Sandia National Laboratories

SAND2019-8715PE

Integrated geomechanical and geophysical processes with machine learning approaches for induced seismicity study

Hongkyu Yoon (Sandia National Lab)

PRESENTED BY

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and 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.

Acknowledgments

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

- Laboratory Directed Research and Development program at Sandia National Laboratories

Collaborators:

- Daniel Lizama, K-Won Chang (SNL)
- Laura Pyrak-Nolte, Liyang Jiang (Purdue)

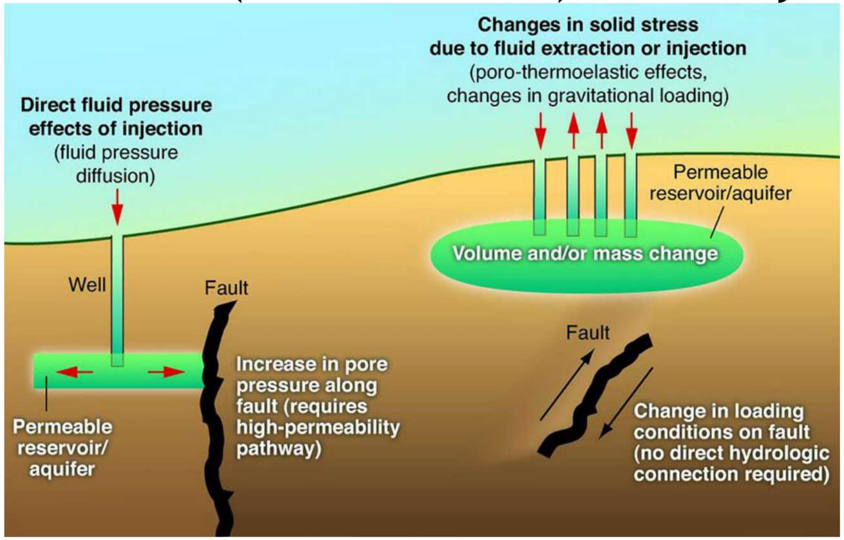
- **Motivations**
- **Linkage between geomechanical and geophysical processes at laboratory scale**
- Machine learning applications at laboratory scale
- Machine learning applications at field scale

Motivations

- Fluid injection or withdrawal causes changes in pore pressure, resulting in stress variations, hydraulic fracturing, fault (re-)activation, and/or fluid saturation changes

- Methodology to reduce risks of induced seismicity and improve modern energy activities in the subsurface:
  - Disposal of water associated with energy extraction (e.g., oil and gas)
  - Geothermal energy production
  - Subsurface carbon storage

Induced (human-caused) seismicity

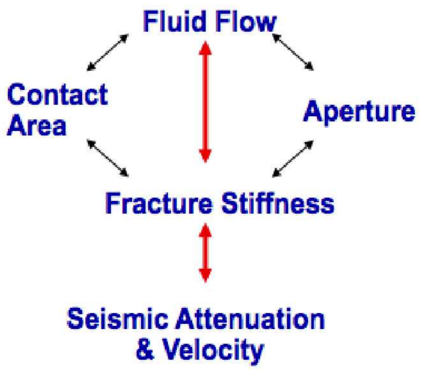


USGS: http://earthquake.usgs.gov/Research/induced/modeling.php

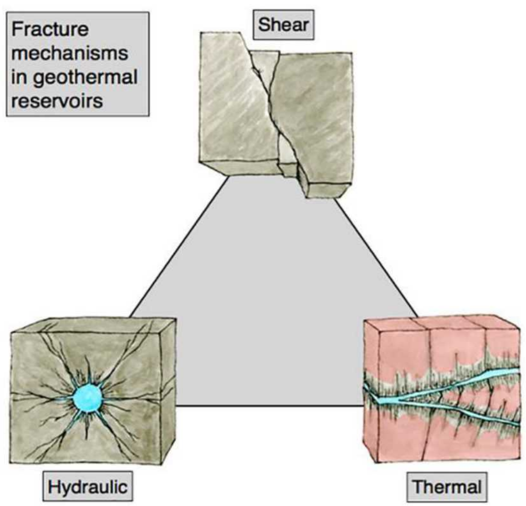
- New groundwork for remote characterization of rock failure by identifying the precursors to the induced seismicity in fractured systems

Linkage between geomechanical and geophysical processes in mechanical discontinuities

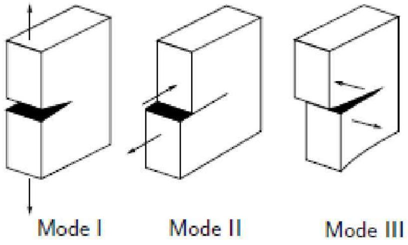
- Precursor(s) to the induced seismicity from existing fault/fracture systems - linking mechanical discontinuities, fracture mechanics, pore pressures and stress to the geophysical signatures is key, yet remains elusive as a result of the heterogeneity (uncertainty) and resulting scale dependence
- Changes 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



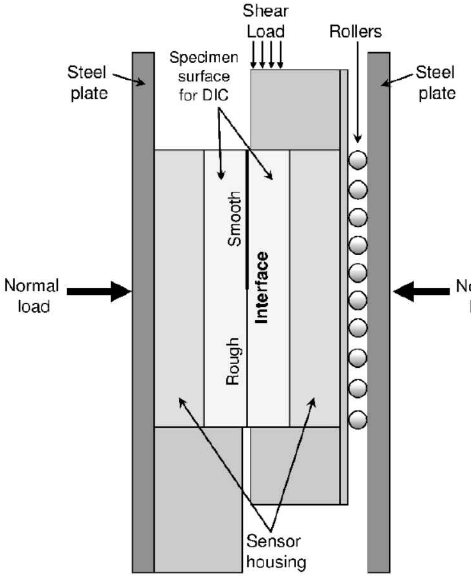
Courtesy from Pyrak-Nolte



Holtzman et al. Sci Adv 2018

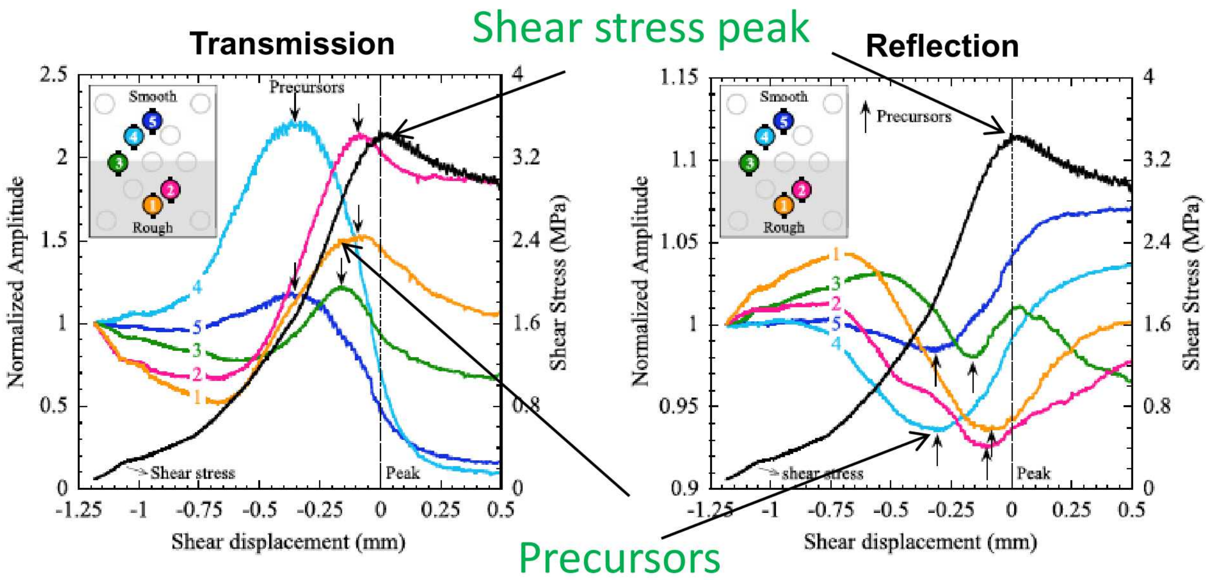


Precursors to Slip along a Mechanical Discontinuity



Bi-axial testing

(Hedayat et al, 2014)



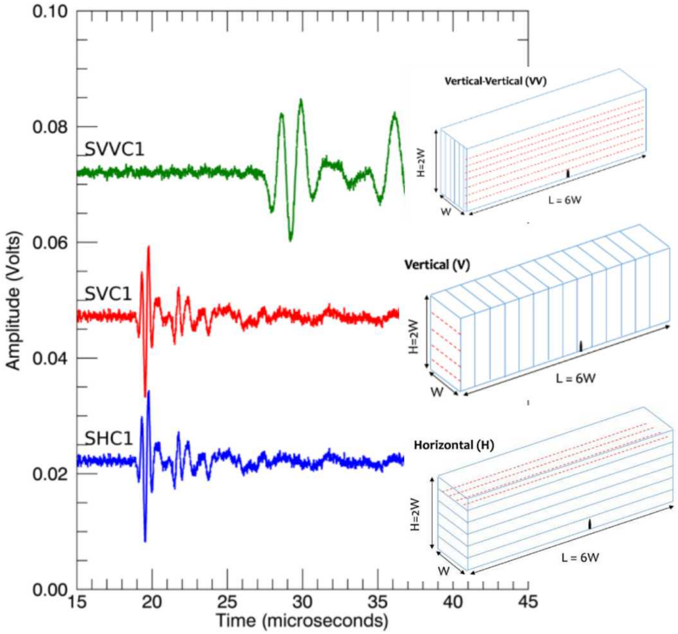
► Increase in transmitted shear wave amplitude prior to achieving the peak shear stress

► Post pre-peak seismic response depends on the frictional characteristics of the interface

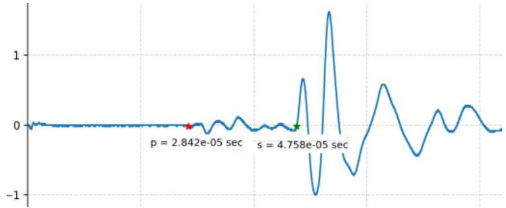
Need to determine how these results apply in a more realistic setting with spatial and temporal variations in pre-existing discontinuities, stress and pressure fields, fluid migration and rock types

Integrated approach for geomechanical and geophysical measurements

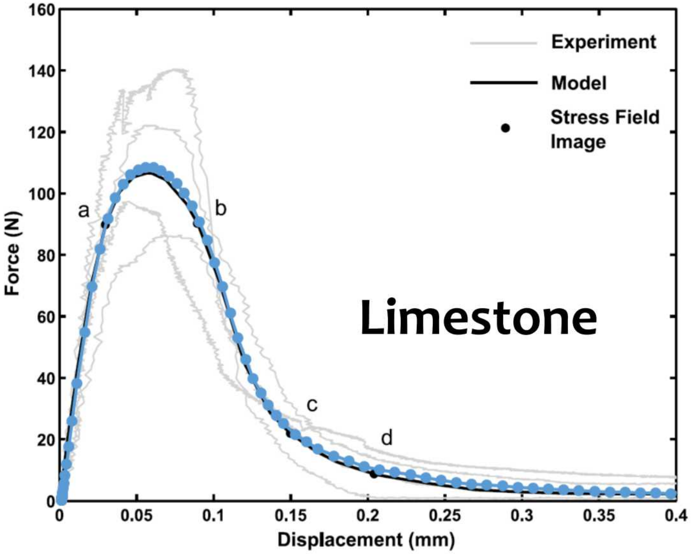
Signals Prior to Failure

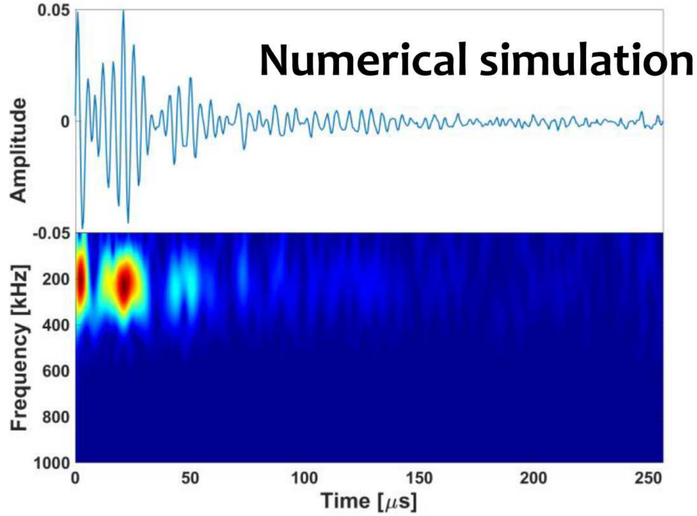


Waveform data analysis (Arrival time picking)



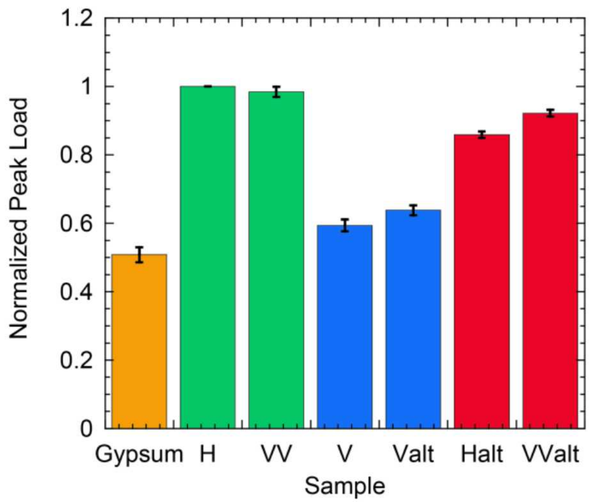
3PB experiments and simulations

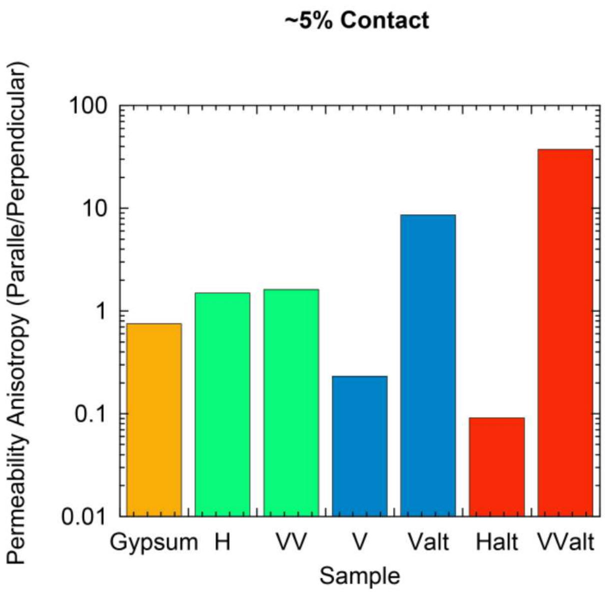


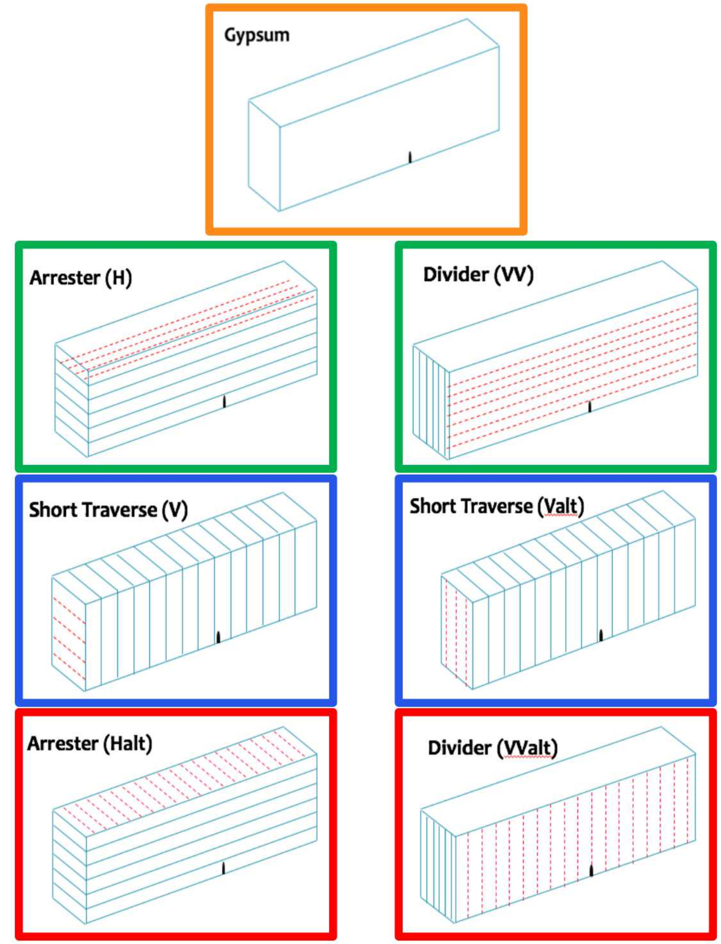


Fracture surface – Flow anisotropy (3D printed)

Load-Displacement Behavior

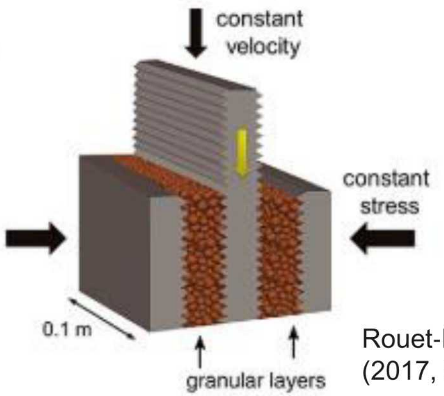






- Motivations
- Linkage between geomechanical and geophysical processes at laboratory scale
- **Machine learning applications at laboratory scale**
- Machine learning applications at field scale

Earthquake Forecasting On Lab Scale Induced Seismic Events



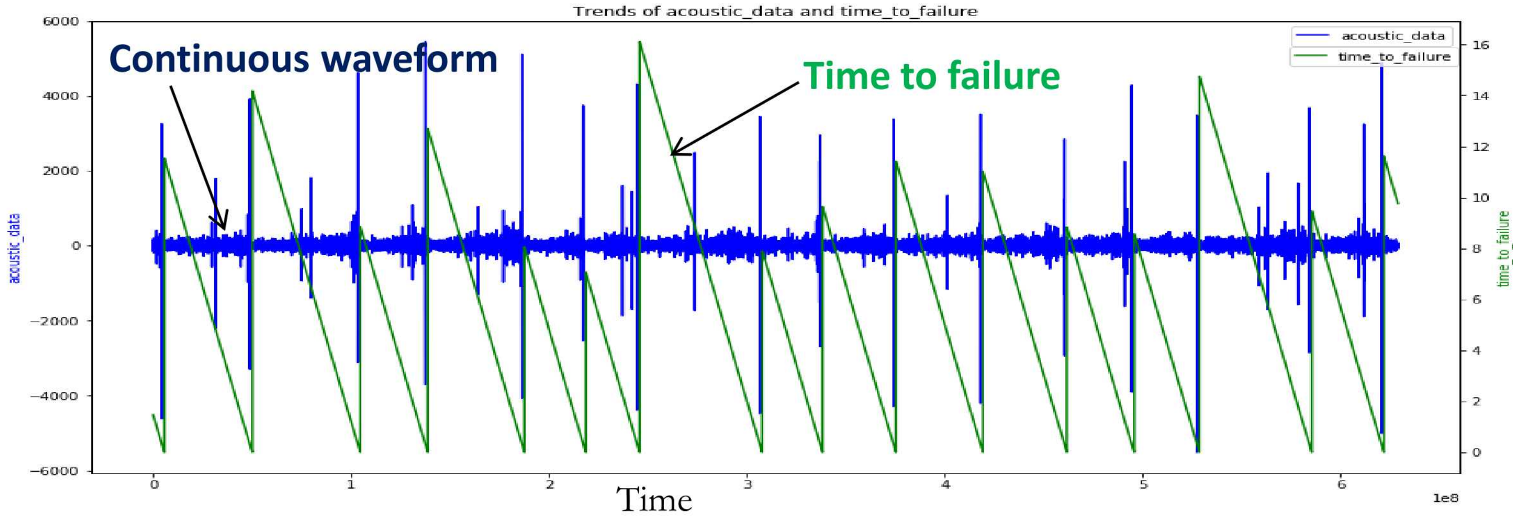
Rouet-Leduc et al. (2017, GRL)

- **Kaggle: LANL Earthquake Prediction**

  Use seismic signals (acoustic emissions) to predict the time remaining for the next earthquake to happen

  - Experimental data: Double direct shear geometry subjected to bi-axial loading

    Aperiodic cycles of stick and slip (loading & failure)
  - Training data: Continuous data containing 16 earthquakes
  - Testing data: Random earthquake cycle segments of 150,000 data-points
  - Approach: Preprocess-> Feature Extraction-> Training->Predictions



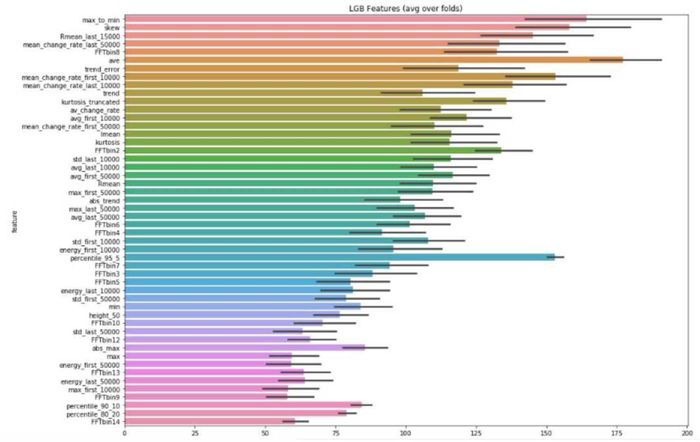
Training sample data file plotted.

Features & Prediction

- Characterize the signal through various measurements
- Features are easily comparable to other signal's features (reduce overfitting)

$\vec{F} =$

Mean	STA/LTA	Maximum
Standard deviation	Correlation	Zero Crossing
Change rate	Kurtosis	Number of peaks
Percentile	Skew	Medians
Quantiles	Energy	Sum
Trend regression	Mel-frequencies	Autocorrelation
FFT	Minimum	Difference



- Data analysis method in which computers learn and autonomously build models based on data patterns.
- **Decision trees**
  Random Forest, Boosting trees (LightGBM)
- **Support Vectors:**
  Support Vector Regressor (SVR)
  Kernel Ridge Regression (KRR)
- **Neural Networks**
  Artificial Neural Networks (ANN)
  Short-Long Term Memmory (LSTM)
  Convolutional Neural Network (CNN)

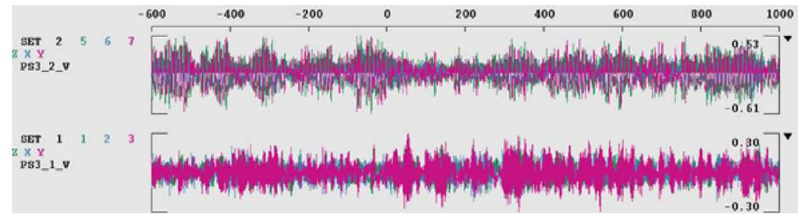
Submissions	CV mean	STD	Public Score	Private Score
LGB	2.0543	0.1198	1.62295	2.65173
XGB	2.0715	0.1196	1.55728	2.64105
KRR	2.0906	0.1078	1.56615	2.52527
Blend KRR XGB	——	——	1.53121	2.56981

https://www.kaggle.com/c/LANL-Earthquake-Prediction/overview

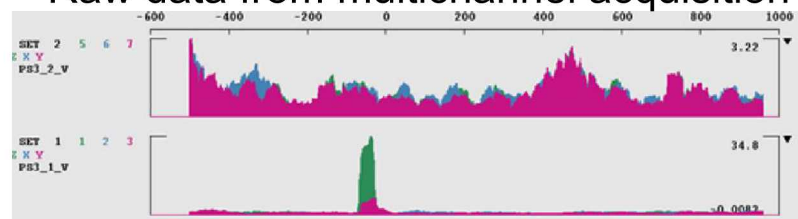
- Motivations
- Linkage between geomechanical and geophysical processes at laboratory scale
- Machine learning applications at laboratory scale
- **Machine learning applications at field scale**

Microseismic Data at IBDP

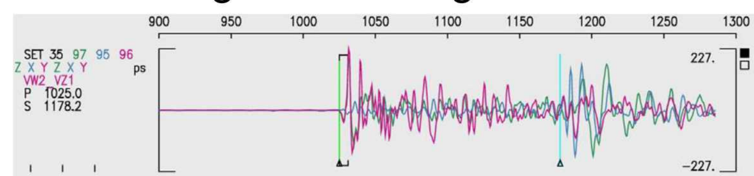
- Raw & processed data (e.g. Will et al., IJGGC 2016)
  - Data acquisition at Injection, monitoring, and verification wells
  - Data analysis for event detection and location
    - Various filters, STL/LTA, and spectral analysis applied
  - Velocity model and MS clustering



Raw data from multichannel acquisition



Short/long term average function

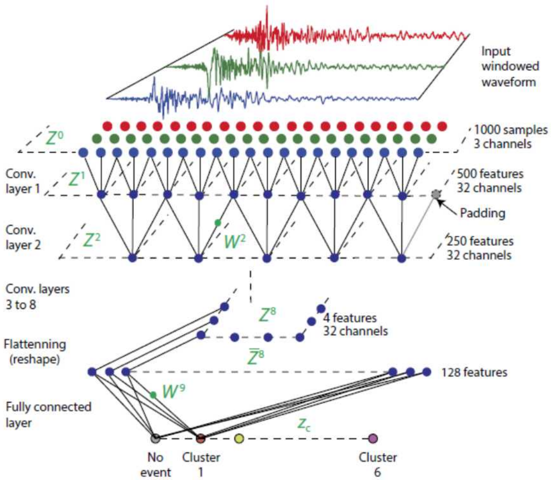


Event waveform

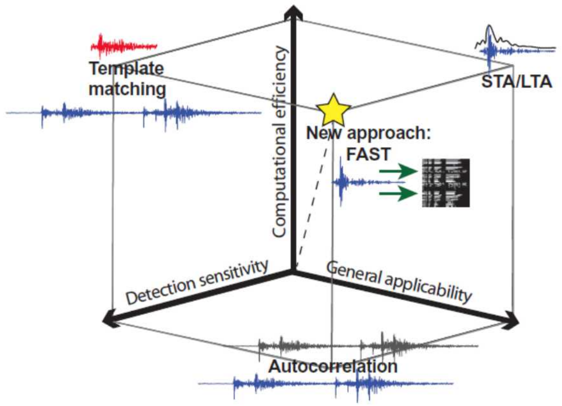
Will et al., IJGGC 2016

ML Approaches

- Supervised ML: Convolutional neural network (CNN) for event detection and location
  - Open source ConvNetQuake (Perol et al., 2018)
    - Processed data from ISGS will be used to train models
    - Trained model will be used to validate again the remaining dataset to develop real-time recognition of events and locations
- Unsupervised ML: Waveform similarity-based event detection methods
  - Fingerprint and Similarity Thresholding (FAST, Stanford FAST group)
    - FAST shows the increase in event detection of low magnitude seismicity by > a factor of 10
    - High efficiency in big data processing time
- Template matching (EQcorrscan)
  - This is a refence case whose results will be compared with ConvNetQuake and FAST for efficiency and interpretability
- Characterization of Microseismic events
  - Spectral clustering and regression-based machine learning analysis (e.g. random forest)
    - Identify seismic phases from successive slip or fracturing stage events and their constitutive wave patterns
    - Extract the salient features present in the data set, such as individual wave types, spectral content, p-s converted waves, and local energy decay
    - Link microseismic data to other measured/simulated quantities (e.g., injection, pressure and stress field)

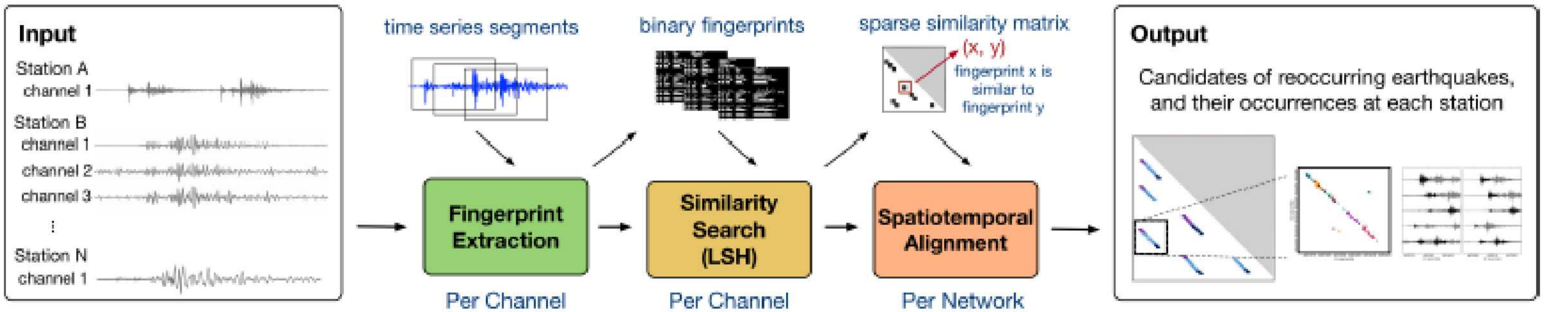


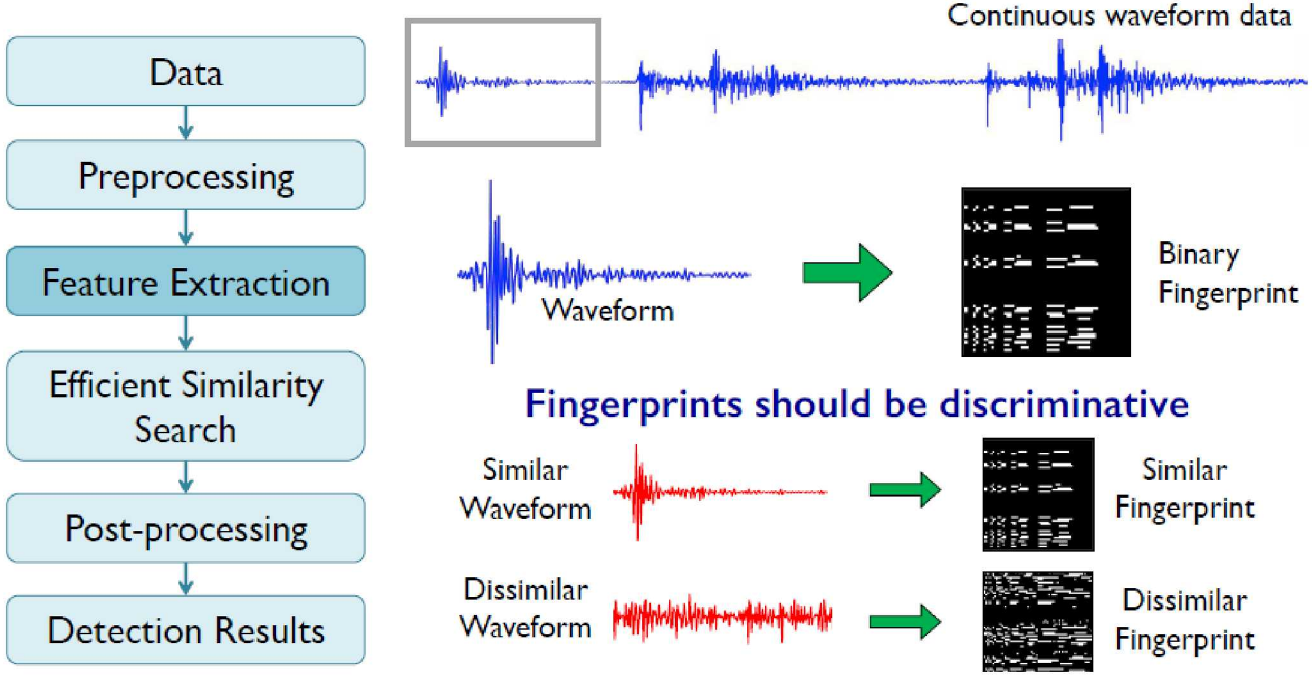
ConvNetQuake CNN Architecture
Perol et al. (2018, SciAdv 2018)



Earthquake detection methods from C.E. Yoon et al. (SciAdv 2015)

FAST Approach





From Bergen et al. 2018 presentation at IRIS
Rong et al. (2018)

ML Applications for Event Detection & Fault System Configulation

- Develop and apply ML/deep learning methods
  - Improve identification of precursors to induced seismicity
  - Improve the detection of 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