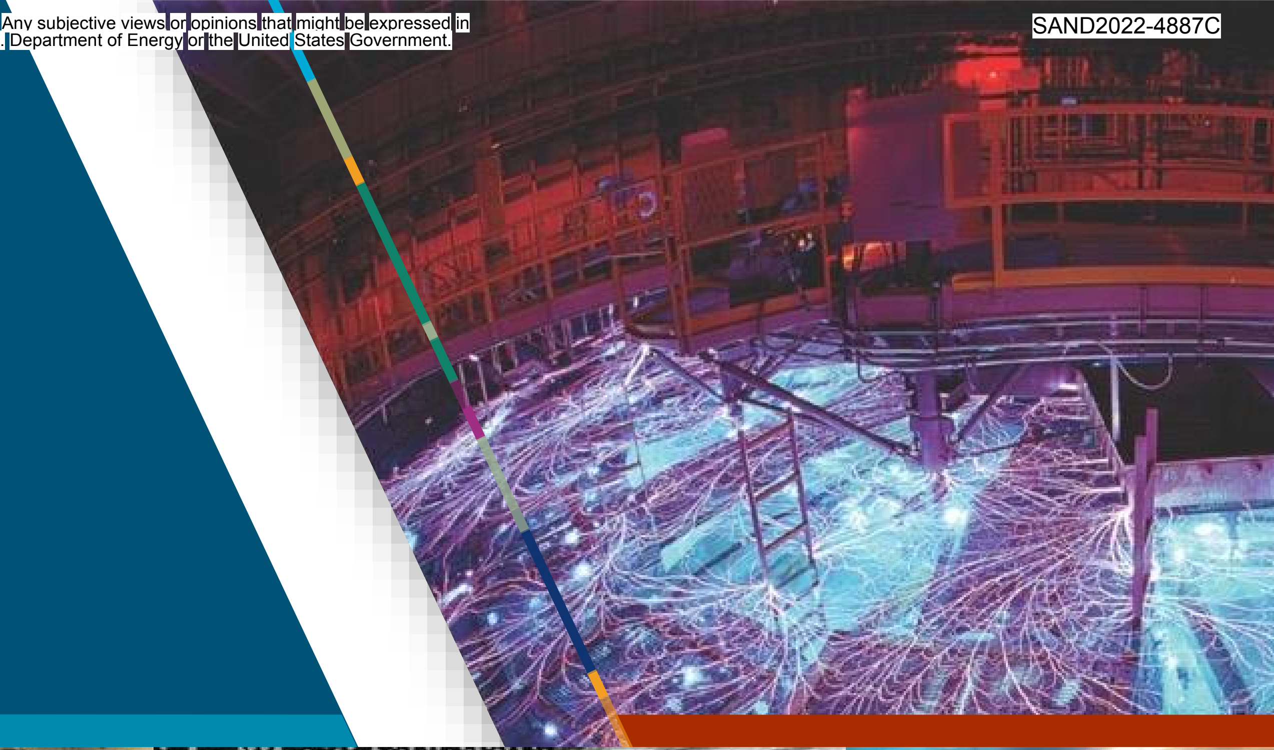
SAND2022-4887C



Sandia National Laboratories

Exceptional service in the national interest

Machine Learning Applications for Induced Seismic Data Analysis at Illinois Basin Decatur Project Site

SSA 2022

Hongkyu Yoon, Daniel Lizama[1], Rachel Willis[2]
Geomechanics Department
Sandia National Laboratories, NM, USA

[1] U of Puerto Rico, , Mayagüez

Acknowledgments

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

- Laboratory Directed Research and Development program at Sandia National Laboratories

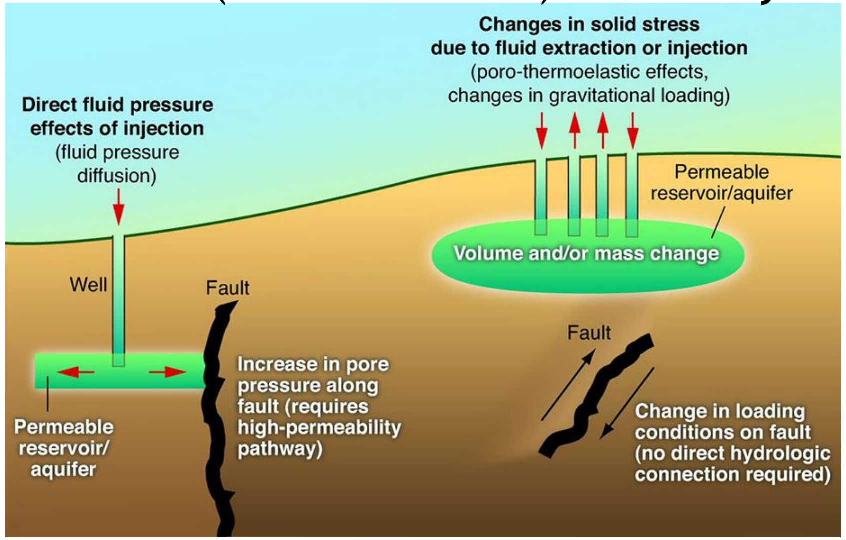
- **Motivations & Illinois Basin Decatur Project (IBDP) data**
- Event detection and phase arrival time estimation
- Fault plane analysis
- Summary

Background & Motivations

♦ Motivations

- Fluid injection or withdrawal causes changes in pore pressure, resulting in induced seismicity (IS) during subsurface energy activities (geologic carbon storage, enhanced geothermal system, wastewater injection, etc.)
- Machine learning (ML) has been successfully developed and applied for data analysis of (micro-)seismic data (e.g., event detection, phase arrival time, source locations)

Induced (human-caused) seismicity



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

♦ Goals

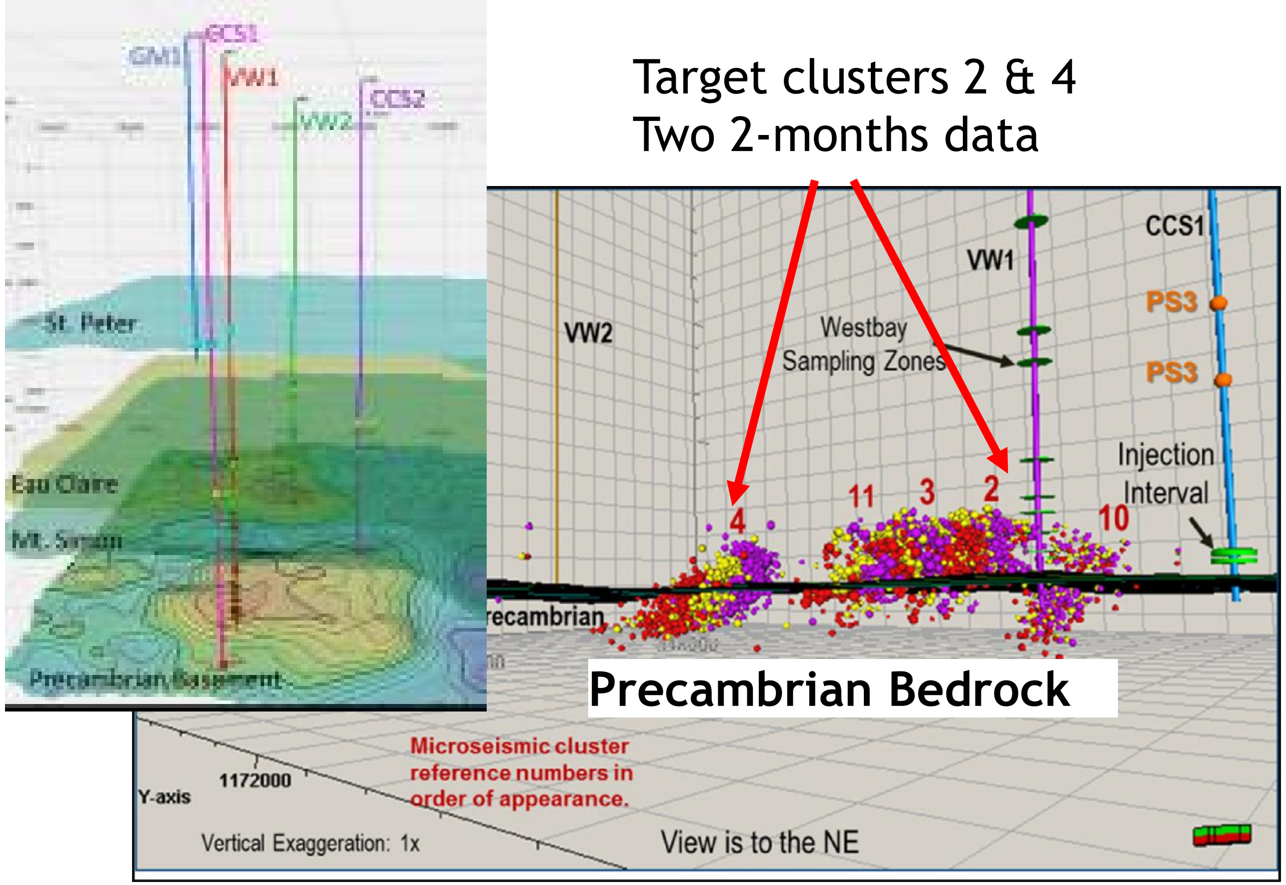
(1) Develop/apply machine-learning techniques for seismic wave data analysis and event detection at Illinois Basin Decatur Project (IBDP) site (geologic carbon storage)

(2) Delineate fracture and failure mechanisms associated with microseismic data

Microseismic data at IBDP

Williams-Stroud et al. (SEG 2019)

Target clusters 2 & 4
Two 2-months data



Note: old (incorrect) located events

Will et al. (IJGGC 2016)

- Illinois Basin Decatur Project (IBDP, 3 yrs): 1 MMT CO_2
- Industrial Carbon Capture & Storage (ICCS, up to 5 yrs): 3-5.5 MMT CO_2
- CarbonSAFE: 50+ MMT CO_2
- Extensive integrated site characterization and monitoring investigations

- Using the initial microseismic data, we aim at improving the detection of low-magnitude, unidentified events & locations to discover undetected/hidden fault/fracture systems
- Characterize microseismic waveforms, the relations among the events, and reliable identification of microseismic sources integrated with forward/inverse modeling

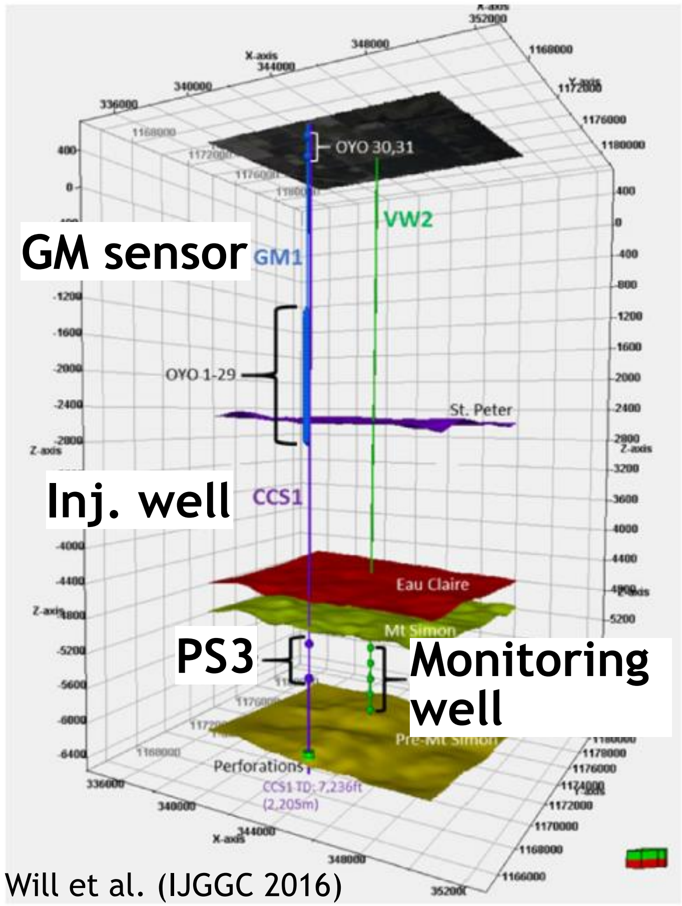
MS Waveform Data at the IBDP Site

Raw (unprocessed) continuous data

- Big data (~ 7TB for 3 months out of a total of 100's TB for 3 yrs)
- 2 kHz sampling rate
- # of traces: 84-94 (inconsistency at an early injection period)
- 4 channel data on two PS3 sensors in injection reservoir formation and 2-3 channels on GM geophones (relatively upper formations)
- Only vertically oriented sensors at an early phase

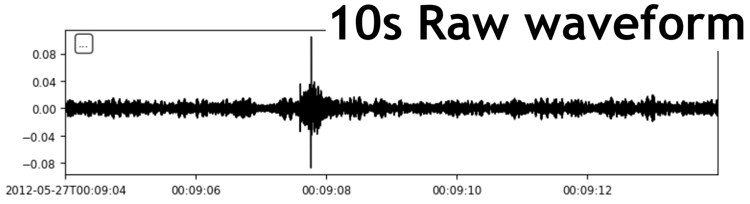
Processed data & catalog (~3 yrs injection)

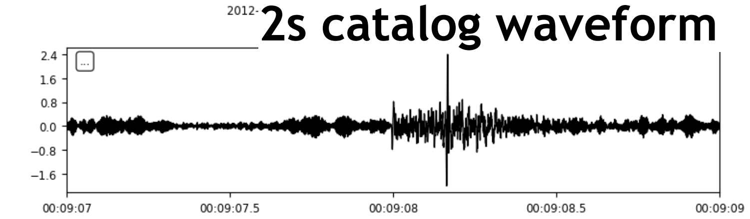
- Detected event (processed 2s window, ~ 19K events, 3 channel (Z,H1,H2)
- A small # of located events (~ 5K events with source locations)
- Relatively low magnitude (mostly <0, max magnitude = ~1.5)
- Processed 2s window data have been shifted from original data (needed to generate event data for machine learning separately)



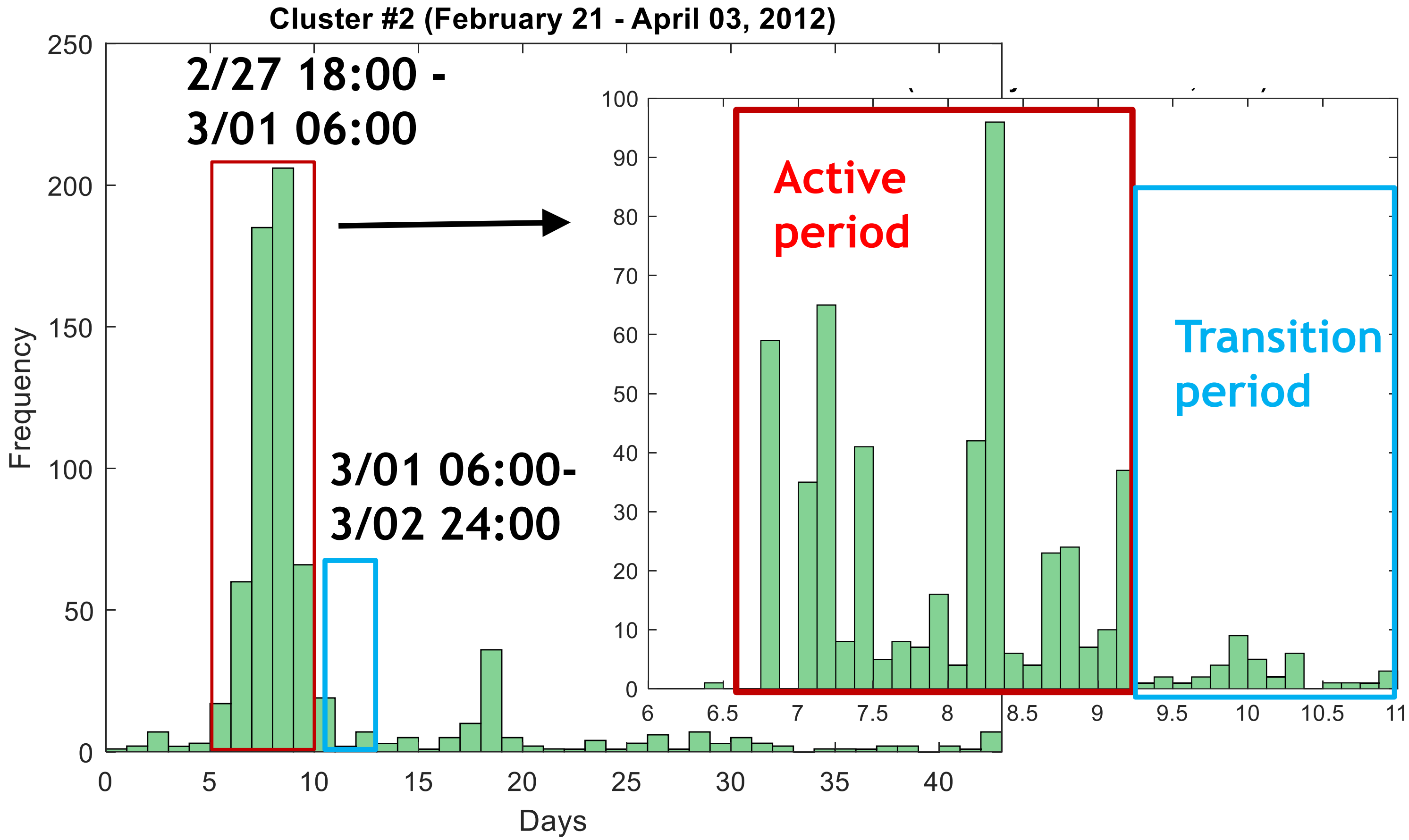
Will et al. (IJGGC 2016)

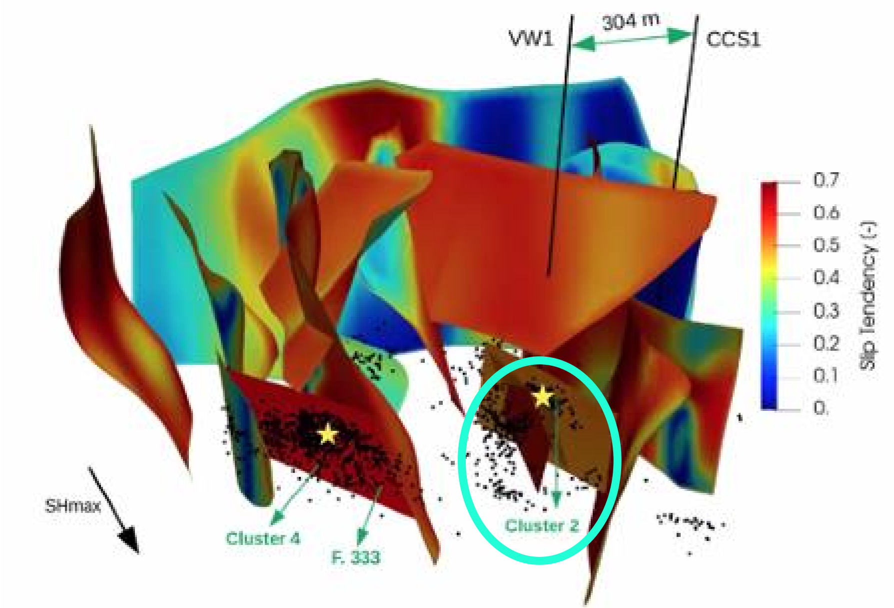
Fig. 1. Subsurface array configuration. Distance units are feet, *Z* axis is referenced to mean sea level.



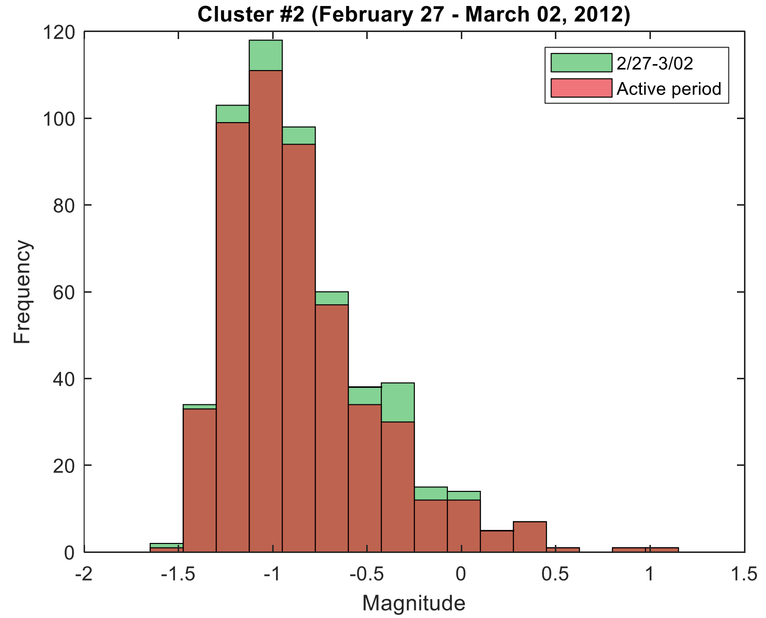


MS Cluster #2 (684 located events)





Josimar Silva et al. (AGU 2021 T22C-02)



- Motivations & Illinois Basin Decatur Project (IBDP) data
- **Event detection and phase arrival time estimation**
- Fault plane analysis
- Summary

Convolutional Neural Networks

Input

$*$

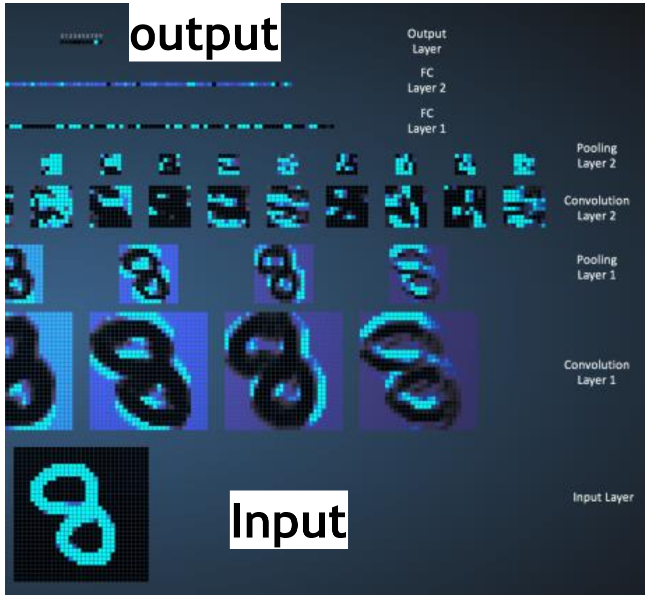
1	0	-1
1	0	-1
1	0	-1

Filter (Kernel)

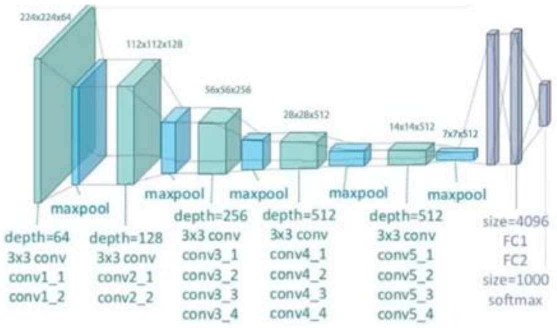
$=$

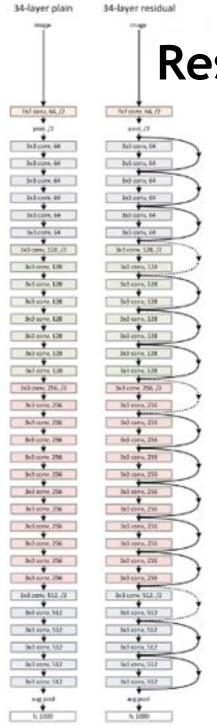
-5	-4	0	8
-10	-2	2	3
0	-2	-4	-7
-3	-2	-3	-16

Feature map



VGG19

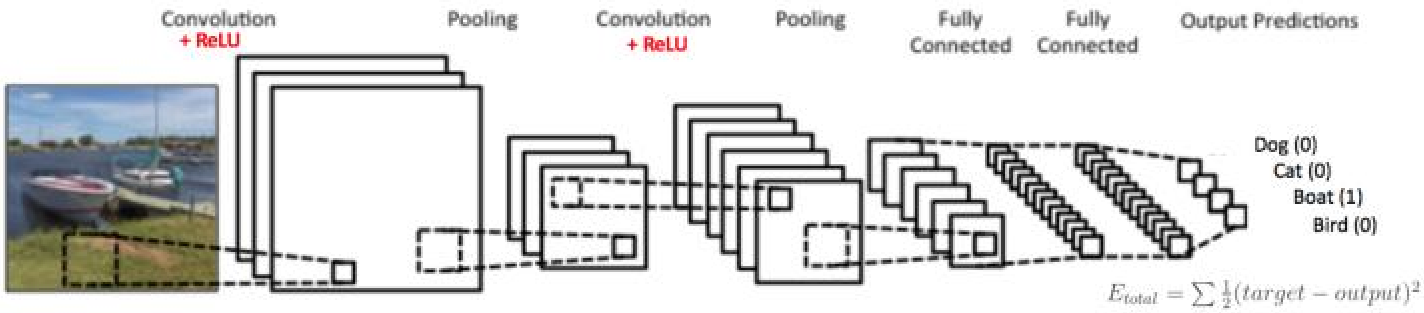




ResNet

Feature Extraction

Classification



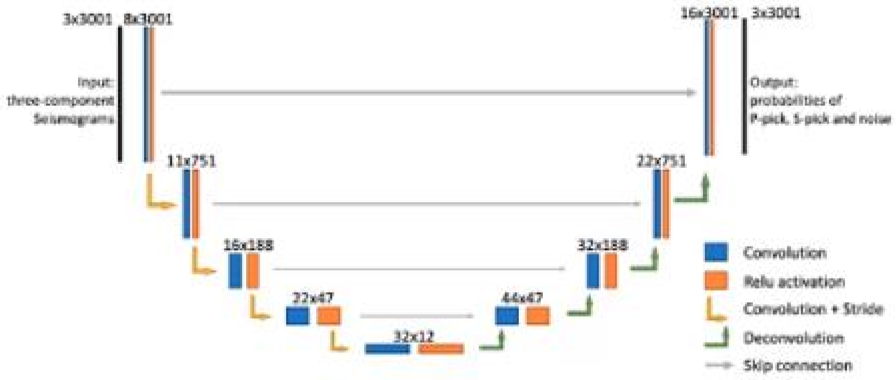
Input
3 RGB channels,
Identical to 3 channel (Z,H1,H2) waveform

Convolution + Pooling layers act as Feature Extractors from the input image, while fully Connected layer acts as a classifier.

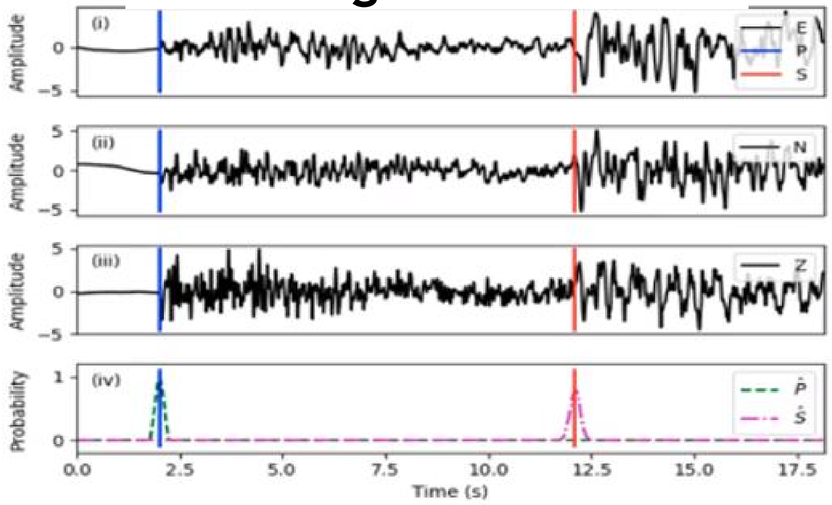
http://cs231n.github.io/convolutional-networks/
https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

Recent Deep Learning Models for Seismic Data

Phase Picking (PhaseNet)



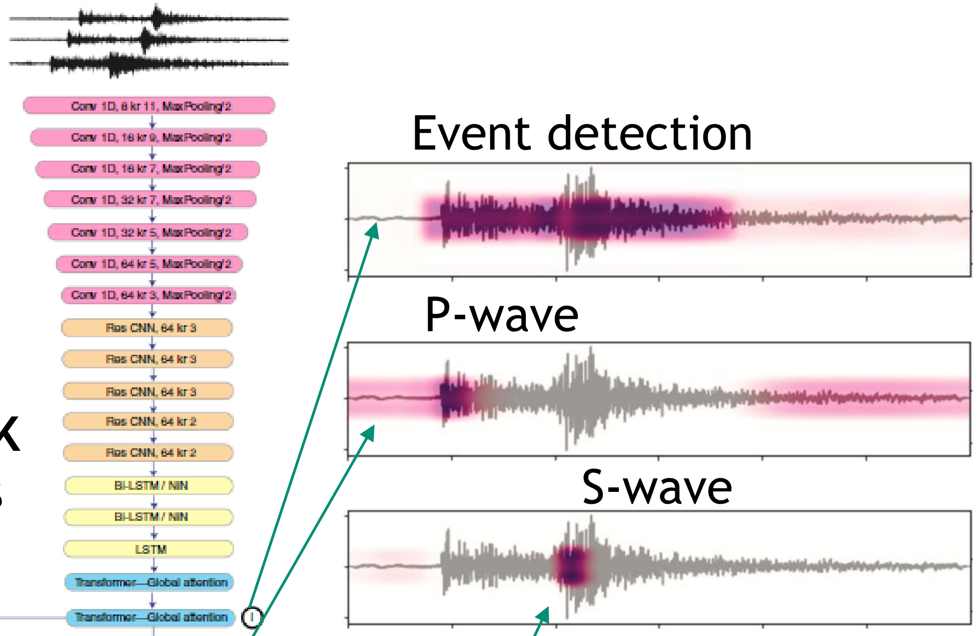
Increasing Pick Precision

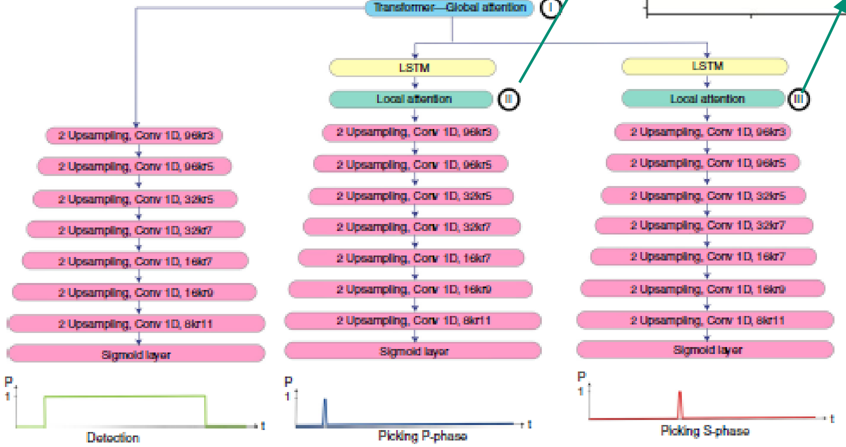


Zhu & Beroza (GJI, 2018)

EQ Transformer

Training data:
1 M EQ and 300 K noise waveforms





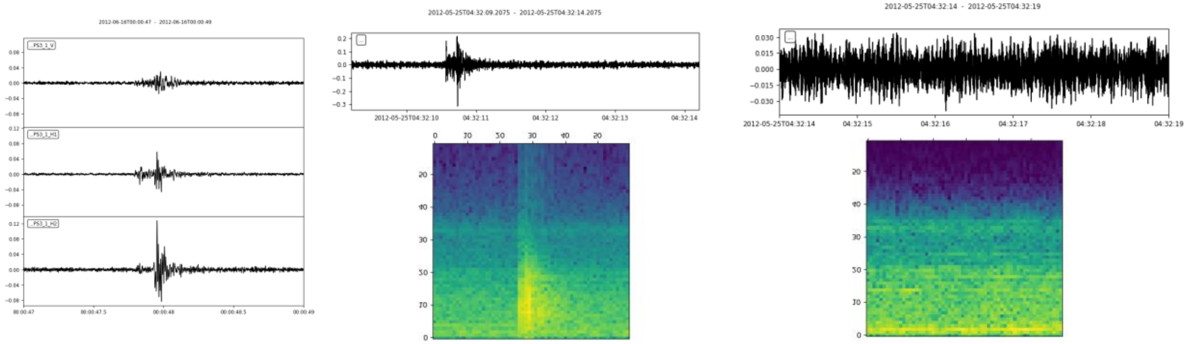
Mousavi et al. (Nat Comm., 2020)

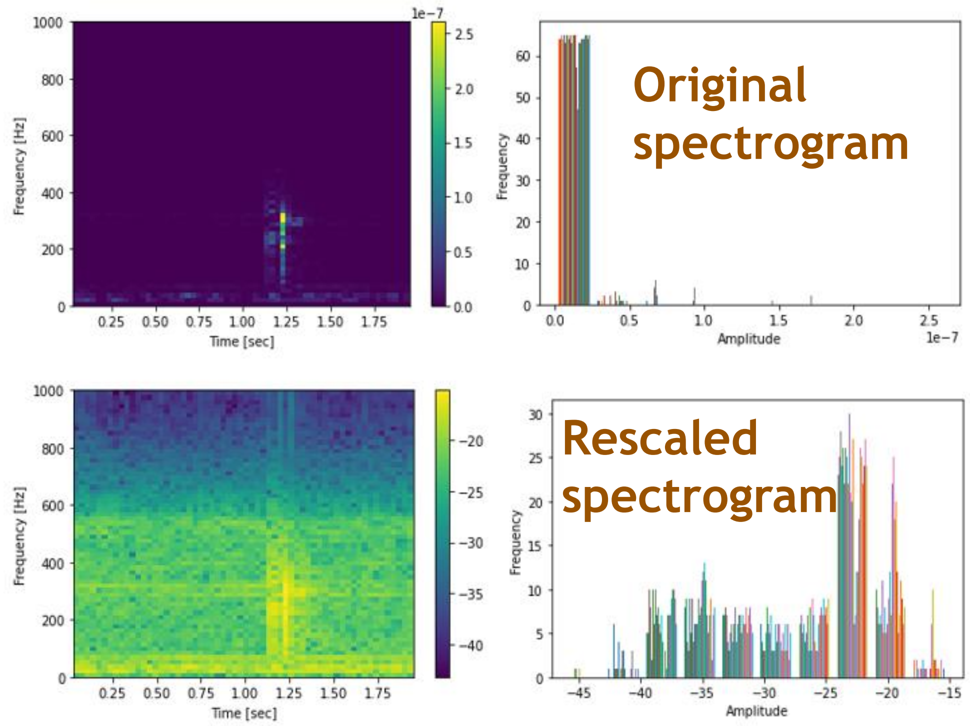
Supervised machine learning – Event detection using CNN

Small data

- Input Data:
  - Three-channel (Z,E,N) waveform data
  - 684 located 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
- Event detection:
  - Continuous waveform data: 1 s moving windows
- Training/validation/testing sets
- Dataset augmentation:
  - Generate additional event windows by shifting 2 sec window to locate signals at varying locations within 2 second window

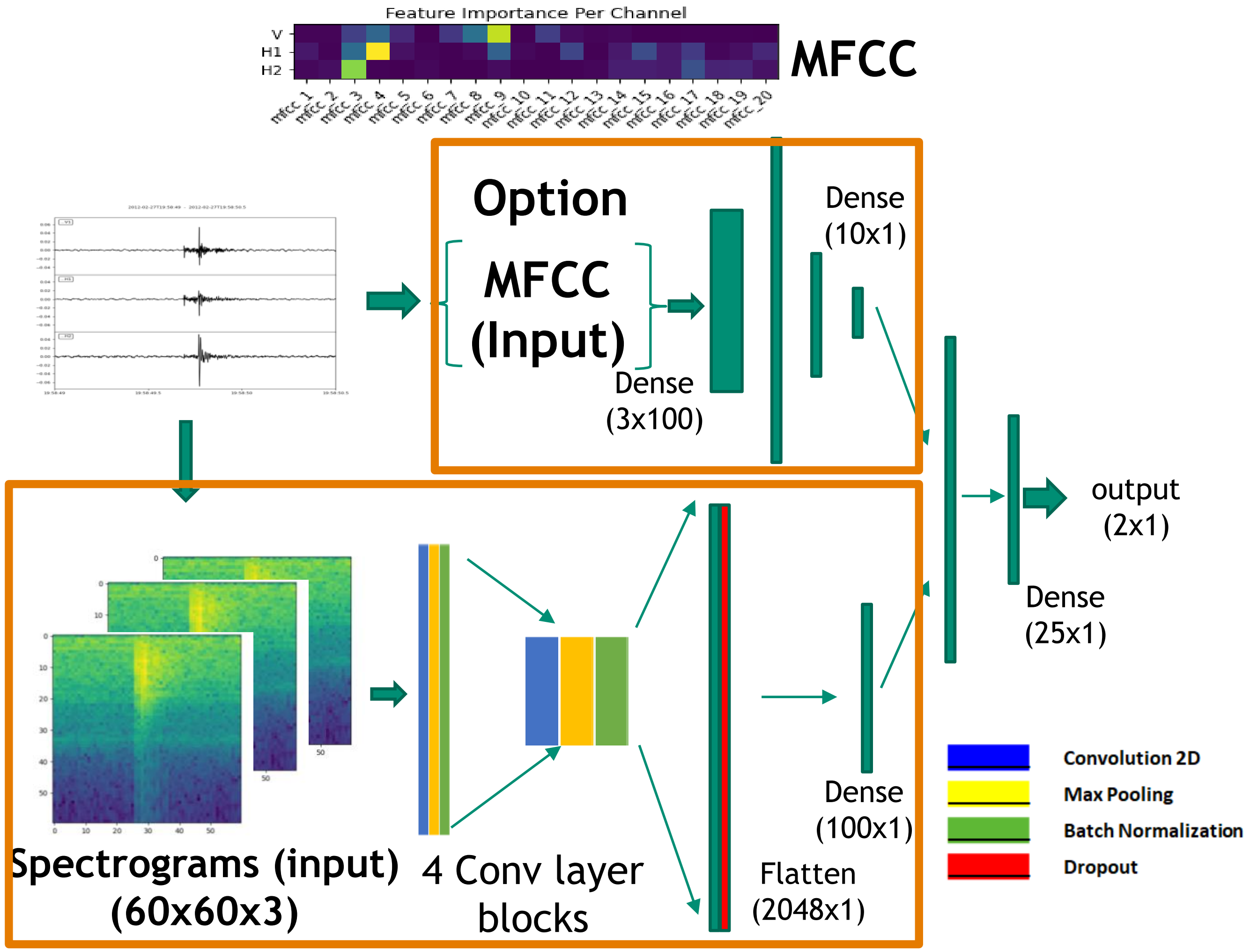
Input 3 channels | Event | Noise





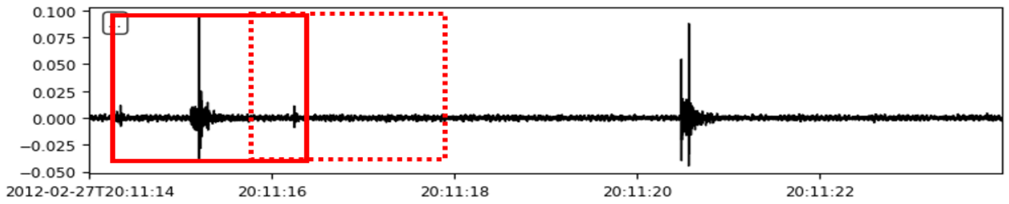
CNN Architecture for Event Detection

Feature vector from Random Forest
on 3 channel data

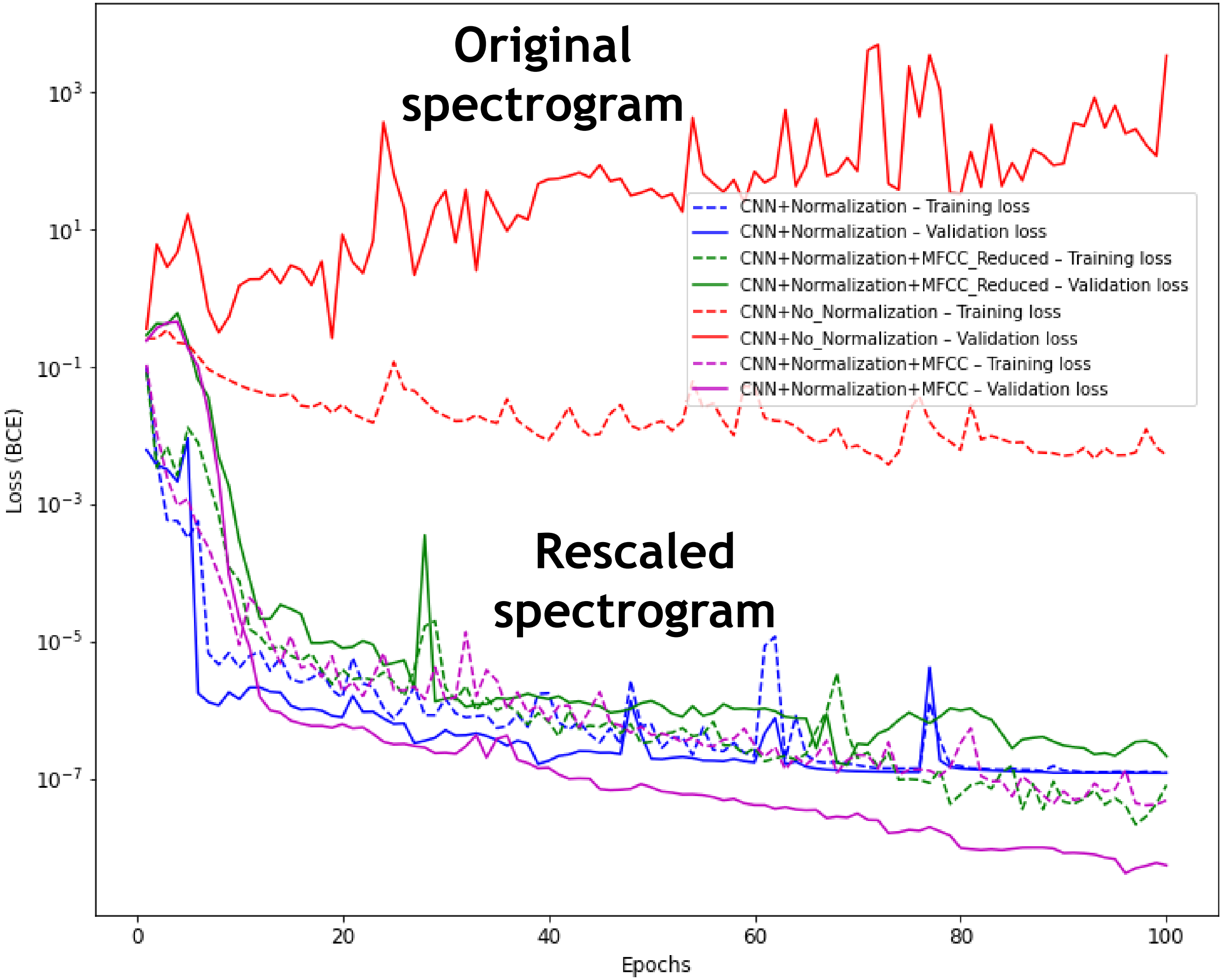


CNN Block

- Input Data:
  - Rescaled spectrogram with log transformation
  - Mel-Frequency Cepstrum Coefficients (MFCC)
- CNN architecture:
  - Simple (good for small training data)
  - MFCC input can be used as physical constraint (Physics-constrained ML framework)
- Model training:
  - The best model based on validation data
- Trained model:
  - detect events for continuous waveform data from Feb to March in 2012 (cluster #2)
  - 1 second moving window

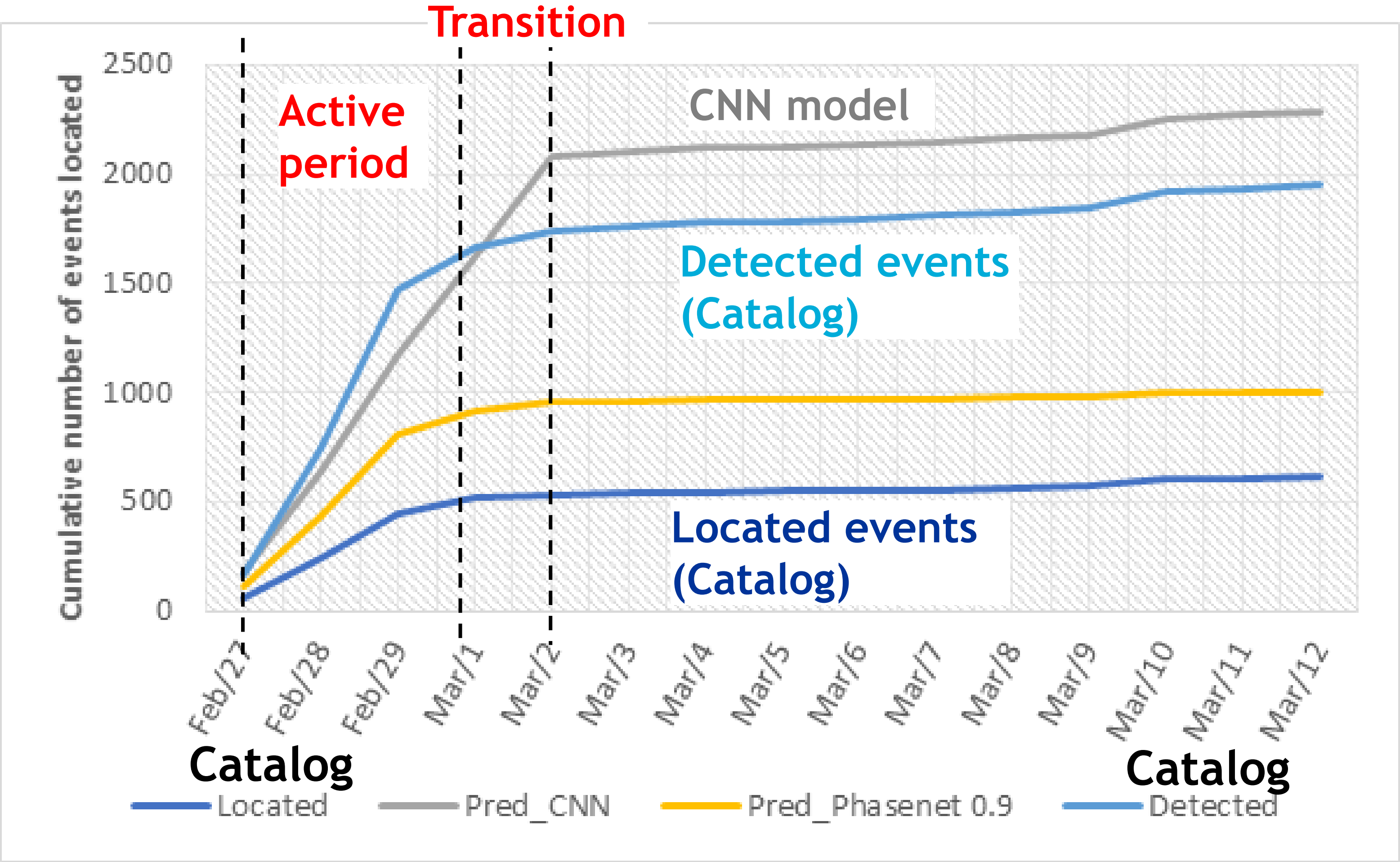


CNN Model Training



- ML models with rescaled spectrogram input dramatically improved model accuracy compared to ML model with original spectrogram
- Training time is super-fast (~15 min on a laptop with one GPU) due to a small CNN architecture (EQTransformer: O(89) hrs using 4 Tesla V100 GPUs)
- CNN only tends to reach a plateau (no more learning) early (epochs = 40-50)
- CNN + full MFCC seems to learn more continuously over 100 epochs
- In this work we used CNN only for event detection

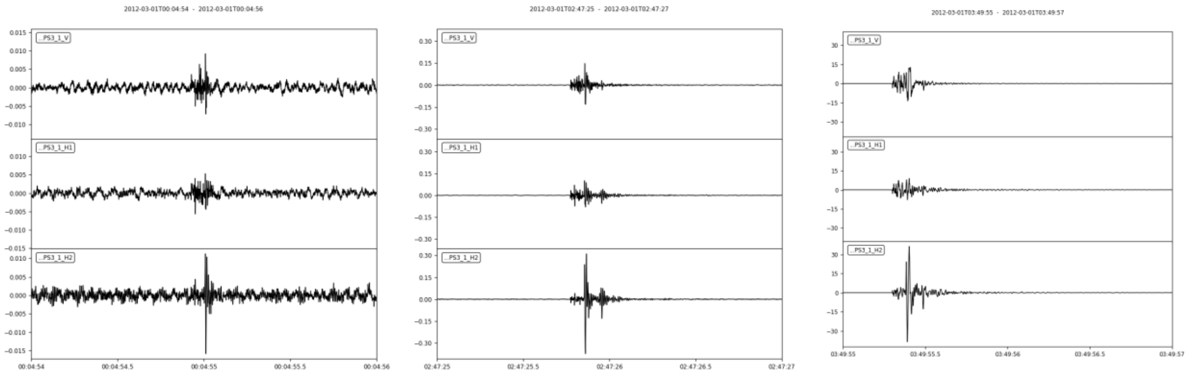
Event Detection



- CNN model tends to pick events more accurately than detected events in catalog
- CNN model detects more events after active event period (02/27/2012-02/29/2012)
- Due to relatively small # of labelled data CNN model performs very well for event detection

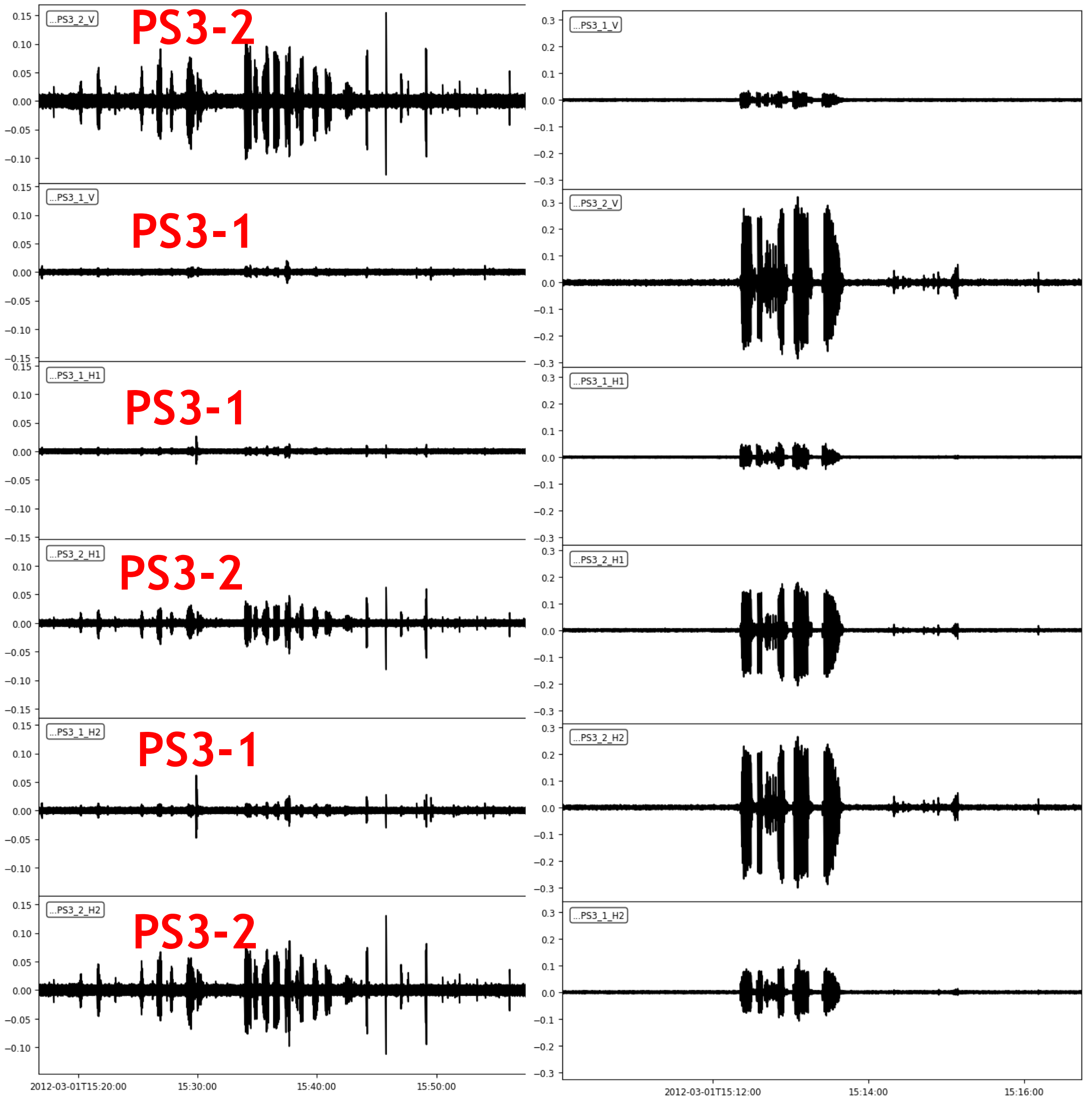
Waveform characteristics: Active vs. transition periods (March 01-02)

Active period (Feb27-29): 2s window events



Transition period

45min window / 17 min window



Long-period long-duration (LPLD) seismic events

- Represent slow shear slip (e.g., hydraulic fracturing)
- Observed in the literature (e.g., Das and Zoback, 2013) where natural fracture density is high, likely caused by high pore pressure and/or high clay contents (i.e., low permeability) => slow slipping
- Tend to be observed “only on faults large enough to produce a sequence of slow slip events”
- This observation needs to be used to parameterize the thickness of fault zone in inverse modeling

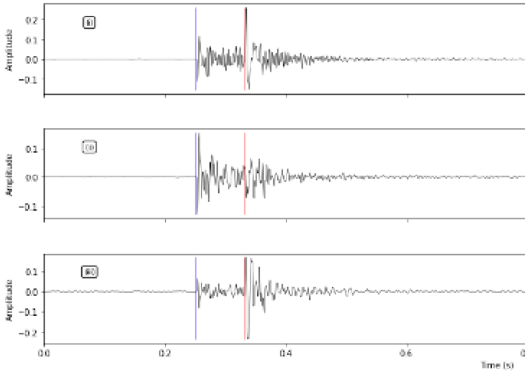
Phase Arrival Time

Training data for arrival times & Transfer learning of PhaseNet

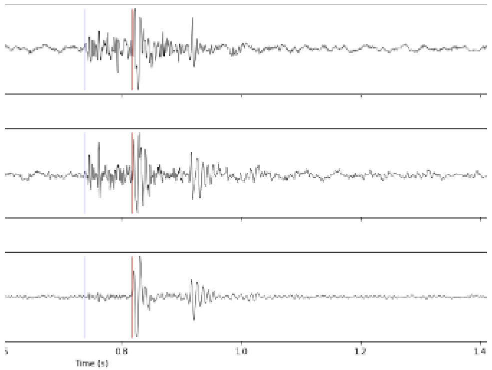
- Arrival time data in Catalog are different from event times of continuous waveform data
- PhasePAPy (Chen & Holland, 2016): P-arrival pick based on AICD
- AR pick (obspy): S-arrival pick based on autoregression-AIC
- These picking results are the best to match manual picking of arrival times of continuous waveform
- From automatic picks, ~80% (419) of Feb-Mar dataset was considered as correct picks and used to re-train the **PhaseNet model**
- A part of the remaining 20% was corrected manually for model validation (mean loss = ~0.02)
- Validation accuracy: P (0.906) and S (0.942)

PhaseNet

P-wave S-wave

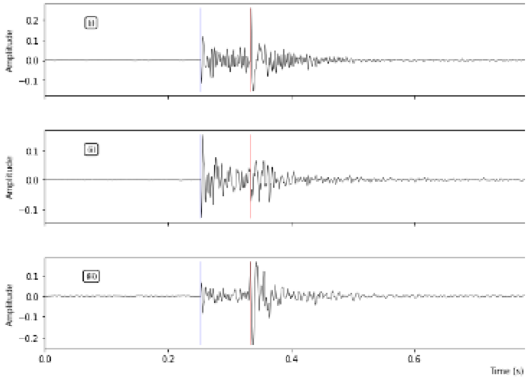


P-wave S-wave

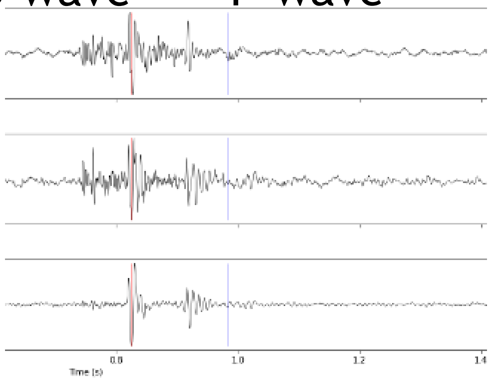


AR picker & PhasePiPy

P-wave S-wave Good

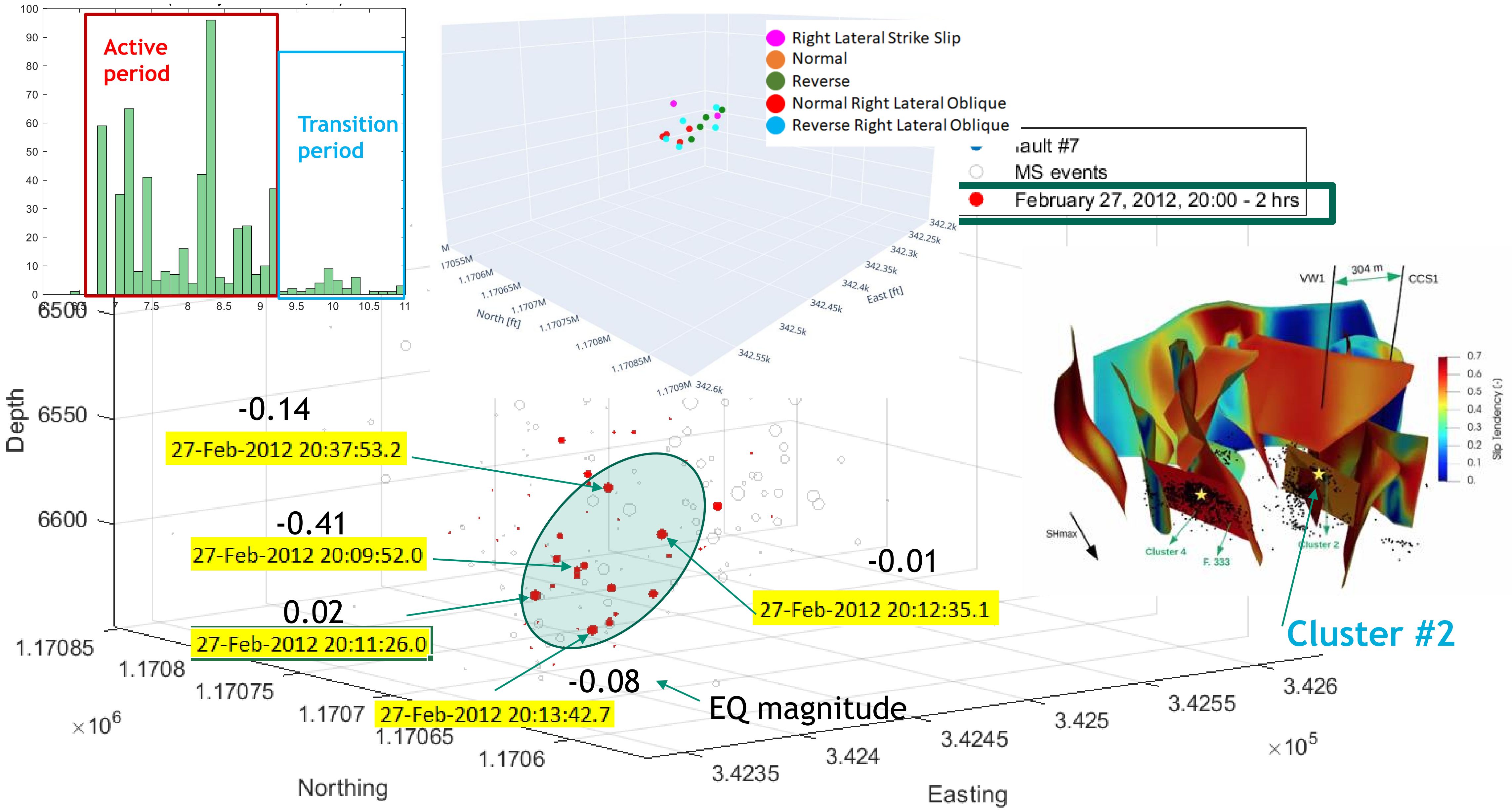


S-wave P-wave Bad

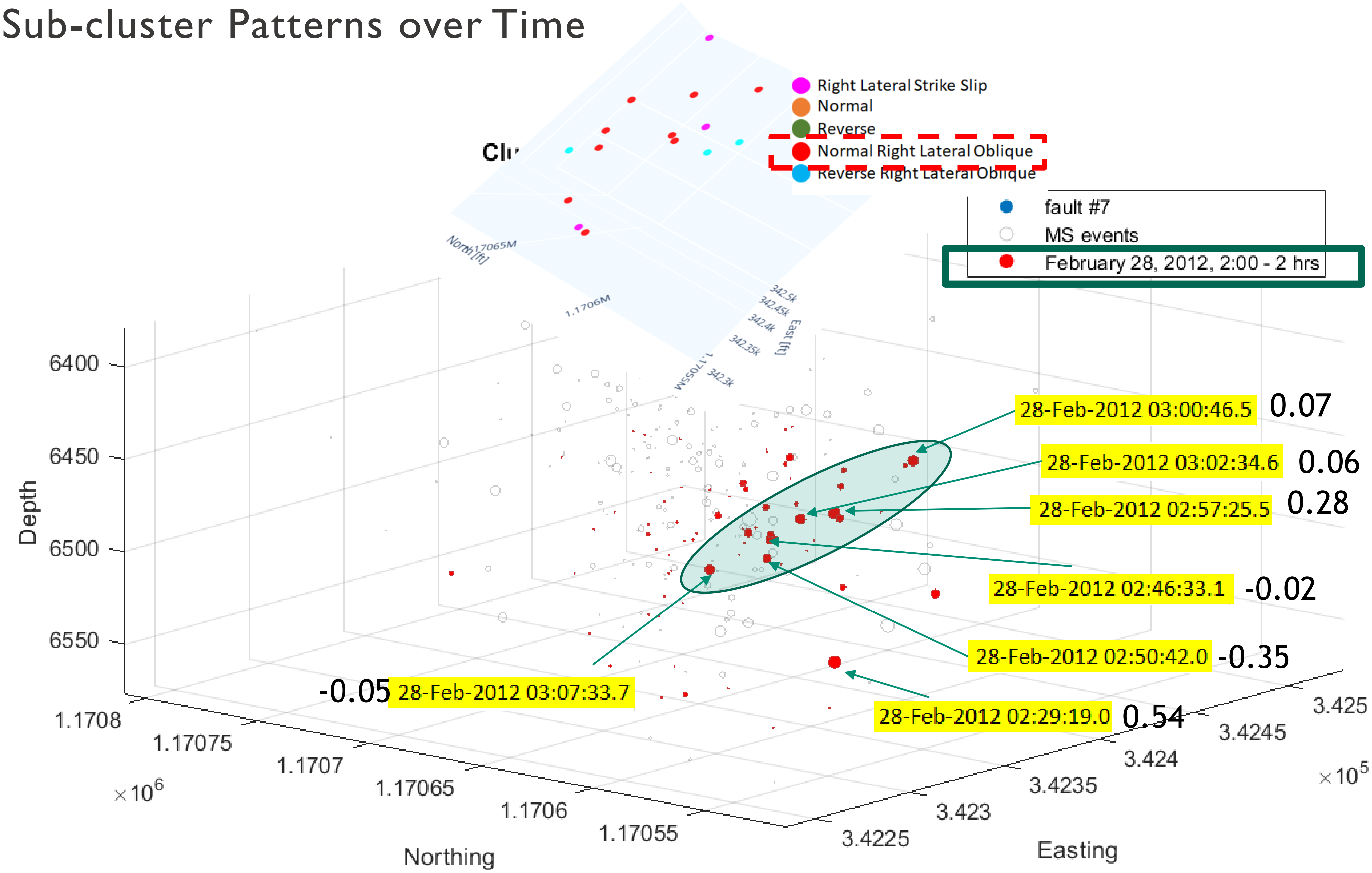


- Motivations & Illinois Basin Decatur Project (IBDP) data
- Event detection and phase arrival time estimation
- **Fault plane analysis**
- **Summary**

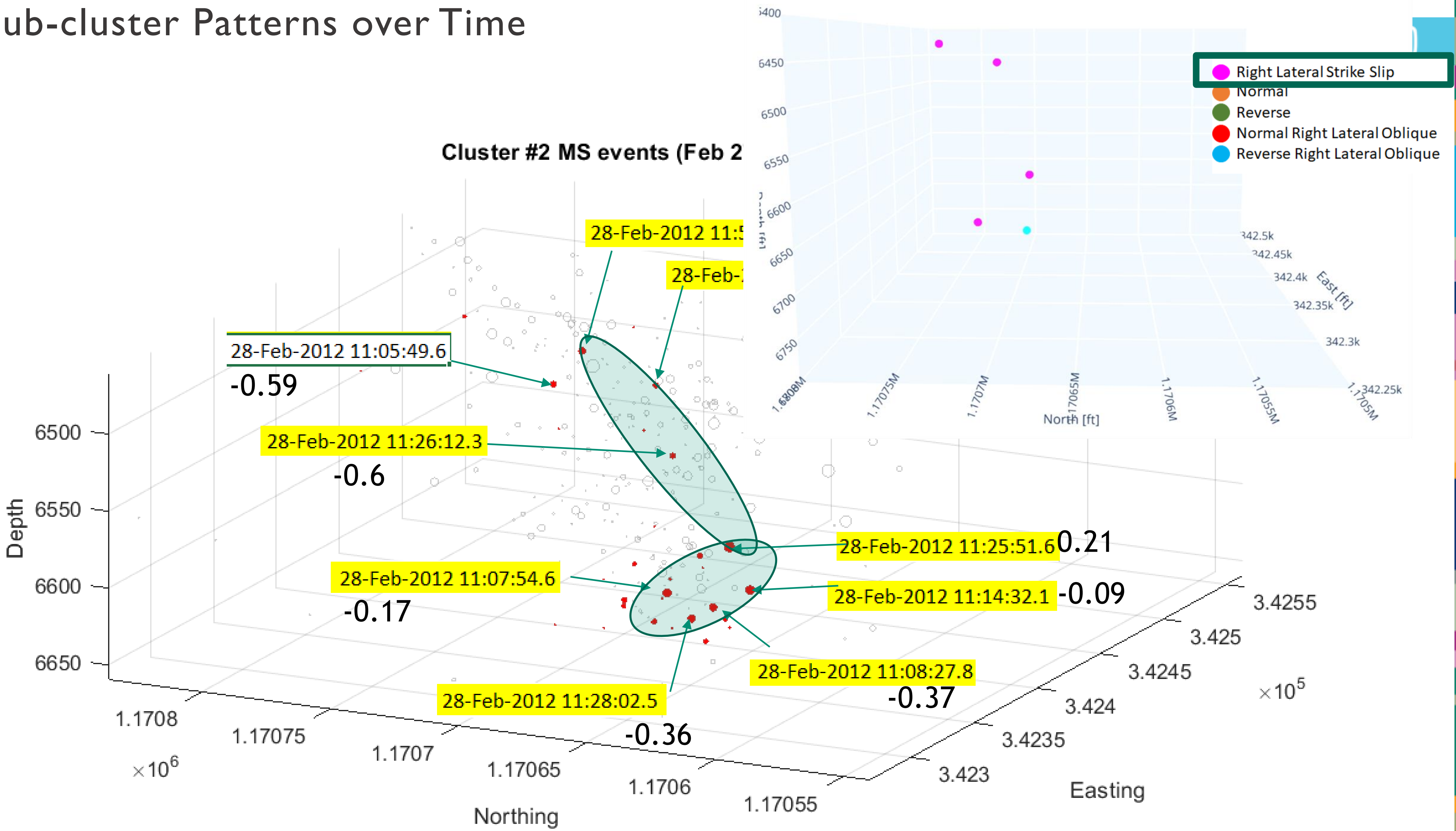
Sub-cluster Patterns over Time & Focal Mechanism Analysis using USGS HASH



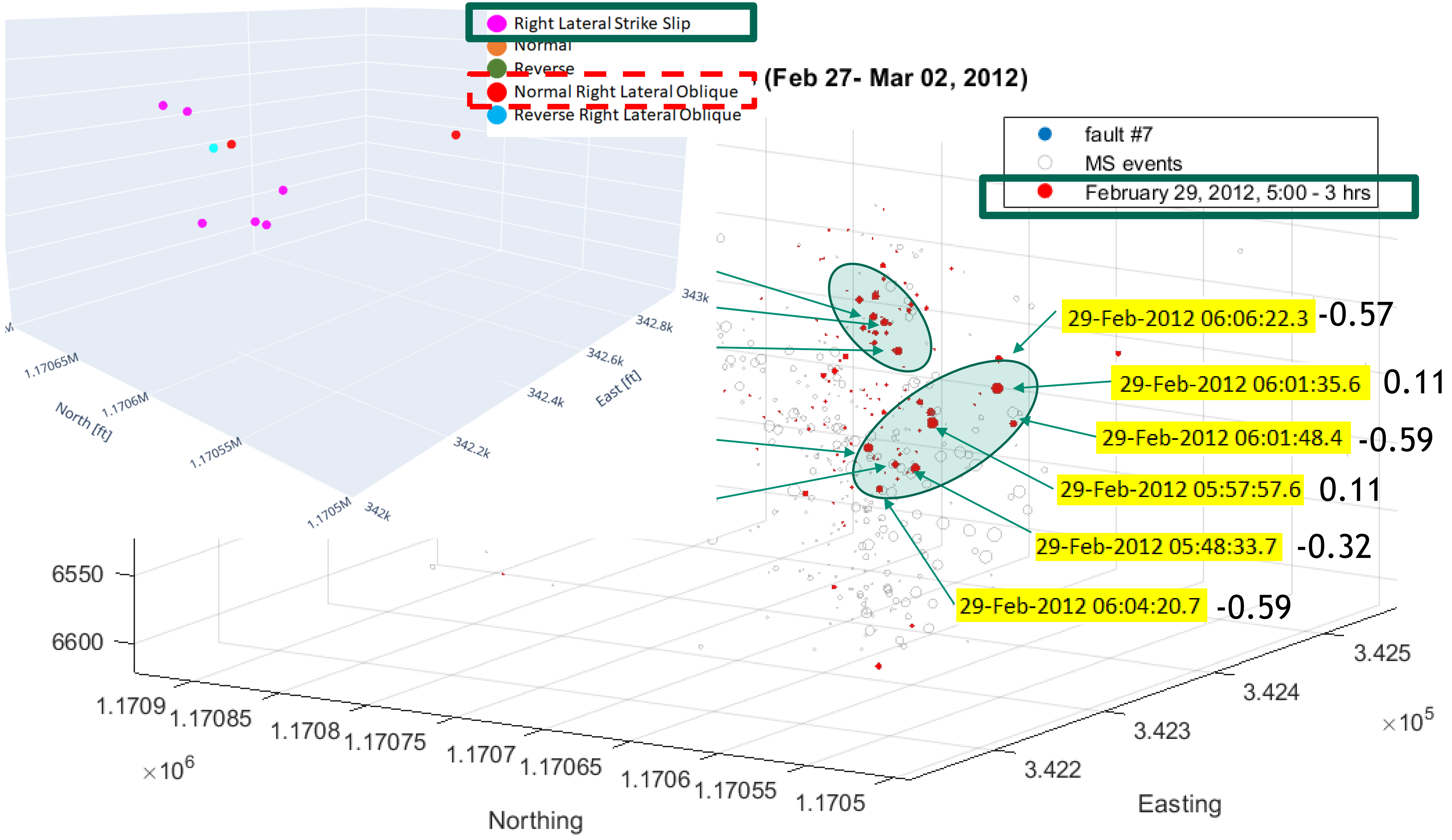
Sub-cluster Patterns over Time



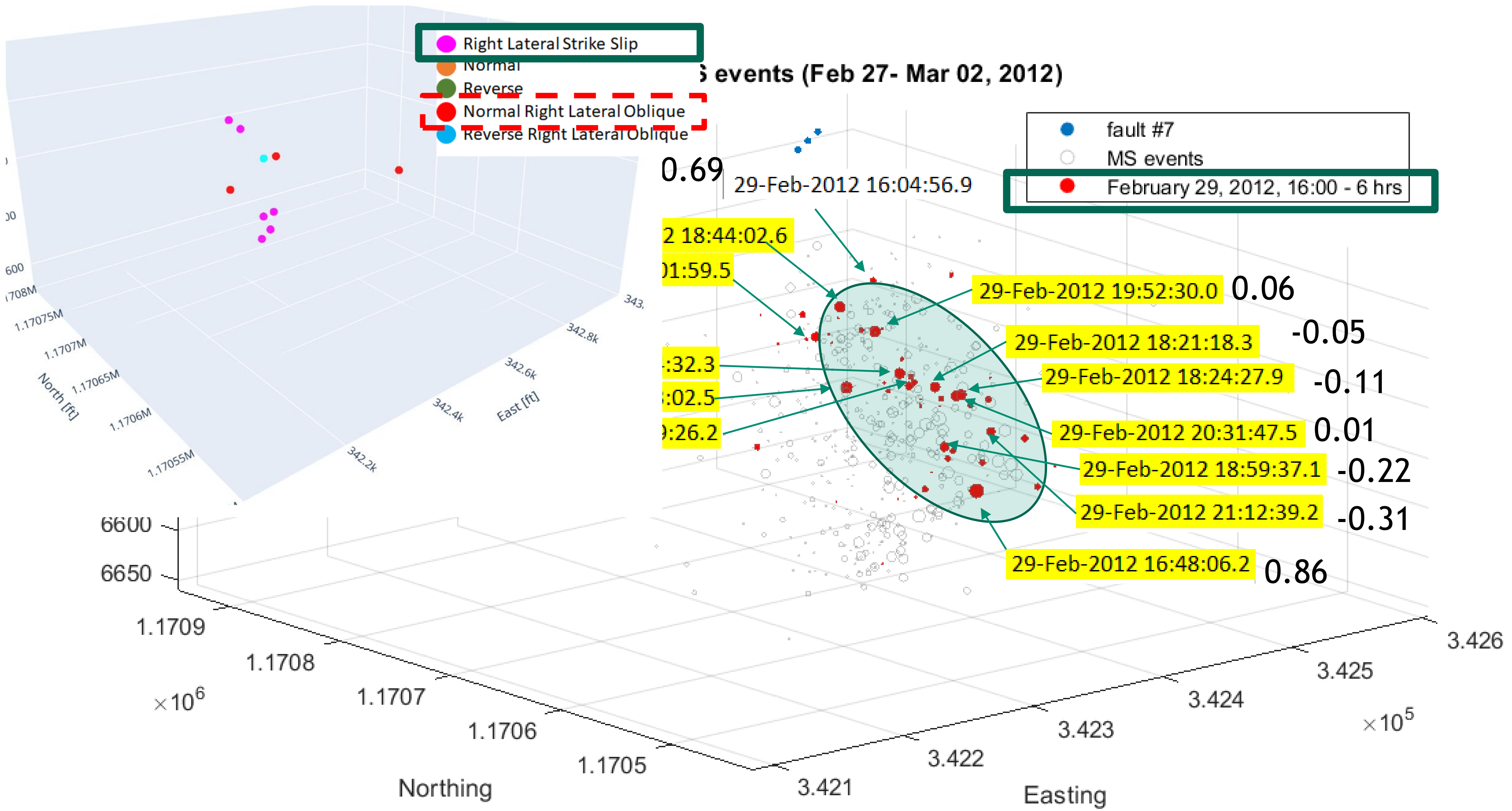
Sub-cluster Patterns over Time



Sub-cluster Patterns over Time



Sub-cluster Patterns over Time



Summary

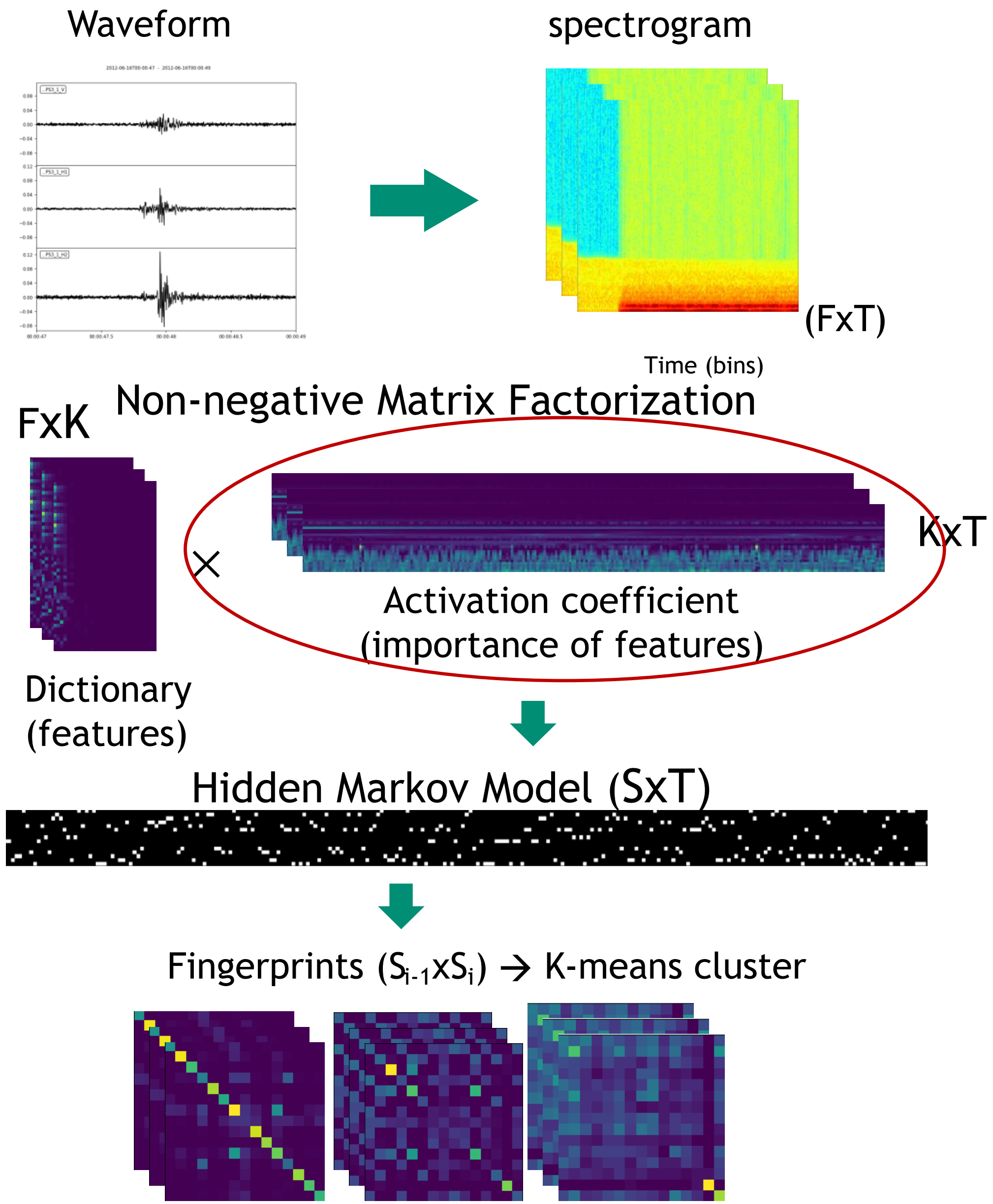
- Rescaled spectrograms as input to ML training dramatically improved ML accuracy
- Simple CNN models trained with located event data only were able to detect events accurately and efficiently
- Re-trained PhaseNet has a relatively high accuracy of phase arrival time picking
- CNN model was able to detect long period long duration patterns (cluster #2)
- During transition period, seismic events tend to be long and overlapped (i.e., slow slip and multiple events) and PS3-2 tends to be higher amplitude than PS3-1 ➔ very distinctive from active and post periods
- Based on LPLD conceptual model, transition waveform characteristics indicate that MS events are likely associated with high density fractures surrounding the main fault after pore pressure increase along the main fault
- Sequence of sub-clusters of MS events indicates the directional stability within the fault architecture, which matches focal mechanism analysis results

Thank You!

?

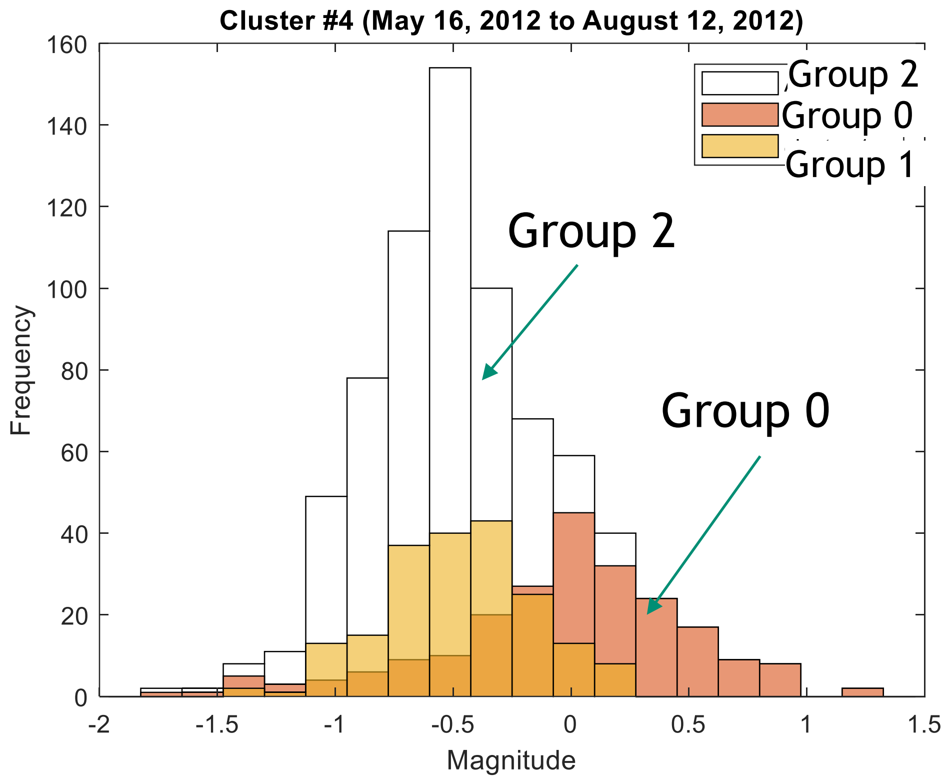
Unsupervised ML – Fingerprint method using NMF-HMM

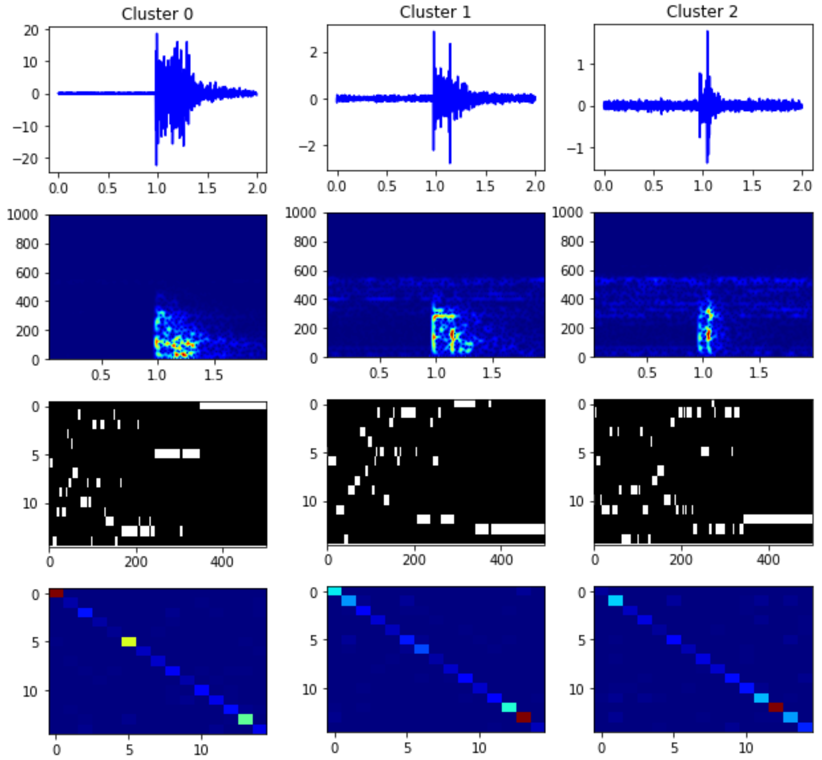
- Fingerprint-based clustering method: Pattern of state sequences forms fingerprints
  - Clustering: acoustic/seismic state ~ mechanical behaviors
  - Spectrogram (Short Time Fourier Transform)
  - Non-negative Matrix Factorization
  - Hidden Markov Model (S states)
  - K-means clustering



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

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)

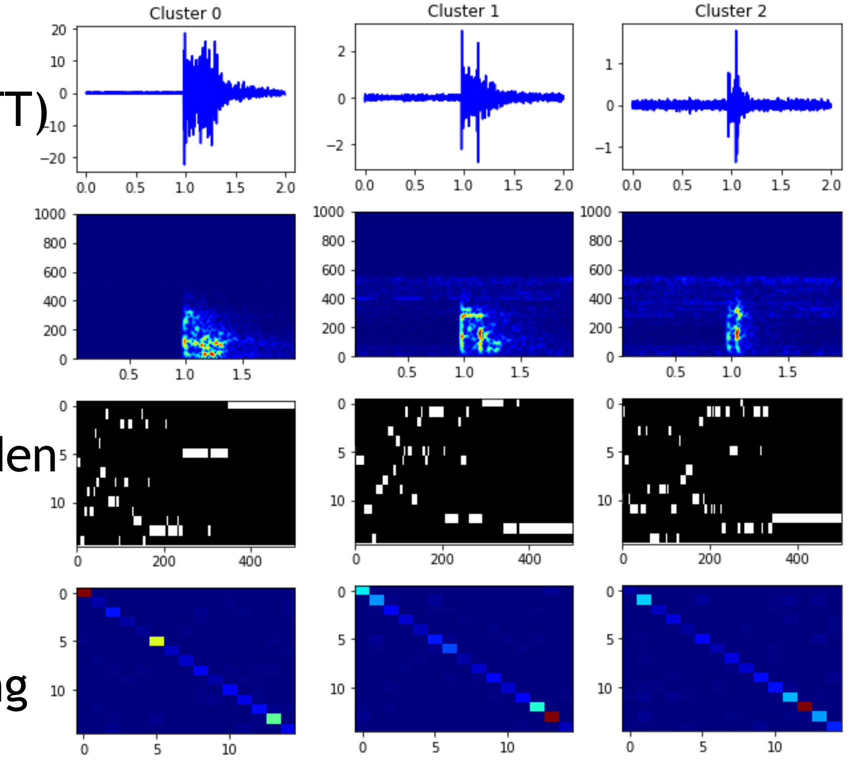
Three groups

1) Waveform
(Bandpass filter, STFT)

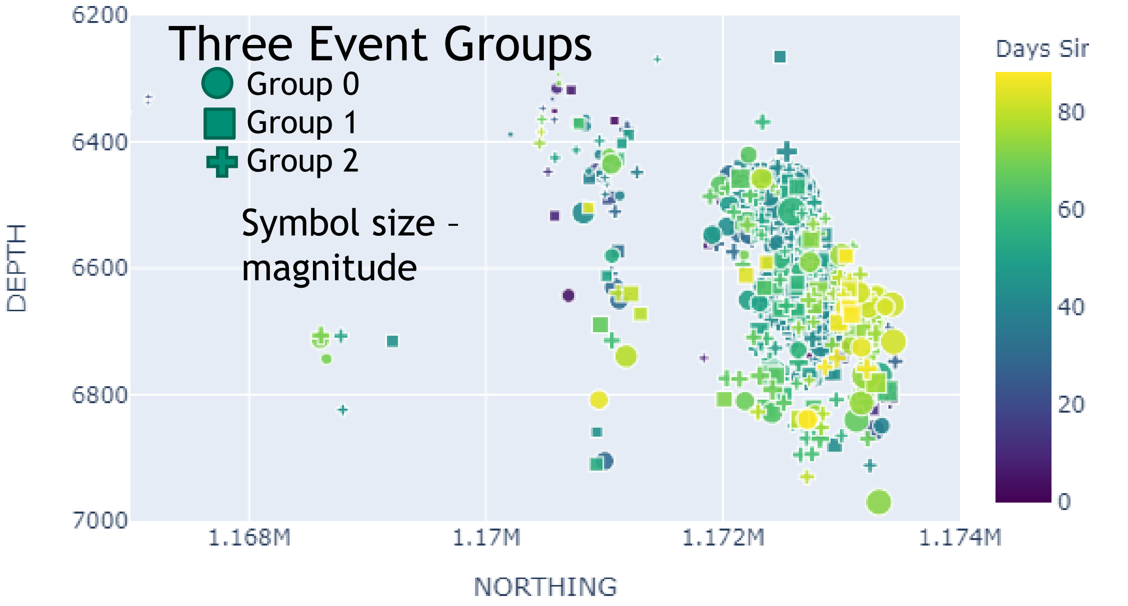
2) Spectrogram
(NMF-> HMM)

3) Transition probabilities of Hidden Markov State

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



Willis & Yoon (In prep for GRL)

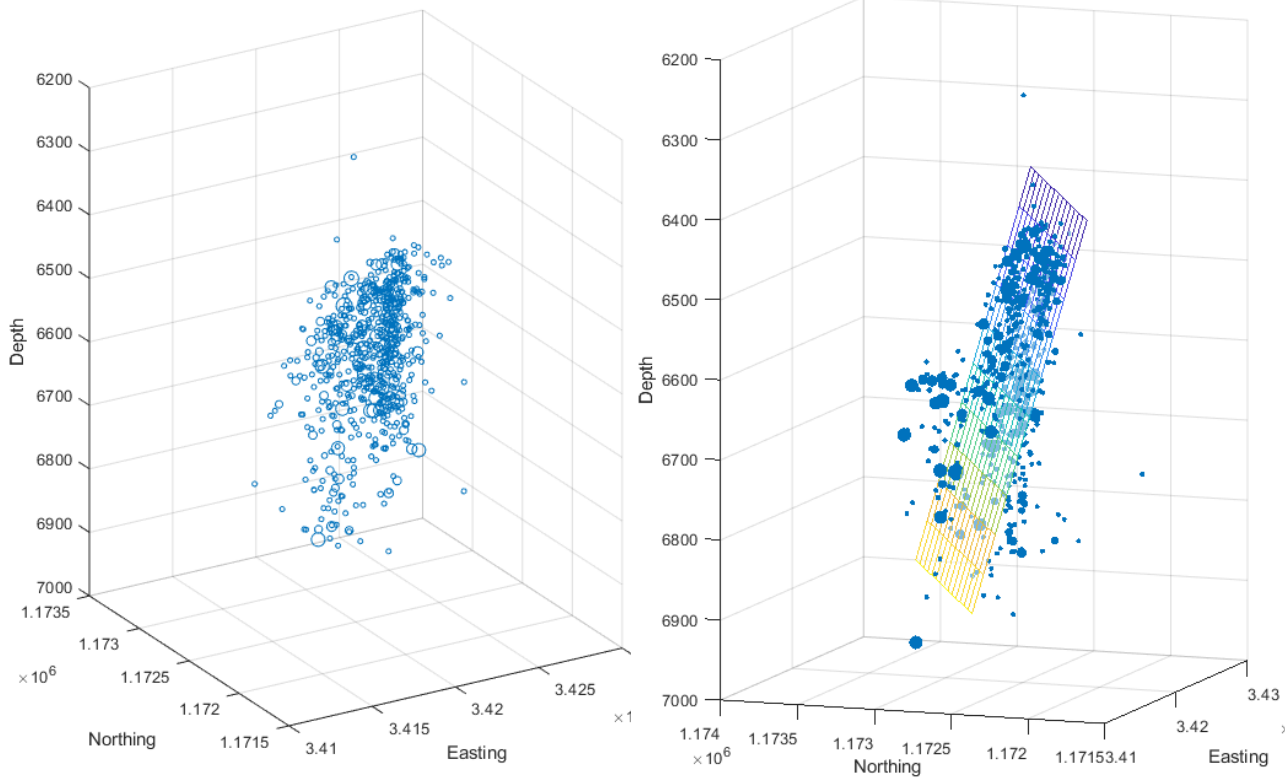


(from 05/16/12)

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

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

All data



Group 0