Revisiting the Timpson Induced Earthquake Sequence: A System of Two Parallel Faults

Kaiwen Wang[1], William Ellsworth[1], Gregory C. Beroza[1]

[1]Department of Geophysics, Stanford University, Stanford, California, USA

Corresponding author: Kaiwen Wang (kaiwenw@stanford.edu)

Key Points:

- Deep learning enables us to enrich the catalog and study detailed fault structure
- Event clustering and relocation reveals complex fault structure with two parallel subfaults
- The two subfaults are better oriented for failure in the local stress field than the overall trend

Abstract

The 17 May 2012 M4.8 Timpson earthquake is the largest known earthquake in eastern Texas. It is suspected to have been induced by wastewater injection from two nearby, high-volume wells. Its cataloged aftershocks form a NW-SE trend, which unlike other induced earthquakes sequences, is unfavorably oriented for failure in the local stress field. To understand this, we enriched the catalog using PhaseNet, a deep-learning-based picker followed by double-difference relocation with cross-correlation-based differential traveltimes. We clustered the aftershocks based on waveform similarity. Most of the seismicity falls into two-clusters, which define a complex fault structure of two parallel subfaults that are more favorably oriented than the overall trend. We inferred from waveform similarity that the sequence initiated on the northern subfault with a M3.9 foreshock and M4.8 mainshock, then extended to the southern subfault with a M4.1 aftershock, and was finally reactivated on the northern subfault with two more M4 events.

Plain Language Summary

The Timpson earthquake is the largest recorded earthquake in east Texas and was likely induced by wastewater injection. The fault would be very unlikely to host an earthquake with previously estimated orientation. To understand why the earthquake occurred, we reanalyzed the data to detect more events using deep learning, and followed that with additional analysis to reduce location uncertainties. We found the main fault was composed of two parallel subfaults. With the newly revealed structure, the fault would be more likely to host an earthquake than previously thought. Our findings suggest that detailed fault structural knowledge could revise the interpretation of a safe fault into a risky fault and is thus important when assessing earthquake hazards.

1 Introduction

The 17 May 2012 Timpson earthquake occurred after five years of wastewater injection in two nearby disposal wells. It is historically the largest recorded earthquake in east Texas, an area with otherwise low seismic hazard (Frohlich et al., 2014). The triggering mechanism for this earthquake sequence is still poorly understood. Frohlich et al. (2014) located the aftershocks and found that they occurred on a NW-SE striking fault. Snee and Zoback (2016) measured the local stress field from borehole breakouts and showed that the SHmax orientation ranged from N68°E to N80°E. They also calculated the Coulomb stresses on the previously determined NW-SE striking fault and found the fault delineated by aftershocks was quite unfavorably oriented for failure. Fan et al. (2016) conducted geomechanical modeling with poroelasticity and showed that the unfavorably oriented fault could be activated under highly elevated pore-pressure; however, their results relied on the strong assumption that no other faults with more favorable orientation were present. Because the notion that induced seismicity occurs on favorably oriented faults is important to characterizing their hazard, we revisited the Timpson induced earthquake sequence to understand the discrepancy.

We first present catalog reconstruction and earthquake relocation results in section 2. Then we describe the two parallel faults structure revealed by waveform-similarity based event clustering (section 3.1). We associated the five major events in the Timpson sequence using a combination of spatiotemporal correlation and waveform similarity with the two sub-parallel faults and resolved the sequence initiation and evolution (section 3.2). Finally, we show the two subfaults

are better oriented in the local stress field, which supports the importance of accurate fault geometry in hazard assessment (section 3.3).

2 Relocation of the Timpson events

Frohlich et al. (2014) built a catalog for the Timpson sequence and found the aftershocks coincide with a mapped fault. In this study, we enriched the catalog by reanalyzing continuous recordings from 26 May 2012 when the local stations were deployed through the end of 2013. We ran PhaseNet (Zhu & Beroza, 2019), a machine-learning-based phase picker to generate P and S phase picks and associated them using REAL (Zhang et al., 2019), a grid search associator. The associated picks were used as input for VELEST (Kissling et al., 1994) with an initial velocity model determined by Frohlich et al. (2014) and subsequently HypoDD (Waldhauser & Ellsworth, 2000; Waldhauser, 2001). We also added cross-correlation differential traveltime measurements in double difference relocation. The relocated events are shown in Figure 1a.

To study the detailed fault structure and evolution of the sequence, we applied hierarchical clustering to waveform similarity using Ward linkage as a clustering metric (Akuhara & Mochizuki, 2014). This is a classical method to group the events by minimizing the variance within clusters. Compared with using simple linkage it better divides the cloud with connecting events between two clusters, which is important in this case as the aftershocks locate close to one another. We filtered the waveforms from 1-20 Hz and calculated pairwise cross-correlation coefficients in 1 second windows around both the P and S wave arrivals. By windowing, we avoided having both the P and S wave in the same window at the cross-correlation step. This retains the sensitivity of cross-correlation coefficients to event waveforms rather than to S-P times, or event location. The distance in clustering space between each event pair was defined as one minus the maximum of cross-correlation coefficients for all stations and phases. We took a maximum to keep the correlation coefficient of the waveform with the highest quality because the station coverage and recording quality changes with time. As similar events tend to have high cross-correlation coefficients, they would have short distances in the clustering space by doing one minus the cross-correlation coefficients. We determined the number of clusters by balancing a tradeoff between distance within clusters and cluster size. One can always divide a cluster into two and make them more tightly connected inside, but splitting into too many small clusters would render it not useful for fault structure interpretation. In balancing this tradeoff, we decided to split the cloud into five clusters, which resulted in two major clusters that contain most events: cluster 1 on the north and cluster 2 on the south (Figure 1a). Events in the two main clusters are marked by dark gray (cluster 1) and light gray (cluster 2). Events from the same cluster are closely located in space even if we only provide waveform similarity information in clustering without using traveltimes.

The two September 2013 M4 events cluster 1 while the January 2013 event is cluster 2. These larger events occurrence time coincided with high seismicity rate in time as well. As shown in Figure 1b, the January 2013 event occurred when cluster 2 was most active and the two September 2013 events occurred during the peak activity of cluster 1. In Figure 1a the 10 May 2012 foreshock and 17 May 2012 mainshock are plotted with original catalog locations from Frohlich et al. (2014). They suggest these early event locations have large uncertainty due to a lack of local stations. We relocated the 10 May 2012 foreshock and 17 May 2012 mainshock using their traveltime differences at the more distant stations relative to other large events. The

relocated early foreshock and mainshock are shown in Figure 1a. Although they seem to locate closer to the northern cluster, considering their location uncertainty it is still not clear how the sequence initialized only from event locations and occurrence time. Below we use early event waveform similarity with the aftershocks as additional information to infer how this sequence initiated and evolved.

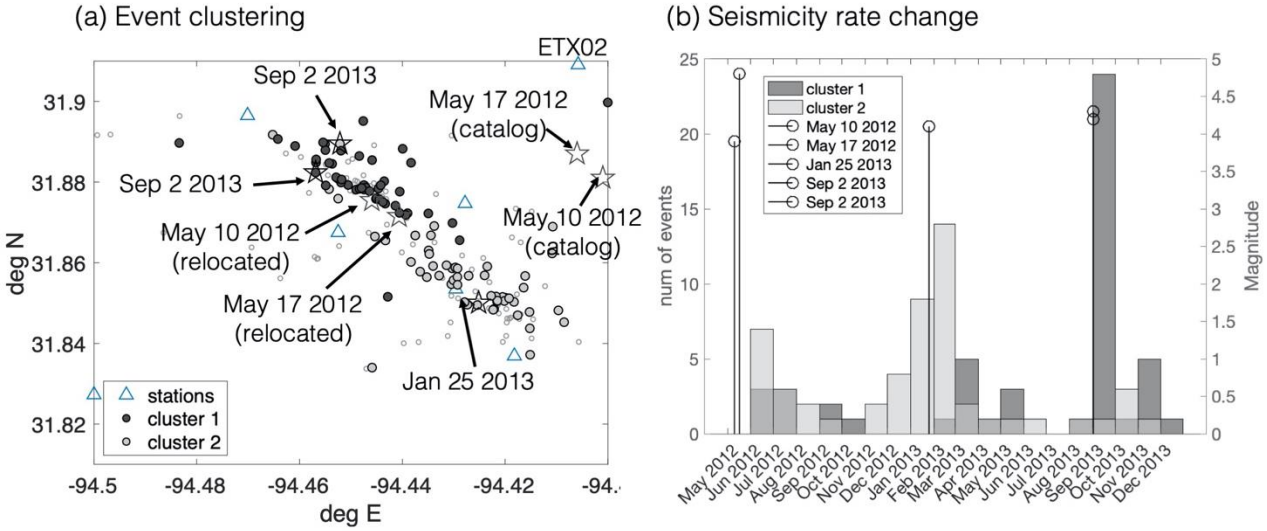

Figure 1. (a) Event clustering based on waveform similarity. Triangles show local stations that were installed after the mainshock. Stars show events over M4. Catalog locations of the M3.9 foreshock on 10 May 2012 and M4.8 mainshock on 17 May 2012 are from Frohlich et al., 2014. We plot Cluster 1 (northern cluster) and Cluster 2 (southern cluster) with dark and light gray circles. Small open circles are events in other clusters. (b) Monthly seismicity rate of the northern (Cluster 1) and southern (Cluster 2) cluster. Open circles indicate large events.

3 Refined fault structure and its implication for fault slip potential

3.1 The case for two parallel faults

We investigated fault structure by first separately relocating cluster 1 and cluster 2. Events in each cluster locate in a unique area and approximate a plane, suggesting that they occurred on distinct fault structures. To verify the geometry of the two clusters, we ran HypoDD (Waldhauser & Ellsworth, 2000; Waldhauser, 2001) in SVD mode with all events from both clusters (Figure 2). Within uncertainties the locations clearly delineate two subparallel faults with about a 1 km right-step offset. The local maximum horizontal stress orientation is N68°E to N80°E (Snee and Zoback, 2016). In this stress field the two faults we delineate are more favorably oriented for slip than the overall trend. The Coulomb stress analysis quantifies this statement in section 3.3 with updated information on the stress orientation. Frohlich et al. (2014) also proposed two groups based on the observation that events S-P time at ETX02(station marked in Figure 1a) formed two groups.

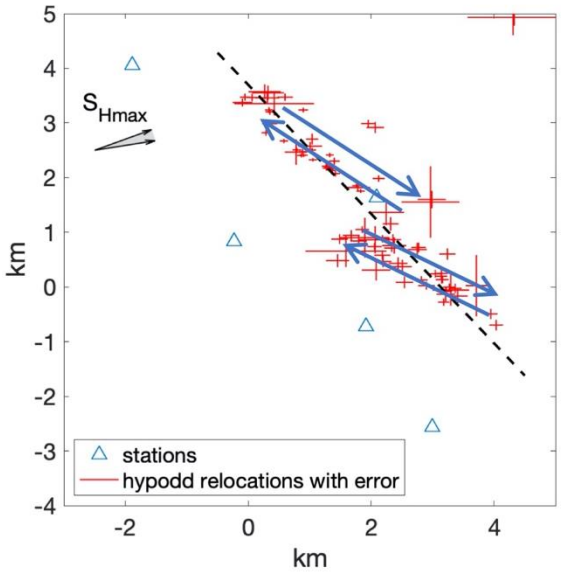

Figure 2. Relocated northern and southern clusters. Crosses are HypoDD relocations of the two clusters with 2σ error bars in SVD mode. The two faults are outlined by arrows and the entire trend found by Frohlich et al. (2014) is marked by dashed line. SHmax orientation measurements are from Lund Snee and Zoback (2016). Dashed line shows the overall trend before clustering. Parallel arrows are schematics of the two subfaults at a more favorable orientation.

3.2 Initiation and evolution of the sequence

Figure 3b visualizes characteristics of the seismograms in the two-clusters. We find systematic differences between the two clusters in the S wave first motion and P coda. Events in the same cluster are similar with one another. We can use this waveform similarity to infer early event locations during the initiation of the sequence and before local stations were deployed.

The Timpson sequence initiated with a M3.9 foreshock on 10 May 2012 and the M4.8 mainshock followed on 17 May 2012. The first temporary local station started operation on 26 May 2012 and the nearest permanent station, NATX is located 25 km to the southwest. As a consequence, the locations of early events, including the M3.9 foreshock and the M4.8 mainshock, are uncertain; however, we can infer their locations from their waveform similarity with the well-located aftershocks at NATX , which is the only regional station available at all times. We computed cross-correlation coefficients between early events and all the aftershocks recorded by NATX in the two well-located clusters. The waveform analysis of the M4.8 mainshock is shown in Figure 3a as an example. The events in the northern cluster and the southern cluster are shaded by their waveform correlation with the M4.8 mainshock. We can see from Figure 3a that the waveform of the M4.8 mainshock is more similar to the northern cluster than the southern cluster. A similar pattern can be seen for the M3.9 foreshock and other early events in 2011 as well. This suggests that the Timpson sequence likely initiated on the northern subfault. Together with the aftershock spatial and temporal pattern show in Figure 1, we infer the sequence evolution history as follows. The Timpson sequence initiated on the northern subfault with a M3.9 foreshock and the M4.8 mainshock in May 2012. It then migrated to the southern subfault with increasing activity leading to the January 2013 M4.1 aftershock. Later, the northern

subfault was reactivated with a peak in seismicity in September 2013 that resulted in two more M4 aftershocks.

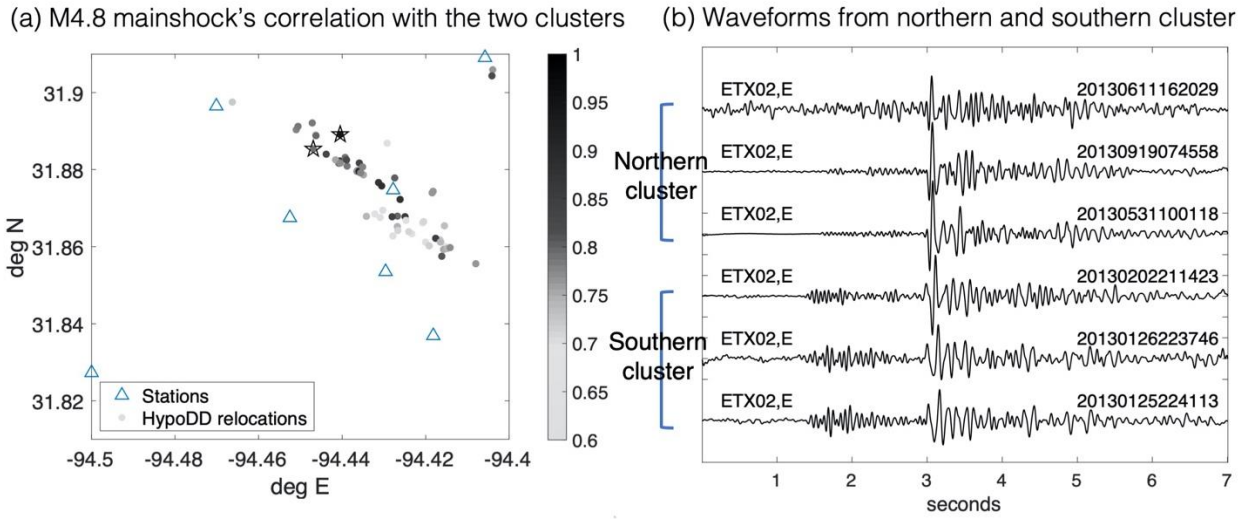

Figure 3. (a) Waveform similarity of the M4.8 mainshock (17 May 2012) with the aftershocks. Stars show the two M4 events in September 2013. The aftershocks on the two subfaults are colored by cross-correlation values between the M4.8 mainshock. Events in darker gray indicate more similar waveform with the M4.8 mainshock. (b) Waveforms examples from the northern and southern cluster. Waveforms are detrended and filtered between 1 to 20Hz, aligned at the S wave arrival time.

3.3 Changes of Coulomb stress with fault structure

We calculated the corresponding Coulomb stresses for the two-parallel-faults versus the single fault interpretation. To estimate the fault orientation with uncertainties, we took the event relocations with errors and simulated 1000 set of earthquake locations using Monte Carlo method (Browaeys, 2020). The fitted fault plane orientations calculated from the mean value of the 1000 slopes are plotted with 2σ errors in Figure 4. We first assumed the entire sequence occurred on a single fault with a best-fit fault orientation of N139.8°E, which agrees with N138°E trend reported by Frohlich et al.(2014). Then we estimated the best fit fault plane orientations from the events in the northern cluster and the southern cluster, respectively. We also checked whether the fault planes outlined by the aftershocks matched the strikes given by focal mechanisms of large events in the sequence. For the four large events with reported moment tensors by the U.S. Geological Survey (USGS) catalog (USGS National Earthquake Information Center, 2016), the reported strikes ranged from 141° to 171°, which is more uncertain than the fault plane orientation. The large uncertainty in the fault strike likely resulted from limited station coverage and velocity model uncertainties. Therefore, we cannot distinguish the two parallel faults scenario from the single fault scenario simply based on reported focal mechanisms. Next we calculated the stress conditions of the faults for each of the two scenarios. Here we applied the updated stress measurements from Lund Snee and Zoback (2020) shown in Figure 4. The local maximum horizontal stress orientation in their updated measurements gradually rotates from N85°E in the southwest to N68°E in the northeast. Given the new stress

measurements, the local maximum horizontal stress orientation would likely to be around N80°E, which we chose for the Coulomb stress calculation.

We visualize the stress conditions in a Mohr circle diagram (Figure 4). As we can see from Figure 4, the stresses acting on both faults, yellow for the northern cluster 1 and blue for the southern cluster 2 are closer to the failure line compared with the single fault scenario (red arrow). The two parallel faults are better oriented in the local stress field. By considering the two-fault structure, we require only half of the pore-pressure perturbation to activate the fault as opposed to the single-fault scenario. We also calculated the required pore-pressure on the conjugate fault plane of the M4.8 mainshock shown with the green arrow in Figure 4. Snee and Zoback(2016) suggested this conjugate plane was almost ideally oriented in the observed stress state. Our results show the required pore-pressure for the two parallel subfaults is comparable to the pressure required to slip the conjugate plane. Our calculated pore-pressure change also agrees with observation of Shirzaei et al. (2016) of little surface uplift, which argues against a large pore pressure increase. According to the geomechanical modeling results of Fan et al. (2016), the two parallel subfaults are within the range where fault slip would occur before hydraulic fracturing. Our findings emphasize that stability of a fault can be very sensitive to its orientation in the stress field. Here, a ten-degree change in the fault orientation or the local stress field orientation turned an unfavorably oriented fault into one plausibly activated by fluid injection. Therefore, precisely determined fault geometry can be a key factor for assessing fault slip potential.

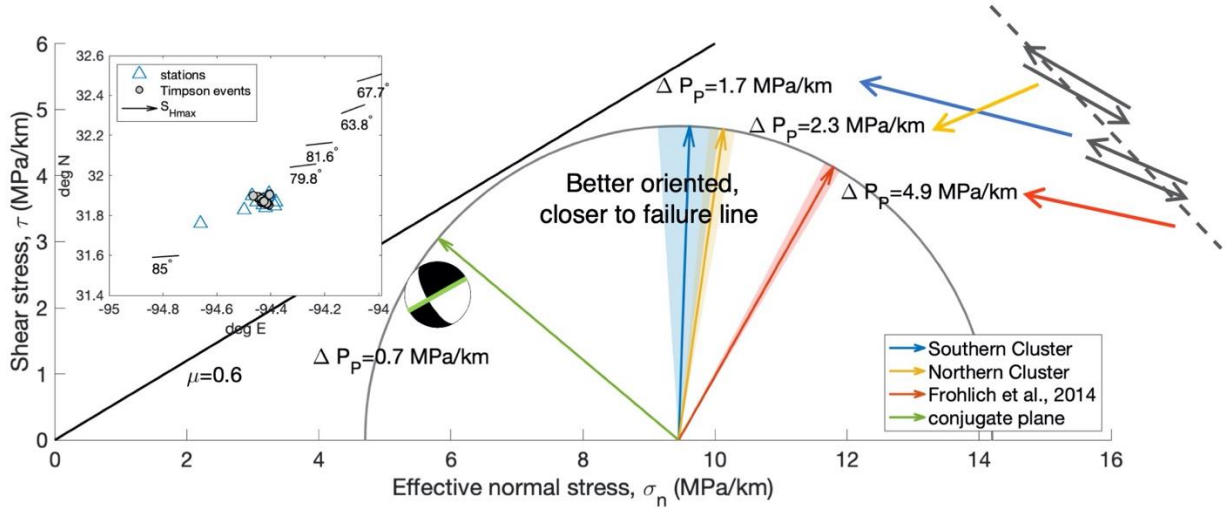

Figure 4. Mohr circle showing stress conditions of faults in the Timpson sequence. Inset figure shows stress measurements from Lund Snee and Zoback (2020). Red arrow shows shear and normal stresses on the plane defined by the overall trend. Blue arrow shows stresses on the southern subfault. Yellow arrow shows stresses on the northern subfault. Shaded area gives 2σ uncertainty in fault plane orientation. The northern and southern subfaults are closer to the failure line than the overall trend and thus better oriented in the local stress field. Green arrow indicates stresses on the conjugate nodal plane from Saint Louis University (SLU) moment tensor (Herrmann, 2016) of the M4.8 mainshock.

4 Conclusions

We revisited the 17 May 2012 Timpson earthquake sequence using machine-learning-based phase detection and uncovered complex fault structure by a combination of precise hypocenter relocation and waveform-similarity-based clustering. The Timpson earthquakes ruptured two nearly parallel subfaults that are separated by a 1 km right step. This structure was verified by double-difference relocation with cross-correlation differential traveltime measurements. We associated the five major events in the Timpson sequence with its aftershocks based on spatiotemporal correlation and waveform similarity and inferred the sequence evolution pattern as follows. The Timpson earthquakes started on the northern subfault which hosted the M3.9 foreshock and the M4.8 mainshock in May 2012. The southern subfault was then activated with increasing seismicity that culminated in a M4.1 event in January 2013. Later in September 2013, the northern fault was reactivated with a peak in seismicity and another two M4 aftershocks. We investigated how the two-subfault structure affected fault slip potential by calculating resolved Coulomb stress with and without this structure and the corresponding pore-pressure perturbation required in each scenario. The calculation suggests that the two subfaults are closer to failure than the single fault scenario and thus better oriented in the local stress field. Our findings provide a possible explanation to the discrepancy between stress and aftershock distribution in the Timpson induced sequence. A key implication for hazard assessment is that detailed fault structural knowledge may significantly change the fault slip potential and the hazard level of a fault.

Acknowledgments

This research was supported by the Stanford Center for Induced and Triggered Seismicity (SCITS). The waveforms can be downloaded from Incorporated Research Institutions for Seismology(IRIS) Data Services(https://ds.iris.edu/ds/nodes/dmc/data/types/waveform-data/). We thank Jens-Erik Lund Snee, Yen Joe Tan and Mostafa Mousavi for helpful discussions and suggestions.

References

Akuhara, T., & Mochizuki, K. (2014). Application of cluster analysis based on waveform cross-correlation coefficients to data recorded by ocean-bottom seismometers: results from off the Kii peninsula. *Earth, Planets and Space*, 66(1), 80.

Browaeys, J. (2020, May). Matlab toolbox linear fit with both uncertainties in x and in y. https://www.mathworks.com/matlabcentral/fileexchange/45711-linear-fit-with-both-uncertainties-in-x-and-in-y. MATLAB Central File Exchange.

Fan, Z., Eichhubl, P., & Gale, J. F. (2016). Geomechanical analysis of fluid injection and seismic fault slip for the mw4. 8 Timpson, Texas, earthquake sequence. *Journal of Geophysical Research: Solid Earth*, 121(4), 2798–2812.

Frohlich, C., Ellsworth, W., Brown, W. A., Brunt, M., Luetgert, J., MacDonald, T., & Walter, S. (2014). The 17 may 2012 m4. 8 earthquake near Timpson, east Texas: An event possibly triggered by fluid injection. *Journal of Geophysical Research: Solid Earth*, 119(1), 581–593.

Herrmann, R. B.(2016).Timpson, Texas earthquakes of May, 2012.http://www.eas.slu.edu/eqc/eqcmt/MECH.NA/20120517081201/TimpsonTX.pdf.

Kissling, E., Ellsworth, W., Eberhart-Phillips, D., & Kradolfer, U. (1994). Initial reference models in local earthquake tomography. *Journal of Geophysical Re- search: Solid Earth*, 99(B10), 19635–19646.

Matlab statistics and machine learning toolbox. (2020).

Snee, J.-E. L., & Zoback, M. D. (2016). State of stress in Texas: Implications for induced seismicity. *Geophysical Research Letters*, 43(19), 10–208.

Snee, J.-E. L., & Zoback, M. D.(2020).Multiscale variations of the crustal stress219field throughout north America. *Nature Communications*,11(1), 1–9.

USGS National Earthquake Information Center, PDE.(2016).234https://earthquake.usgs.gov/earthquakes/eventpage/usp000jkhb/moment-235tensor.

Waldhauser, F. (2001). HypoDD–a program to compute double-difference hypocenter locations.

Waldhauser, F., & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the northern hayward fault, California. *Bulletin of the Seismological Society of America*, 90(6), 1353–1368.

Zhang, M., Ellsworth, W. L., & Beroza, G. C. (2019). Rapid earthquake association and location. *Seismological Research Letters*, 90(6), 2276–2284.

Zhu, W., & Beroza, G. C. (2019). PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1), 261–273.