SAND2020-12585C

Sandia National Laboratories

Exceptional service in the national interest

Machin Learning Applications for Acoustic Emission and Induced Seismic Data

CouFrac 2020

Hongkyu Yoon
Geomechanics Department
Sandia National Laboratories, NM, USA

Acknowledgments

- Laboratory Directed Research and Development program at Sandia National Laboratories

- U.S. DOE, Office of Fossil Energy, Fossil Energy Research and Development Program

Collaborators:
Daniel Lizama, Rachel Willis (Sandia)
Laura Pyrak-Nolte, Antonio Bobet, Liyang Jiang (Purdue Univ.)

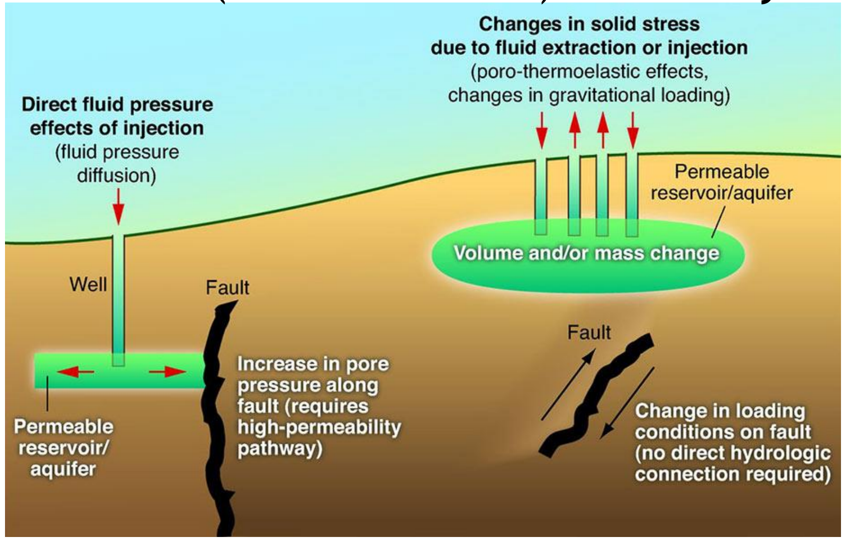
- **Motivations**
- Machine learning applications at laboratory scale
- Unsupervised machine learning for microseismic data at a field scale CO_2 injection site
- Supervised machine learning for microseismic data
- Summary

Background & Motivations

♦ Motivations

- Fluid injection or withdrawal causes changes in pore pressure, resulting in induced seismicity during subsurface energy activities
- Reduce risks of induced seismicity and improve subsurface energy activities (unconventional resource recovery, geological carbon storage, geothermal energy recovery)

Induced (human-caused) seismicity



USGS: http://earthquake.usgs.gov/Research/induced/modeling.php

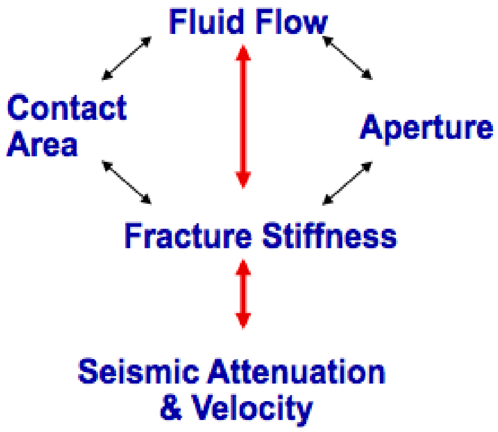
♦ Goals

(1) Delineate fracture and failure mechanisms using well-controlled experiments

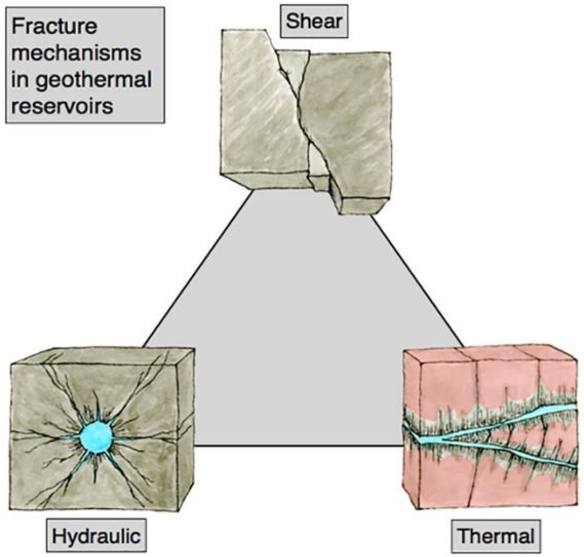
(2) Develop/apply machine-learning techniques for seismic wave data analysis and event detection

Linkage between geomechanical and geophysical processes in mechanical discontinuities

- Changes in the spectral contents of waveforms are likely due to wave propagation + faulting processes - initiation, propagation and coalescence of pre-existing discontinuities loaded in mixed mode I-II-III (**Damage Mechanics Challenge**)
- Machine learning has been increasingly applied for seismic data analysis



Courtesy from Pyrak-Nolte

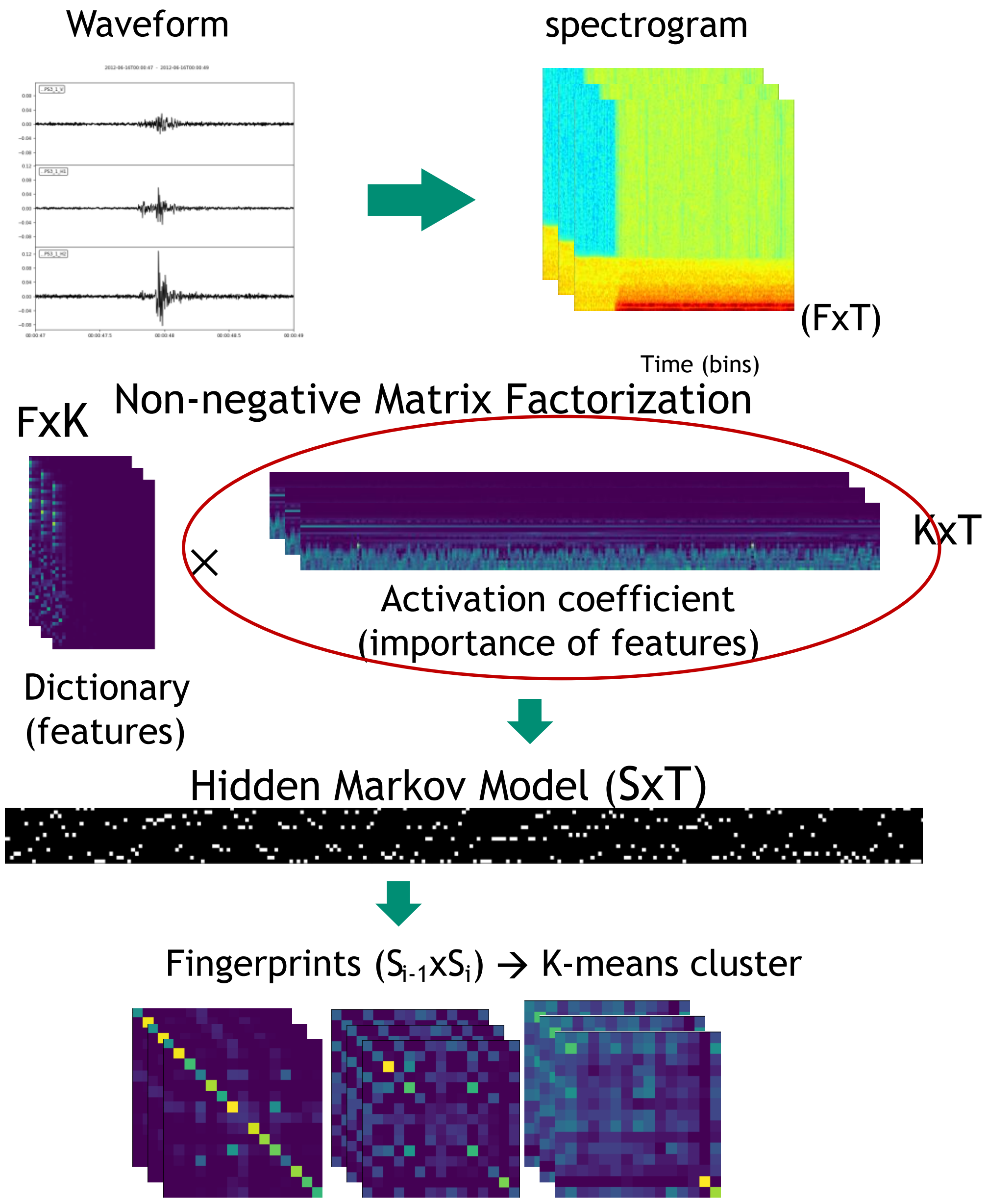


Holtzman et al. Sci Adv 2018

- Motivations
- **Machine learning applications at laboratory scale**
- Unsupervised machine learning for microseismic data at a field scale CO_2 injection site
- Supervised machine learning for microseismic data
- Summary

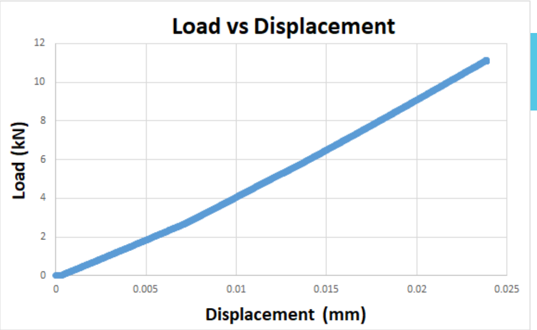
Unsupervised ML – Fingerprint method using NMF-HMM

- Fingerprint-based clustering method: Pattern of state sequences forms fingerprints
  - Clustering: acoustic/seismic state ~ mechanical behaviors
  - Spectrogram (Short Time Fourier Transform)
  - Non-negative Matrix Factorization
  - Hidden Markov Model (S states)
  - K-means clustering

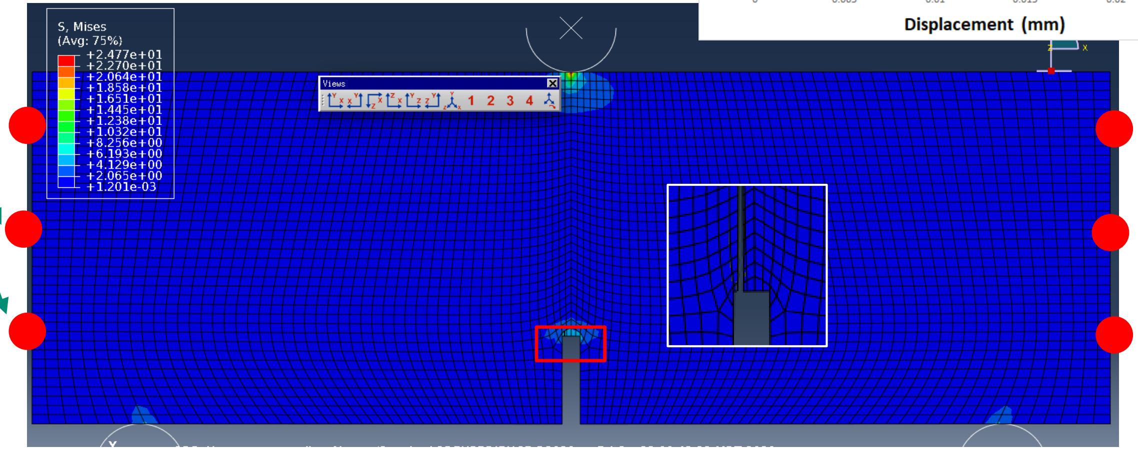


Ref: Holtzman et al. (Sci. Adv. 2018)

Simulation of fracture failure and acoustic emission

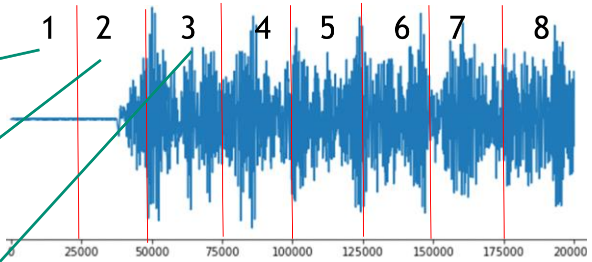


Sensors

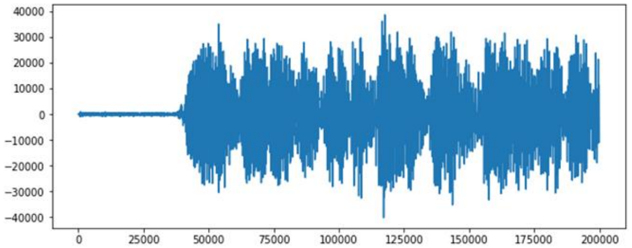


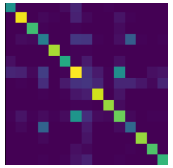
- 3 point bending simulation
  - 3 point bending with a central notch (ABAQUS)
  - Crack propagation and acoustic emission (XFEM)
  - 6 sensor nodes sampling at 8 MHz
  - Velocity and acceleration data from each sensor
  - #1-#3 sensor data for training, 4-6 for testing
  - Limestone material properties (homogeneous)

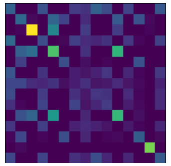
Waveform

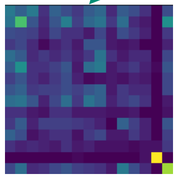


Acceleration





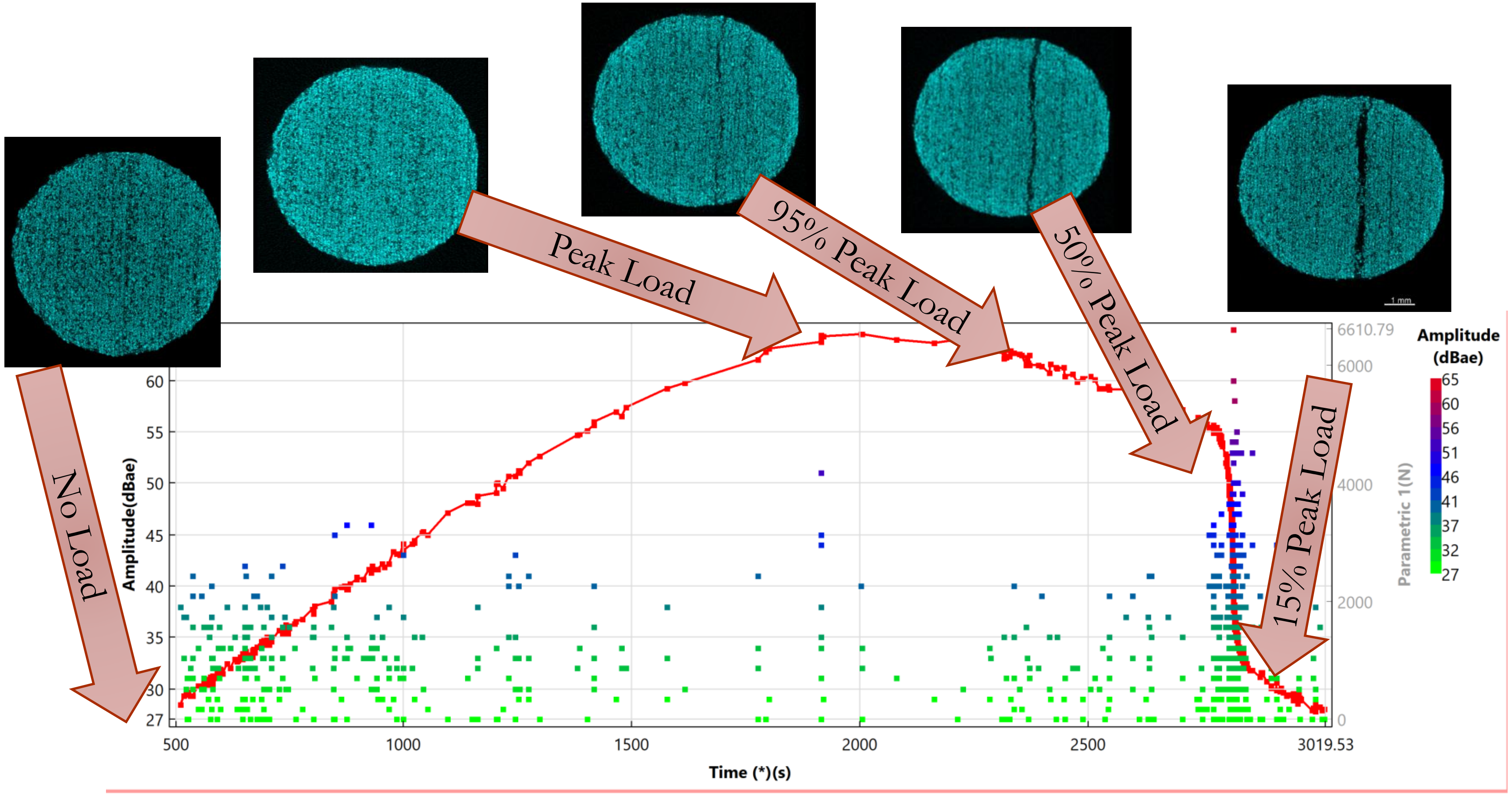




Acoustic state clustering by slicing simulated data into 8 window sets containing pre-crack (step 1), crack initiation and propagation (step 2) and post crack (steps 3-8) waveforms

Experimental Setup & Acoustic Data Acquisition

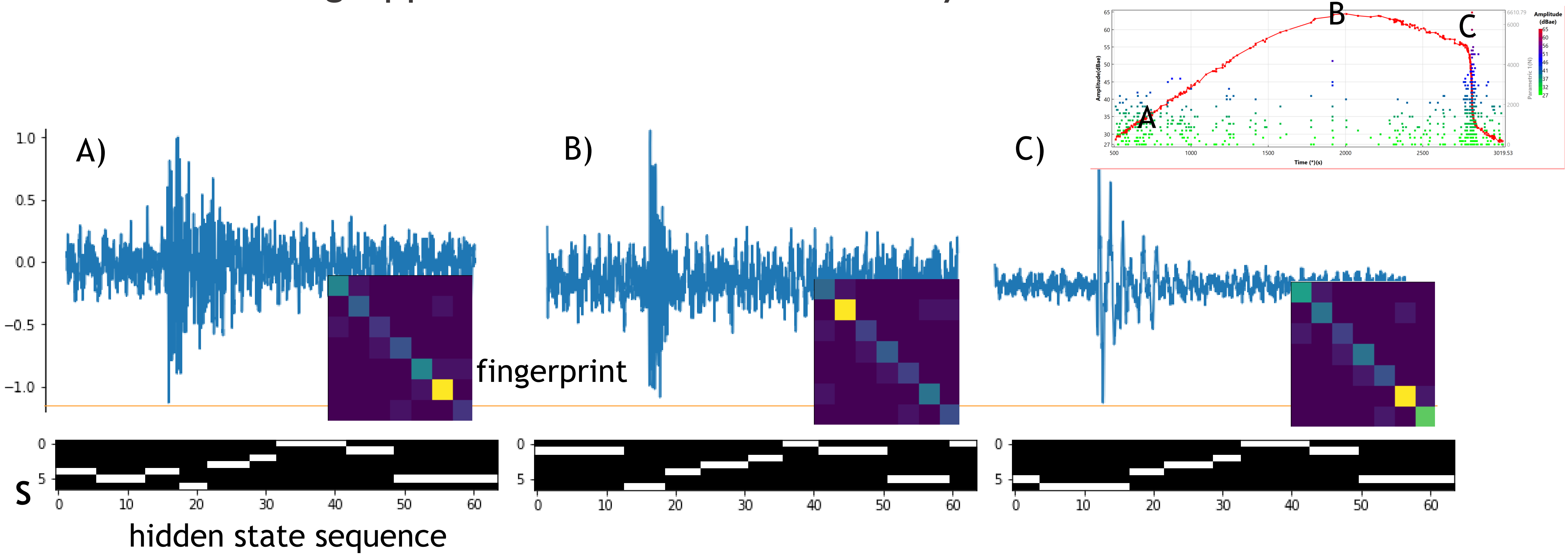
- MicroCT images



- Six sensors (Channels)
- 200-400 kHz filter

Jiang et al. (Sci. Rep. 2020)

Machine Learning Applications for Waveform Analysis



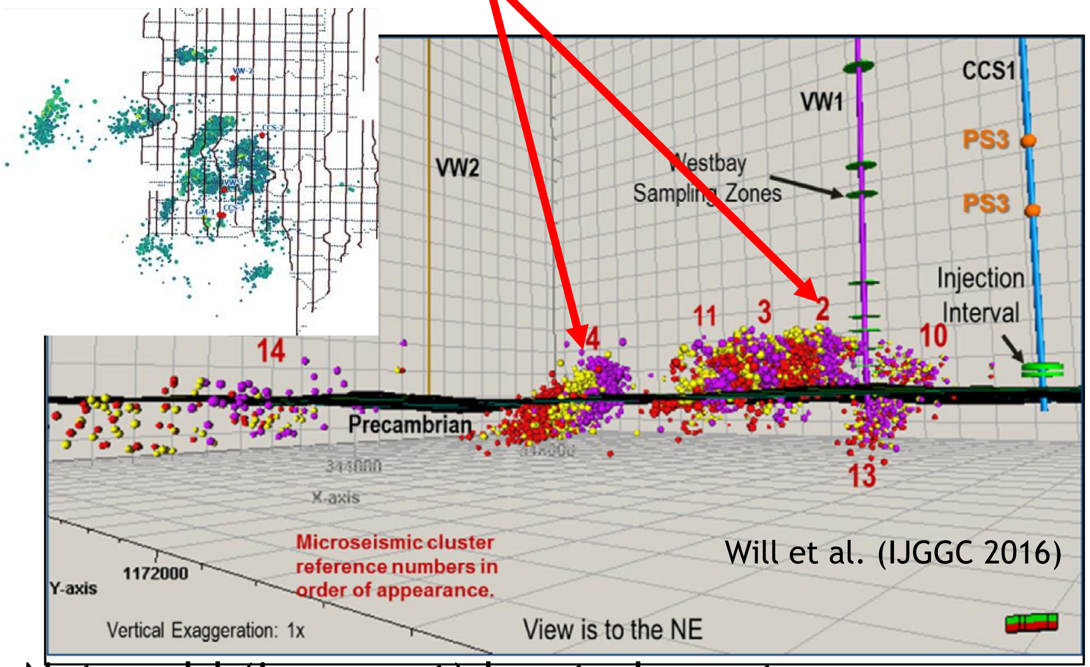
- Fingerprint based mechanical state clustering by comparing waveform evolution in the UCS test.
- Waveforms correspond to 3 stages in the loading curve, namely A) initial loading slope, B) max load region, C) post-failure region.

- Motivations
- Machine learning applications at laboratory scale
- **Unsupervised machine learning for microseismic data at a field scale CO_2 injection site**
- Supervised machine learning for microseismic data
- Summary

Machine learning applications for fault detection and the presence of hidden faults/fracture networks (in collaboration with ISGS and MIT)

Target clusters 2 & 4
Two 2-months data

Illinois Basin Decatur Project (IBDP)



Note: old (incorrect) located events

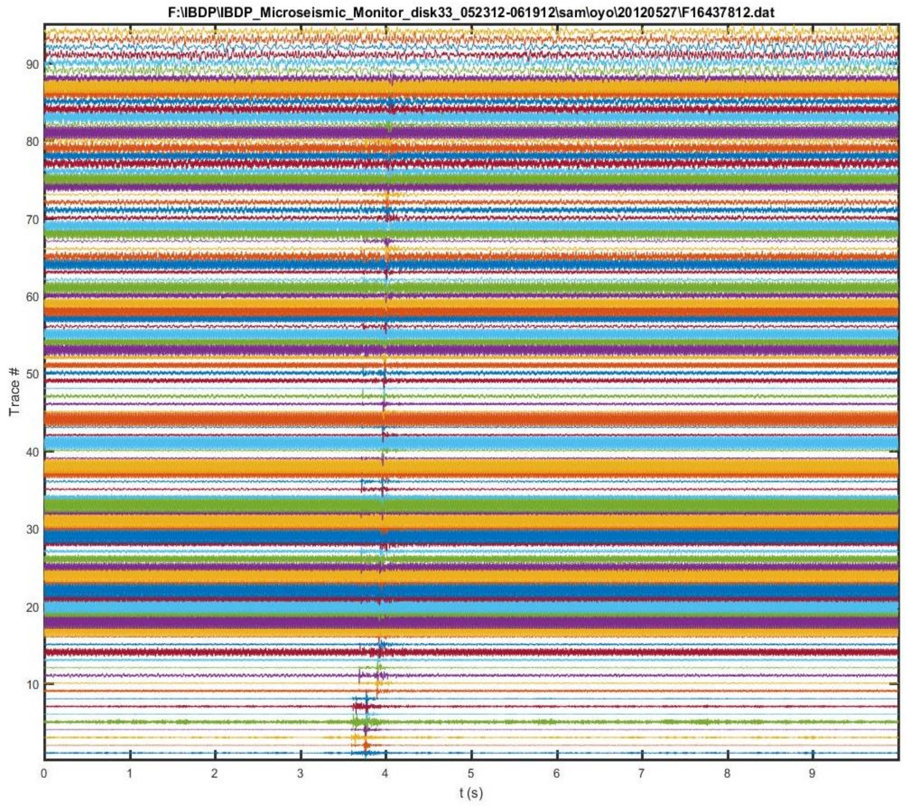
- Raw continuous data (100’s TBs) – 10s windows
- Detected event data (~19K data) – 2s windows

  ↓ manual selection
- Located event data (~5K data) - 2s windows

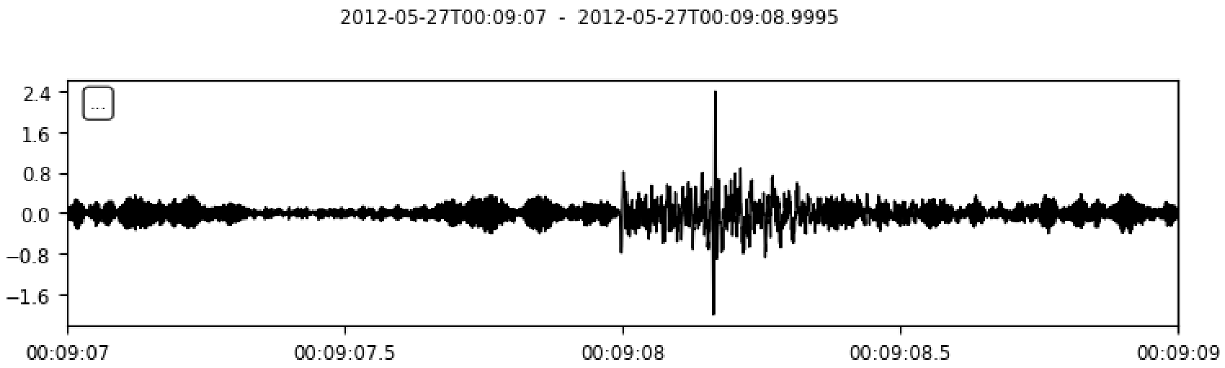
- Improve the detection of low-magnitude, unidentified events & locations to discover undetected/hidden fault/fracture systems
- Rapid recognition of the presence of faults/fault interactions
- Characterize microseismic waveforms, the relations among the events, and reliable identification of microseismic sources integrated with forward/inverse modeling

Waveform Data at the IBDP Site

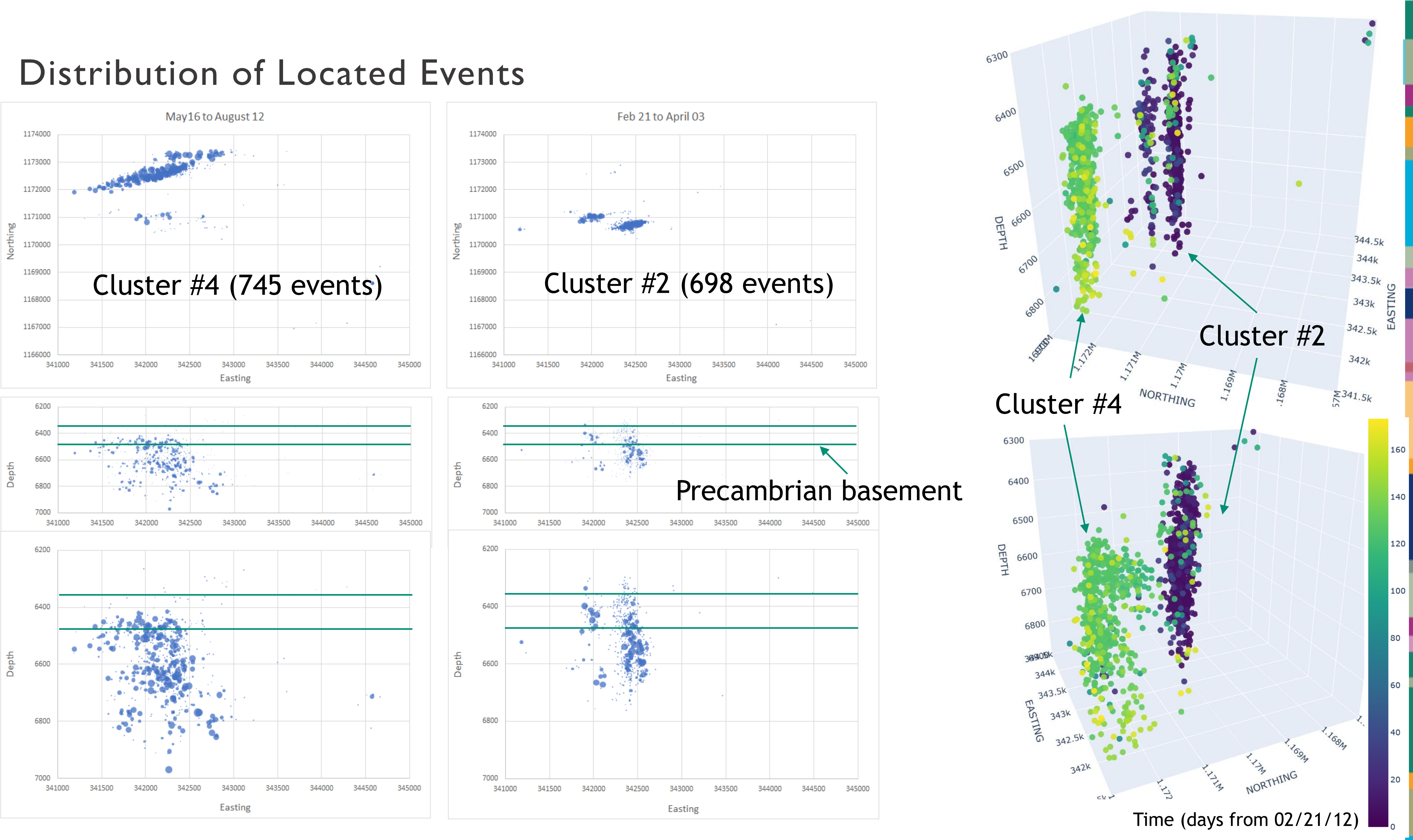
- Raw (unprocessed) data (~ 7TB for ~4 months)
- Length: 10 seconds window
- 2 kHz sampling rate
- Traces: 84-94
- 8640 sequential files for 24 hr
- 4 channel on PS3 (injection reservoirs) and 2-3 channels on GM geophones



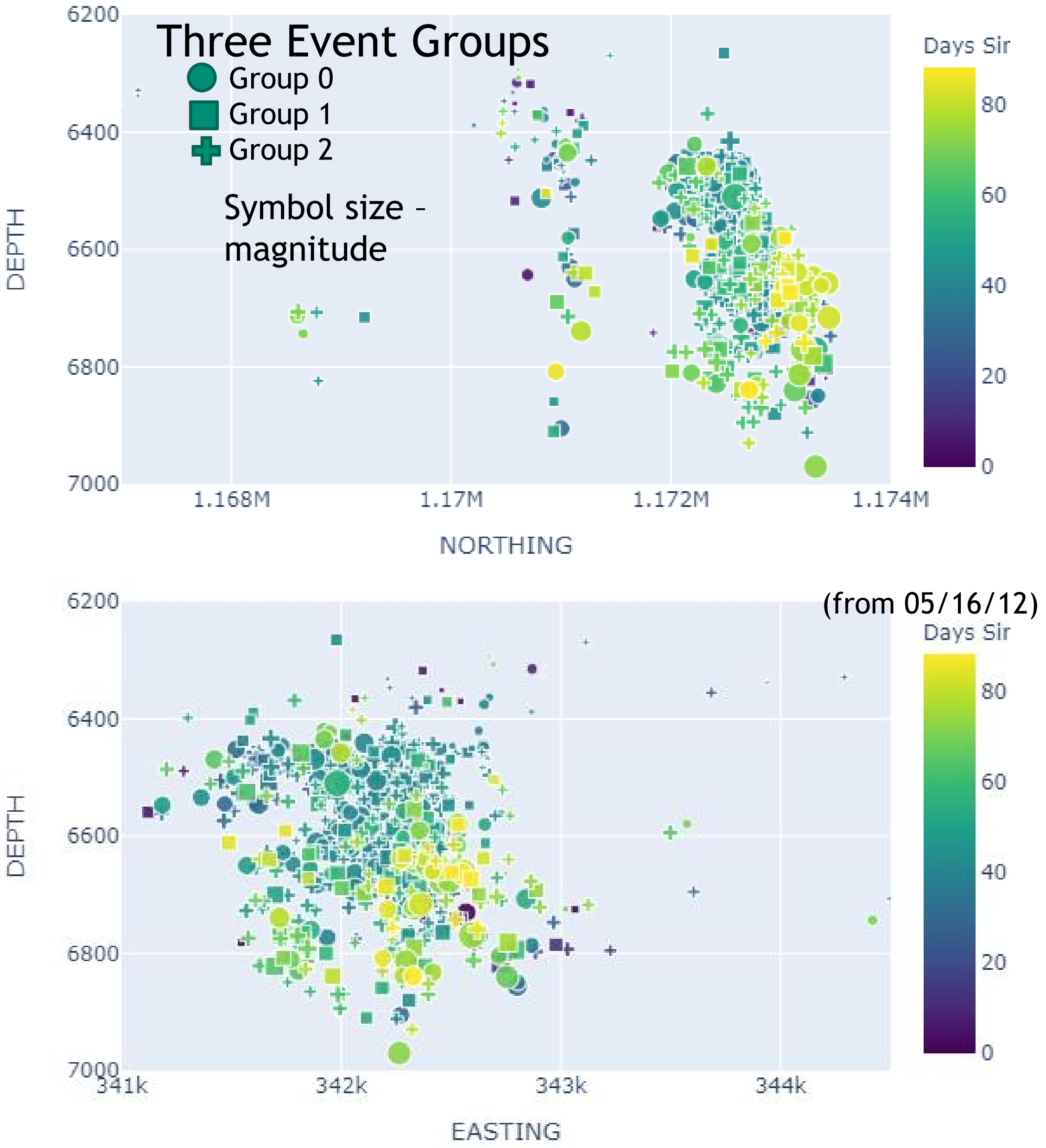
- Located events with processed data of 3 channels over a 2 second window



Distribution of Located Events



Unsupervised machine learning – fingerprint based clustering (cluster #4)



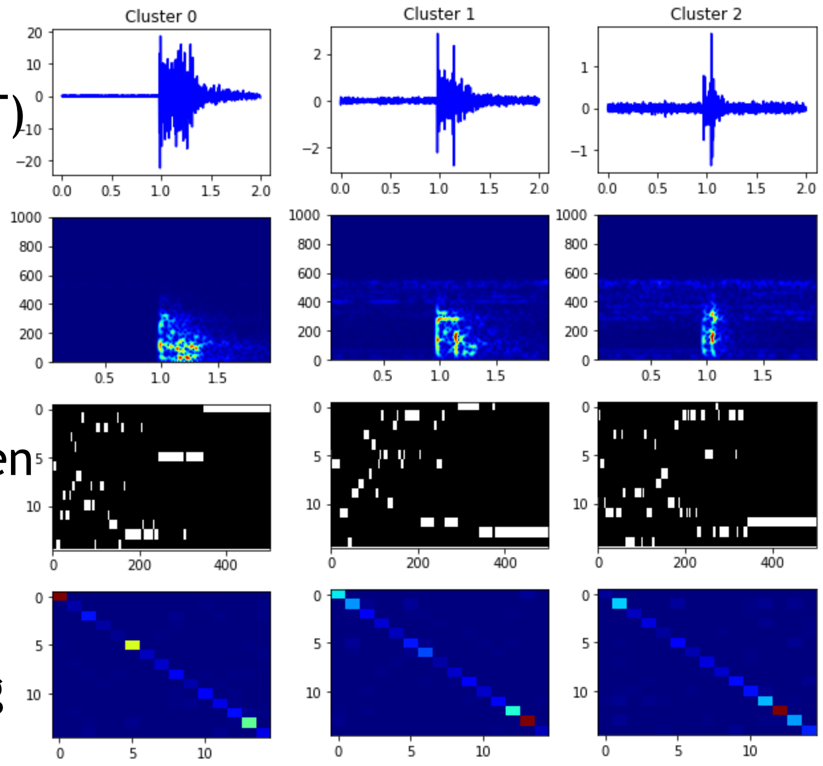
Three groups

1) Waveform (Bandpass filter, STFT)

2) Spectrogram (NMF-> HMM)

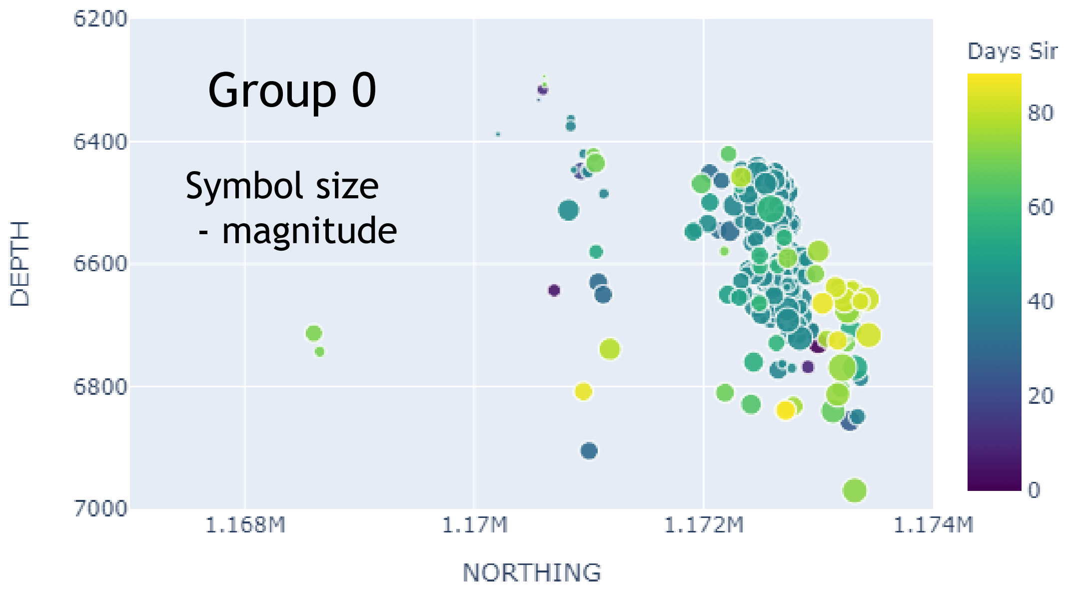
3) Transition probabilities of Hidden Markov State

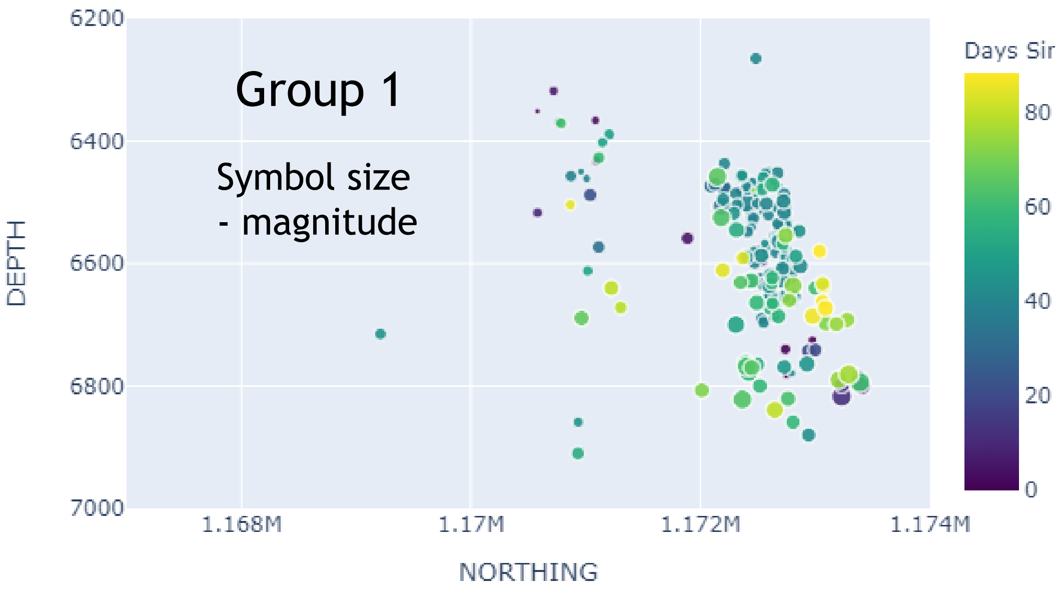
4) Fingerprint map -> k-means clustering (grouping)

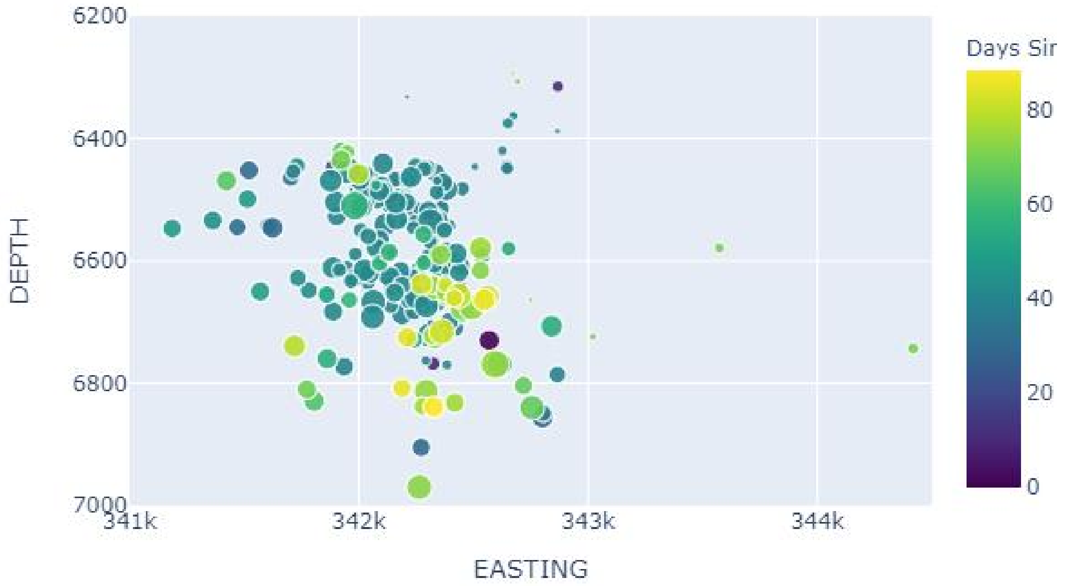


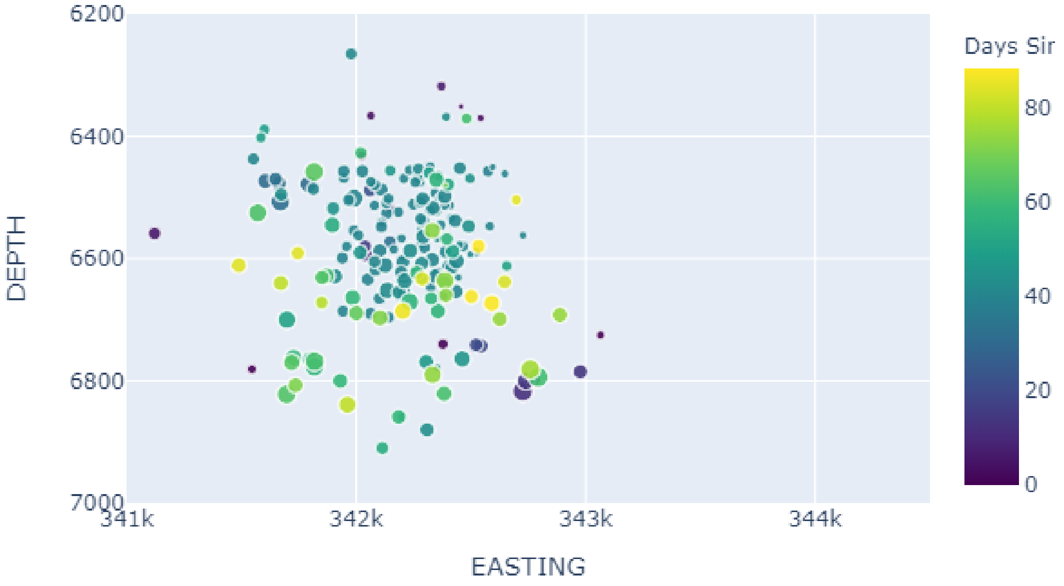
- Group 0: high signal to noise ratio
  transition probabilities from one high state to one low state
- Group 1: low to intermediate signal to noise ratio
  intermediate change in transition probabilities
- Group 2: low signal to noise ratio
  high fluctuation in transition probabilities

Spatio-temporal Evolution of Event Groups

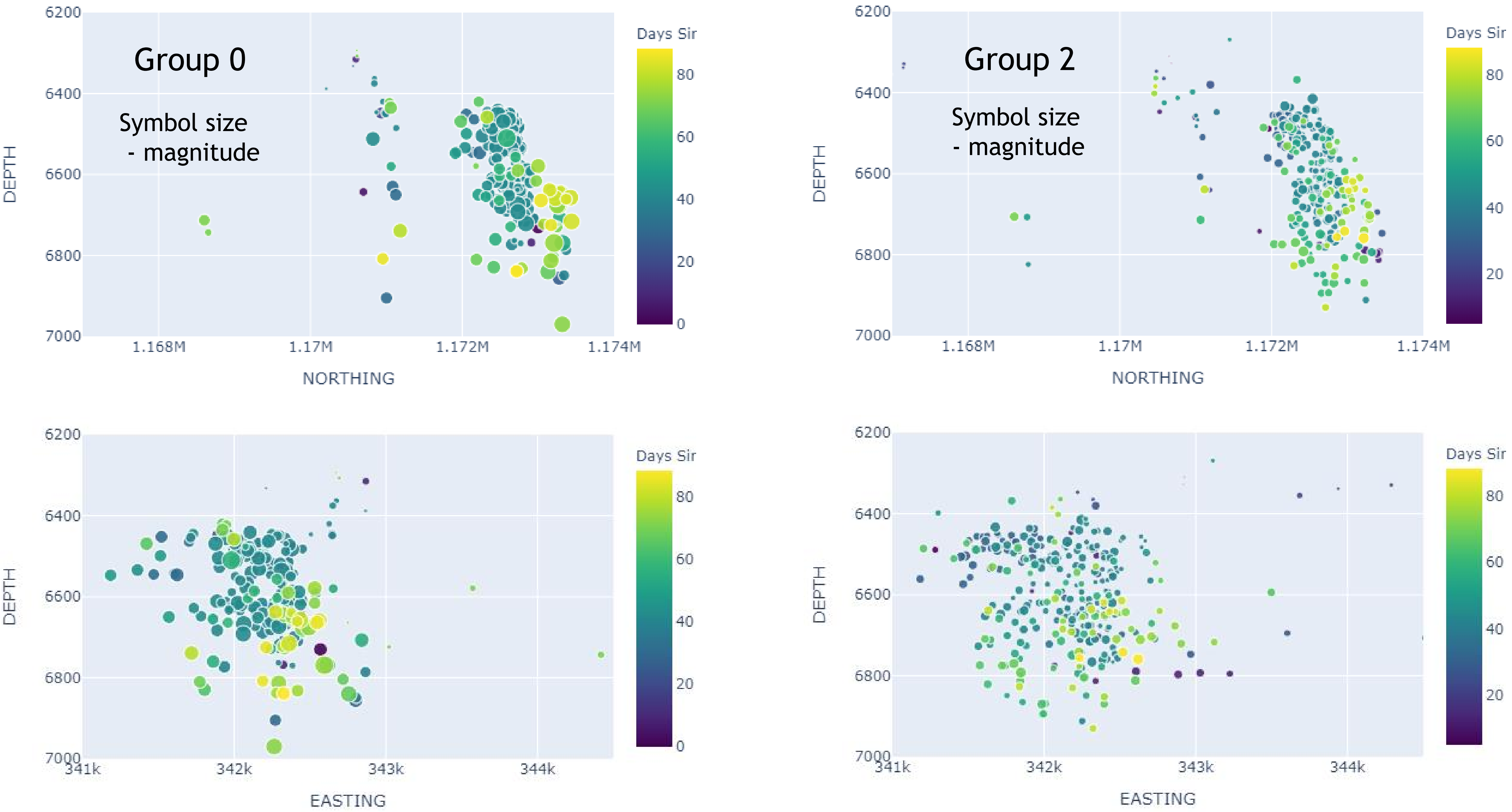




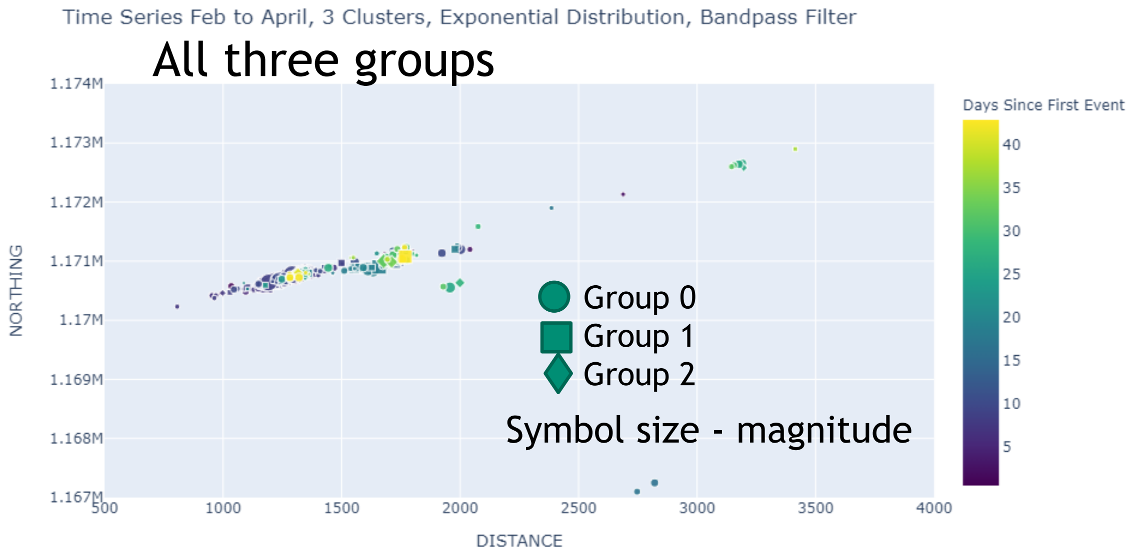


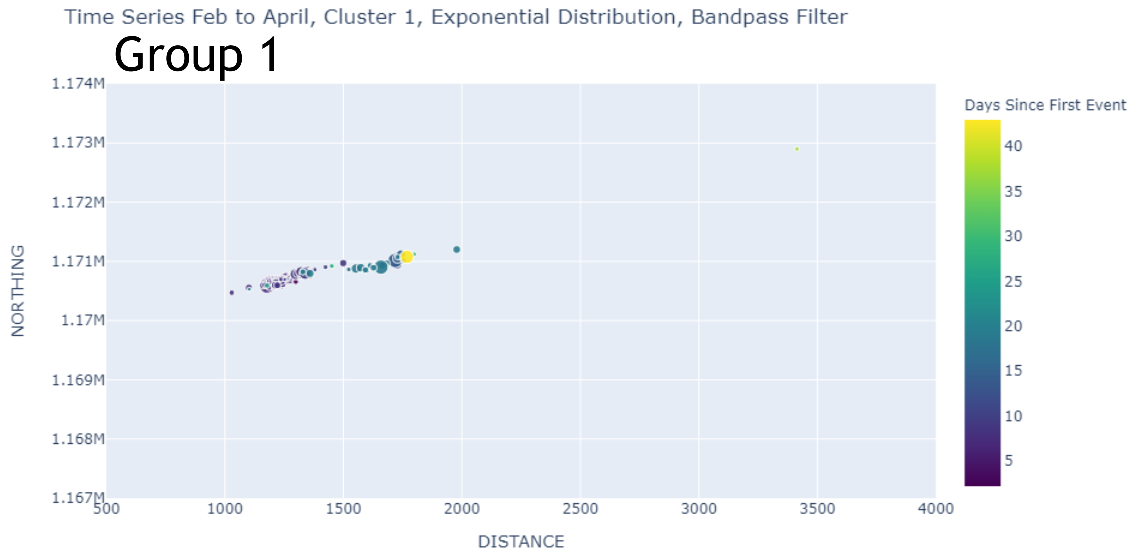


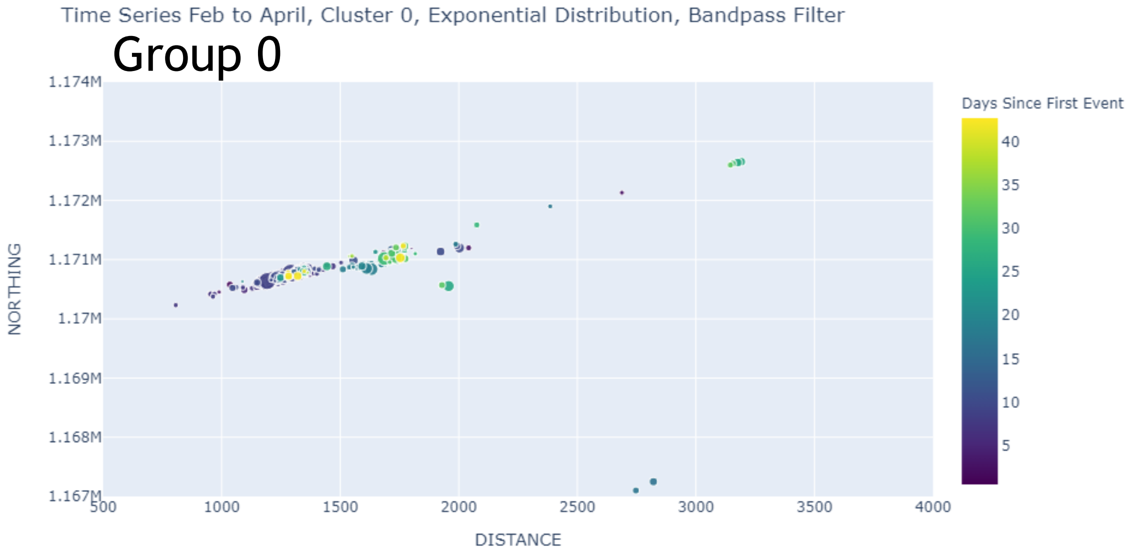
Spatio-Temporal Evolution of Event Groups

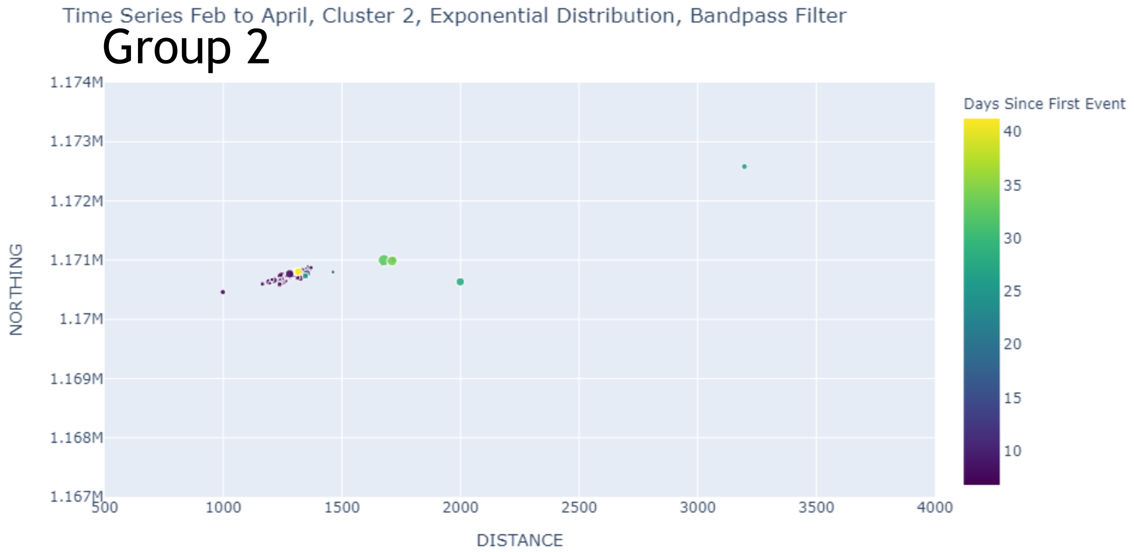


Spatio-Temporal Evolution of Event Groups

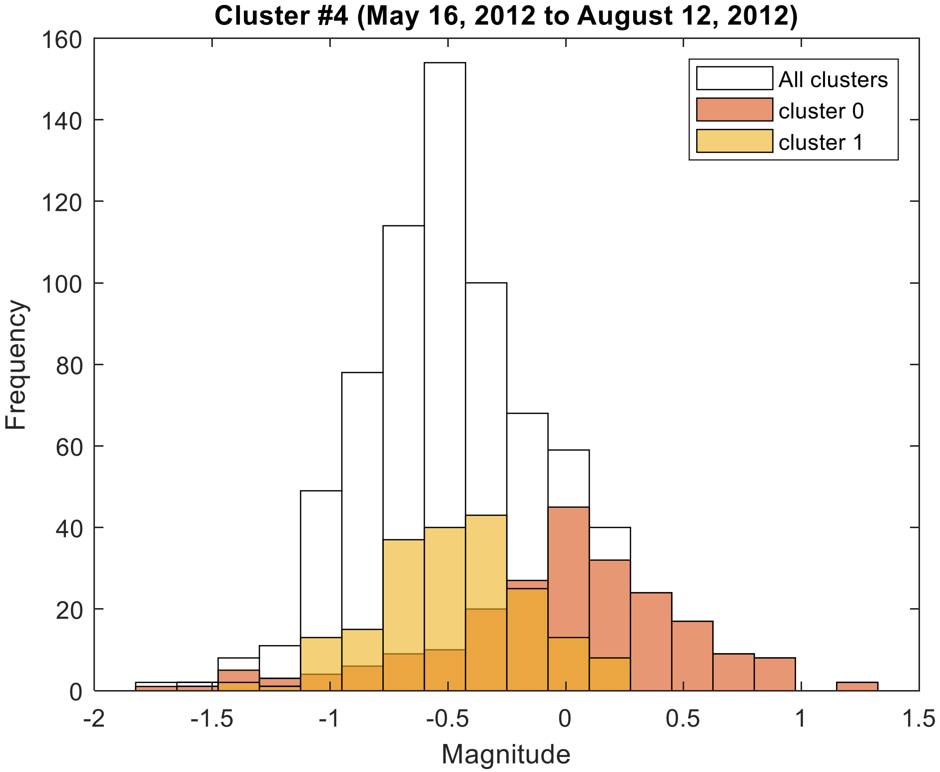


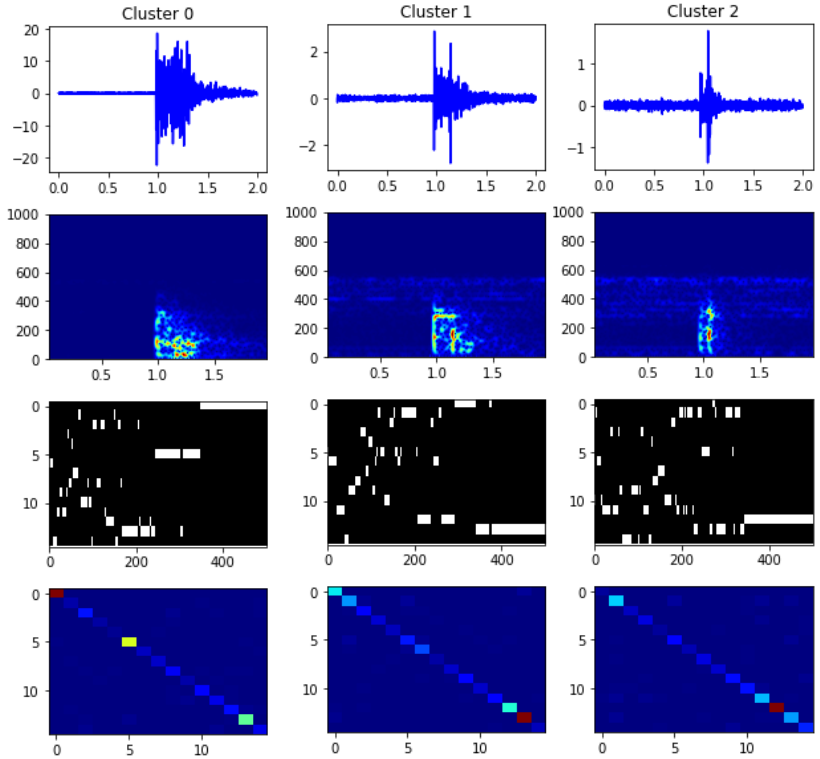






Unsupervised machine learning – fingerprint based clustering results



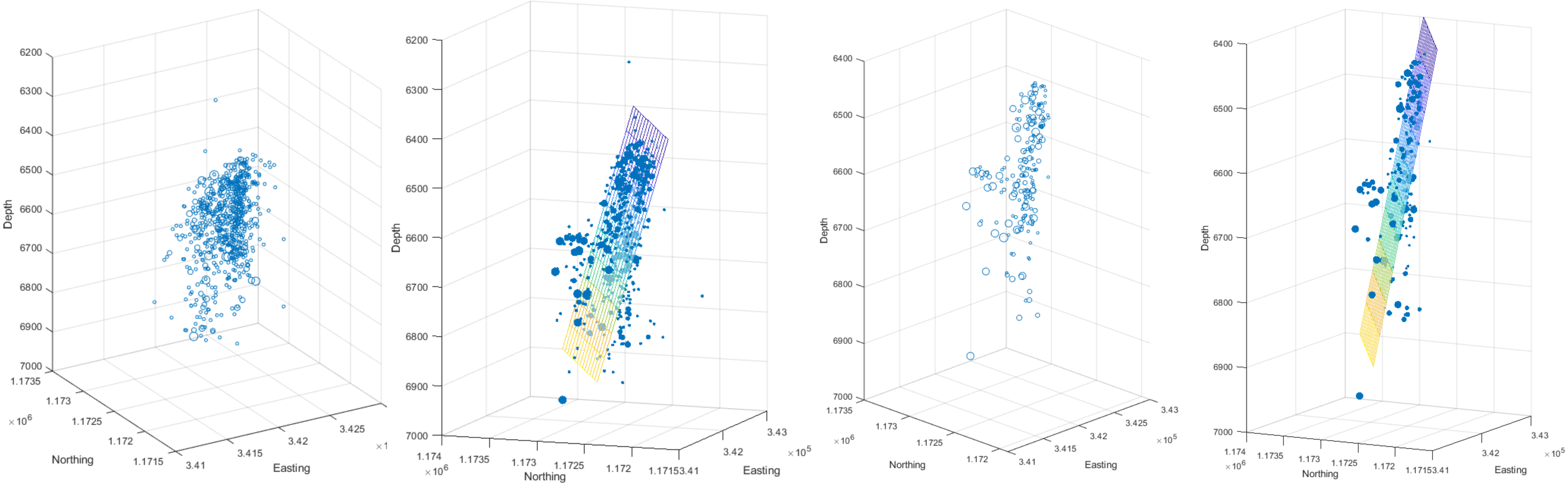


Magnitude

	All events	Group 0	Group 1	Group 2
Mean	-0.42	0.05	-0.48	-0.65
Std	0.46	0.46	0.32	0.25

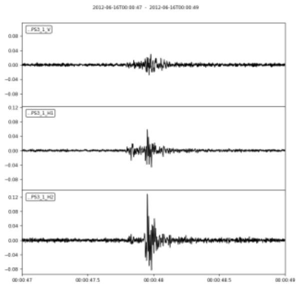
- Group 0: Dominantly high magnitude events
- Group 1: Intermediate magnitude events
- Group 2: Low magnitude events

Unsupervised machine learning – fingerprint based clustering (cluster #4)



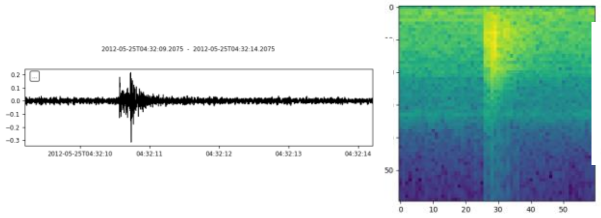
- Motivations
- Machine learning applications at laboratory scale
- Unsupervised machine learning for microseismic data at a field scale CO_2 injection site
- **Supervised machine learning for microseismic data**
- Summary

Supervised machine learning – Event detection using CNN

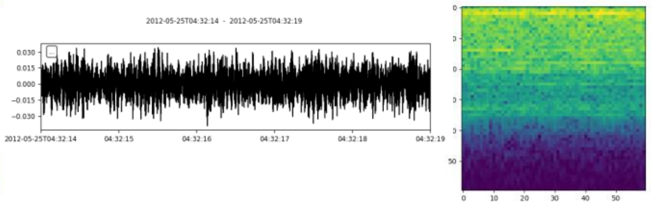


Input 3 channels waveform

Event: Waveform-Spectrogram



Noise: Waveform-Spectrogram



- Input Data:
  - Three-channel (Z,E,N) waveform data
  - 684 events samples
  - Located events cataloged for Feb to April, 2012
  - ~15300 noise data
- Data Processing:
  - Bandpass filter (10 – 400 Hz)
  - Waveform to spectrogram in frequency
  - Rescaled spectrogram with log transformation
- Windows for classification:
  - Continuous waveform data into daily or 8 hrs streams
  - 1 second moving windows for event detection
- Training/validation/testing sets
- Dataset augmentation:
  - Avoid overfitting with deep classifiers
  - Generate additional event windows by shifting 2 second windows to locate signals at varying locations within 2 second windows
- Trained ML model used to detect events for continuous waveform data from Feb to March

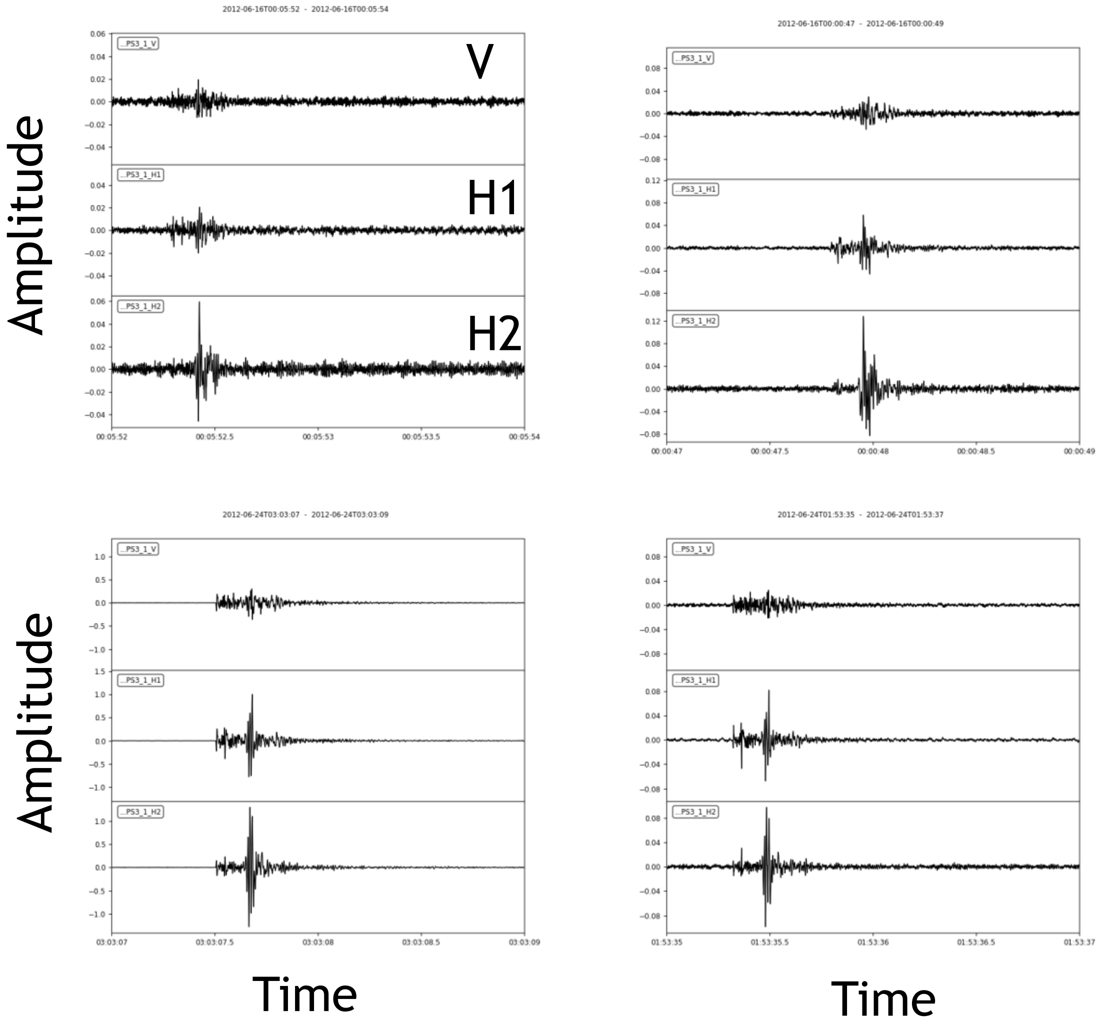
- Convolutional neural network:
  - Input (3 channel with 2D spectrogram image)
  - Four conv2d layers
  - Batch normalization and max pooling
  - Two dense layers (100 & 2)
  - Binary crossentropy classification (event or noise)
  - GPU with cuda 10.1

Supervised machine learning – Results of Event detection

Date	Events in Catalog	Events Predicted	Newly Detected Events
Feb 27-29, 2012	451	1341	766
Mar 1-2, 2012	85	1102	991
Mar 3-8, 2012	23	91	56
Mar 9-12, 2012	53	130	69

- Event detection:
  - 1,882 new events were identified (684 catalog events)
  - A high number of events were detected during a short time period (e.g., Mar 1-2 followed by Feb27-29)
  - Many of newly detected events are real events that are missed in the original catalog

Examples of newly detected events



Summary

- ► Laboratory data and numerical simulations are used to generate for testing machine learning algorithms

- ► Both unsupervised and supervised machine learning methods were applied for clustering of event waveforms and event detection

- ► This study demonstrates that ML can be a great tool for scaling and transfer learning (e.g., validated approach with lab scale data for field scale data)

- ► Integration of multiphysics (geomechanical and geophysical approaches) and multiple tools (controlled experiments, simulations, machine learning) for sensing and data analysis will be able to advance both fundamental and practical study of a wide range of geoscience applications.

Thank You!