# Project report for SCEC proposal "Using deep learning to probe the spatiotemporal evolution and fault structure of the four-year Cahuilla, California earthquake swarm"

## Project objectives

The primary goals of this project were to study the spatiotemporal evolution and fault structure behind the 2016-2020 earthquake swarm near Cahuilla, California. This swarm was among the most prominent—both in number of events, and duration—observed in California over several decades. We proposed to use recently developed deep learning algorithms to build a very detailed earthquake catalog and use it to study the physical processes behind this swarm. Beyond seismicity, we examined earthquake source properties over the duration of the sequence. The results from this study were published in *Science* in May 2020 (doi:10.1126/science.abb0779).

## Methods

<u>Seismicity catalog construction</u>

We used deep learning methods to build our seismicity catalog by detecting and locating earthquakes over the entire continuous waveform dataset. The procedure first uses the generalized phase detection approach (Ross et al., 2018) to perform single station phase detection on all 3-component stations and then associates these detections across the seismic network to specific event detections.

For the phase detection, we built a training dataset of 1 million P-wave and S-wave seismograms from southern California. The datasets were assembled by randomly selecting 1 million P- and S-wave picks each from the SCSN archives (see Data), along with their associated 3-component seismograms. The seismograms were 16 seconds long, with the pick randomly located somewhere within the 16 second window. The data were filtered between 1-20 Hz for this part of the procedure.

The neural network uses a U-net architecture (Ronneberger et al., 2015) to predict a binary segmentation mask 0.2 sec in duration, centered on the real picks, which was motivated by the segmentation network of (Pardo et al., 2019) that uses DeepLab v3+. The model was trained using the Adam optimization algorithm (Kingma & Ba, 2014) with default parameters and early stopping criteria. Picks were made by taking the peak sigmoid probability whenever a value of 0.5 was exceeded for either phase type. The trained model was applied to all continuous waveform data to build a database of tentative phase arrivals. In total, 37 million phase detections were made.

Next, we applied the PhaseLink algorithm (Ross et al., 2019) to associate the phase detections to earthquakes and build an initial catalog. PhaseLink is a deep learning algorithm for phase association and uses Gated Recurrent Units to sequentially predict and link together phase detections. A synthetic training dataset was created for the region shown in Fig S1 by placing hypocenters randomly throughout the region. This procedure is described in detail in Ross et al. (2019). We processed windows of 500 picks at a time with a maximum sequence duration of 120 sec. We required 12 phase detections to nucleate a cluster, merged clusters with at least 1 phase detection in common, and required a minimum of 12 detections left after removing duplicates to save an event detection. These parameters were chosen to ensure the detections were of very high quality and also to keep the false positive rate well below 1%, which was performed by visual inspection of random detections. In total 22,700 events were detected.

The detected events were then located with NonLinLoc, a probabilistic non-linear hypocenter inversion algorithm (Lomax et al., 2000). We used the 1D velocity model of Hadley and

Kanamori (Hadley & Kanamori, 1977) and equal differential time likelihood function. The results were refined using station correction factors from the mean residuals over all events. Local magnitudes were estimated using the waveform processing workflow, attenuation relations, and station terms used by the Southern California Seismic Network (Uhrhammer et al., 2011). The calculated local magnitudes range from -0.27 to 4.56. We use a goodness of fit test at the 95% level (Wiemer & Wyss, 2000) to estimate the magnitude of completeness of 0.56 and a b-value of 1.05. The local magnitude range for this dataset corresponds to a moment magnitude range of 0.7 to 4.4.

We then cross-correlated the earthquakes to measure precise differential times for relocation purposes. We used seismograms starting 0.1 sec before all picked phases and lasting a total of 1.0 seconds. For each earthquake, we correlated it with the 200 nearest neighbors. We required a minimum cross-correlation coefficient of 0.8 and band-pass filtered the data between 1 and 15 Hz beforehand. This resulted in 89 million differential times.

Finally, the catalog was relocated with GrowClust (Trugman & Shearer, 2017), a cluster-based double-difference relocation algorithm. We used a minimum *r* value of 0.82 and required at least 10 differential times for relocation. We performed 100 bootstrap resamples to estimate location errors. To prevent any potential issues from cluster splitting, we lowered the minimum connectivity fraction (0.0001) until all events in the swarm were joined into a single cluster.

Stress drop estimation

We use the spectral decomposition technique detailed in (Trugman et al., 2017) to estimate Brune-type stress drops for a well-recorded subset of full event catalog. Spectral decomposition is an iterative inversion technique that takes as input a vector containing observed displacement spectra in log-units from many earthquakes recorded at many stations, and solves for relative source, station, and path terms at each frequency point. We apply a multitaper approach to compute P-wave spectra on vertical component channels from time windows that are 1.5 s in length, beginning 0.1 s before the phase arrival. In the inversion, we only consider spectra from that have signal-to-noise (SNR) ratio of 3.0 in the 3-30Hz frequency band. We further exclude events that have local magnitude less than 0.5 or have fewer than 6 spectral recordings that meet the SNR criteria. We also implement an automated procedure to screen for and exclude spectra from clipped waveforms, which are few in number but are sometimes observed on broadband recordings of the larger events.

Spectral decomposition is effective in isolating the relative shapes of the source spectrum, but to calculate stress drop, we first need to apply an empirical correction to remove path and station effects that are common to all sources, such as average near-source and near-station attenuation. We use the approach of (Trugman et al., 2017) to do this, which assumes that on average, the source spectra in waveform dataset conform to the Brune model with $f^{2}$ falloff (Brune, 1970), but makes no assumptions regarding self-similarity. With the corrected source spectra in hand, we estimate corner frequency $f_c$ and compute stress drop using the usual relation:

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

where $M_0$ is the estimated seismic moment (which is proportional to the low-frequency spectral moment), $\beta$ is the shear wavespeed, which we assume to be depth-dependent and estimate from the 1D velocity model, and $k$=0.38 is a model-dependent constant (Kaneko & Shearer, 2014) that has a significant impact on the absolute level of stress drop but not the local variability, which is our focus in this work. We obtain measurement uncertainties for the stress drop of each event by bootstrapping the set of stations that go into the source spectral measurements, obtaining median values (1-sigma) of 0.15 in log10 MPa units. In total, we obtain stress drop estimates for 3041 events in the main cluster of the Cahuilla swarm (Figure S2).

**Results**

A high-resolution perspective of the 3D geometry of the fault zone is observed in the relocated seismicity. The swarm occurred within a single non-planar right-lateral strike-slip fault surface that dips 70-80 degrees toward the northeast (Fig. 1). In the fault-normal direction, the seismicity is just tens of meters thick, which is comparable to the relative horizontal location errors (38 meters); the seismogenic portion of the fault zone may therefore be even narrower. The fault geometry is non-planar both along-strike and along-dip, most notably at the northwestern edge, where the fault surface undergoes multiple reversals in the dip direction as a function of depth.

Perhaps most strikingly, within the fault zone, there is a stack of elongated strike-parallel channels of relatively high seismicity density with vertical separation of about 200-400 meters. This distance is larger than the relative vertical error (87 meters). These channels have a weakly undulating geometry and persist over most of the fault zone. They are also prominently observed in the aggregate distribution of focal depths (Fig. 1), where they are represented by five peaks.

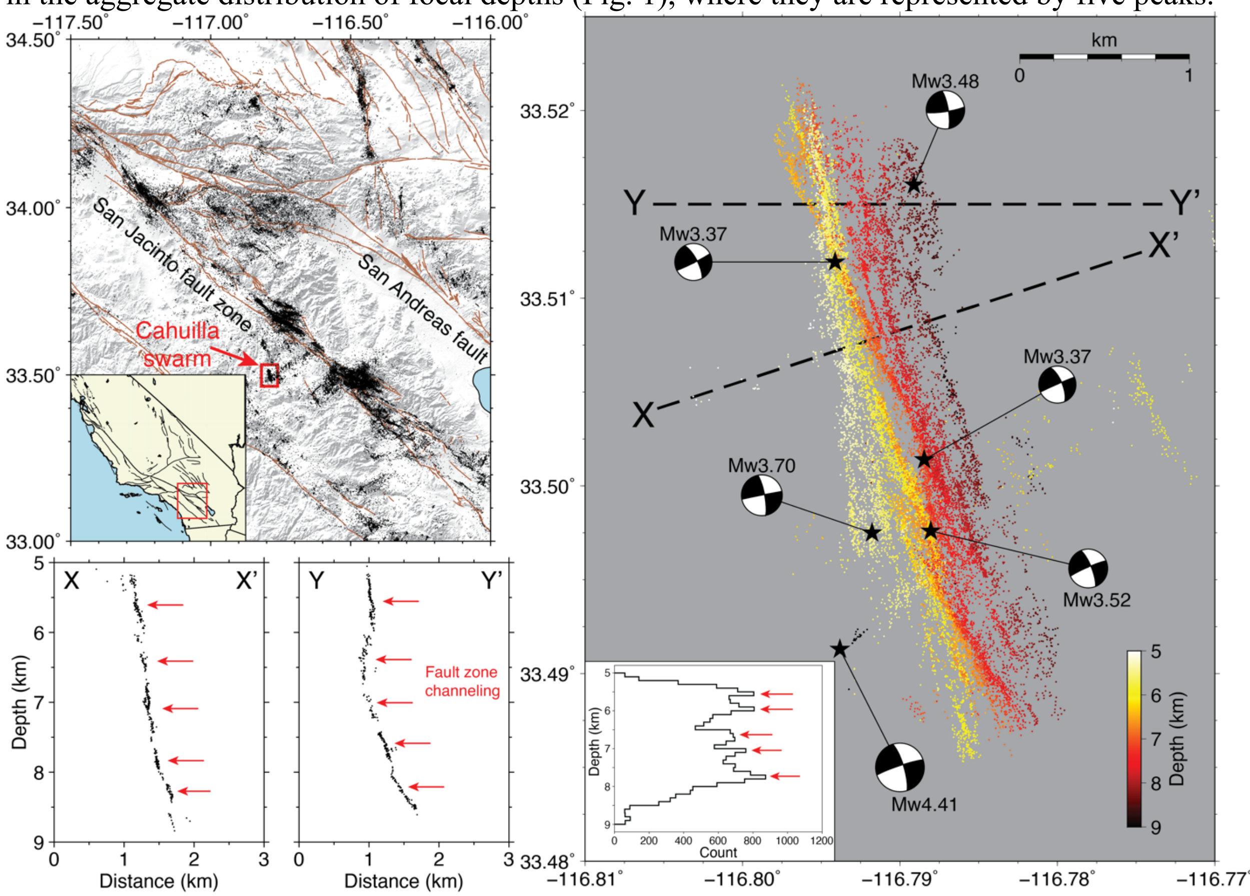


**Fig. 1.** Relocated seismicity during the Cahuilla swarm. The fault zone is less than 50 m wide and exhibits prominent channeling along-strike. Cross sections X-X' and Y-Y' show that there is considerable curvature to the fault geometry over short distances.

The 3D fault geometry is critical for understanding the spatiotemporal evolution of the swarm (Fig. 2). We can trace the origin of the sequence to a 100m-wide region near the base of the fault zone (8 km depth). The swarm expanded slowly within the fault zone over several years, ultimately extending across a 4 km by 4 km area. We observed a prominent front to the migration, with sustained earthquake activity behind it (Fig. 2). However, virtually no earthquake triggering occurs ahead of the front, which is consistent with expectations for a sustained point injection source (Shapiro et al., 1999). Interpreting the migration of seismicity as due to fluid flow, the slow diffusion front suggests that the swarm is driven predominantly by fluid migration (Hainzl & Ogata, 2005), rather than static or dynamic triggering as is more typical of earthquake sequences. The migration speeds are extremely slow (1-5 m/day), which is much slower than the propagation speed of slow slip events (tens of km/day) (Roland & McGuire, 2009), and also slower than many natural injection-driven swarms (Cox, 2016). We estimate the permeability of this fault zone to be ~$10^{-17}$-$10^{-18}$ $m^2$ ("Materials and Methods are described in the Supplementary Material.," n.d.). The observations bear similarities to earthquake sequences induced by deep fluid disposal (Shapiro et al., 1999), although here fluid injection is occurring directly into the fault zone from a deep, natural source.

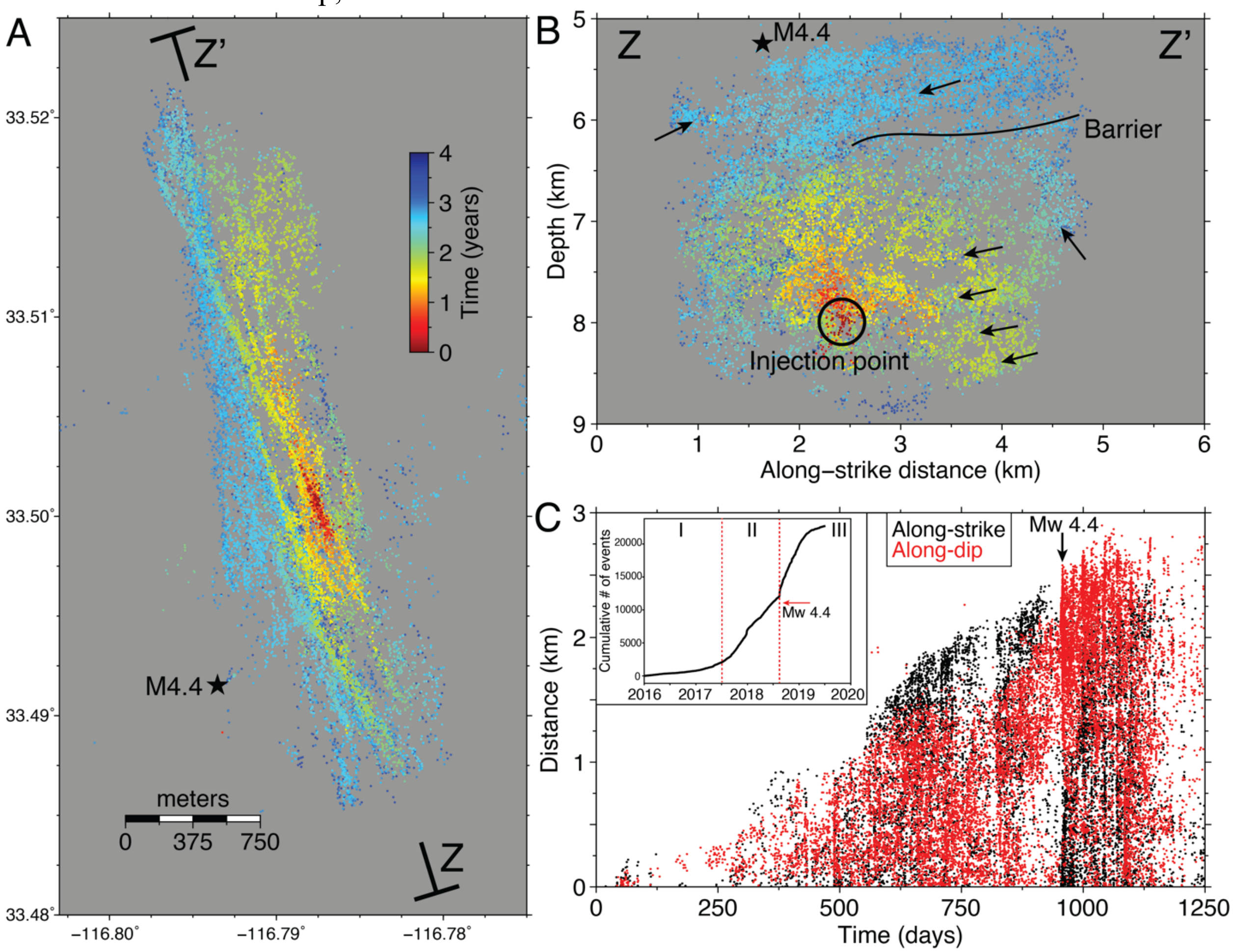


**Fig. 2.** Spatiotemporal evolution of the sequence. Fluids are naturally injected into the base of the fault zone and diffuse within it through channels (black arrows). A) Map view of seismicity evolution, with events color-coded by occurrence time. B) Fault-parallel cross-section along Z-

Z', note the inferred permeability barrier and injection point. C) Absolute distance of seismicity relative to injection point in B projected into along-strike and along-dip components.

The spatial expansion of the swarm is not uniform, indicating that the permeability architecture of the fault zone is variable (Fig. 2A,B). There may be a permeability barrier 0.5 km below the injection source at around 8 km depth, as only a minimal amount of deeper seismicity occurs during the nearly four years of activity. This may alternatively indicate that the fault zone itself does not extend to greater depths, or that there is a change in rheology at this point. Laterally, the swarm is constrained on both sides after expanding to an along-strike length of 4 km, which is reached about 3 years into the sequence. This lateral expansion is asymmetric, migrating about 1.8 times faster to the northwest than the southeast. Further expansion may be limited by permeability barriers, or the fault zone itself may not extend beyond the final spatial extent of the sequence. Fluid pressures may also no longer be high enough to produce additional events. In the fault-normal direction, the swarm is limited to a few tens of meters in width with no evidence of expansion (Fig. 1). Our observations are consistent with models of fault architecture that have a damage zone embedded within relatively impermeable host rock (Bense et al., 2013).

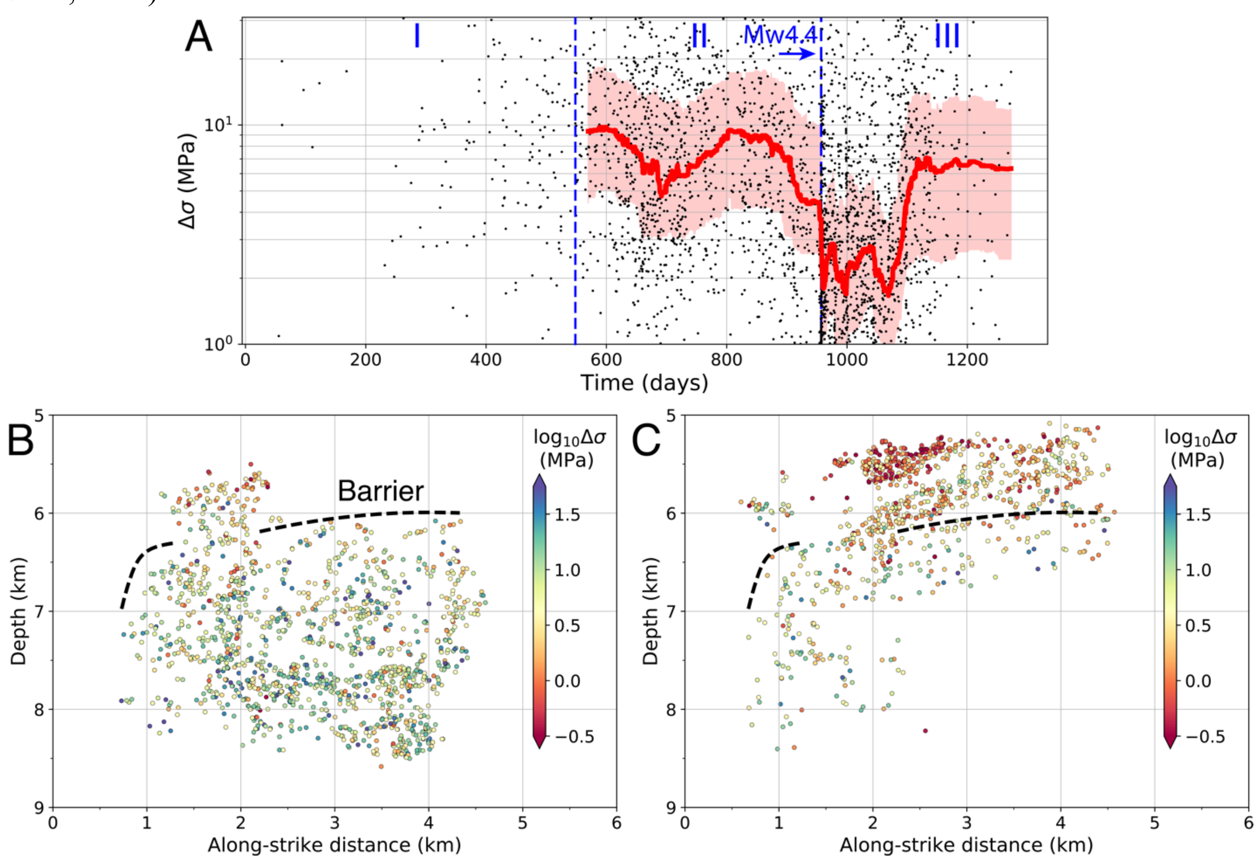


**Fig. 3.** Evolution of stress drop (Δσ). A) Time evolution. Solid red line marks the median Δσ in a causal, 200-event moving window. Shading denotes inter-quartile range. B) Along-strike cross-sectional view of events occurring before $M_w$4.4. Bluer colors indicate events that are enriched in high-frequency energy. Δσ decreases as the swarm circumvents the barrier at 6 km depth. C) Along-strike cross-sectional view of events occurring in the first 50 days after the $M_w$4.4. For events shallower than 6 km, Δσ is an order of magnitude lower.

## References


Bense, V. F., Gleeson, T., Loveless, S. E., Bour, O., & Scibek, J. (2013). Fault zone hydrogeology. *Earth-Science Reviews*, *127*, 171–192. https://doi.org/10.1016/j.earscirev.2013.09.008

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

Cox, S. F. (2016). Injection-Driven Swarm Seismicity and Permeability Enhancement: Implications for the Dynamics of Hydrothermal Ore Systems in High Fluid-Flux, Overpressured Faulting Regimes—An Invited Paper. *Economic Geology*, *111*(3), 559–587. https://doi.org/10.2113/econgeo.111.3.559

Hadley, D., & Kanamori, H. (1977). Seismic structure of the Transverse Ranges, California. *GSA Bulletin*, *88*(10), 1469–1478. https://doi.org/10.1130/0016-7606(1977)88<1469:SSOTTR>2.0.CO;2

Hainzl, S., & Ogata, Y. (2005). Detecting fluid signals in seismicity data through statistical earthquake modeling. *Journal of Geophysical Research: Solid Earth*, *110*(B5). https://doi.org/10.1029/2004JB003247

Kaneko, Y., & Shearer, P. M. (2014). Seismic source spectra and estimated stress drop derived from cohesive-zone models of circular subshear rupture. *Geophysical Journal International*, *197*(2), 1002–1015. https://doi.org/10.1093/gji/ggu030

Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *ArXiv:1412.6980 [Cs]*. Retrieved from http://arxiv.org/abs/1412.6980

Lomax, A., Virieux, J., Volant, P., & Berge-Thierry, C. (2000). Probabilistic Earthquake Location in 3D and Layered Models. In C. H. Thurber & N. Rabinowitz (Eds.), *Advances* *in Seismic Event Location* (pp. 101–134). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-015-9536-0_5

Materials and Methods are described in the Supplementary Material. (n.d.).

Pardo, E., Garfias, C., & Malpica, N. (2019). Seismic Phase Picking Using Convolutional Networks. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(9), 7086–7092. https://doi.org/10.1109/TGRS.2019.2911402

Roland, E., & McGuire, J. J. (2009). Earthquake swarms on transform faults. *Geophysical Journal International*, *178*(3), 1677–1690. https://doi.org/10.1111/j.1365-246X.2009.04214.x

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234–241). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Ross, Z. E., Meier, M.-A., Hauksson, E., & Heaton, T. H. (2018). Generalized Seismic Phase Detection with Deep Learning. *Bulletin of the Seismological Society of America*, *108*(5A), 2894–2901. https://doi.org/10.1785/0120180080

Ross, Z. E., Yue, Y., Meier, M.-A., Hauksson, E., & Heaton, T. H. (2019). PhaseLink: A Deep Learning Approach to Seismic Phase Association. *Journal of Geophysical Research: Solid Earth*, *124*(1), 856–869. https://doi.org/10.1029/2018JB016674

Shapiro, S. A., Audigane, P., & Royer, J.-J. (1999). Large-scale in situ permeability tensor of rocks from induced microseismicity. *Geophysical Journal International*, *137*(1), 207–213. https://doi.org/10.1046/j.1365-246x.1999.00781.x

Trugman, D. T., & Shearer, P. M. (2017). GrowClust: A Hierarchical Clustering Algorithm for Relative Earthquake Relocation, with Application to the Spanish Springs and Sheldon, Nevada, Earthquake Sequences. *Seismological Research Letters*, *88*(2A), 379–391. https://doi.org/10.1785/0220160188

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*(10), 8021–8034. https://doi.org/10.1002/2017JB014649

Uhrhammer, R. A., Hellweg, M., Hutton, K., Lombard, P., Walters, A. W., Hauksson, E., & Oppenheimer, D. (2011). California Integrated Seismic Network (CISN) Local Magnitude Determination in California and VicinityCalifornia Integrated Seismic Network Local Magnitude Determination in California and Vicinity. *Bulletin of the Seismological Society of America*, *101*(6), 2685–2693. https://doi.org/10.1785/0120100106

Wiemer, S., & Wyss, M. (2000). Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan. *Bulletin of the Seismological Society of America*, *90*(4), 859–869. https://doi.org/10.1785/0119990114