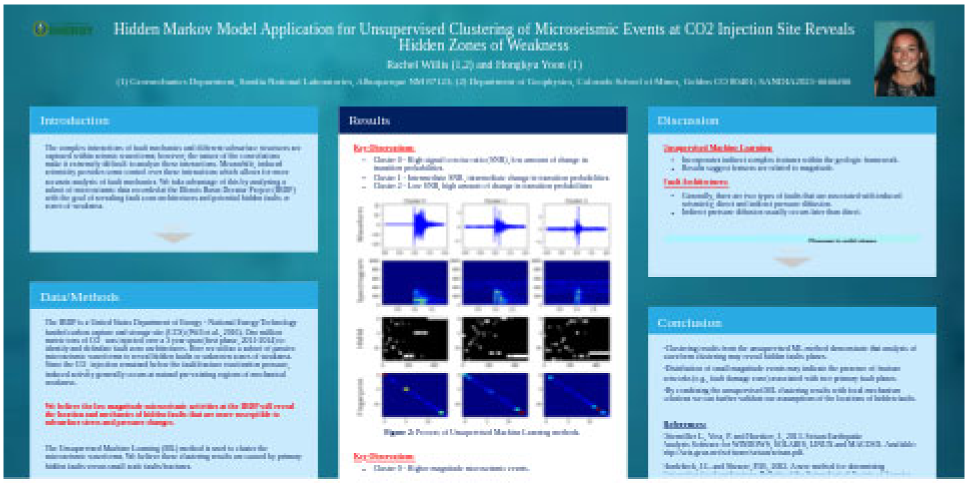
SAND2021-15414C

Hidden Markov Model Application for Unsupervised Clustering of Microseismic Events at CO2 Injection Site Reveals Hidden Zones of Weakness



Rachel Willis (1,2) and Hongkyu Yoon (1)

(1) Geomechanics Department, Sandia National Laboratories, Albuquerque NM 87123; (2) Department of Geophysics, Colorado School of Mines, Golden CO 80401; SANDIA2021-#######

PRESENTED AT:

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

INTRODUCTION

The complex interactions of fault mechanics and different subsurface structures are captured within seismic waveforms; however, the nature of the convolutions make it extremely difficult to analyze these interactions. Meanwhile, induced seismicity provides some control over these interactions which allows for more accurate analysis of fault mechanics. We take advantage of this by analyzing a subset of microseismic data recorded at the Illinois Basin Decatur Project (IBDP) with the goal of revealing fault zone architectures and potential hidden faults or zones of weakness.

Although traditional waveform analysis has been comprehensively studied, microseismic waveform analysis has not. This is mainly due to small scale features of the Earth associated with microseismicity. We adopt an Unsupervised Machine Learning method to identify and characterize subtle patterns within the microseismic waveforms. This reduces the computational expense generally associated with waveform analysis and allows for more complex patterns to be found that may have been overlooked by human analysis.

We compare waveform characterizations with focal mechanism solutions computed with the HASH (Hardebeck and Shearer, 2002) implemented in SEISAN software (Ottemöller et al. 2013) and present a novel method for location and characterizing hidden faults using microseismic data from CO_2 injection.

DATA/METHODS

The IBDP is a United States Department of Energy - National Energy Technology funded carbon capture and storage site (CCS) (Will et al., 2016). One million metric tons of CO_2 was injected over a 3 year span (first phase, 2011-2014) to identify and delineate fault zone architectures. Here we utilize a subset of passive microseismic waveforms to reveal hidden faults or unknown zones of weakness. Since the CO_2 injection remained below the fault/fracture reactivation pressure, induced activity generally occurs at natural pre-existing regions of mechanical weakness.

We believe the low magnitude microseismic activities at the IBDP will reveal the location and mechanics of hidden faults that are more susceptible to subsurface stress and pressure changes.

The Unsupervised Machine Learning (ML) method is used to cluster the microseismic waveforms. We believe these clustering results are caused by primary hidden faults versus small scale faults/fractures.

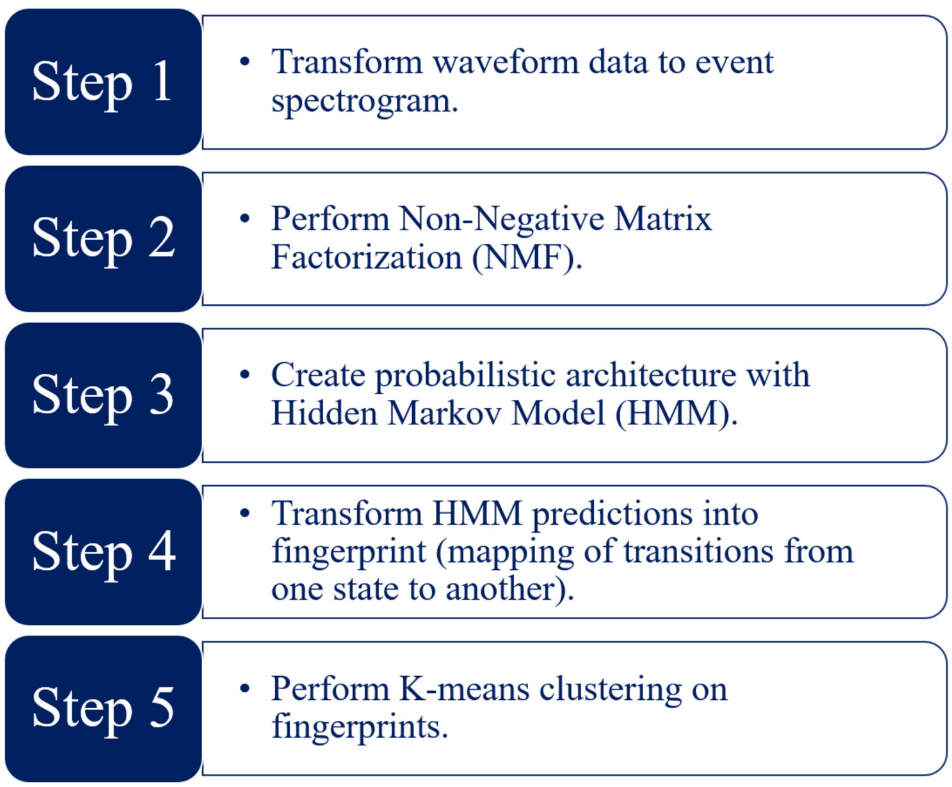


Figure 1: Processing and Machine Learning methods used to create clusters of microseismic waveforms.

We evaluate clustering results by comparing spatio-temporal patterns and focal mechanism solutions to determine locations of hidden faults.

RESULTS

Key Observations:

- Cluster 0 - High signal to noise ratio (SNR), low amount of change in transition probabilities.
- Cluster 1 - Intermediate SNR, intermediate change in transition probabilities.
- Cluster 2 - Low SNR, high amount of change in transition probabilities

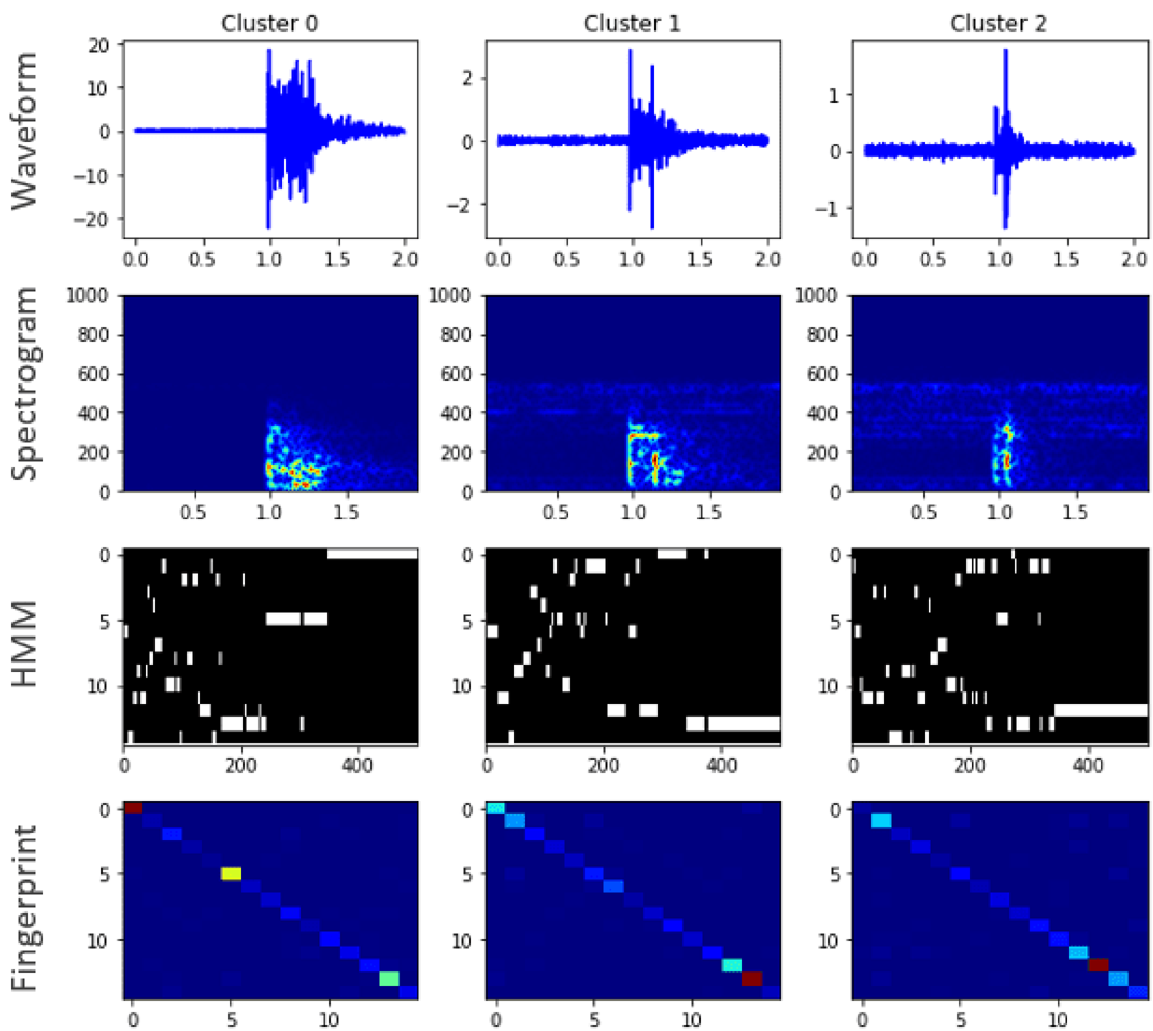


Figure 2: Process of Unsupervised Machine Learning methods with 2 second window waveform data.

Key Observations:

- Cluster 0 - Higher magnitude microseismic events.
- Cluster 1 - Intermediate to lower magnitude events.
- Cluster 2 - Lower magnitude events.

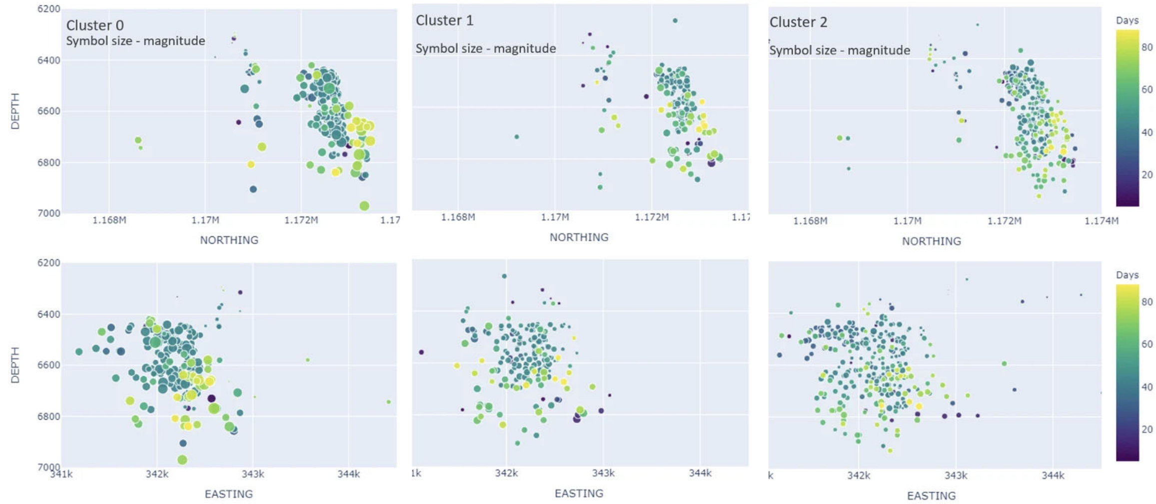


Figure 3: Spatio-temporal evolution of each cluster and the corresponding event magnitudes.

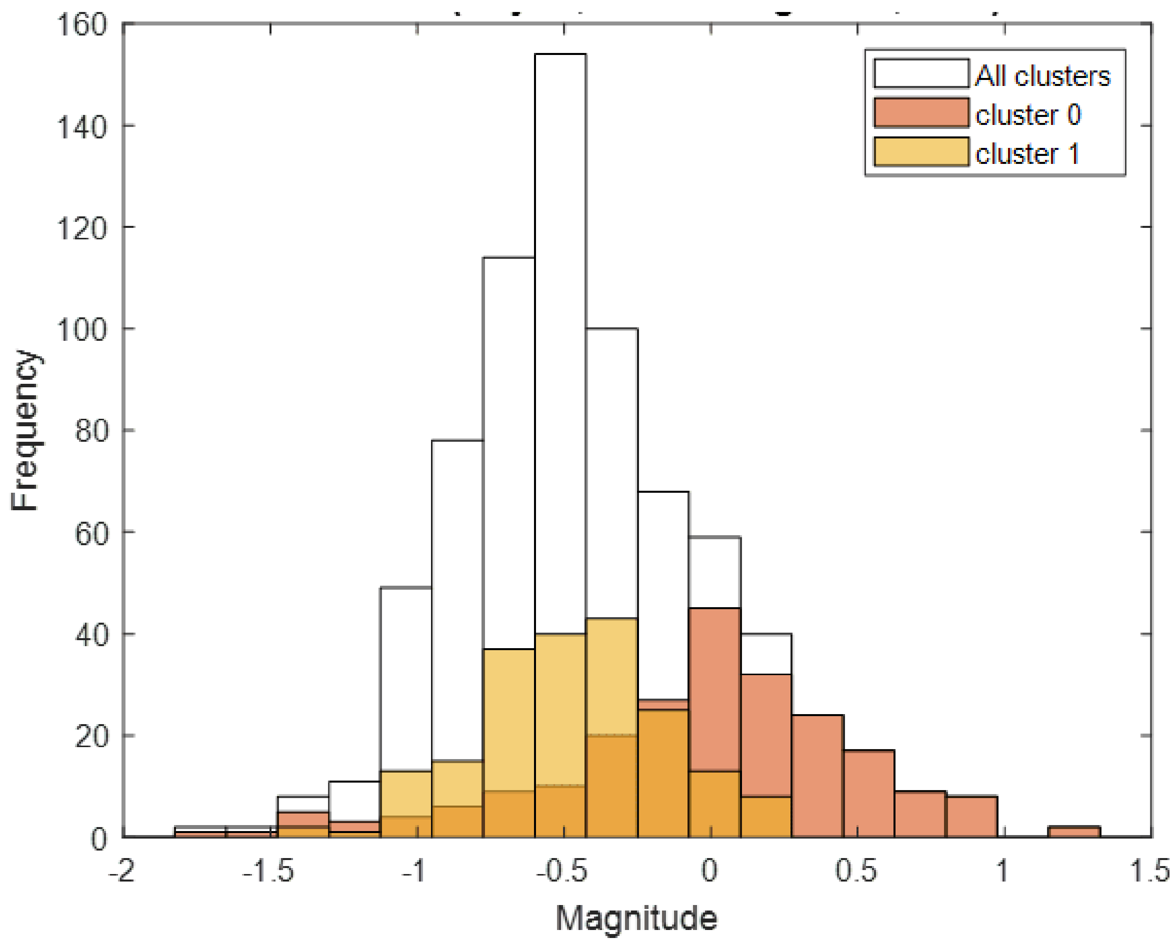


Figure 4: Frequency distribution of microseismic event magnitudes.

Cluster 0 has the highest magnitude and highest SNR; therefore, we believe these events occur along hidden faults. We compute focal mechanism solutions for Cluster 0.

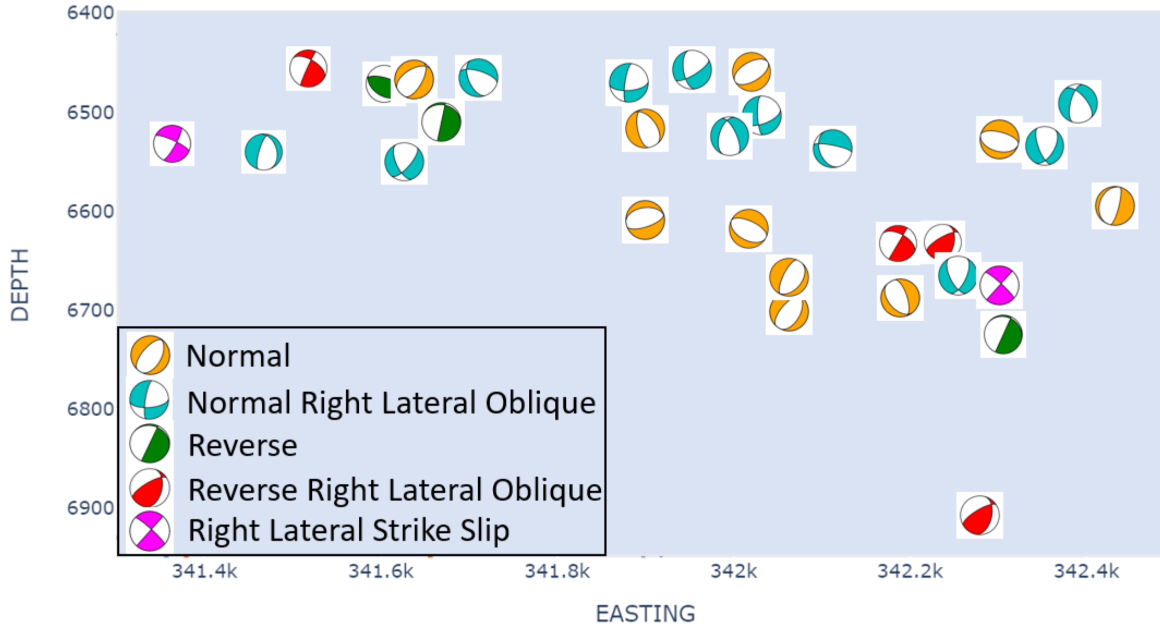


Figure 5: Focal mechanism solutions for highest magnitude events in Cluster 0.

DISCUSSION

<u>Unsupervised Machine Learning:</u>

- Incorporates indirect complex features within the geologic framework.
- Results suggest features are related to magnitude.

<u>Fault Architectures:</u>

- Generally, there are two types of faults that are associated with induced seismicity, direct and indirect pressure diffusion.
- Indirect pressure diffusion usually occurs later than direct.

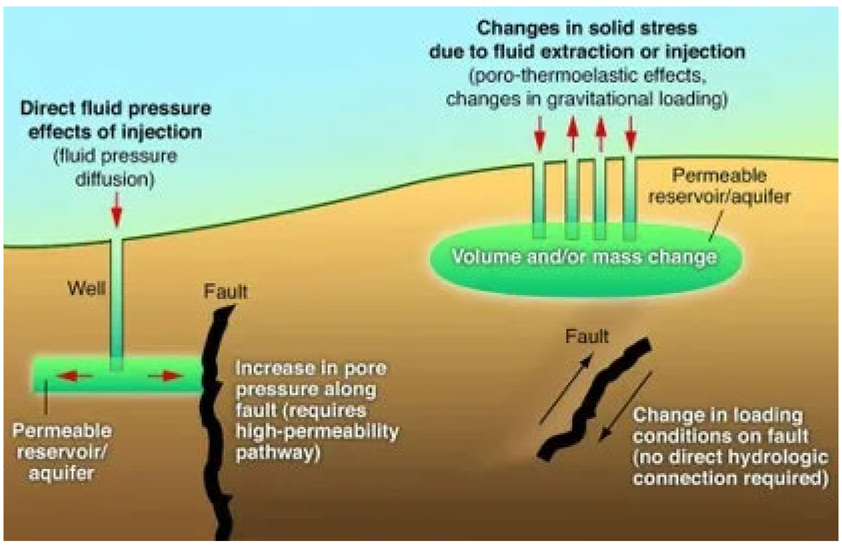


Figure 6: Direct and indirect pressure diffusion results in the two primary faults associated with induced seismicity (copied from USGS: (http://earthquake.usgs.gov/research/induced/modeling.php) http://earthquake.usgs.gov/Research/i (http://earthquake.usgs.gov/Research/i)nduced/modeling.php)

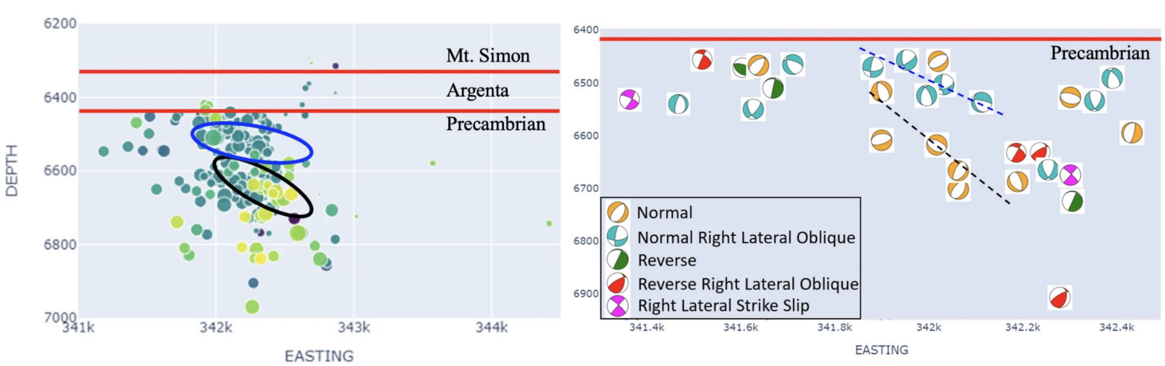


Figure 7: Right: Focal mechanism solutions. Bue and black dotted lines represent potential faults. Left: Spatio-temporal evolution with geologic layers. Mt. Simon is a reservior formation, Argenta is a layer in between a reservior and a basement, and Precambrian is a basement formation. The blue and black circles represent corresponding potential faults.

We see that there are common focal mechanism solutions that exhibit a roughly linear relation which may represent hidden faults. These linear relations also match up with common spatio-temporal patterns furthering the validity of the hidden faults locations.

CONCLUSION

•Clustering results from the unsupervised ML method demonstrate that analysis of waveform clustering may reveal hidden faults planes.

•Distribution of small magnitude events may indicate the presence of fracture networks (e.g., fault damage zone) associated with two primary fault planes.

•By combining the unsupervised ML clustering results with focal mechanism solutions we can further validate our assumptions of the locations of hidden faults.

References:

Ottemöller L., Voss, P. and Havskov, J., 2013. Seisan Earthquake Analysis Software for WINDOWS, SOLARIS, LINUX and MACOSX. Available: http://seis.geus.net/software/seisan/seisan.pdf.

Hardebeck, J.L. and Shearer, P.M., 2002. A new method for determining first motion focal mechanisms, Bulletin of the Seismological Society of America, 92, 2264-2276.

Will et al. "Microseismic data acquisition, processing, and event characterization at the Illinois Basin–Decatur Project." International Journal of Greenhouse Gas Control 54 (2016): 404-420.

Acknowledgements:

This work is supported by U.S. DOE, Office of Fossil Energy, Fossil Energy Research and Development Program.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Enery's National Nuclear Security Administration under contract DE-NA0003525.

ABSTRACT

In many geoscience applications, a mechanistic understanding of the seismic process is essential due to its significance in energy recovery from the subsurface, public safety, and mitigation of induced and natural seismic hazards. This research is analyzing two groups of microseismic data collected from the Illinois Basin Decatur Project (IBDP) site where the CO2 injection process lasted three years with the goal of locating and characterizing fault zone architecture. During subsurface injection activities there is a risk of activating hidden/unknown faults with stress and pore pressure changes associated with the injection. At the IBDP site, CO2 injection is maintained below fracture pressure meaning induced activities generally develop at natural pre-existing zones of mechanical weakness. Thus, the low magnitude microseismic data may reveal hidden/unknown faults that are prone to changes in stress and pore pressure.

Unsupervised learning techniques can be used to classify data based on complex relations, such as varying waveform characteristics resulting from different zones of mechanical weakness. We begin by pre-processing the data with filtering and a Short Time Fourier Transform (STFT) to reduce dimensionality and highlight rupture motion. Then we use an Unsupervised Machine Learning (ML) approach which combines the Nonnegative Matrix Factorization (NMF) and the Hidden Markov Model (HMM) to construct a time dependent probabilistic architecture. We represent the resulting spatio-temporal patterns as fingerprints of waveform characteristics. Unsupervised clustering is performed on the fingerprints to identify similarities and reveal time dependent patterns of the microseismic data associated with hidden/unknown faults. We present how the spatio-temporal patterns are related to changes in pore pressure and stress caused by CO2 injection. We also compare clustering results with the focal mechanism solutions of the microseismic data which are calculated with the open source HASH implementation in SEISAN and reveal a new methodology for locating and characterizing hidden/unknown faults using microseismic data from CO2 injection.

SNL is managed and operated by NTESS under DOE NNSA contract DE-NA0003525.

REFERENCES

[1] Dando, B. D. E., Goertz-Allmann, B., Iranpour, K., Kühn, D., & Oye, V. (2019). Enhancing CO2 monitoring at the Decatur CCS site through improved microseismic location constraints. In *SEG Technical Program Expanded Abstracts 2019* (pp. 4893–4897). Society of Exploration Geophysicists. https://doi.org/doi:10.1190/segam2019-3208059.1

[2] Havskov, J., Voss, P.H. and Ottemoller, L. (2020). Seismological Observatory Software: 30 Yr of SEISAN. Seismological Research Letters, 91 (3): 1846-1852. DOI: https://doi.org/10.1785/0220190313

[3] 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), eaao2929. https://doi.org/10.1126/sciadv.aao2929

[4] Will, R., Smith, V., Lee, D., & Senel, O. (2016). Data integration, reservoir response, and application. *International Journal of Greenhouse Gas Control*, *54*, 389–403. https://doi.org/https://doi.org/10.1016/j.ijggc.2015.12.020