self-similar model first proposed by Aki (1967).

We can quantify this scaling trend by performing a weighted regression analysis of stress drop and moment, fitting a linear model of the form

$$\log_{10} \Delta \sigma = \epsilon_0 + \epsilon_1 \log_{10} M_0,$$
Figure 4. The distribution and magnitude scaling of corner frequency and stress drop for southern Kansas earthquakes. (a) The total stress drop distribution, marginalized over all magnitudes and (b) distributions corresponding to different magnitude ranges. Median and log$_{10}$ standard deviation values are listed in each histogram inset. The gray bars in Figure 4b correspond to the total histogram (for all magnitudes) shown in Figure 4a. The scaling of (c) corner frequency $f_c$ and (d) stress drop $\Delta \sigma$ with seismic moment $M_0$. Black dots correspond to measurements of source properties for individual events, and the median $f_c$ and $\Delta \sigma$ in $M_0$ bins of width 0.4 (log$_{10}$ Nm units) are marked with orange squares. The best fitting scaling parameter $\epsilon_1$ for the binned data (obtained from weighted regression analysis, see text for details) is plotted with a solid orange line, and its numerical value and 2 sigma uncertainty is listed in the Figure 4d inset. The dashed black lines in Figure 4c correspond to constant $\Delta \sigma$ contours of 0.1, 1, 10, and 100 MPa. Open circles denote events with poorly resolved corner frequencies due to bandwidth limitations ($f_c > 25$ Hz or bootstrap interquartile uncertainty $> 5$ Hz).

where the parameter $\epsilon_1$ measures the slope of the increase of log$_{10}$ $\Delta \sigma$ with log$_{10}$ $M_0$. For the results presented here, we compute median stress drop in bins of width 0.4 in log$_{10}$ $M_0$, and apply weights based on the median uncertainty in $\Delta \sigma$ and number of observations in each bin, but obtain comparable results for both unbinned and unweighted regression. The scaling parameter $\epsilon_1 = 0.25$ (2sigma uncertainty $\pm 0.035$) is clearly positive and therefore inconsistent with the null hypothesis of self-similar, constant-$\Delta \sigma$ data, which would have $\epsilon_1 = 0$ to within the uncertainties. These results for the southern Kansas events are within the $\epsilon_1 \sim 0.1-0.4$ range of scaling results obtained for California earthquakes by Trugman and Shearer (2017a) and by other studies that have quantified an increase in stress drop or scaled energy with moment (e.g., Agurto-Detzel et al., 2017; Calderoni et al., 2013; Izutani & Kanamori, 2001; Mayeda and Walter, 1996; Mayeda et al., 2005, 2007; Mori et al., 2003; Pacor et al., 2016; Takahashi et al., 2005). However, our Kansas data set contains few higher-magnitude events (20 $M_1 > 3.5$ events, which is less than 1% of the total count), and as such we have poor resolution of the distribution and scaling of $\Delta \sigma$ above $M \sim 3.5$. Based on the data we do have, it would be reasonable to expect earthquakes with $\Delta \sigma$ in the 1–10 MPa range for the $M \geq 4$ that are of fundamental interest to hazard calculations.
Another point of interest from a hazard perspective is the systematic increase in median \( f_c \) and \( \Delta \sigma \) with hypocentral depth (Figures 5a and 5b). In contrast, we do not observe a comparable depth-dependent trend in \( M_0 \) that could potentially account for the scaling results presented above (Figure 5c). The depth dependence in both \( f_c \) and \( \Delta \sigma \) (but not \( M_0 \)) implies an increase in rupture velocity \( v_r \) that outpaces the expected increase in shear wave speed \( \beta \) with depth (i.e., an increase in the ratio \( v_r / \beta \) with depth). Because earthquakes in southern Kansas are characterized by a shallower depth distribution (2–8 km) compared to seismicity in the western and eastern United States (which typically extends to a significantly greater maximum depth), it is plausible that the lower median \( \Delta \sigma \) values we observe are due in part to the shallowness of the seismicity (Aguirto-Detzel et al., 2017; Boyd et al., 2017; Hardebeck & Aron, 2009; Pacor et al., 2016; Sumy et al., 2017). However, the shallower depth distribution of the events in our data set does not fully explain their anomalously low \( \Delta \sigma \), as can be seen quantitatively by comparing median \( \Delta \sigma \) for Kansas and California seismicity within a fixed depth range. For example, in Kansas we observe a median \( \Delta \sigma \) of 0.4 MPa at 5 km depth (Figure 5), compared to the 1–4 MPa range observed at this depth in five different regions of California (Trugman & Shearer, 2017a).

**Figure 6.** Temporal evolution of source parameters in southern Kansas. Magnitude (\( M_L \)) versus time, with events color coded by magnitude-adjusted, normalized stress drop \( Z_{\Delta \sigma} \) (equation (5)), with bluer colors indicating higher \( Z_{\Delta \sigma} \). Limited station coverage is sparse before September 2014, resulting in fewer events with resolvable source parameter estimates during this time.
Figure 7. Variability of source parameters within and between different earthquake sequences in southern Kansas. (a) Joint probability density of rescaled time $T_{ij}$ and rescaled distance $R_{ij}$ (log-log scale) for the nearest-neighbor cluster identification method (Zaliapin & Ben-Zion, 2013). The black dashed line corresponds to the threshold distance $\eta_{ij}$ used to partition events into individual sequences (see text for details). (b) Boxplot distribution of $\Delta \sigma$ (MPa) for prominent earthquake sequences in southern Kansas. Sequences medians are denoted with a solid horizontal line, while the box and whiskers denote the interquartile range (50% confidence interval) and 90% confidence interval, respectively. $\Delta \sigma$ values for the largest event in each sequence (main shocks) are marked with stars, and the corresponding magnitudes are listed along the x axis. (c) Similar to Figure 7b but for the distribution of magnitude-adjusted, normalized stress drop $Z_{\Delta \sigma}$ (equation (5)).

We next turn our attention to the temporal variability of $\Delta \sigma$ for earthquakes in our study region. It is useful at this stage to introduce a normalized, magnitude-adjusted $\Delta \sigma$:

$$Z_{\Delta \sigma} = \frac{\log_{10} \Delta \sigma - E[\log_{10} \Delta \sigma | M_0]}{\text{STD} \{\log_{10} \Delta \sigma - E[\log_{10} \Delta \sigma | M_0]\}},$$

where $E[\log_{10} \Delta \sigma | M_0]$ is expected $\Delta \sigma$, given $M_0$ (equation (4)), and STD $\{\cdot\}$ refers to the standard deviation. Thus, the metric $Z_{\Delta \sigma}$ is a normalized measure of the deviation from the expected $\Delta \sigma$ value of each event, conditioned on the observed $M_0$. This framework allows us to better isolate significant temporal variations in $\Delta \sigma$ by accounting for the magnitude-scaling trend observed in Figure 4. In Figure 6, we plot local magnitude as a function of time, with events color coded by $Z_{\Delta \sigma}$ such that bluer colors correspond to events with higher than expected $\Delta \sigma$ (i.e., events enriched in high-frequency energy). Prior to September 2014, we have few well-resolved source parameter estimates due to the sparsity in local station coverage before the ISMP network was fully installed. Following this, in the weeks preceding the 12 November 2014 Milan earthquake ($M_L 5.2, M_w 4.9$) (Choy et al., 2016), we observe elevated levels of $\Delta \sigma$. The Milan event itself has a slightly lower estimated $\Delta \sigma$ (3.6 MPa) than would be expected by extrapolating the scaling trend of equation (4) though still nearly an order of magnitude greater than the median value of the data set as a whole. The early aftershocks of Milan have lower than expected $\Delta \sigma$ values that appear to increase with time, consistent with observations of the Prague, OK, sequence (Sumy et al., 2017; Yenier et al., 2017). In early 2015, we observe another temporal cluster of high $\Delta \sigma$ events, with several other sequences later in 2015 and 2016 exhibiting analogous behavior.

To examine the variability in $\Delta \sigma$ within and between prominent earthquake sequences, we use the method described by Zaliapin and Ben-Zion (2013) that partitions events into individual sequences based upon nearest-neighbor space-time distances. The nearest-neighbors method has been shown to be effective in characterizing the space-time clustering statistics of both tectonic and induced earthquake sequences (Schoenball et al., 2015; Zaliapin & Ben-Zion, 2016). It defines the distance $\eta_{ij}$ between an event pair (parent $i$, daughter $j$) to be the product of a rescaled time $T_{ij} = dt_{ij} 10^{-M_i/2}$ and rescaled distance $R_{ij} = dr_{ij} 10^{-M_i/2}$, where $dt_{ij}$ is the difference in time in years, $dr_{ij}$ is the spatial distance in kilometers, $M_i$ is the magnitude of the parent event, and $d = 1.6$ is the assumed fractal dimension. Events are then linked to their nearest neighbors,
and individual sequences are defined by selecting a threshold distance $\eta_{ij}$ such that the sequences are sufficiently clustered in space and time (Figure 7a). With sequences thus defined, we then compare the variability in $\Delta \sigma$ both within and between the most prominent sequences in our data set, each of which contains at least one $M_L > 3.0$ event (the largest of which we classify as the main shock of the sequence). Although $\Delta \sigma$ typically varies by slightly more than an order of magnitude within each sequence, median values can vary by as much as a factor of 3 between sequences (Figure 7b). Main shock values of $\Delta \sigma$ tend to be higher than the median value of their respective sequences, as expected given the observed magnitude scaling. However, we do not observe comparable systematic differences for the magnitude-corrected values of $Z_{\Delta \sigma}$ (Figure 7c).

These temporal and sequence-specific variations in $\Delta \sigma$ could be caused by a number of factors, including local variations in geologic properties, crustal stress heterogeneity and its redistribution during and between individual earthquake sequences, and time-dependent changes in anthropogenic stressing from oil production and fluid injection. To gain insight into the latter, in Figure 8, we plot $Z_{\Delta \sigma}$ in map view for earthquakes within our study region and compare to the locations of active wastewater injection and enhanced oil recovery wells (section 2), both of which are thought to influence seismicity rates (e.g., Rubinstein & Mahani, 2015). We do not observe a significant correlation between $Z_{\Delta \sigma}$ and radial distance to the nearest active well (Figure 9), although it is interesting that clusters of events with the highest $Z_{\Delta \sigma}$ values tend to be near active wells. This weak or nonexistent dependence suggests that the presence of temporally and spatially coherent clusters of events with similar stress drop are caused primarily by factors unrelated to a localized influence injection activity of the nearest wells, such as local differences in fault strength, the distribution of geological or geometric asperities, or lithology. However, we note that because we do not have access to daily injection records for each well, it is difficult to perform a truly quantitative analysis in this regard.

5. Discussion
In this study, we use $P$ wave spectra from earthquakes in southern Kansas to provide source parameter and uncertainty estimates for seismic
moment ($M_0$), corner frequency ($f_c$), and Brune-type stress drop ($\Delta\sigma$). However, we emphasize that the absolute values of the source parameters are valid only under the assumption of the assumed source model, which in this case is a Brune-type spectrum (Brune, 1970) with $\omega^{-2}$ high-frequency falloff (equation (2) with $n = 2$). Our uncertainty estimates are therefore lower bounds because they neglect the epistemic uncertainties associated with this parameterization of the source spectral model and with the assumption that a circular crack rupture (equation (3)) adequately describes the relevant source physics. As discussed in Trugman and Shearer (2017a), the strength of the inferred scaling ($c_r$) of $\Delta\sigma$ with $M_0$ is correlated with the assumed high-frequency falloff rate $n$. If, for example, the true average $n$ for these earthquakes is less than the canonical $\omega^{-2}$ value of 2 (Brune, 1970), then the scaling parameter $c_r$ will be lower, and more generally, if $n$ varies on an event-to-event basis, this will bias estimates of $f_c$ and $\Delta\sigma$ for events in which $n$ differs markedly from 2. The spectra do not contain adequate signal bandwidth or precision to provide independent estimates of $f_c$ and $n$, only their joint influence on the spectral shape.

Despite these concerns, we can still draw useful conclusions by (i) focusing on relative variations in the source parameters, which tend to be robust with respect to the model parameterization, and (ii) by comparing to other data sets of earthquakes analyzed using the same methodological assumptions. Even from this more cautionary perspective, we can conclude that the southern Kansas earthquakes are characterized by relatively low $\Delta\sigma$ values compared to naturally occurring seismicity within tectonically active regions of California. The observed increase in $\Delta\sigma$ with hypocentral depth cannot fully account for this discrepancy, which suggests that the nucleation or rupture processes of these events may differ in some more fundamental way. While these results are consistent with several recent studies of likely induced earthquakes (Agurto-Detzel et al., 2017; Boyd et al., 2017; Suny et al., 2017), others (Huang et al., 2016, 2017; Zhang et al., 2016) have suggested that tectonic and induced events have comparable source parameters. These disparate conclusions may indicate that differences in local faulting conditions, tectonic stress regime, or history of anthropogenic activity may all play an important role. For example, the 25 earthquakes in the Guy-Greenbrier, Arkansas, earthquakes analyzed by Huang et al. (2016) differ substantially in both their tectonic setting and exposure to widespread regional injection compared to the southern Kansas earthquakes. Boyd et al. (2017) and Huang et al. (2017) both suggest that variations in $\Delta\sigma$ within the U.S. can be understood in the context of Mohr-Coulomb theory, in which the failure stress depends on hypocentral depth, fault style, and coefficient of friction (a proxy for fault strength). The results we present here are broadly consistent with this framework, as the shallower hypocentral depths and transtensional tectonic regime of southern Kansas would portend lower median $\Delta\sigma$ than in the deeper, transpressional regime of Southern California. Local variations in fault strength and its dependence on fluid injection may have an additional modulating effect that could be explored further in future studies.

We also observe an increase in $\Delta\sigma$ with magnitude for the smaller events that comprise the majority of our data set ($M < 3.5$). Although the inferred scaling is comparable to that observed in California (Trugman & Shearer, 2017a), it is sensitive to the modeling assumptions as discussed above. Bandwidth limitations can in some circumstances hinder the resolution of $f_c$ for smaller magnitude events (Abercrombie, 2015; Ide, 2003; Huang et al., 2016), but we believe that this effect does not significantly influence the results presented in this study. The low median stress drop of the southern Kansas earthquakes, when combined with the low regional attenuation and wide available signal bandwidth, provides a near-optimal setting for source spectral analyses, and indeed only a small fraction of our data set is poorly resolved (Figure 4). Further, the observed scaling is controlled fundamentally by the shape of the stacked relative source spectra (Trugman & Shearer, 2017a) and not the inferred source parameter values of individual events. It is, however, possible that this scaling trend may not extrapolate linearly to the larger magnitudes ($M4$ and $M5$) that are poorly sampled by our data set. Still, the larger events in our data set, including the $M_{L}5.2$ Milan earthquake, typically have $\Delta\sigma$ in the $1–10$ MPa range, which is $2.5–25$ times greater than the median value of the data set as a whole.

We observe coherent spatial and temporal variations and clustering of stress drop within our study region, but these variations do not appear to have a clear relation with distance to nearby injection and enhanced oil recovery wells. This is perhaps not surprising, as we lack adequate spatial and temporal resolution within the publicly available data to disentangle its effect from the other factors controlling source properties. Further, while the first-order influence of injection on seismicity rate is apparent based on comparison to the historical record (Buchanan, 2015; Buchanan et al., 2014; Choy et al., 2016; Weingarten et al., 2015), its immediate influence on source properties is more nebulous both from an observational perspective due to the lack of historical precedent and from a geophysical perspective due to the complex, nonlinear interactions.
between anthropogenic stressing and the rupture dynamics of triggered events. Sumy et al. (2017) likewise observe both significant temporal changes in $\Delta \sigma$ and a lack of correlation with injection well distance for aftershocks of the 2011 Prague, Oklahoma sequence and attribute the low observed $\Delta \sigma$ values to the more widespread effects of regional injection patterns that weakened basement fault structures on regional rather than local length scales. Fluid injection on the spatial scale of that observed in southern Kansas during this time period may generate significant stress perturbations at distances of tens of kilometers or more through a combination of pore pressure increase and poroelastic stressing (Goebel et al., 2017; Segall & Lu, 2015). It is also notable that Boyd et al. (2017) observe both a mild increase of $\Delta \sigma$ with $M_0$ for their data set composed entirely of $M > 3$ events, as well as anomalously low $\Delta \sigma$ for select aftershock sequences in the central United States like that of the 2014 Milan event.

In addition to their utility in understanding source dynamics, source parameter estimates may also provide observational constraints for seismic hazard assessment. Because ground motion intensities at high frequencies are controlled primarily by stress drop (Baltay et al., 2013, 2017; Boore, 2003; Douglas & Edwards, 2016; Yenier & Atkinson, 2014), its characterization is of fundamental interest to studies that aim to develop ground motion prediction equations for induced events (Atkinson & Assatourians, 2017; Atkinson et al., 2016; Yenier et al., 2017). In this study, we observe quantifiable time-dependent and depth-dependent variations in stress drop, both of which are in accord with the conclusions of Yenier et al. (2017) and Atkinson and Assatourians (2017) for ground motions of recent seismicity in Oklahoma. This consistency suggests that the results we present could potentially serve as a basis for future studies focused on quantifying the influence that spatiotemporal and depth-dependent variations in earthquake source properties may have on the observed ground motion amplitudes of induced earthquakes in the central United States.

### 6. Summary

We estimate seismic moment, corner frequency, and Brune-type stress drop for 2,069 $M_{1.5} - 5.2$ earthquakes occurring from 2014 to 2016 in an active area of wastewater injection and oil and gas production in southern Kansas. We find that these earthquakes have relatively low stress drop values that increase with hypocentral depth. We observe an increase in median stress drop as a function of magnitude for the $M_{1.5} - 3.5$ earthquakes that comprise the majority of our data set. However, this scaling trend may partially slow or saturate at higher magnitudes, and its strength is sensitive to the parameterization of the assumed source model. We find coherent temporal and spatial variations in the source parameters of earthquakes in southern Kansas, but these variations are not systematically related to the activity of nearby wastewater injection and enhanced oil recovery wells.

### References


