Unsupervised learning tracks spatiotemporal evolution of hydraulic fractures

Aditya Chakravarty^ and Siddharth Misra^*

^Harold Vance Department of Petroleum Engineering, College of Engineering, Texas A&M University, College Station, Texas, USA
**Department of Geology and Geophysics, College of Geosciences, Texas A&M University, College Station, Texas, USA*

Abstract

Enhanced geothermal systems can provide a substantial share of the global energy demand if certain hurdles are overcome. One such hurdle is the accurate imaging of the fracture networks created in subsurface through hydraulic stimulation of these systems. Microseismicity associated with the stimulation is the primary means to locate the event hypocenters for estimating the stimulated rock volume. Using the data from a single three-component accelerometer in a monitoring well, the polarization features viz. azimuth, incidence, rectilinearity, and planarity are used as inputs for the unsupervised manifold approximation followed by clustering. We show that density-based clusters in the projected 3D space correspond to distinct types of hydraulically fractured zones around the injection point, thereby refining the interpretation of the microseismic cloud. The temporal evolution of these clusters can be used to track fracture creation and propagation for the various types of fracture zones.

Plain language summary

Geothermal energy is harnessed by injecting fluids in hot dry, rocks underground. For safe and efficient energy extraction, it is essential to have a clear picture of fractures created by pumping liquids at high pressures. Locating the earthquake events generated by cracking of rocks is the primary method of imaging fractures from seismic measurements. However, a lot of data is still left unused. We developed a method that uses blind machine learning methods (i.e., methods which assume no prior knowledge of the system) that uses the complete three-dimensional wave motion generated from such earthquakes. Our method aids in refining the interpretation of the earthquake locations. We show that this method can be effective in identifying distinct fracture sets from the cloud of microseismic events in a deep underground mine. This generalized workflow can be adapted to improve fracture characterization for both geothermal, as well as hydrocarbon resource production.

Introduction

Enhanced geothermal systems (EGS) are geothermal energy resources that require reliable circulation of fluids between the injection and production boreholes for maximizing the heat recovery. Detailed characterization of fracture networks is crucial for EGS development and energy production. Field-scale experiments to develop reliable fracture characterization techniques for EGS need heavy capital investment and large-scale instrumentation. The technical and financial challenges in field-scale experiments are driving the rise of intermediate-scale experiments, typically in the range of tens to hundreds of meters. The intermediate-scale experiments are more realistic representation of large-scale heterogeneous rock volume, in comparison to laboratory scale samples, while offering the possibility of dense instrumentation. Several such experiments have been carried out to characterize the response of hydraulic stimulation in crystalline rock, such as Aspo Hard rock laboratory in Sweden (Kwatiek et al., 2018) and Grimsel test site in Switzerland (Amann et al., 2018; Villiger et al., 2020).

Microseismicity (passive seismic) due to hydraulic stimulation is the key to map induced fractures in the subsurface (Fisher et al, 2008, Chen et al, 2018, McKean et al., 2019). The methods of picking phase arrivals for locating the microearthquake hypocenters consider only the largest amplitude signals that simultaneously register on multiple sensors within the array. This restricts the use of fine-scale details present in the continuous stream of signal for purposes of reliable fracture characterization. The present study lays out a data-driven workflow that analyzes the continuous signal of a sensor using unsupervised manifold-approximation learning followed by clustering. In doing so, we are using the high resolution spatiotemporal information about fracture growth during fluid injection.

Unsupervised learning methods have been applied to study a variety of seismological phenomena covering different length scales. Holtzmann et al. (2017) applied non-negative matrix factorization to extract features from a set of earthquakes recorded at a geothermal field and determined clusters which corresponded to distinct periods and rates of fluid injection. Mousavi et al. (2019) used deep learning features based on earthquake spectrograms to distinguish between local and tele-seismic earthquake signals. Bolton et al. (2019) used unsupervised clustering on statistical features of continuous acoustic emissions recorded during a laboratory-scale friction stick-slip experiment. The clustering indicated that the statistical features acoustic emission signal like variance and kurtosis changed systematically with the stress states. Directivity can be defined as the focusing of wave energy along a discontinuity in the direction of rupture. Ross et al. (2020) showed an unsupervised method of estimating directivity od large populations regional earthquakes, using frequency spectra. Chakravarty et al. (2021) showed that machine learning methods can improve fracture imaging from measurements in laboratory-scale experiments using acoustic emissions and ultrasonic transmission.

Johnson et al. (2020) used the spectral characteristics of continuous geophone signal combined with k-Means clustering to determine five types of signals recorded over the San Jacinto fault. They concluded that the non-tectonic signals primarily consist of distinct type of noise. The area under study was isolated and thus recoded minimal anthropogenic signals. Volcanic earthquake signals have been studied using machine learning to interpret the signals associated with different stages of eruptive cycles. For example, volcano-seismicity has been studied based on the infrasound signals recorded during Mt. Etna eruption (Watson, 2020). Time-domain and statistical features were reduced in dimension using principal component analysis and k-means clustering applied to assign labels. Shi et al. (2021) used array-signal processing features to obtain the covariance matrix-based features like entropy, coherency, and variance to determine clusters in the principal component space, and showed that the clusters were well correlated to the temporal evolution of the events.

Novelty of the proposed study

Conventional physics-based modelling approaches have certain limitations for purposes of fracture characterization under realistic dynamic subsurface environments. Our study is based on unsupervised learning to provide an alternative approach to map fracture growth and its properties during hydraulic stimulation by processing the three-component signal features have been used for microseismic based fracture characterization. The proposed workflow allows spatial tracking of the fracture growth at a high temporal resolution. The three-component information has hitherto not been used. Although there have been attempts to improve the microseismic signal quality through unsupervised learning, e.g. Zhou (2020); to the best of authors knowledge, no reference exists in published literature which address the issue of fracture characterization through analysis of three component microseismic signal. The specific questions attempted to answer in this article are: 1) what information about hydraulic fracture process can be extracted from the three-component particle motions of microseismic signals?

2) how such features can be applied towards unsupervised machine learning applications? and
3) what are the advantages and limitations of using such an approach?

Methods: Data

The EGS Collab project aims to deepen our understanding of the processes to create subsurface heat exchangers at realistic depth (Kneafsey, 2020). The test site is the Sanford Underground Research Facility (SURF) located at Leads, South Dakota. The nominal depth is 1500 meters. The setup consists of eight sub-horizontal boreholes for creating and monitoring the growth of fracture networks. The monitoring boreholes are densely equipped with active seismic sources, three-component accelerometers, hydrophones (pressure transducers), electrical resistivity probes, fiber optics (temperature, strain and acoustic) and a borehole displacement sensor. All sensors are cemented in place at the boreholes. We focus on the data recorded during the stimulation of a notch at depth of 50 m from the wellhead in the well E1-1 from 22 May to 24 May. A table showing brief description of the physical parameters over different experiments is given in Table S1.

The continuous microseismic signal is recorded by an array of three-component accelerometers and hydrophones (pressure transducers) installed in six monitoring boreholes surrounding the stimulated volume. The raw data recorded by the accelerometer/hydrophone arrays consists of 32-second-long continuous signals sampled with a frequency of 100 kHz. There is a gap of 1.5 seconds between successive continuous recordings. The signal is linearly detrended and a bandpass filter between 3 kHz to 10 kHz applied to isolate the frequencies of interest. All the accelerometers and hydrophones were grouted in place with cement within the monitoring boreholes.

Microseismic event catalog creation and determining fracture planes

Using a minimum of 10 accelerometer and hydrophone detections to qualify a microearthquake event, Schoenball et al (2020) used the software ‘Hypoinverse’ to locate the microearthquake event hypocenters. The list of event hypocenter locations and their time of occurrence will henceforth be referred as the ‘original catalog’. Chai et al. (2020) used a pre-trained deep learning neural network ‘PhaseNet’ (Zhu et al., 2019) and double difference relocation to refine the hypocenter locations. The revised locations henceforth will be referred as the ‘revised catalog’.

On 22 May, 10 minutes of continuous signal was measured, and 37 events were located. The injection rate was maintained at 0.2 liter/minute. The injection point was situated at depth of 164 feet from the well head. Considering the cumulative information gathered about the microseismicity at the end of the intermediate field-scale experiments in December 2018, 4 locations out of 37 located hypocenter locations were assigned fracture planes. To assign fracture labels, the distribution of the event hypocenters with a location uncertainty of less than 1.5 meters was considered, and the events that were interpreted to be associated with planar fractures were selected. 10 fractures were outlined from the complete original catalog in this manner. The location and orientation of the fracture ellipse were determined from the covariance matrix of the shortlisted hypocenters (Schoenball 2020).

Feature extraction

We assume that only impulsive signals are generated due to the microseismic events. Hence, one needs to isolate such signals from the continuous seismic record. To isolate impulsive signals from the continuous stream, we use a short-time-average/long-term-average (STA/LTA) filter to obtain triggers for impulsive signal in a single component (X-component). The length of each trigger is 8 ms. The

margins at the start and end of the triggers are adjusted to maintain a uniform size for each trigger. The seismic data processing was done on Obspy package (Beyreuther, 2010).

Not all accelerometers have the same sensitivity to ground motion. Figure S1 shows the differences in sensitivity of different accelerometers under study. We define signal-to-noise ratio (SNR) as the ratio of signal energy to noise energy. Noise samples having length identical to triggers are extracted, and the ratio of the energy of each trigger to that of the noise sample is defined as the SNR of a trigger. Figure S2 shows the calculated SNR for the six accelerometers. OT16 has exceptionally high SNR compared to other sensors. The others have a significantly lower SNR most likely due to improper coupling with the ground. We thus focus on OT16 sensor for the subsequent analysis.

The three-dimensional particle trajectory (i.e., hodogram) recorded in the accelerometer OT16 is shown in Figure 1. A hodogram is simply a cross plot of the three components of particle motion over time. All hodograms have the same length of 800 time steps (8 ms). The particle motion is representative of the seismic wavefield incident on the sensor at discrete time intervals. The accelerations show high degree of ellipticity in the shear phases. This ellipticity in the motion is caused by shear wave splitting. The anisotropy of the propagation medium splits the shear waves into orthogonal fast and slow shear waves that manifest into elliptical particle trajectories.

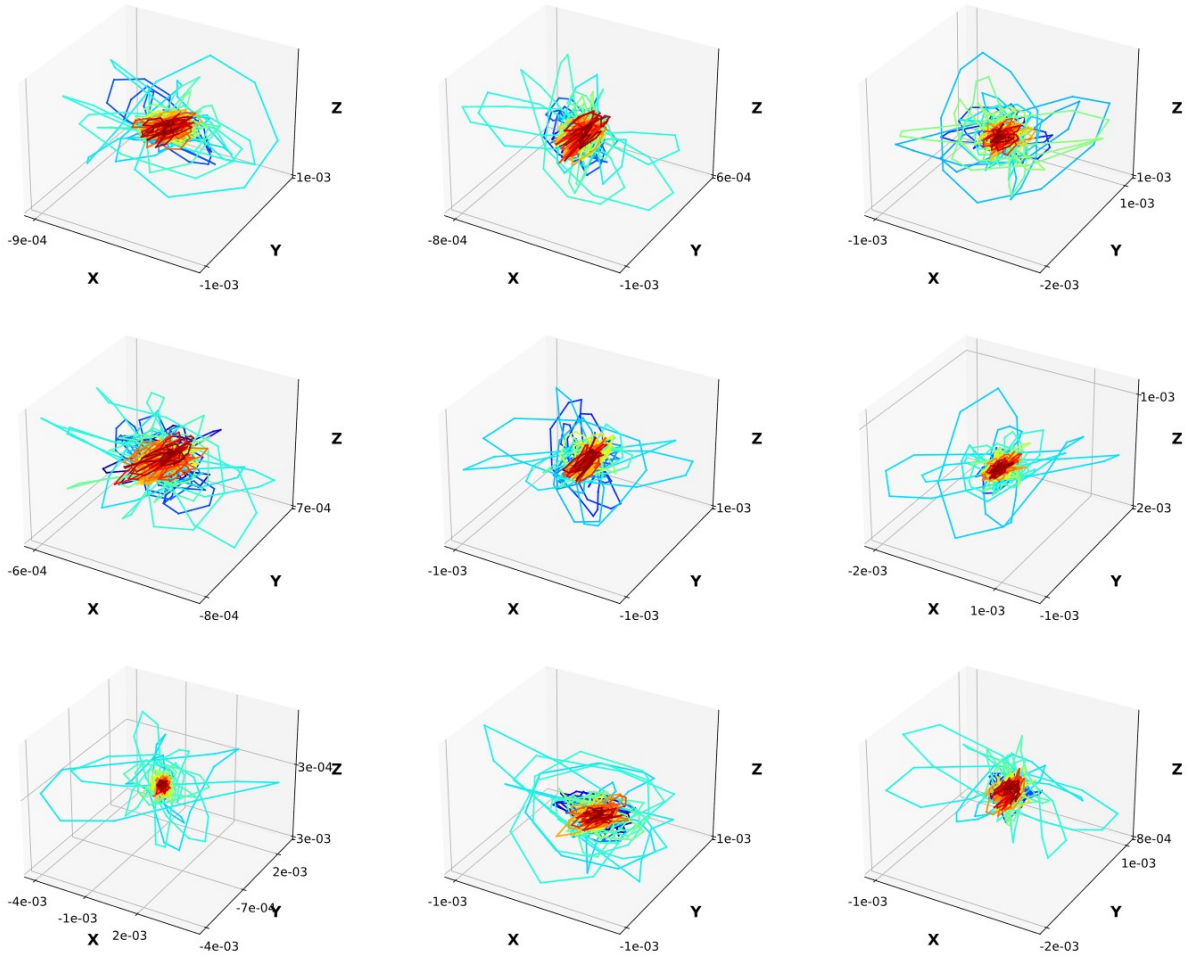

Figure 1: Three-dimensional particle acceleration as recorded by accelerometer OT16. Nine random samples out of 815 are shown. Cooler color represents early time and warmer color (red) represents late time. Note the pronounced changes in direction in the ellipticity caused due to shear wave splitting. All triggers have the same time length (8 ms). The axis ticks show the signal amplitude in Volts.

The covariance matrix of the three-component trigger signal is a 3×3 matrix. Since the input is real-valued signals, the matrix is positive definite, implying that the eigenvalues are non-negative and real. The covariance matrix is well suited for our application because scattering distortions and seismic noise are usually uncorrelated among the three components (Jurkervics, 1988). This property renders the covariance matrix-based approach particularly suitable for processing the noisy signals in our dataset. Matrix factorization of the covariance matrix yields eigenvalues (3×1 matrix) and eigenvectors (3×3 matrix). The eigenvalues are represented as λ_1, λ_2 and λ_3, and the eigenvectors are represented as u_1, u_2 and u_3. Based on the eigenvalues and eigenvectors thus obtained and using the feature definitions presented in Montalbetti and Kanasewich (1970), the polarization parameters are expressed as:

$$\text{Azimuth} = \arctan\left(\frac{u_{21}}{u_{11}}\right) \quad (1)$$

$$\text{Incidence} = \arccos\left(\sqrt{u_{11}^2 + u_{21}^2}/u_{31}\right) \quad (2)$$

$$\text{Rectilinearity} = 1 - \sqrt{\frac{\lambda_1}{\lambda_2}} \quad (3)$$

$$\text{Planarity} = 1 - \frac{2\lambda_3}{\lambda_1 + \lambda_2} \quad (4)$$

In summary, the 800 timesteps in each trigger of 8 ms in duration is processed to obtain the 4 above-mentioned features. These four features are used as inputs for the unsupervised learning. Using STA/LTA filter on the 10 minutes of continuous data recorded on May 22, 815 triggers were detected. Therefore, the extracted feature matrix for 22 May has 815 rows (each representing one trigger) and four columns (each representing one extracted feature), which is then scaled using MinMax scaler for optimal performance of unsupervised methods. The four features are either bound between 0 to 1 or -180 to +180, so MinMax scaler is suited for such a dataset with well-defined upper and lower bounds.

Unsupervised learning

After feature extraction, we apply manifold-learning methods to project the 4-dimensional feature data on to a 3-dimensional space. This aids the visualization and interpretation of data. Manifold is a complex multidimensional mathematical surface. Manifold learning methods are suited when there exist non-linear relationships within the data. The goal is to represent the distribution of data in the high-dimensional space using a low-dimensional surface (manifold), or in other words, the data points are considered as samples from the low-dimensional manifold that is embedded in the high-dimensional space.

Figure S3 compares the performance of two manifold-learning methods, namely uniform manifold approximation projection (UMAP) and t-distributed stochastic neighbor embedding (t-SNE), on the 4-

dimensional extracted feature space. Both methods use graph layout algorithms that aim to group similar samples together in a low-dimensional space, with a focus on preserving the balance between local and global structures as observed in the high-dimensional space. The comparison between t-SNE and UMAP reveals that UMAP results in tighter data clusters in reduced dimension space (Figure S3). In this paper, our focus is on the use of UMAP followed by clustering. UMAP is designed to be faster than t-SNE. Consequently, UMAP is suitable for large-sized data. UMAP uses fuzzy simplicial complex to build the high-dimensional graph wherein the edge weights represent the likelihood of connection between two points. In doing so, UMAP first learns the manifold in high dimensional space. UMAP then optimizes the layout of a graph in the low-dimensional space that is structurally as similar as possible to the graph identified in the high-dimensional space.

After obtaining the three-dimensional UMAP embedding space, we used density-based spatial clustering of applications with noise (DBSCAN) to assign cluster labels. The use of feature extraction followed by UMAP and then DBSCAN generates 4 distinct clusters indicative of the 4 distinct types of fracture systems produced due to the hydraulic stimulation on May 22. The sensitivity of cluster assignments to variations in UMAP hyperparameters, namely the number of neighbors (*nn*) and minimum distance (*min_d*), is shown in Figure 2. In the example shown, the number of neighbors varies between 6 and 20, and the minimum distance varies between 0.01 and 0.001. For one order variation in *nn* and *min_d*, the clustering tendency of the samples remains largely unchanged, attesting to the relative insensitivity of the proposed method to variations in UMAP hyperparameters.

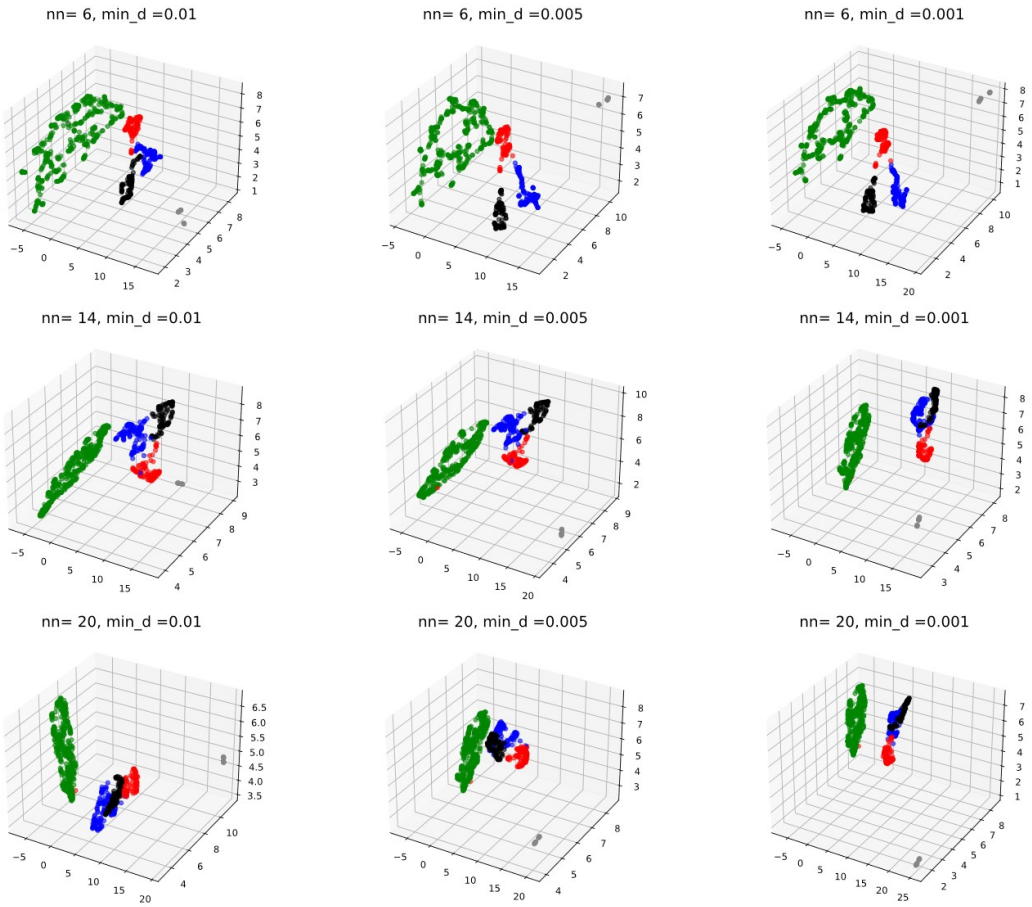

Figure 2: Variations in cluster distribution in UMAP embedding space due to the variations in the UMAP hyperparameters, number of neighbors (*nn*) and minimum distance (*min_d*). Colors represent different cluster labels, and axes represent the three embedding axes in the UMAP space.

The UMAP hyperparameter, number of neighbors (*nn*), is the most important that governs the construction of the high-dimensional graph. Low value of *nn* emphasizes the local structure in comparison to the global structure; thereby, focusing on the fine details in the high-dimensional space. The next important hyperparameter is minimum distance (*min_d*) that governs the compactness of the groups in the low-dimensional UMAP projection. Low value of *min_d* leads to tightly packed groups with more emphasis on the broad topological structure. In the proposed study, the optimum value of *nn* is 8 and *min_d* is 0.001.

Results and discussion

Using the same STA/LTA thresholds, we detected 815 triggers in the 10 minutes of continuous recording on May 22 (Day 1), 4800 triggers in the 90 minutes of continuous recording on May 23 (Day 2), and 14000 triggers in the 25 minutes of continuous recording on May 24 (Day 3). Figure S4 compares the rate of injection against the trigger rate for the three days. The hypocenters were located through inversion of arrival times of high SNR events detected simultaneously on several accelerometers and hydrophones.

Day 1: Figure S5 shows the distribution of 815 triggers in the 3D UMAP space for May 22 2018 (day 1) stimulation. In May 22 dataset, four dominant clusters are observed. The minor fifth cluster near embedding axis 3 corresponds to the set of electronic noise signals. Due to the high planarity and rectilinearity of electronic noise, the corresponding cluster lies relatively distant from the other clusters in UMAP space. The projection of polarization features in the UMAP space is an efficient unsupervised method to filter the high amplitude electronic noise from the continuous record. Out of 815 triggers detected on OT16 using STA/LTA, only 37 of those triggers correspond to located events. Figure S5 shows 4 of the 37 events that have been assigned a fracture plane based on the cumulative information of the event hypocenter locations at the end of stimulation in December 2018. Notably, the four events lie on distinct clusters in the UMAP space with no two different fractures sharing the same clusters. This is primarily because the events located in distinct fracture planes tend to have distinct polarization features because of surrounding media and the location of the event in space and time. Based on the original catalog, three fracture planes had been identified on May 22, namely 1003, 1005 and 1011. The fracture plane (1011) was assigned to 2 of the 4 events. The UMAP results indicate that, at the time of creation, these two events were in the distinct fractured zones that got connected over the 3 days of stimulation. When the data of only May 22 is analyzed, the two events related to the fracture plane 1011 lie on two distinct clusters, which suggests that at the end of May 22 stimulation, the two locations corresponded to two distinct fracture sets. The results and discussions of Day 2 and Day 3 are detailed in the supplementary information and include Figures S6 and S7.

The temporal variations in the clusters thus obtained are shown in Figure 3. Being a subset of the triggers, the located event hypocenters are also assigned clusters (second panel from bottom). The height of the stemplots is proportional to the log of energy of the trigger sample. Considering the located event hypocenters, cluster 4 is dominant at the early time and cluster 3 is dominant at the late time. Clusters 1 and 2 are predominant at the middle time of the injection. Cluster 1 (blue) has the largest dimensions among the four fracture sets defined henceforth. Although cluster 2 (green) has

relatively high activity throughout the injection, the extent of cluster 2 (Figure 4) is small, and limited around the injection point.

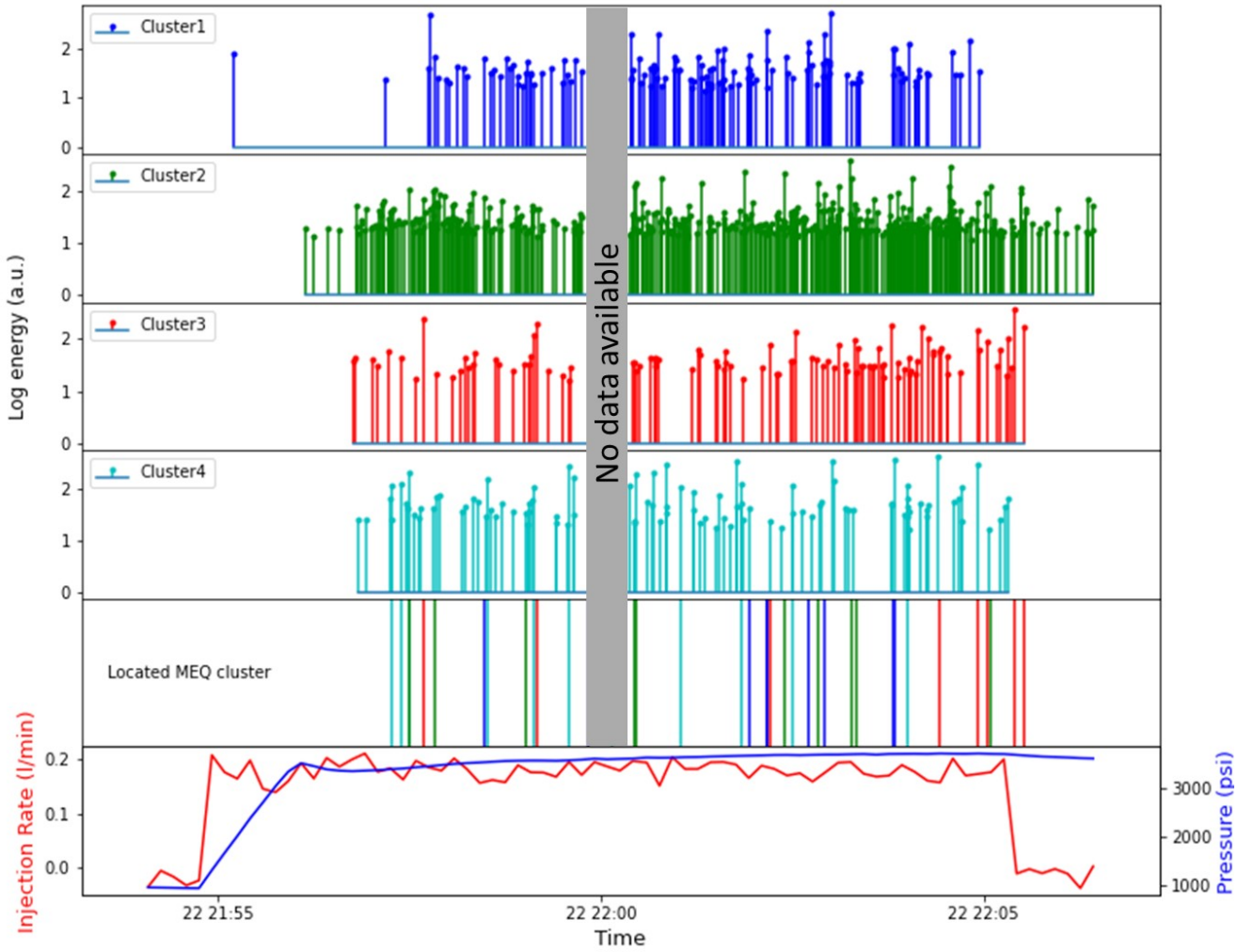

Figure 3: Temporal distribution of the fracture clusters created in the May 22 (Day 1) injection cycle (top four panels). Second last panel shows the time distribution of the 37 event hypocenters that were located out of 815 triggers. Bottom panel shows the hydraulic stimulation parameters, namely injection rate and pressure.

The spatial distribution of the 37 event hypocenters in the stimulated volume around the injection point that were produced on May 22 (Day 1) is shown in Figure 4. Each event is assigned a cluster label and color based on the proposed unsupervised learning workflow. There are 4 clusters in the figure. In Figure 4, events with identical cluster labels have similar locations in space. Clusters 1 and 3 (blue and red, respectively) form the two main branches of the hydraulic fracture. Cluster 2 (green) is limited to the zone beneath the injection points as shown in Figure 4. Cluster 4 (cyan) corresponds to a fracture which has grown along the wellbore.

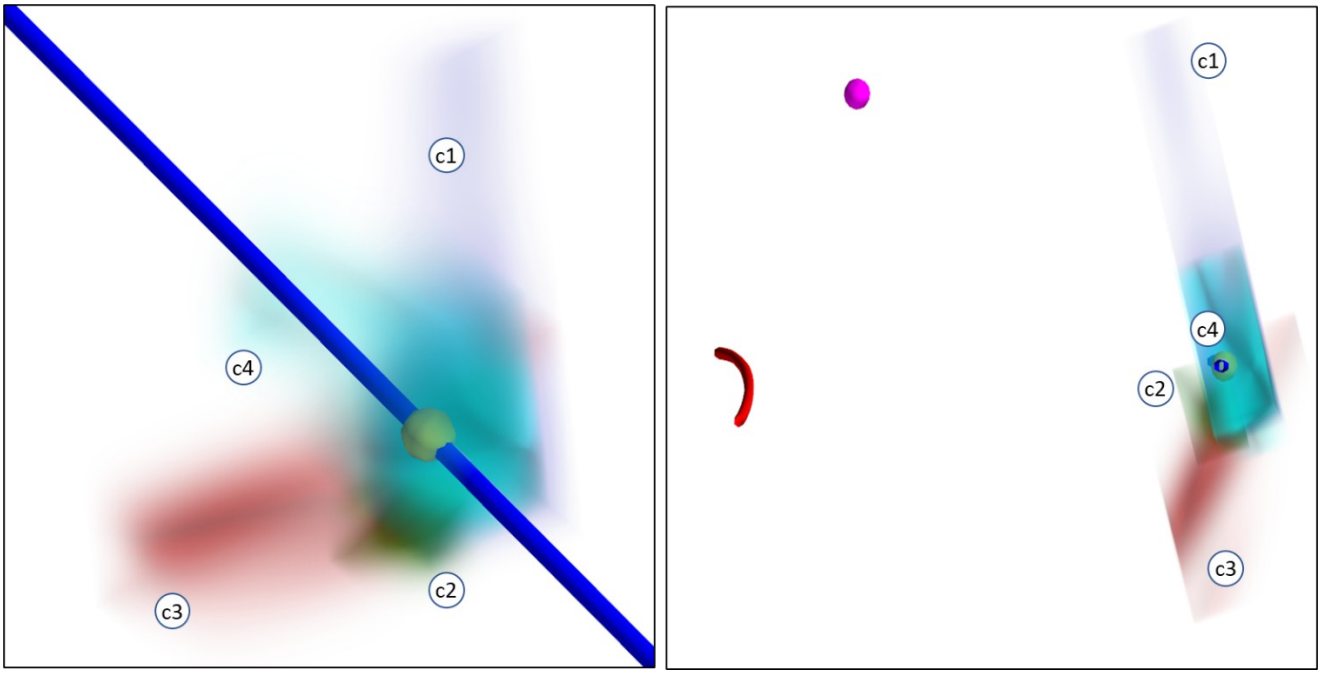

Figure 4: Close-up (left) and gun barrel (right) view of the fractured volume colored by cluster labels. The cluster labels are marked over their corresponding fracture branches. Red and blue indicate production and injection well, respectively. Pink sphere is OT 16 sensor. Yellow patch over the well represents the injection point. The average distance between the injection and production well is 9 meters.

For a single sensor, the recorded signal can be represented as a convolution of a source function and a medium transfer function. Since the particle motion (and hence, the polarization features) are representative of the seismic wavefield incident at discrete times, points having same cluster label impose similar seismic wavefield on the accelerometer. Our conclusion is that the microseismicity from distinct strands of fractures represents statistically different seismic wavefields. The examples from Days 1, 2 and 3 presented in Figures S5, S6 and S7 suggest that different strands of hydraulic fractures show different polarization signatures. As distinct clusters represent the different fracture branches around the injection point, the fine scale temporal variation in the cluster energy (as shown in Figure 3) is a strong proxy for the incremental growth of these fracture branches over time.

Conclusions

The present study demonstrates the application of unsupervised manifold-learning method on multi-component accelerometer measurements acquired in a miniaturized field-scale experiment, designed to study the spatiotemporal evolution of hydraulically induced fractures in enhanced-geothermal system. Based on signal-to-noise consideration, we use the three-dimensional particle motion measured by a single accelerometer installed on the monitoring borehole surrounding the stimulated volume. Individual triggers within the continuous data stream were isolated to identify the signals representing the hydraulically induced microseismic events.

Four polarization features (azimuth, incidence, rectilinearity, and planarity) were then derived from the identified signals. These features were processed using uniform manifold-approximation methods to project the raw signals corresponding to microseismic events on to a 3D space. Our study shows that the density-based clusters in the projected 3D space correspond to distinct types of hydraulically

induced fracture zones in the reservoir volumes around the injection points. The temporal evolution of these clusters is used to track the intensity and duration of reservoir stimulation (fracture creation and propagation) for the various types of fracture zones. The projection of polarization features in the UMAP space is an efficient unsupervised method to filter the high amplitude electronic noise from the continuous record. The proposed unsupervised learning indicates that, during the stimulation, there are many small secondary fractures with diverse orientation created, and not a single large fracture as often pictured through modelling and simulation efforts. Based on the microseismicity, the geometry of the hydraulic fracture created on May 22 is not penny-shaped, as modelled through conventional analytical methods, but irregular due to the stress and lithological heterogeneity of medium.

Such a workflow can be adapted to track the spatial and temporal seismic wavefield properties of a stimulated rock volume with a high time resolution. The information thus generated, along with microseismic locations, can improve the characterization of connected fracture pathways and yield a better understanding of fracture geometry.

Acknowledgements

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences Geosciences, and Biosciences Division, under Award Number DE-SC0020675. We acknowledge the efforts of the EGS collab team for conceptualizing and executing the experiments. We are grateful to the EGS collab team for providing public access to datasets of experiment 1. We are grateful to the seismology team at Lawrence Berkeley National Laboratory for fruitful discussions on microseismic analysis. Data availability - the microseismic data used in this study is in public domain and can be accessed from the Geothermal Data Repository using the link: https://gdr.openei.org/submissions/1166.

References

Amann, F., Gischig, V., Evans, K., Doetsch, J., Jalali, R., Valley, B., et al. (2018). The seismo-hydromechanical behavior during deep geothermal reservoir stimulations: Open questions tackled in a decameter-scale in situ stimulation experiment. Solid Earth, 9(1), 115–137. https://doi.org/10.5194/se-9-115-2018

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. (2010), ObsPy: A Python Toolbox for Seismology, Seismological Research Letters, 81 (3), 530-533.

Bolton, D. C., Marone, C., Shokouhi, P., Rivière, J., Rouet-Leduc, B., Hulbert, C., & Johnson, P. A. (2019, May 1). Characterizing acoustic signals and searching for precursors during the laboratory seismic cycle using unsupervised machine learning. *Seismological Research Letters*. Seismological Society of America. https://doi.org/10.1785/0220180367

Chai, C., Maceira, M., Santos-Villalobos, H. J., Venkatakrishnan, S. V., Schoenball, M., Zhu, W., … Thurber, C. (2020). Using a Deep Neural Network and Transfer Learning to Bridge Scales for Seismic Phase Picking. *Geophysical Research Letters*, *47*(16). https://doi.org/10.1029/2020GL088651

Chakravarty, A., Misra, S., & Rai, C. S. (2021). Visualization of hydraulic fracture using physics-informed clustering to process ultrasonic shear waves. International Journal of Rock Mechanics and Mining Sciences, 137, 104568. https://doi.org/10.1016/j.ijrmms.2020.104568

Chen, H., Meng, X., Niu, F., Tang, Y., Yin, C., & Wu, F. (2018). Microseismic Monitoring of Stimulating Shale Gas Reservoir in SW China: 2. Spatial Clustering Controlled by the Preexisting Faults and Fractures. *Journal of Geophysical Research: Solid Earth*, *123*(2), 1659–1672. https://doi.org/10.1002/2017JB014491

Fischer, T., Hainzl, S., Eisner, L., Shapiro, S. A., & Le Calvez, J. (2008). Microseismic signatures of hydraulic fracture growth in sediment formations: Observations and modeling. *Journal of Geophysical Research*, *113*(B2), B02307. https://doi.org/10.1029/2007JB005070

Holtzman, B. K., Paté, A., Paisley, J., Waldhauser, F., & Repetto, D. (2018). Machine learning reveals cyclic changes in seismic source spectra in Geysers geothermal field. *Science Advances*, *4*(5). https://doi.org/10.1126/sciadv.aao2929

James F. Montalbetti, Ernest R. Kanasewich, Enhancement of Teleseismic Body Phases with a Polarization Filter, *Geophysical Journal International*, Volume 21, Issue 2, November 1970, 119–129.

Johnson, C. W., Ben-Zion, Y., Meng, H., & Vernon, F. (2020). Identifying Different Classes of Seismic Noise Signals Using Unsupervised Learning. *Geophysical Research Letters*, *47*(15). https://doi.org/10.1029/2020GL088353

Jurkevics A., 1988, Polarization analysis of three-component array data, *Bull. seism. Soc. Am.*,78,1725–1743.

Kneafsey, T. J., Blankenship, D., Dobson, P. F., Morris, J. P., White, M. D., Fu, P., Schwering, P. C., Ajo-Franklin, J. B., Huang, L., Schoenball, M., Johnson, T. C., Knox, H. A., Neupane, G., Weers, J., Horne, R., Zhang, Y., Roggenthen, W., Doe, T., Mattson, E., … Collab, E. (n.d.). The EGS Collab Project: Learnings from Experiment 1. In PROCEEDINGS. Retrieved April 7, 2021, from https://www.energy.gov/eere/forge/forge-home

Kwiatek, G., Martínez-Garzón, P., Plenkers, K., Leonhardt, M., Zang, A., von Specht, S., et al. (2018). Insights into complex subdecimeter fracturing processes occurring during a water injection experiment at depth in Äspö hard rock laboratory, Sweden. Journal of Geophysical Research: Solid Earth, 123, 6616–6635. https://doi.org/10.1029/2017JB014715

McKean, S. H., Priest, J. A., Dettmer, J., & Eaton, D. W. (2019). Quantifying Fracture Networks Inferred from Microseismic Point Clouds by a Gaussian Mixture Model With Physical Constraints. *Geophysical Research Letters*, *46*(20), 11008–11017. https://doi.org/10.1029/2019GL083406

Mousavi, S. M., Zhu, W., Ellsworth, W., & Beroza, G. (2019). Unsupervised Clustering of Seismic Signals Using Deep Convolutional Autoencoders. *IEEE Geoscience and Remote Sensing Letters*, *16*(11), 1693–1697. https://doi.org/10.1109/LGRS.2019.2909218

Ross, Z. E., Trugman, D. T., Azizzadenesheli, K., & Anandkumar, A. (2020). Directivity Modes of Earthquake Populations with Unsupervised Learning. Journal of Geophysical Research: Solid Earth, 125(2), e2019JB018299. https://doi.org/10.1029/2019JB018299

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., & Johnson, P. A. (2017). Machine Learning Predicts Laboratory Earthquakes. *Geophysical Research Letters*, *44*(18), 9276–9282. https://doi.org/10.1002/2017GL074677

Schoenball, M., Ajo-Franklin, J. B., Blankenship, D., Chai, C., Chakravarty, A., Dobson, P., … Wood, T. (2020). Creation of a Mixed Mode Fracture Network at Mesoscale Through Hydraulic Fracturing and Shear Stimulation. Journal of Geophysical Research: Solid Earth, 125(12), e2020JB019807. https://doi.org/10.1029/2020JB019807

Shi, P., Seydoux, L., & Poli, P. (2021). Unsupervised Learning of Seismic Wavefield Features: Clustering Continuous Array Seismic Data During the 2009 L'Aquila Earthquake. *Journal of Geophysical Research: Solid Earth*, *126*(1), e2020JB020506. https://doi.org/10.1029/2020JB020506

Villiger, L., Gischig, V. S., Doetsch, J., Krietsch, H., Dutler, N. O., Jalali, M. R., et al. (2020). Influence of reservoir geology on seismic response during decameter scale hydraulic stimulations in crystalline rock. Solid Earth, 11(2), 627–655. https://doi.org/10.5194/se-11-627-2020

Watson, L. M. (2020). Using unsupervised machine learning to identify changes in eruptive behavior at Mount Etna, Italy. *Journal of Volcanology and Geothermal Research*, *405*, 107042. https://doi.org/10.1016/j.jvolgeores.2020.107042

Zhou, Y., & Wu, G. (2020). Unsupervised machine learning for waveform extraction in microseismic denoising. Journal of Applied Geophysics, 173, 103879. https://doi.org/10.1016/j.jappgeo.2019.103879

Zhu, W., & Beroza, G. C. (2019). PhaseNet: A deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, *216*(1), 261–273. https://doi.org/10.1093/gji/ggy423