DOE Award number: DE-FE-0026517
Recipient of award: Archer-Daniel-Midland Company, 1001 N. Brush College Rd., Decatur, IL 62521
Title: Intelligent Monitoring System for Real-time Geologic Storage, Optimization, and Reservoir Management
Principal Investigator: J. Ole Kaven
Period covered in this progress report: 10/1/2015 – 9/30/2021
Submitted: September 2021

Task 4.0 Passive Microseismicity Monitoring

Executive summary:

The objective of the subtask was to develop a near-real-time monitoring system for seismic data at the Decatur, IL, geologic carbon sequestration (GCS) site and specifically include fiber-optic cable derived distributed acoustic signal (DAS) data in the process. Owing to the large volumes of data, we opted to utilize existing deep borehole conventional seismic sensors for detection and pull DAS and shallow borehole seismic data once a detection has been made. Unfortunately, the horizontal fiber-optic cables did not yield microseismic signals for use in locating events near the GCS site. Various stacking and filtering approaches were tested without any coherent detection becoming apparent. We attribute the insensitivity to local microseismic events to a lack of coupling in downhole, vertical cable and the non-ideal alignment of the horizontal fiber-optic cable to the vertically polarized seismic energy. Despite the inability to detect and utilize the DAS data from the fiber-optic cables, we developed a general processing framework that enables easy adaptation for future deployment of fiber-optic cables with better suited alignment at Decatur or elsewhere.

Subtask 4.1 Develop near-real-time extraction of DAS derived data into detection and location processing system.

1. General layout of data flow/processing/etc.

While passive microseismic monitoring on distributed acoustic sensing (DAS) can be successfully employed for microseismicity near the sensor string (<~1km), e.g., hydraulic fracturing operations (Molenaar at el., 2011), its applicability to monitoring more distant microseismicity is still in question (Karam et al., 2013). These difficulties are due to the low signal-to-noise ratio of small seismic events recorded on DAS, which is a consequence of the sensitivity of the DAS and often

due to the noise of the injection itself (e.g., Maxwell and Urbancic, 2001). However, DAS also offers advantages over conventional instruments, primarily a very large sensor count (4000+) and extensive linear coverage, both of which present unique opportunities for passive microseismic analyses.

DAS systems generate vast amounts of data given the large number of individual (virtual) sensors and the high sampling frequencies. To avoid major data transmission volumes, we iterated on the data acquisition and transmission design until dependable transmission of data from the field site to our servers was achieved. In all there are three different types of passive seismic data that we use in our near real-time processing: 1) the deep, conventional borehole seismometer network installed in wells GM1 and CCS1, hereon referred to as deep borehole array SCSDBA; 2) the shallow borehole and surface seismometer network deployed by the U.S. Geological Survey, referred to as GSDEC; and finally, 3) the fiber optic cable derived DAS data deployed in this project, hereon referred to as DECDAS.

The overall data acquisition, transmission, and event detection is shown in the data flow diagram in Figure 1. We start the processing by transmitting the SCSDBA data every ten minutes. This is done on the ADM server side using a code called `sendFileII`, which requires firewall exemptions on both the ADM- and USGS-server sides. Once the SCSDBA data is sent to our project specific server (ISDPXL), the data file is unzipped and individual file names are time stamped to Coordinated Universal Time (UTC time) based on the start times recorded in the headers of the files. Each file is moved into a daily (UTC date and time) directory on the project server (ISDPXL). The separate steps are depicted in the top three boxes in the diagram in Figure 1.

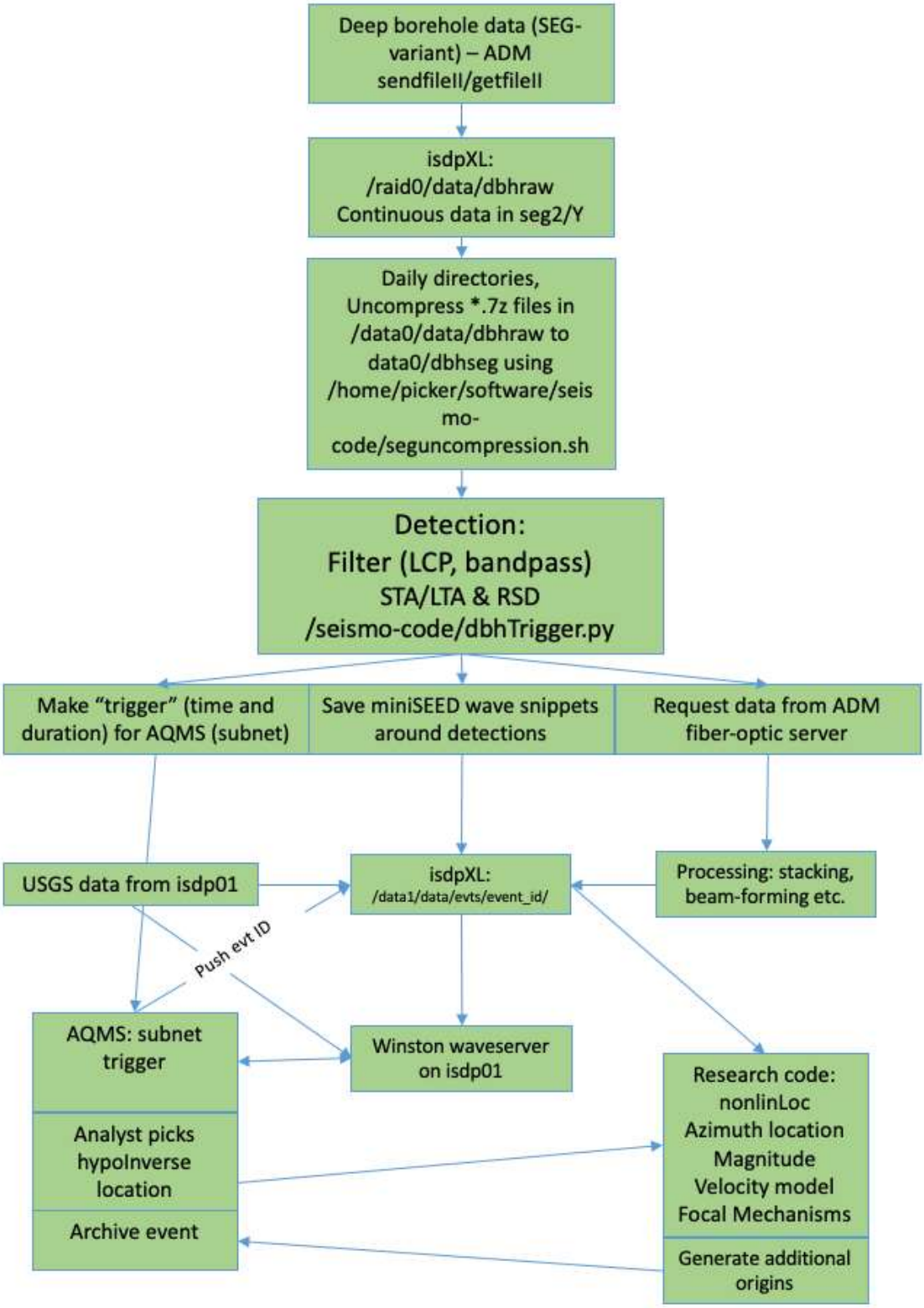
Deep borehole data (SEG-variant) – ADM sendfileII/getfileII
isdpXL: /raid0/data/dbhraw Continuous data in seg2/Y
Daily directories, Uncompress *.7z files in /data0/data/dbhraw to data0/dbhseg using /home/picker/software/seismo-code/seguncompression.sh
Detection: Filter (LCP, bandpass) STA/LTA & RSD /seismo-code/dbhTrigger.py
Make "trigger" (time and duration) for AQMS (subnet)
Save miniSEED wave snippets around detections
Request data from ADM fiber-optic server
USGS data from isdp01
isdpXL: /data1/data/evts/event_id/
Processing: stacking, beam-forming etc.
Push evt ID
AQMS: subnet trigger
Winston waveserver on isdp01
Analyst picks hypoInverse location
Archive event
Research code: nonlinLoc Azimuth location Magnitude Velocity model Focal Mechanisms
Generate additional origins

Fig 1. Processing flow chart for near-real-time detection and location of microseismic events at the Decatur, IL, GCS site.

2. Data transmission details

Continuous data transfer from the SCSDBA to the USGS server (ISDPXL) is done using the programs `sendfileII, fgetfileII`, and `makebhfile`. The programs run continuously, so they are appropriate for real-time and near-real-time applications. These programs are distributed as part of Earthworm, but they do not depend on any other parts of Earthworm. The `sendfileII` program gets files from a queue directory, transmits the files via a socket connection, and deletes the files from the queue directory. The `getfileII` program receives data files from `sendfileII` via a socket connection and saves them in specified directories. The `makehbfile` program periodically places a heartbeat file in the queue directory used by `sendfileII`. If the network connection between computers is down for any reason, files will accumulate in `sendfileII`'s queue directory until the network comes back up. Then all files will be sent immediately. However, `getfileII` will handle only a single file transaction at a time, that is, `getfileII` is a single-threaded, non-forking process.

3. DAS data acquisition and archival

Owing to the large amount of data generated by the DAS system employed at Decatur (DECDAS), we retrieve the DAS data only for events detected on the SCSDBA array. When a trigger is initiated, the trigger routine (`dailyTrigger.py`) will remote copy DECDAS data from the ADM external server (IP address: 96.92.207.122) into the individual triggered event directory.

4. Details of translator code for proprietary SEG variant

In order to facilitate the processing of the Silixa proprietary file format ‘TDMS’, we translated a MatLab script provided by Romain Pevzner at Curtin University, Australia. The translation of the TDMS-MatLab converter was done using python to enable accessing data in our python environment. We tested our translated code against the original MatLab conversion script and conclude that both data format readers yield identical results. The original MatLab script and our derivative code in python remain proprietary tools that Silixa is intent on keeping software proprietary from the public and or any funding agency

Subtask 4.2 Event detection and phase identification

1. STA/LTA detection

The advent of digital seismic data required automated techniques to detect earthquakes and determine automatic estimates of the phase onsets. Phase onsets in microseismic monitoring largely consist of downward compressional waves (P), upward/direct compressional (p) waves, and secondary (downward first) shear waves (S). Standard algorithms for event detection and phase arrival estimates are plentiful, with the most common being a short-term average over long-term average energy density squared ratio (STA/LTA) (Allen, 1978). We employ a recursive STA/LTA detection algorithm as it performs very well when tuned to local settings. For data from the deep borehole sensors (SCSDBA), we select 4 individual channels: PS1_1H, PS1_2Z, GM1_02Z, GM1_04Z, where PS1 indicates two sensors within injection well CCS1 and GM1 indicates sensors in geophysical monitoring well GM1 (Couëslan et al., 2014); the last letters refer to either vertical (Z) or horizontal (H) orientation of the individual sensor. We choose these channels as they show the least noise of the deep sensors. The data is demeaned, notch-filtered at 60Hz, and bandpass filtered (20 Hz < f < 100 Hz) based on observed noise patterns and microseismic energy distribution of the data. We then require a minimum of 3 channel triggers to coincide to generate a detection when the STA/LTA window lengths are 0.5/5 seconds (Withers et al., 1998).

2. Template matching

A well recorded seismic signal is largely a product of the 1) earthquake source, 2) medium the energy travels through, and 3) how the receiver records the energy. If an earthquake source "repeats" (i.e. multiple earthquakes in close proximity to one another along a similar fault patch), the event waveforms recorded at a sensor will be similar. Template matching is an important technique for the characterization of these repetitive seismic sources as it mitigates the path uncertainty (e.g., Schaff and Waldhauser, 2010). This procedure takes the waveform from a known earthquake and cross-correlates it against continuously recorded seismic data to identify other similar, smaller amplitude signals that might have been missed by other earthquake detection methods. As induced seismicity commonly occurs as swarms containing similar events, template matching aids our understanding of induced earthquakes by detecting smaller magnitude earthquakes that are generally missed in routine processing. While traditional earthquake detection techniques rely on relatively large amplitudes to be recorded, template matching can identify smaller amplitude events that have a much lower signal-to-noise ratio.

Waveform template matching using regional seismic networks commonly increases the number of detected earthquakes by roughly an order of magnitude. With the increase in the number of continuously recorded seismometers and computational processing capabilities, the approach of applying template matching to regional stations has grown in popularity over the past decade. However, template matching has rarely been applied to borehole networks and DAS. While template matching using DAS may be feasible, the signal-to-noise ratio of

microseismic events recorded by the DAS installation was too low to resolve the waveform for template events (described in Subtasks 4.2 & 4.3). Instead, we evaluate the potential to apply template matching to local borehole instruments.

We applied template matching to 85 earthquakes identified using the deep borehole sensors (PS1 & PS2) to evaluate the benefit of using borehole instruments to monitor for microseismicity (Fig. 2). Templates were two seconds in length, bandpass filtered between 50-90 Hz, and cross-correlated between 1 September 2016 - 30 November 2017. The template matching catalog contained 349 events, roughly four times as many earthquakes than were in the original catalog.

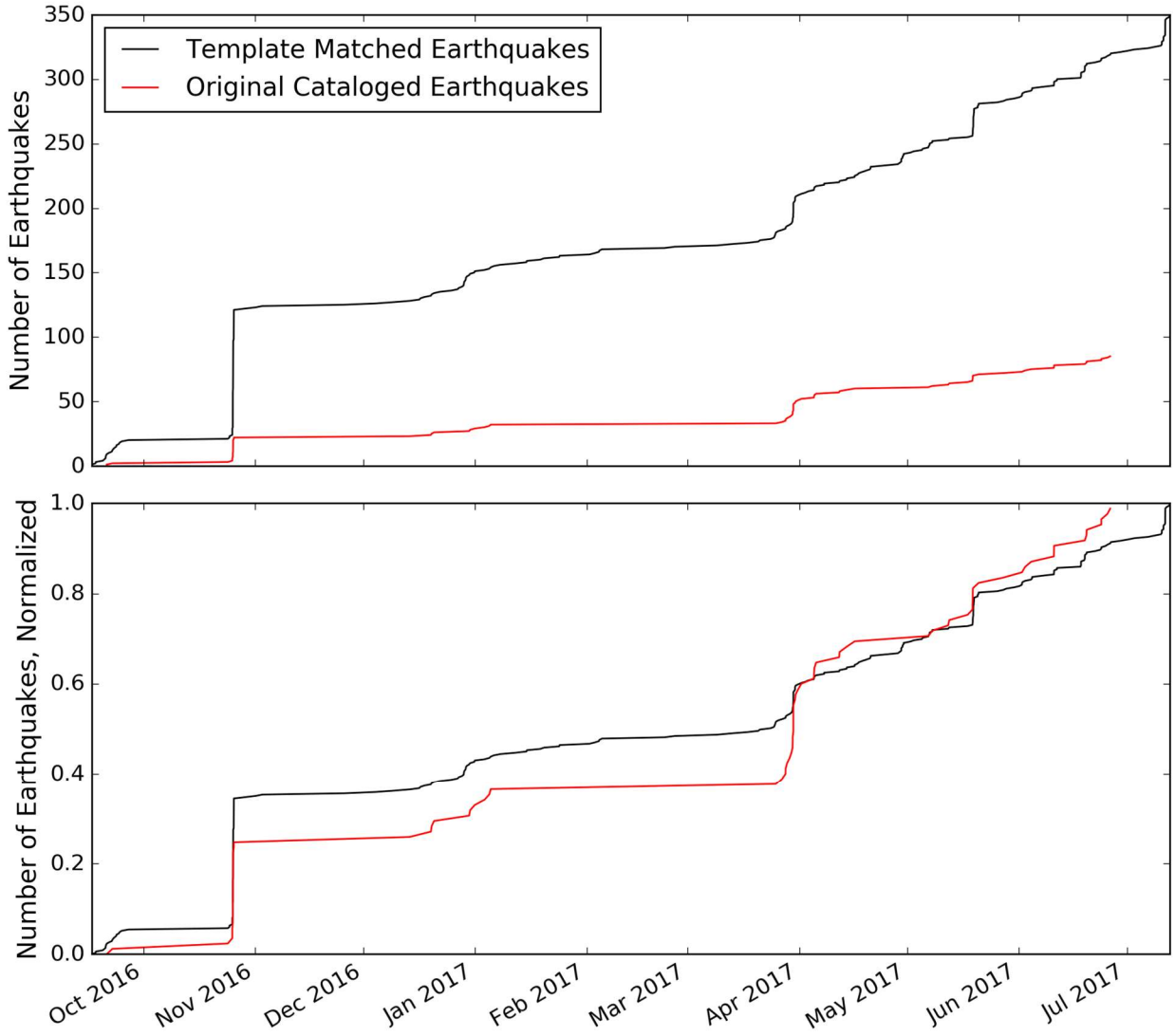

Fig 2. Demonstration of template matching applied to the deep borehole sensors (PS1 & PS2) showing the number of events in the original catalog (red) and detected events (black).

3. Filtering/stacking/cross-correlation of DAS data

Many microseismic events that were well-recorded on both the deep borehole and shallow borehole networks. However, the signal-to-noise ratio of microseismic events recorded by the DAS installation was too low to resolve the waveform for the microseismicity (Figure 3). We verified that this was not due to a time zone issue (Figure 3) and demonstrate some of the different filtering and stacking steps we tried to improve the signal-to-noise ratio in Subtask 4.3.

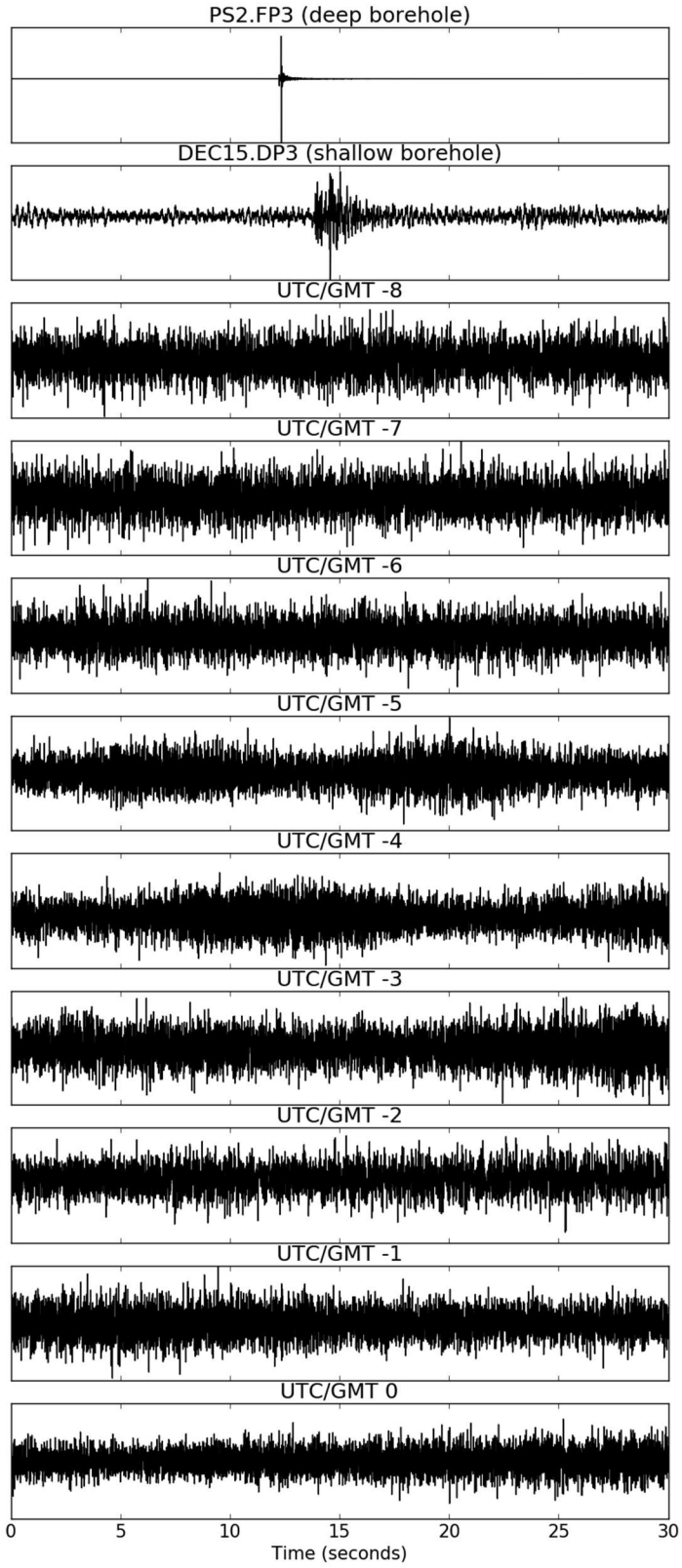
PS2.FP3 (deep borehole)
DEC15.DP3 (shallow borehole)
UTC/GMT -8
UTC/GMT -7
UTC/GMT -6
UTC/GMT -5
UTC/GMT -4
UTC/GMT -3
UTC/GMT -2
UTC/GMT -1
UTC/GMT 0
0
5
10
15
20
25
30
Time (seconds)

Figure 3. Example records for the 21 October 2018 [~10:59:30] local event for a single channel from the deep borehole, shallow borehole, and a representative DAS channel. DAS records are shown at different time zones to confirm the lack of an observed signal.

4. Description of manual picking interface

Analyst picked phase arrivals are crucial in determining high accuracy hypocentral parameters, including the origin time and location. To facilitate the use of all three data streams we built a custom user interface that ingests data from SCSDBA, GSDEC, and DECDAS sources. The interface is based on the open source `streampick` package (Isken, 2013). The framework depends on the event detection and archiving in event directories described earlier. Once an event is detected, instantiated on ISDPXL, and data are gathered, an analyst can access data from all individual channels per station (Fig. 4). In the wavetrace panel, phase arrivals can be added for both P- and S-wave arrivals using the cursor and determining the phase using specified keys. Those data are automatically saved to a QUAKEML file in the event directory. The function bar includes options to plot all arrivals for the particular event on a Wadati plot (i.e. P-arrival plotted against the difference of S- and P-wave arrivals) to ensure erroneous phase arrivals can be easily identified; the panel also includes the option to locate the event directly using either standard Geiger method location techniques ("Locate hypo") or the probabilistic location method ("Locate Nlloc"), see section 4.3.3 for more details on the location method.

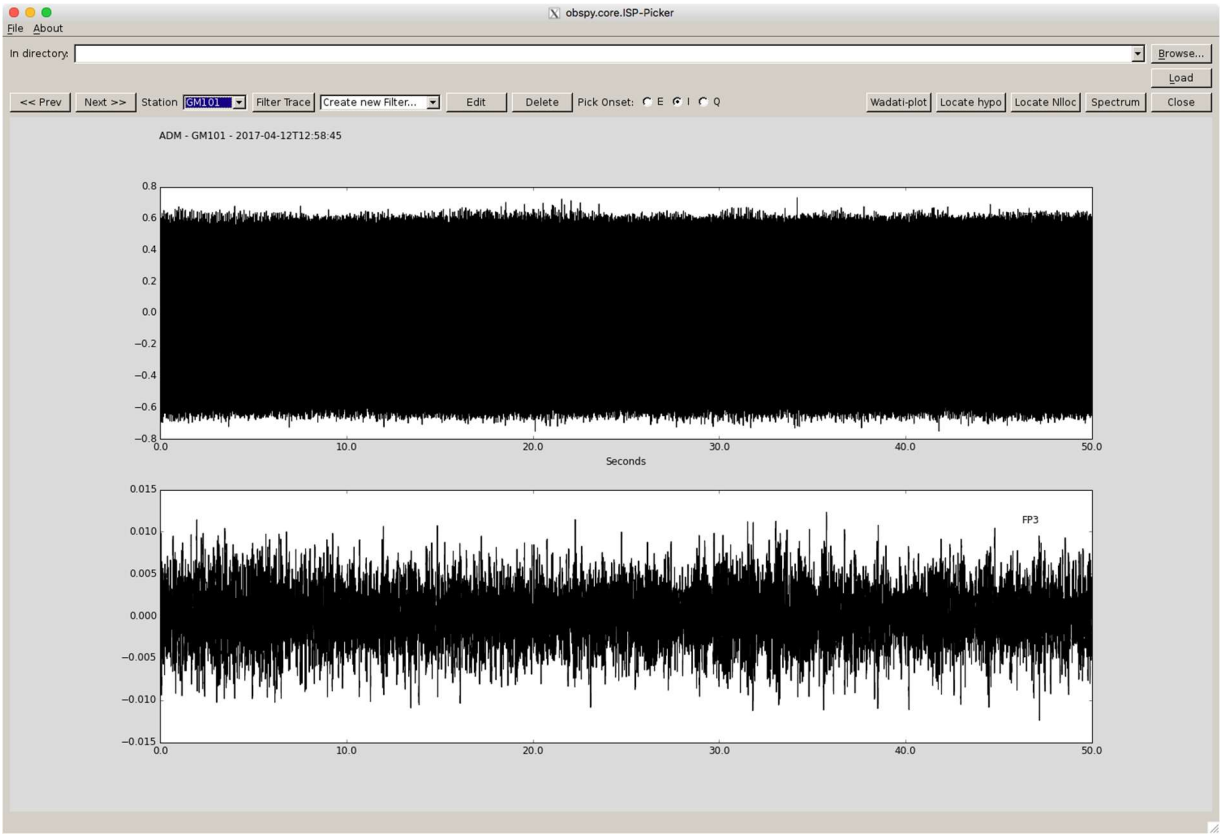

Figure 4. Custom built user interface depicting two channels from one station in the wavetrace panel (GM1_1). Specific channel names are shown in the upper right-hand corner of each trace panel. Mouse and keyboard are used to pick phase arrivals

Once event hypocenters are located, the manual picking interface displays a window with the event hypocenters and uncertainty regions displayed in a separate window (Fig. 5). The map and two cross-section views aid in determining the quality of the location, specifically, if the uncertainty region is skewed and whether reductions in uncertainty can be made by improving phase picks.

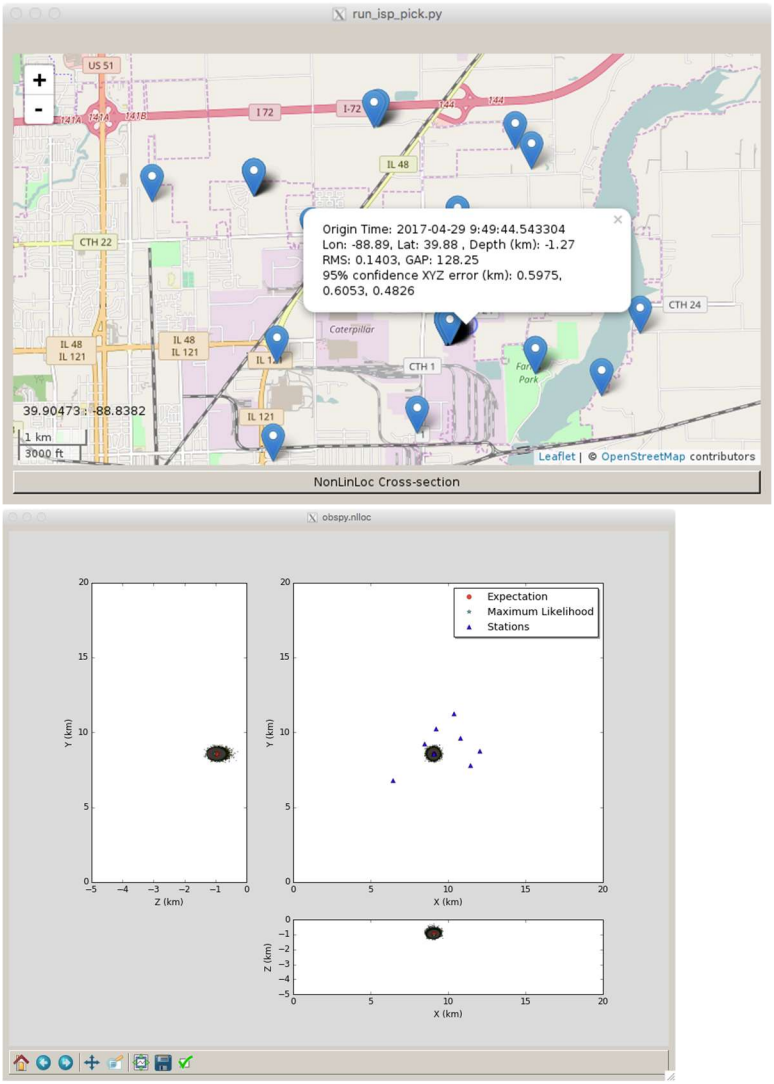

Figure 5. Custom-built hypocenter location map and cross-section interface for analyst review. Each panel can be zoomed to verify locations and hypocentral uncertainties.

The flexibility of the custom-built interface can easily be extended to ingest all DAS derived data or stacks of DAS derived data. Additionally, extension to various hypocentral location and magnitude estimation methods can and has been added.

Subtask 4.3 Improving location methods and weighting for improved event location uncertainties

1. DAS data on local event

If the vertical DAS had been coupled to the borehole, it would have likely had larger signal-to-noise ratios than the horizontal DAS strings due to a smaller source-sensor distance. However, since the vertical DAS was not coupled to the borehole, it was not possible to utilize it to detect microseismic events. As a result, we were limited to recordings on the horizontal DAS.

We attempted various types of low (15-100 Hz in intervals of 5 Hz), high (1, 5-80 Hz in intervals of 5 Hz), and bandpass (1-100, 5-15, 5-20, 5-40, 10-50, 20-40, & 50-100 Hz) filters along with channel stacking approaches (bin and rolling stacks using 5, 10, 50, & 100 channels) to improve the signal-to-noise ratio of the DAS records. Filtering the data did not clearly resolve any signals, and stacking channels caused other noise signals to be amplified. We show example waveforms of this for two local events (Figs. 6 & 8) along with the spectra for the records (Figs. 7 & 9). While the borehole networks recorded these microseismic events well, the DAS appears to be dominated by noise. Note, the linear feature decreasing in frequency with time (Figure 7) that coincides with the arrival time of the local event is part of the active orbital source described in other portions of the final report.

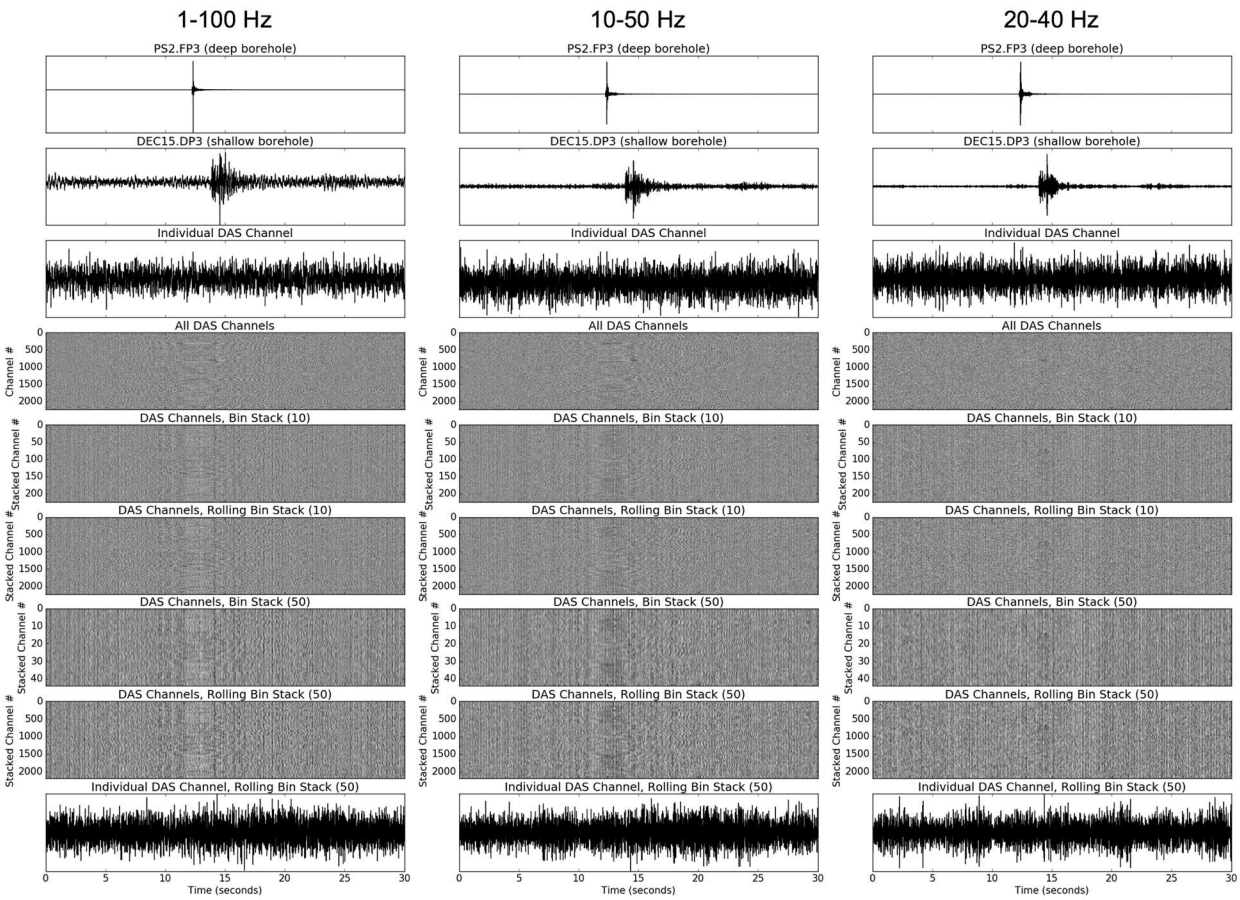

Figure 6. Example records for the 21 October 2018 [~10:59:30] local event on the deep borehole, shallow borehole, and DAS. Waveforms bandpass filtered between 1-100 Hz (left), 10-50 Hz (middle), and 20-40 Hz (right).

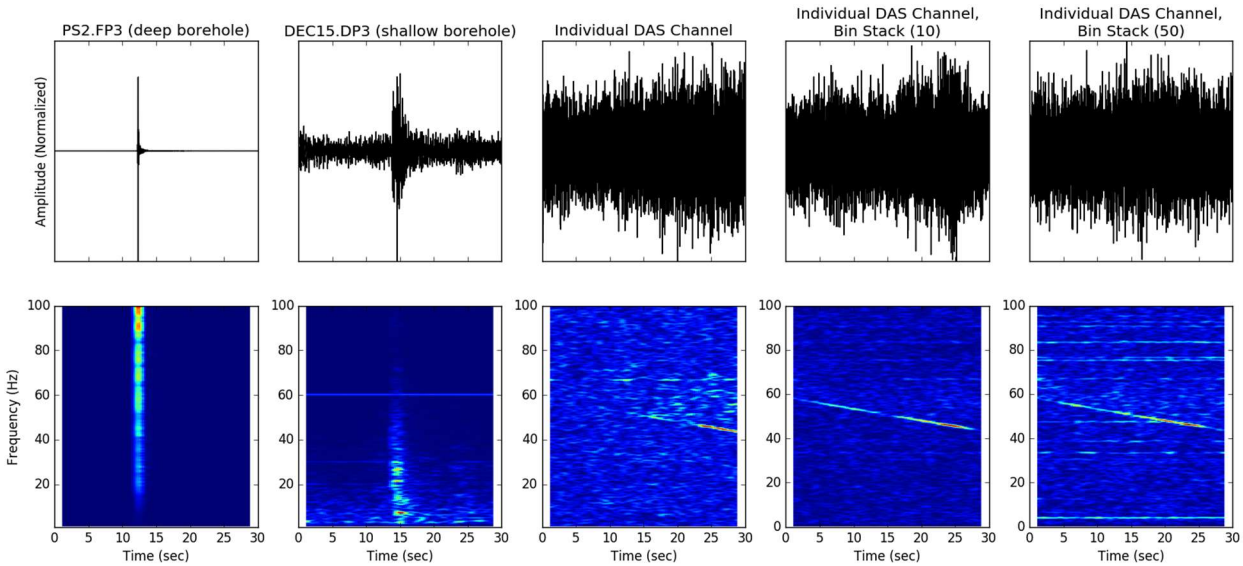

Figure 7. Example records for the 21 October 2018 [~10:59:30] local event on the deep borehole, shallow borehole, and DAS in the time (top) and frequency (bottom) domains.

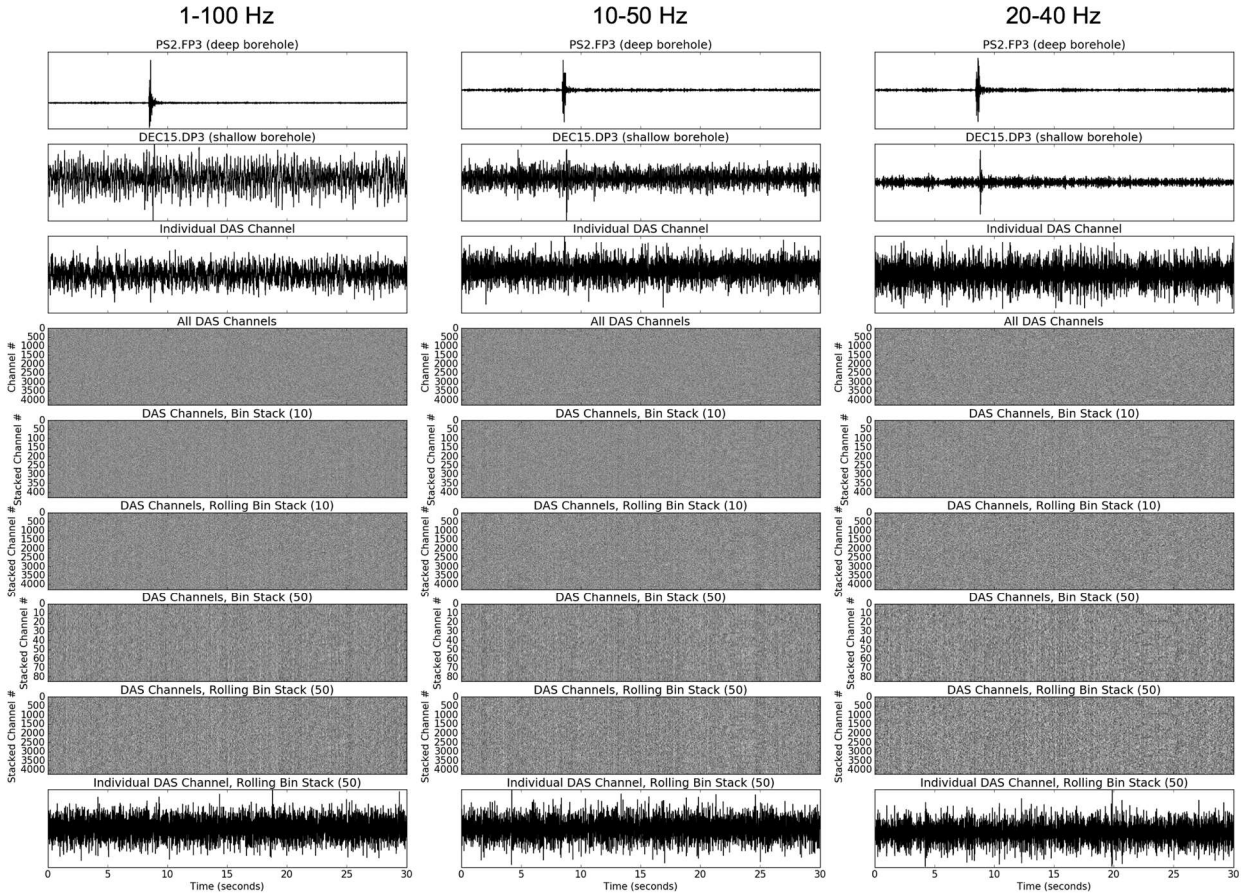

Figure 8. Example records for the 28 July 2017 [~00:36:57] local event on the deep borehole, shallow borehole, and DAS. Waveforms bandpass filtered between 1-100 Hz (left), 10-50 Hz (middle), and 20-40 Hz (right).

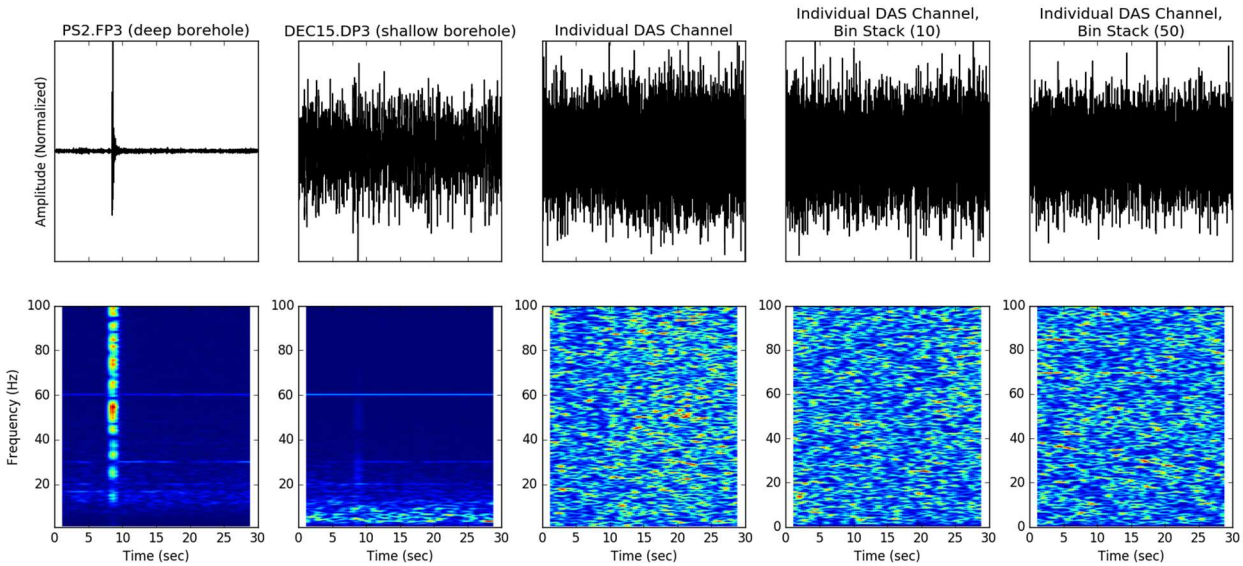

Figure 9. Example records for the 28 July 2017 [~00:36:57] local event on the deep borehole, shallow borehole, and DAS in the time (top) and frequency (bottom) domains.

2. DAS data on regional event

Similar to the results from the microseismic records, regional events were also not well-recorded on the DAS. The various types of filtering applied to local microseismic data (15-100 Hz in intervals of 5 Hz), high (1, 5-80 Hz in intervals of 5 Hz), and bandpass (1-100, 5-15, 5-20, 5-40, 10-50, 20-40, & 50-100 Hz) filters along with channel stacking approaches (bin and rolling stacks using 5, 10, 50, & 100 channels) were also used for the regional events. We show example waveforms of this for two regional events (Figs. 10 &12) along with the spectra for the records (Figs. 11 & 13). While the regional earthquakes are well recorded on the borehole networks, the DAS appears to be dominated by noise.

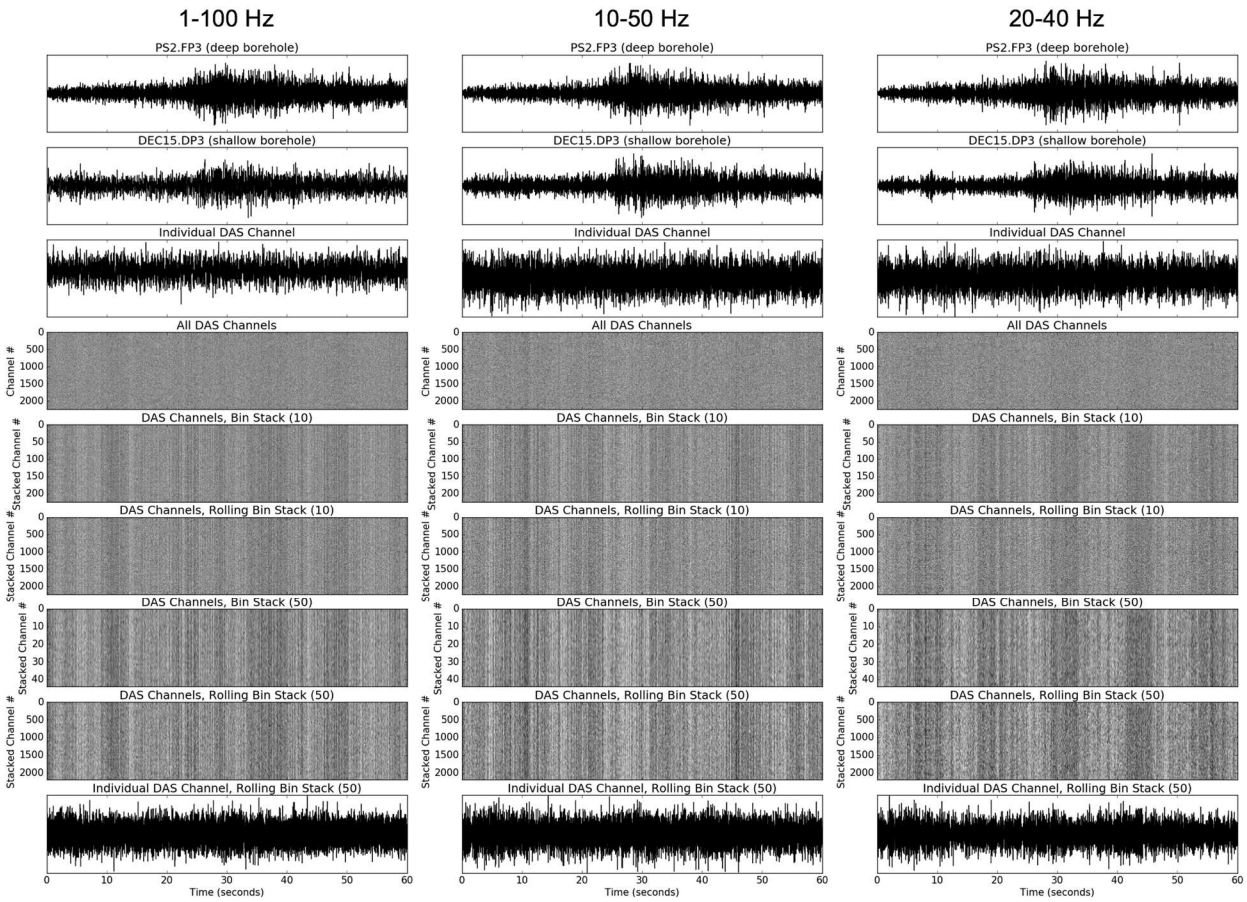

Figure 10. Example records for the 20 September 2018 Md 2.3, ~170 km south of Decatur on the deep borehole, shallow borehole, and DAS. Waveforms bandpass filtered between 1-100 Hz (left), 10-50 Hz (middle), and 20-40 Hz (right).

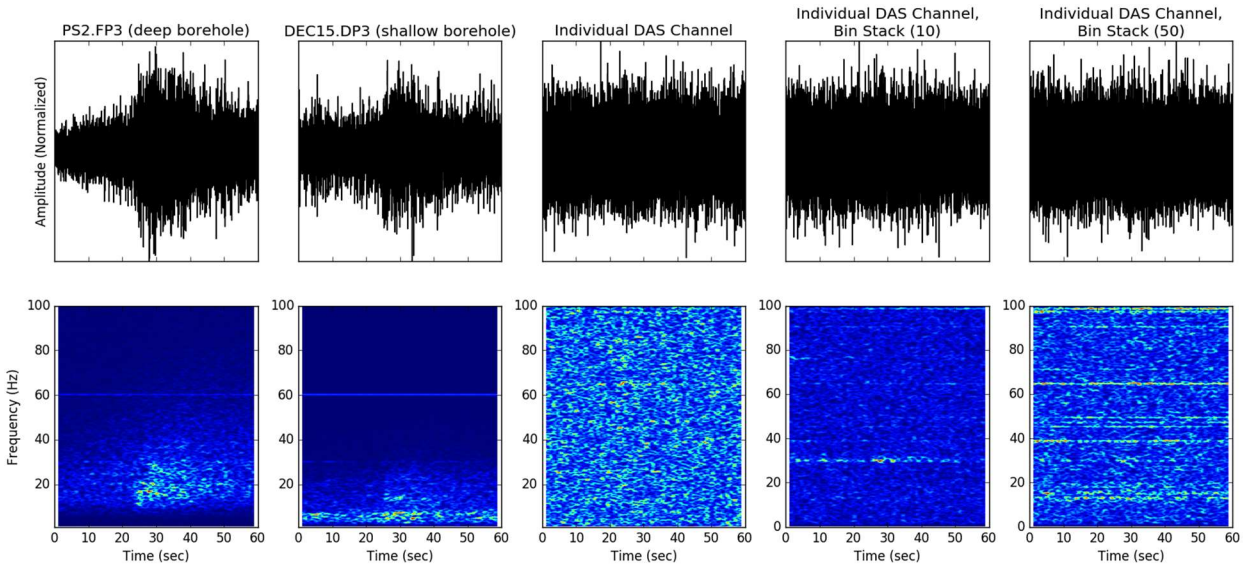

Figure 11. Example records for the 20 September 2018 Md 2.3, ~170 km south of Decatur on the deep borehole, shallow borehole, and DAS in the time (top) and frequency (bottom) domains.

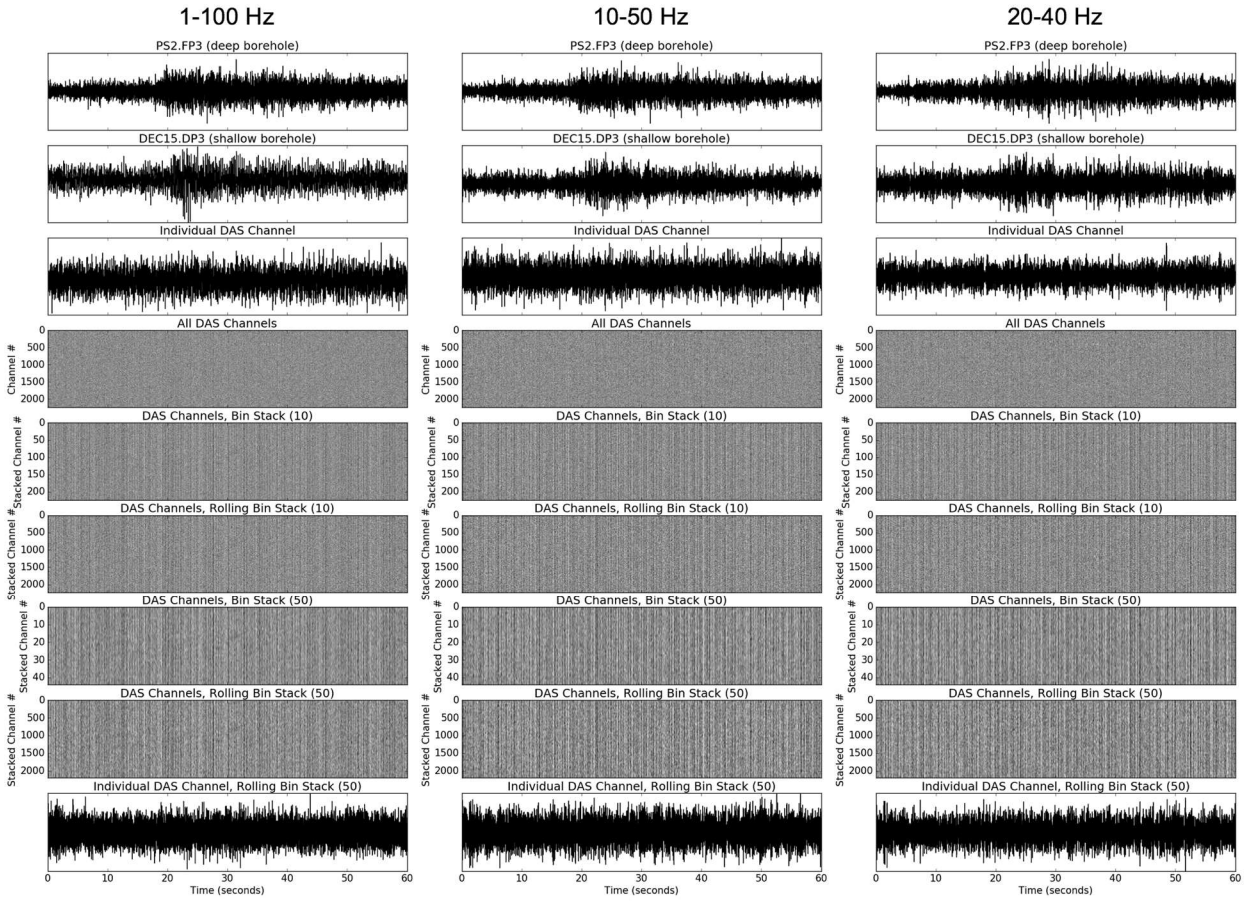

Figure 12. Example records for the 15 October 2018 Md 2.1, ~185 km south of Decatur on the deep borehole, shallow borehole, and DAS. Waveforms bandpass filtered between 1-100 Hz (left), 10-50 Hz (middle), and 20-40 Hz (right).

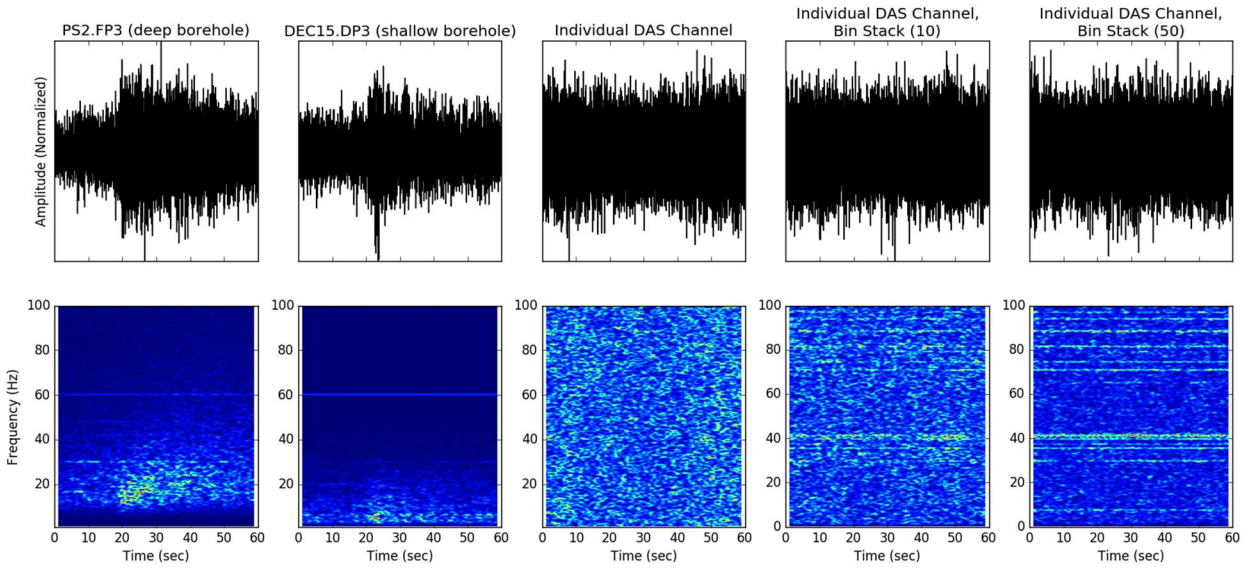

Figure 13. Example records for the 15 October 2018 Md 2.1, ~185 km south of Decatur on the deep borehole, shallow borehole, and DAS in the time (top) and frequency (bottom) domains.

3. Location method (NonLinLoc)

All earthquakes were independently relocated using a probabilistic formulation of the nonlinear inverse problem governing hypocentral location methods (NonLinLoc; Lomax et al., 2000). We assumed that uncertainties in a priori information (e.g., phase arrival pick times and theoretical arrival times) are Gaussian and independent of one another, which permits calculation of the posterior probability density function (PDF). We used the OctTree sampling algorithm in NonLinLoc, which efficiently calculates the PDF for hypocentral parameters (Husen and Smith, 2004). The velocity model consisted of a 38-layer P- and S-wave 1D model interpolated from the borehole velocity survey conducted by Schlumberger Carbon Services. The travel time model volumes were determined for 10x10x10 m grids.

This location method solves for travel times in the velocity model. This method completely maps the PDF through the model domain. For each event, we obtain a cloud of possible locations within the prescribed probability density function bounds and we select the maximum likelihood location from the point cloud to arrive at the preferred event location. However, given the vertical arrangement of borehole seismic sensors this generally results in a donut-shaped distribution of possible locations that fit the uncertainty conditions due to the azimuthal ambiguity (Fig. 14).

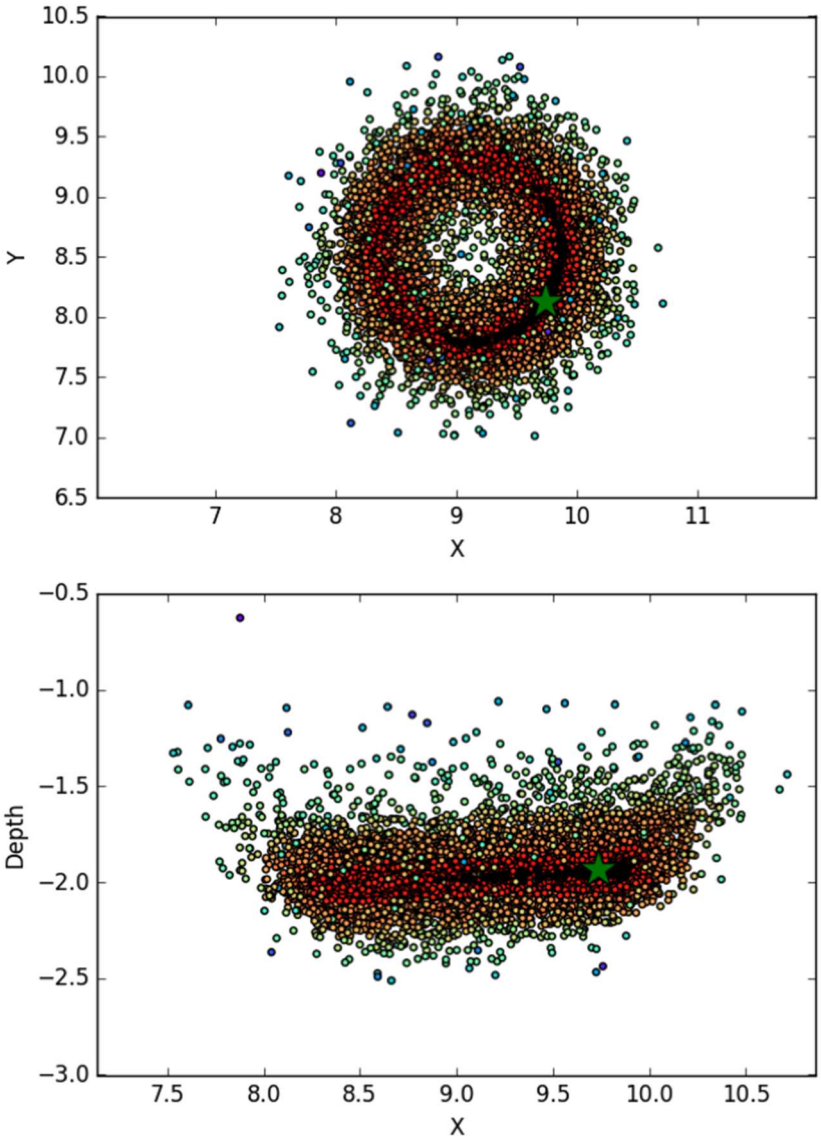

Figure 14. Map (top) and cross-sectional (bottom) view of PDF solutions for one microseismic event. Black dots indicate preferred locations based on probability density bounds. Green star indicates maximum likelihood location. All coordinates are in a local reference frame and in km.

To further reduce the number of probable event locations we use additional waveform data to discern the incident polarization direction of recordings (Fig. 15). The polarization direction of recorded S-waves is perpendicular to the back-azimuth from the station to the event location. Figure 15 shows the polarization direction of S-waves determined using a 0.1 second window following the S-pick after band-pass filtering from 20 to 40Hz (Vidale, 1986). Note that planarity and ellipticity of all borehole sensors are increasing in quality with depth, but can confidently be deduced for all sensors, thus providing robust estimates of S-wave azimuths for all sensors.

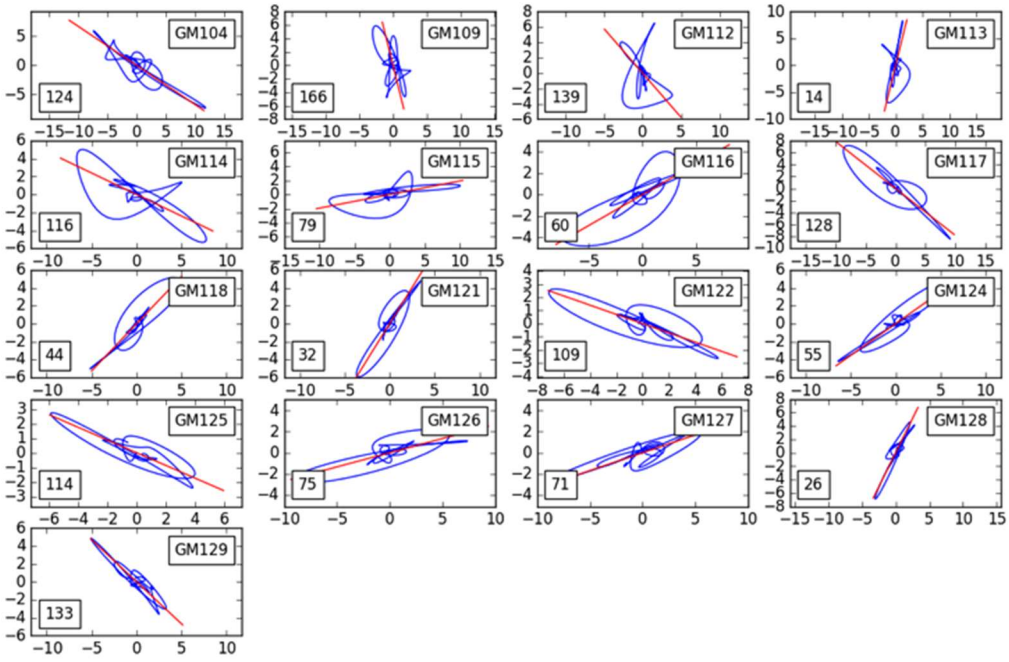

Figure 15. S-wave polarity (blue curve) and best fitting azimuth (red line) at each of the GM1 borehole sensors. Resulting S-wave azimuth plotted in boxes in lower left corner. GM1 borehole array sensors IDs decrease in number with depth (e.g. GM129 is the deepest channel depicted).

Using the estimates of S-wave polarization, we can then weigh probable event locations according to how well the azimuths fit the locations (Fig. 16). This method does not resolve the 180-degree ambiguity of azimuths (Fig. 16), but given the PDF of the probable event locations, a weighted maximum likelihood event location can be robustly deduced (see red dots in Fig. 16). This weighting of the PDF further improves our estimates of absolute location of each event located and includes formal estimates of the location uncertainty.

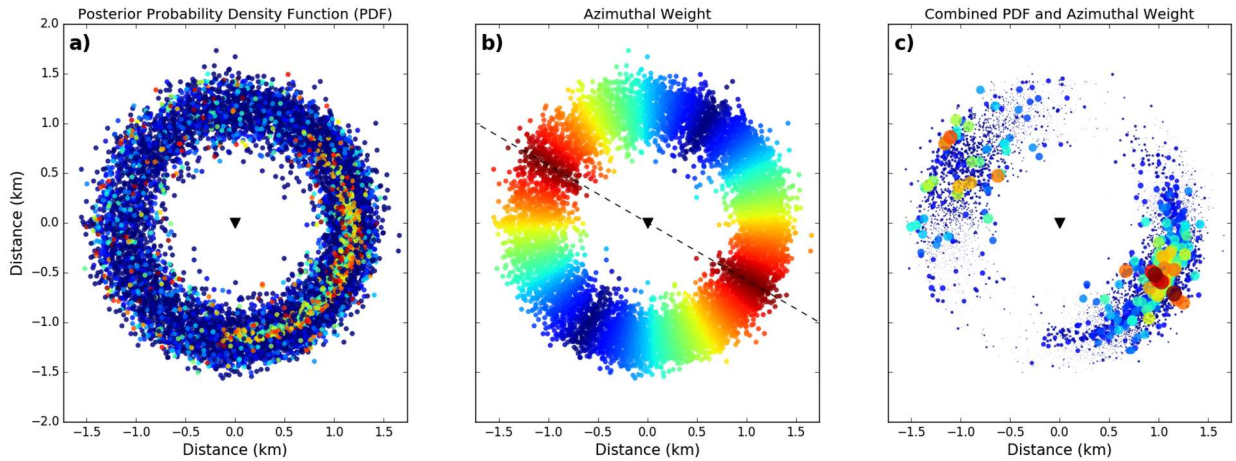

Figure 16. a) Posterior probability density function (PDF) of the non-linear grid search location method. b) Azimuthal weighting based on S-wave polarity with warm colors indicate higher likelihood of location. c) Combined PDF and azimuthal weighting for best fitting locations (red dots). Color and size indicate highest PDF and azimuthal weight.

4. Test shot performance on borehole array

We tested our microseismic event location algorithm using perforation shots carried out in well VW#1. We locate events by first getting an initial location and then using the residuals for each channel as individual weights in a second location. Using this approach without any further calibration we manage to locate three of the four perforations shots very well in depth (Table 1).

Table 1. Actual and location depths and differences in km for three individual perforations shots fired in VW#1.

	Actual depth (km)	**NonLinLoc depth (km)**	**Difference (km)**
Perf #2	1.953	1.9584	0.005
Perf #3	1.517	1.5327	0.015
Perf #4	1.514	1.5220	0.008

The horizontal coordinates were all within roughly 100m of the actual horizontal position with the larger uncertainty owing to the polar ambiguity inherent in the vertical array. The polarity calculation cannot be carried out at this point, as several channels appear to be misreported and we did not receive confirmation on the actual orientations from Schlumberger Carbon Services. If those sensor orientations issues had been addressed, we could have reduced the horizontal uncertainty to within a few tens of meters, albeit still subject to the 180-degree ambiguity (see Subtask 4.3.3).

Summary

The goal of this subtask was to build a workflow for DAS generated seismic data to be integrated at the Decatur, IL, geologic sequestration site's microseismic data processing. The main objectives were to: 1) ensure near-real-time data transmittal and processing, 2) improve event detections, 3) employ location techniques that rely on all data streams, including DAS-derived data. Owing to the multitude of available data, we focused on optimizing data flows, testing detection methods for each stream of data, i.e., deep borehole, DAS, shallow borehole, and extending location methods for the specific needs at Decatur. Unfortunately, the DAS-derived

data at Decatur did not prove useful in detecting or locating seismicity. Various stacking and filter optimization schemes were employed to overcome the insensitivity of the installed fiber-optic cables to seismic waves but no avail. However, we have developed a framework for the inclusion of fiber-optic cable derived DAS data in the standard workflow in microseismic monitoring. Our data flow, analyst interface, and extended location methods have been employed in other settings, e.g., West Texas (Skoumal et al, 2020), with little change from the outlined workflow.

References

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, Bull. Seismol. Soc. Am. 68, no. 5, 1521–1532

Couëslan, M.L., Butsch, R., Will, R. and Locke,. R. A (2014). Integrated reservoir monitoring at the Illinois Basin – Decatur Project, Energy Procedia, 63, 2836-2847.

Husen, S., and R. B. Smith (2004). Probabilistic earthquake relocation in three-dimensional velocity models for the Yellowstone National Park region, Wyoming, Bull. Seismol. Soc. Am. 94, no. 3, 880–896.

Isken, M., StreamPick, https://github.com/miili/StreamPick, last accessed September 2013.

Lomax, A., J. Virieux, P. Volant, and C. Berge-Thierry (2000). Probabilistic earthquake location in 3D and layered models, in Advances in Seismic Event Location, Springer, Dordrecht, The Netherlands, 101–134.

Schaff, D. P., & Waldhauser, F. (2010). One magnitude unit reduction in detection threshold by cross correlation applied to Parkfield (California) and China seismicity. B. Seismol. Soc. Am., 100(6), 3224-3238.

Skoumal, R. J., Kaven, J.O., Barbour, A. J., Wicks, C., Brudzinski, M., Cochran, E.S., and Rubinstein, J. (2021). The induced Mw 5.0 March 2020 West Texas seismic sequence. J. Geophys. Res., 126. https://doi.org/10.1029/2020JB020693.

Vidale, J.E. (1986) Complex polarization analysis of particle motion, Bulletin of the Seismological Society of America, 76 (5), 1395-1405.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., and Trujilo, J. (1998). A comparison of selected trigger algorithms for automated global seismic phase and event detection, B. Seismol. Soc. Am., 88, 95-106.