AGU PUBLICATIONS

Journal of Geophysical Research – Solid Earth, special issue on Machine Learning for Solid Earth

Supporting Information for

Monitoring Fracture Saturation with Internal Seismic Sources and Twin Neural Networks

D.D. Nolte[1] and L.J. Pyrak-Nolte[1,2,3]

[1]Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana.
[2]Department of Earth, Atmospheric and Planetary Sciences, Purdue University, West Lafayette, Indiana.
[3]Lyles School of Civil Engineering, Purdue University, West Lafayette, Indiana.
Corresponding author: Laura and Pyrak-Nolte (ljpn@purdue.edu)

Contents of this file

Additional Supporting Information

The data sets are available on the Purdue University Research Repository (PURR) that provides a data-sharing platform (https://purr.purdue.edu). On PURR, the data are listed and can be obtained from the following citation:

Nolte, D. D. and Pyrak-Nolte, L. J. (2022), "Data for Monitoring Fracture Saturation with Internal Transportable Seismic Sources and Twin Neural Networks." (DOI: 10.4231/DXGB-KW29).

Captions for Datasets S1 to S4

Dataset S1: This data set contains the cleaned and normalized signals from the moving source experiments for the 2 class analysis. The S1.zip file contacts an information file (READ_ME_Moving_2_20210411_Infomation) about the number of signals, time per points and column information in the ".txt" files in the 6 directories. The directory names are based on the group numbers listed in Supplement Table S5.

Dataset S2: This data set contains the cleaned and normalized signals from the moving source experiments for the 4 class analysis The S2.zip file contacts an information file

(README_5_W0W1R2P3W4P5_Information.rtf) about the number of signals, time per points and column information in the ".txt" files in the 4 directories. The directory names are based on the group numbers listed in Supplement Table S5.

Dataset S3: This data set contains the original signals from the stationary source experiments for the 4 class non-prospective analysis The S3.zip file contacts 4 directories that each contain a README file with information file about the number of signals, time per points. The directory names are based on the group numbers listed in Supplement Table S5. The first column of the data file contains the test number, the second column represents the event number for a particular test and the 3rd column is set to 0. The rest of the columns contain the signals.

Dataset S4: This data set contains the original signals from the stationary source experiments for the 4 class -prospective analysis. The S4.zip file contacts 8 directories that each contain a README file with information file about the number of signals, time per points. The directory names are based on the group numbers listed in Supplement Table S5. The first column of the data file contains the test number, the second column represents the event number for a particular test and the 3rd column is set to 0. The rest of the columns contain the signals.

Introduction

Here, additional information is provided on the order in which the experiments were performed, the number of tests per fracture condition, the methods use to order and clean the data, the number of signals pre- and post- cleaning, the source event location method and an analysis of the potential error in locations. We also include the original data in binary format for the experiments referred to S1.1-S1.3 that includes all of the channels that were recorded (see manuscript section 2.2.2 and Table 4).

S.1 Experimental Notes

This section provides additional information on the experimental measurements.

S1.1 Order of Experiments for the Moving Source for 4 Class Analysis

In the moving-source experiments with the 4 fracture conditions described in the manuscript (Section 2.1), data were collected in 9 separate groups in the order shown in Table S1. First, 5 tests were performed for the saturated case (Group 1), followed by 5 tests for S3 (Group 2), 5 tests for S5 (Group 3) and 5 tests for S3S5 (Group 4). Then, the order of the tests was repeated (Groups 5-9). Between each group, the water in the release fracture was refreshed and replaced by water with 1% Clorox™ to prevent bacteria colonies.

Table S1. Order of the experiments for the moving-source experiments for the 4 different fracture set conditions and the number of tests performed in each group.

Group/Order	Condition	Number of Tests
1	Saturated	5
2	S3	5
3	S5	5
4	S3S5	5
5	Saturated	5
6	S3	5
7	S5	5
9	S3S5	9

S1.2 Order of Experiments for the Moving Source for 2 Class Analysis

The analysis of saturation with depth (Section 3.3 in the manuscript) was based on data from moving source experiments with 2 fracture conditions. Data were collected in 6 separate groups in the order shown in Table S2. Initially, all fractures were saturated. Prior to collecting data, roughly 45-48 cc of water were drained from F0 and F5 using a syringe (referred to as condition S5). After 50 minutes, F0 and F5 were re-saturated. First, 5 tests were performed for the saturated case (Group 1) followed by 5 tests for S5 (Group 2). Then, the order of the tests was repeated (Groups 3-6). Between each group, the water in the release fracture was refreshed and replaced by water with 1% Clorox™ to prevent bacteria colonies.

Table S2. Order of the experiments for the moving source experiments for the 2 different fracture set conditions and the number of tests performed in each group.

Group/Order	Condition	Number of Tests
1	Saturated	5
2	S5	5
3	Saturated	5
4	S5	5
5	Saturated	5
6	S5	5

S1.3 Order of the Experiments for the Stationary Source

As mentioned in the manuscript section 2.2.1, a wire ledge was placed at a depth of 30 mm in the release fracture, F2, to control the location of the dust. The order of the experiments for stationary sources for non-prospective and prospective analyses is shown in Tables S3 and Table S4, respectively. Between each group of experiments, the water in the release fracture was refreshed and replaced by water with 1% Clorox™ to prevent bacteria colonies. The number of tests performed for a given condition depended on the number of signals recorded by channel 5.

Table S3. Order of the experiments for the stationary (non-prospective) experiments for the 4 different fracture set conditions and the number of tests performed in each group.

Group/Order	Condition	Number of Tests
1	S3S5	10
2	S3	10
3	S5	13
4	Saturated	4

Table S4. Order of the experiments for the stationary (prospective) experiments for the 4 different fracture set conditions and the number of tests performed in each group.

Group/Order	Condition	Number of Tests
1	S3S5	25
2	Saturated	5
3	S3	10
4	S5	10
5	S3S5	28
6	Saturated	6
7	S3	10
8	S5	10

S2. Data Handling

This section describes the steps for processing the data prior to application of the machine learning method for discrimination of the condition of the fracture system.

S2.1 Sorting Recorded Signals into Events

The Mistra software generates a file that is referred to as a line file (Figure S1). The ID of a signal equals 1 (1st column in Figure 1) if a signal is part of an event. The file lists all of the channel numbers that recorded a signal for an event and the time of the event. The time of recording for each channel in an event is given by DD:HH:MM:SS:mmmuuun where D

```
 2   0 00:00:35.1099000
 2   0 00:00:35.1199000
 2   0 00:00:35.1299000
 2   0 00:00:35.1399000
 2   0 00:00:35.1499000
 2   0 00:00:35.1599000
ID DDD HH:MM:SS.mmmuuun  CH  RISE  COUN  ENER  DURATION AMP A-FRQ    RMS ASL SIG STRNGTH  ABS-ENERGY P-FRQ
 1   0 00:00:35.1658660   2     0     2     0        70  40    29 0.0028  12    2.098E+03  10.625E+00    48
 1   0 00:00:35.1658730   5     0     1     0         0  40     0 0.0028  11   61.000E+00 654.736E-03    48
ID DDD HH:MM:SS.mmmuuun
 2   0 00:00:35.1699000
ID DDD HH:MM:SS.mmmuuun  CH  RISE  COUN  ENER  DURATION AMP A-FRQ    RMS ASL SIG STRNGTH  ABS-ENERGY P-FRQ
 1   0 00:00:35.1712127   2     1    16     5       326  56    49 0.0030  12   32.815E+03 690.564E+00    48
 1   0 00:00:35.1712153   5     1    10     3       238  56    42 0.0028  12   24.568E+03 552.141E+00   107
 1   0 00:00:35.1712155   1     4    13     4       257  52    51 0.0024  11   25.306E+03 404.125E+00    39
 1   0 00:00:35.1712187   3    76    11     3       261  50    42 0.0026  11   20.243E+03 271.067E+00    48
 1   0 00:00:35.1712193   6     3     7     2       269  48    26 0.0024  10   13.661E+03 117.810E+00    19
 1   0 00:00:35.1712207   7     3     8     2       264  47    30 0.0022  10   13.789E+03 123.945E+00    19
 1   0 00:00:35.1712227   4    81    11     3       246  51    45 0.0024  11   20.206E+03 280.272E+00    48
 1   0 00:00:35.1712310   8   107     7     2       313  45    22 0.0022  10   15.061E+03 116.618E+00    29
ID DDD HH:MM:SS.mmmuuun
 2   0 00:00:35.1799000
ID DDD HH:MM:SS.mmmuuun  CH  RISE  COUN  ENER  DURATION AMP A-FRQ    RMS ASL SIG STRNGTH  ABS-ENERGY P-FRQ
* Gp# 1[2,5,1,3,6,7,4,8] x,y,z = 0.02829, 0.02548, -0.01951, q = 1 dT[3, 3, 6, 7, 8, 10, 18] Src Amplitude = 56.0
*    0 00:00:35.1712127   2     1    16     5       326  56    49 0.0030  12   32.815E+03 690.564E+00    48
ID DDD HH:MM:SS.mmmuuun
 2   0 00:00:35.1899000
 2   0 00:00:35.1999000
```

Figure S1. Example of information in a line file generated by the Mistra software related to the recorded acoustic emissions.

= day, H = hour, M = minute, m = millisecond, u = microsecond and n = nanosecond. The difference in arrival times, dT, among the channels for an event was found by subtracting the minimum time recorded for that event from the times recorded by all of the other channels. Events were not used in the study if any of the dTs among the channels for the same event exceeded 15 microseconds. The expected maximum dT between signals recorded by transducers CH1 and CH5 for a dust grain resting on the bottom of the sample is ~4.4 microseconds when the dust grain is in fracture F2 (see Figure 1 in manuscript for fracture locations). Also, for a moving source, only events were used that included signals recorded by CH5 and that were recorded by at least a total of 5 different channels (sensors) to minimize errors in location.

S2.2 Data Cleaning and Normalization Approach

For the moving-source datasets, the dTs were used to locate the source (see section 3 below for more information on the location method). The next step in the signal processing was to take the signals from CH5 (the first 2600 points ~ 260 microseconds) and align the first arrivals to the same time. Specifically, the first positive peak of each signal within the first 800 points was found using a peak detection code (IDL code lclxtrem.pro with badpar.pro written by M.W. Buie, Lowell Observatory in 1997, modified in 1999) and then the signals were shifted to the same starting point (point = 200). Next, signals with large amplitudes or all zero were identified and removed. Large-amplitude signals were defined as signals with a summed absolute value of the first 200 points in the shifted signals that were greater than 5. Signals with sums greater than 5 were found to contain multiple events or were a late arrival from a previous event. For signals generated by the moving source, the next step involved sorting the remaining signals by depth location.

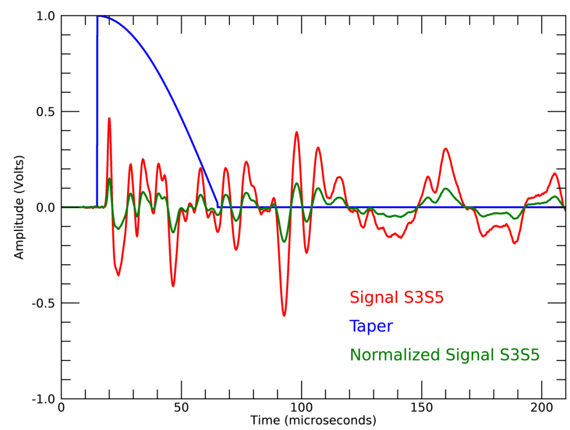

Figure S2. A signal from a moving source for the S3S5 condition (red), the taper (blue) and the resulting normalized signal (green) which is smaller in amplitude than the original signal.

The final signal length was set at 2200 points (i.e. the first 2200 points). The signals were then normalized by the energy in first 500 points (~50 microseconds). The energy was determined by tapering the signal with an open cosine taper (Figure S2 blue curve). The open cosine taper was selected to account for the variation in the source amplitude by weighting the amplitudes of the first arrivals more than the coda, and to give no weight to arrivals after 50 microseconds to suppress energy from reverberations from the walls of the sample. In a separate analysis we did not normalize by energy, but instead normalized all signals by the maximum of the first arrival. This first-arrival normalization was less stable, and led to weaker classification, because the uncontrolled sources have a wide variability. The energy normalization, by being an integral over the waveform, averages out peak-by-peak variability, while the taper removes sample wall reverberations.

The square root of the sum of the squares of the tapered signal amplitudes was taken and then used to normalize the amplitude of the signal. The normalized signals were used in the machine learning analysis. The total number of signals from CH5 used in the ML analysis for the moving and the stationary sources is given in Table S5 along with the number of signals prior to cleaning.

Table S5. Number of signals after data cleaning for CH5 from experiments with moving and stationary sources. The number in the parentheses is the number of signals from CH5 prior to cleaning. The bold-faced numbers indicate the number of signals after cleaning that used in the ML analysis for the moving source. The group number corresponds to the order of the experiments listed in Tables S1-S3 for the moving and stationary sources.

	Saturated	S3	S5	S3S5
Moving Source 4 Class	*Group 1* **865** (876) *Group 5* **1078** (1095)	*Group 2* **753** (762) *Group 6* **1253** (1262)	*Group 3* **781** (789) *Group 7* **871** (880)	*Group 4* **711** (711) *Group 9* **1231** (1236)
Moving Source 2 Class	*Group 1* **1347** (1461) *Group 3* **1458** (1472) *Group 5* **1223** (1290)		*Group 2* **949** (1022) *Group 4* **909** (1019) *Group 6* **1385** (1585)	
Stationary Source Non-Prospective	*Group 4* **2062** (2102)	*Group 2* **1553** (1669)	*Group 3* **1174** (1341)	*Group 1* **1192** (1505)
Stationary Source Prospective	*Group 2* **2065** (2150) *Group 6* **2025** (2110)	*Group 3* **1899** (1972) *Group 7* **1839** (1949)	*Group 4* **1922** (2010) *Group 8* **2139** (2227)	*Group 1* **1084** (1321) *Group 5* **1556** (1794)

S2.3 Distribution of Data among Training, Validation and Testing Sets

The data were split into training, validation and testing data sets. The validation data set was used to estimate the performance of the TNN while tuning hyperparameters. The test data sets are not used in the training or the tuning and was held back to provide insight into the overall performance of the tuned TNN.

For the non-prospective 4 Class case with stationary or moving sources, the data were split 20% for testing and 80% for training and validation (TV) analysis. The TV set of signals were balanced among the classes in terms of the number of signals in each class. If a class was underrepresented, signals were sampled with replacement from the original class. After balancing, the TV set of signals was split into 70% for training and 30% for validation. The validation set was not use in the training but was monitored while hyperparameters were tuned. The trained model with the best performance on the validation set was selected, and then the test data were run on that model. The 20% test set was not balanced.

For the prospective 4 Class stationary data analysis, the training and validation set came from Groups 1, 2, 3 and 4 (see Table S5) which we will term Group A. After balancing, Group A was split such that 70% of the data was used for training and 30% for validation. The test set came entirely from Groups 5, 6, 7, and 8 (see Table S5) with no balancing.

For the 2 Class moving source data analysis, the data set was split 80% for training and validation (TV), and 20% for testing. The TV set was split into 70% for training and 30% for validation. The validation set was not use in the training but was monitored while hyperparameters were tuned. The trained model with the best performance on the validation set was selected, and then the test data were run on that model. The 20% test set was not balanced.

S3. Source Location

S3.1 Event Location Method

For a moving source, event location was performed using the measured differences in arrival times, dT, of an event at the different sensors that recorded the event (see section 2.1). The dTs were used as the input to a custom-code in IDL that used a least-squares iterative method [*Geiger*, 1912] to determine the *x, y* and *z* position of the source for an assumed system velocity. Events were required to have been received by 5 or more sensors to minimize errors in location. An assumed acoustic velocity for the acrylic-based fracture system of V_{system} = 1910 m/s was used based on the arrival times measured by a controlled source for a wave transmitted from through the entire sample (see Figure 4 in manuscript). To determine if the V_{system} =1910 m/s was reasonable, a study was performed that takes

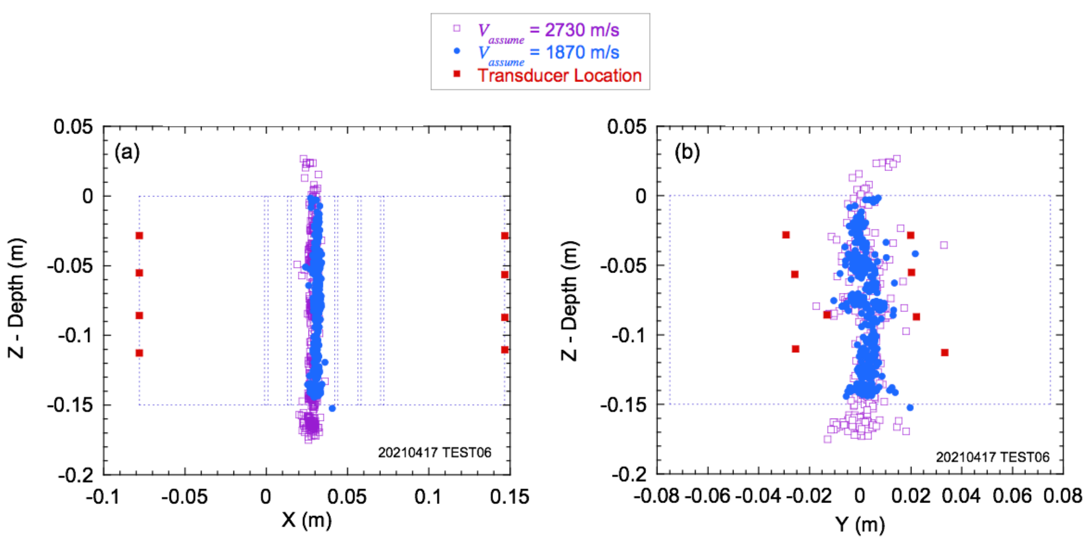

Figure S3. Comparison of event locations as a function of depth and (a) *x* and (b) *y* positions for V_{system} = 2730 m/s (purple open squares) and 1910 m/s (solid blue circles). The dash blue lines indicate the sample boundaries and the location of the fractures. The red squares represent the transducer locations (see Table 3 for channel locations).

advantage of the known positions of the dust. For instance, when using chattering dust, two known calibration points exist: (1) $z = 0$ m when the dust floats on the surface of the water prior to descending into the fracture and (2) when the dust rests on the bottom of the fracture ($z = -0.15$ m). These calibration points are useful for estimating the apparent velocity of a system that contains discontinuities because two source locations are known (Pyrak-Nolte et al., 2020). Figure S3 provides a comparison of the event locations for a moving source using the measured V_{system} = 1910 m/s versus the speed of sound in acrylic V_{system} = 2730 m/s (i.e. the matrix material). The use of the measured V_{system} limits the event locations between the top and bottom of the sample (given by dashed blue lines in Figure S3).

S3.2 Error Analysis of Event Location

For a single fracture, Pyrak-Nolte et al. (2020) examined the effect of refraction of an emitted acoustic signal from chattering dust in a single water-filled fracture. Upon emission, an acoustic wave is transmitted from the source through water and is refracted at the water-acrylic interface. They found that the deviation in the interpreted location and speed of descent increased with increasing aperture. For a fracture aperture of 2 mm, the error in descent speed was 1.3% with a 0.554 mm $\pm$0.146 mm shift in the *x*-location of the source.

Here, a conservation of transverse momentum (CoTM) approach was used to examine the errors in the *x*, *y* and *z* location from a moving source that was released in fracture F2 of a computational model containing 6 fractures similar to our laboratory fracture test bed. Travel times for an acoustic wave propagated from the dust location (source) to each of the 8 sensors (receivers) locations were calculated by using the CoTM to calculate the length of the refracted paths in the matrix and the water-filled fractures. The sensors were placed in the same locations as listed in Table 3 in the manuscript. As a wave travels through a fracture system (Figure S4), the wave is refracted at the interfaces between the water-filled fractures and the acrylic matrix. The refraction at the interfaces results in a difference in path length relative to the length of the direct ray path between the dust source and each receiver.

The CoTM approach is applied to individual unit cells. A unit cell consists of two media and the interface between them. The number of unit cells depends on the number of fractures in a system. For the 6-fracture system presented here with the dust grain released in F2, 12 unit cells were used: 5 unit cells between F2 and the left-hand side of the sample (see Figure 1 in the manuscript for fracture location) and 7 unit cells between F2 and the right-hand side of the sample.

The 1st step in the CoTM method is to initially calculate a direct path between the source (dust) location <x_d, y_d, z_d> and receiver <x_r, y_r, z_r> (dashed line in Figure S4). Then the point of intersection from the receiver location to the fracture wall is found (working from the end of the sample inward). In this analysis, $x_d = 0$ is located at the center of F2 (the source fracture) and assumes the dust is released at the center of the fracture plane $y_d=0$. The depth of the dust, z_d, range from 0 mm to 150 mm, and was incremented in 1 mm steps to simulate a moving source. The point of intersection, <x_{p1}, y_{p1}, z_{p1}> with the fracture wall is found using

$$x_{p1} = x_r + t(x_d - x_r)$$
$$y_{p1} = y_r + t(y_d - y_r)$$
$$z_{p1} = z_r + t(z_d - z_r)$$
$$t = \frac{x_{p1} - x_r}{x_d - x_r}$$

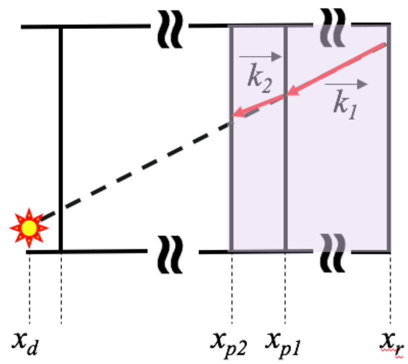

Figure S4. Exaggerated sketch of 1st unit cell with medium 1 and medium 2 to show refracted wave paths (red arrows). Not drawn to scale. Wavy lines indicate breaks. Source (star burst) is located in center of release fracture. All fractures are not shown. Dashed line represents the direct path between the source and receiver location (on x_r interface).

The wave vector, $\overrightarrow{k_1}$, is calculated from

$$x_{k1} = x_{p1} - x_r$$
$$y_{k1} = y_{p1} - y_r$$
$$z_{k1} = z_{p1} - z_r$$
$$|k_1| = \sqrt{x_{k1}^2 + y_{k1}^2 + z_{k1}^2}$$

and points from the receiver to the point of intersection on the 1st fracture wall. To find the direction of the wave in the second medium, $\overrightarrow{k_2}$, in this case water, conservation of transverse momentum is assumed, namely:

$$y_{k_2} = y_{k_1}$$
$$z_{k_2} = z_{k_1}$$

with $x_{k2} = x_{p2}$ as the known position of the second fracture wall. Then the point of intersection of $\overrightarrow{k_2}$ with the 2nd fracture wall is found by starting with

$$|k_2| = \sqrt{k_{2x}^2 + k_{2y}^2 + k_{2z}^2}$$
$$|k_2| = \frac{v_1}{v_2}|k_1|$$
$$k_{2x}^2 = \left(\frac{v_1}{v_2}|k_1|\right)^2 - k_{2y}^2 - k_{2z}^2$$

where v_1 and v_2 are the velocity for medium 1 and medium 2, respectively. For the 1st unit cell and other odd-numbered unit cells, medium 1 is the acrylic cell and medium 2 is the water. For the 2nd unit cell and other even-numbered unit cells, medium 1 is water and medium 2 is acrylic. In the numerical calculation "signum(x_{k1})" is used to isolate the correct root (real root). The point of intersection of $\overrightarrow{k_2}$ with the 2nd fracture wall $<x_{p2}, y_{p2}, z_{p2}>$ is found from

$$y_{p2} = y_{p1} + \left[y_{k_2} \frac{\left(x_{p2} - x_{p1}\right)}{x_{k_2}} \right]$$

$$z_{p2} = z_{p1} + \left[z_{k_2} \frac{\left(x_{p2} - x_{p1}\right)}{x_{k_2}} \right]$$

and x_{p2} is the known x-location of the second fracture wall. When accounting for refraction, the initial refracted ray may not "hit" the initial assumed position of the dust. Therefore, an iterative approach is used to find the z_d position needed to be within 100 micrometers of the center of the transducer.

After finding the refracted path length in each medium, the total travel time for the refracted ray path was calculated assuming compressional wave velocities of 1480 m/s and 2730 m/s for the water and acrylic, respectively. Source location was performed using the Geiger method [*Geiger*, 1912], an iterative approach that minimizes the difference between the observed and the predicted differences in arrival times among the sensors (dT). System velocities of V_{system} = 2430, 2530, 2630, 2730 and 2830 m/s were used to exam the errors in location that can arise from the assumption of composite or effective velocity for a system composed of a matrix with 6 water-filled fractures. In the simulations an idealized 6-fracture sample was used and the effect of dust released in any of the fractures on source location was also examined. For all release points, the dust was located in the center of the fracture plane ($y = 0$) in the y-direction and for the x-locations centered at x = 0, 0.0145, 0.029, 0.0435, 0.058 and 0.0705 m for release in fractures F0, F1, F2, F3, F4 and F5, respectively. All fractures were assumed to have an aperture of 2 mm. The idealized system has two matrix blocks (0.077 m and 0.078 m) for the large blocks on the ends of the sample (left &right), the thin slabs between fractures were 0.0125 m thick. A sample height of 150 mm was used.

For a moving source released in any of the fractures, the error in location caused by the additional path length from refraction depends on the depth in the sample (Figure S5). In the experimental study presented in the manuscript, the moving source was located in fracture F2. The accuracy of the interpretation of fracture location (x) was found to depend on the symmetry in fracture arrangement about the fracture plane containing the source, on the system velocity, V_{system}, and on depth. The error in x-, y- and z- locations depends on the depth of the moving source (Figure S5). The error in x-location (Figure S5a) for a source in F2 was small (~ -34 to 12 μm) compared to the error in x-location for a source released in either F0 or F5 (-3 to +5 mm) (Figure S5a). When a source fracture is close to the symmetry point of a fracture system, (i.e. central slab with an equal number of fractures on either side), the interpreted x-location from the dust AE is within 10s of micrometers of

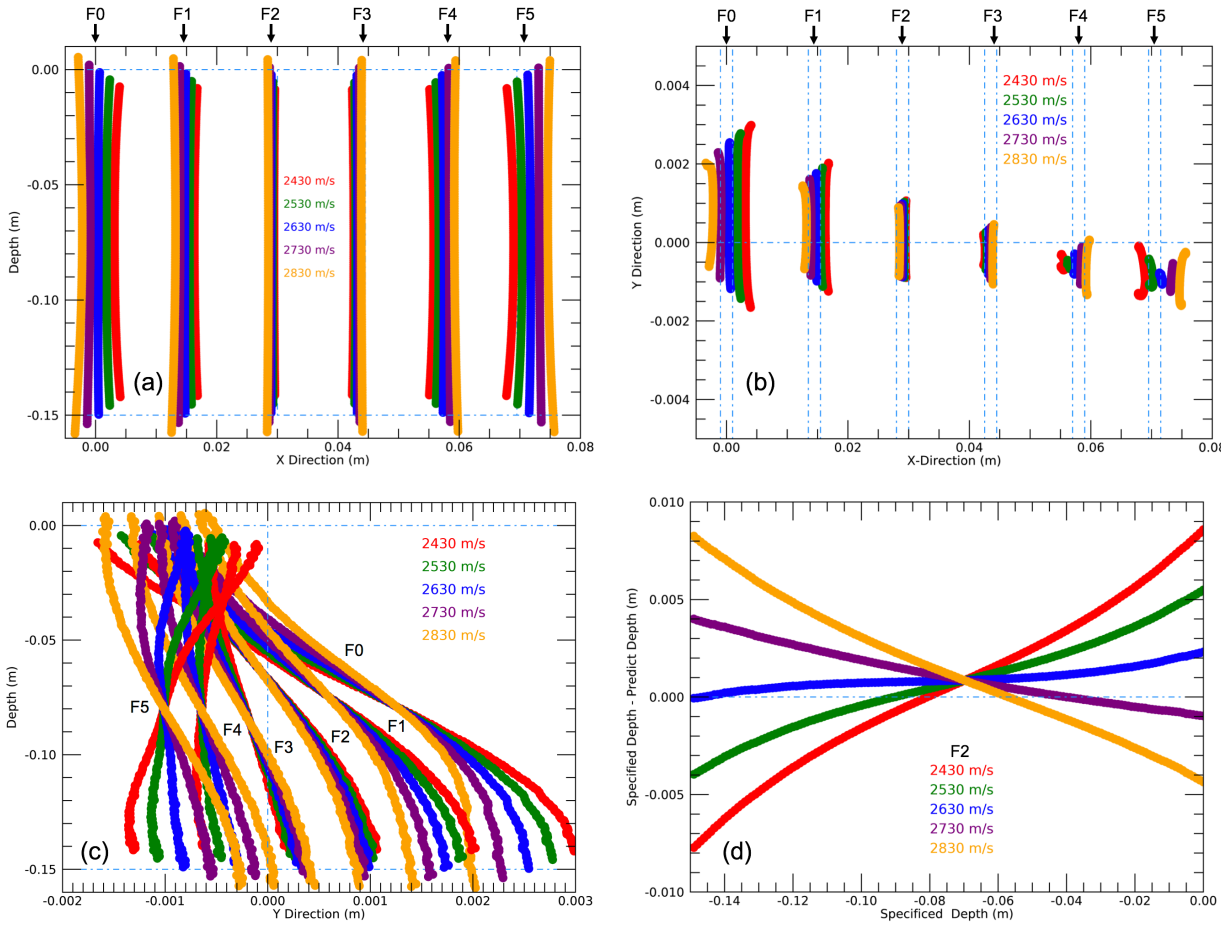

Figure S5. Comparison of computed source location for a moving source in an idealized set of 6 fractures for the source locations in the different fractures (released in F0, F1, F2, F3, F4, or F5) and for different values of V_{assume}. (a) Predicted depth of the event versus *x*-location of the fracture (see Figure 1 manuscript for coordinate system) with the center of the fracture where the source is released indicated by the black arrow, (b) Predicted *x*-location versus the *y*-location that is parallel to a fracture plane with the fracture walls given by the blue dashed lines. (c) Predicted depth of the event versus *y*-location of the fracture. (d) The difference between the specified depth and the predicted depth of the source location as a function of specified depth for the source released in fracture F2.

the actual the *x*-location (e.g. Figure S5a for fractures F2 or F3). However, a shift in *x*-location occurs when the fracture arrangement is not symmetric around the source (Figure S5a). This shift arises from the additional time delays from the presence of more fractures between the source fracture and one of the sensor planes. For example, when a source is released in either fracture F2 or F3, waves traveling in one direction will experience the additional delay caused by traveling through 3 water-filled fractures but only 2 water-filled fractures in the other direction. The interpreted *x*-locations of fractures F2 and F3 are shifted towards the side of the sample with fewer fractures between the source and the receiver. In addition, though the simulated dust was restricted in location to $y = 0$, the located y-positions all deviated from 0 m (Figure S5b) and the amount of deviation varied with depth (Figure S5c). The minimum differences in simulated versus specified locations are observed when sources are close to the point of symmetry. This suggests that accurate location of fractures from acoustic emissions or induced seismicity are affected by other fractures in the system as well as the location of the sensors relative to the source location.

As mentioned in supplemental information section 3.1, two calibration points are used to help defined V_{system}, $z = 0$ and $z = -0.15$m. In Figure S5a & d, interpretation of these calibration locations from dTs depends on the value of V_{system}. If V_{system} is too large (e.g. V_{system} = 2830 m/s in Figure 5a), the interpreted locations from the dTs exceed the dimensions of the samples (horizontal dashed blue line in Figure 5a). If V_{system} is too small, the interpreted locations for the calibration points from the dTs move inwards and away from the sample boundaries (e.g. V_{system} = 2430 m/s Figure 5a). The use of the calibration locations is critical to minimizing the error in location caused by the water-filled fractures. Figure S5d shows the difference in the specified depth versus the predicted depth for different V_{system}. A V_{system} =2630 m/s exhibited the smallest errors in z-location with depth in this comparison on an idealized system.

For the experiments reported in the manuscript, the dust source was always released in fracture F2 (Figure 1 in manuscript). Based on the simulations for moving source released in fracture F2, the error in location (x_e, y_e, z_e) because of refraction is on the order of $\{ x_e, y_e, z_e \} = \{\pm$ 30 μm, $\pm$1.5 mm, $\pm$2 mm$\}$. The error from refraction is less than deviations observed in the source z-location in the data from the stationary dust (i.e. at a fixed depth). The dust sat on a wire at a measured depth of -30 mm. Using the measured V_{system} = 1910 m/s, the averaged interpreted depth for the fully saturated condition from 46 events with at least 5 sensors was z = -30.7 mm with a standard deviation of 5.5 mm and a standard error of 2.0 mm.

Data Set S1. Original unprocessed data for the 2 class analysis of moving source data. (ds01)

Data Set S2. Original unprocessed data for the 4 class analysis of moving source data. (ds02)

Data Set S3. Original unprocessed data for the 4 class analysis of stationary source. (ds03)

Data Set S4. Original unprocessed data for the 2 class analysis of stationary source. (ds04)