SMART-CS Initiative

Science-informed **M**achine Learning to **A**ccelerate **R**eal **T**ime (SMART) Decisions in Subsurface Applications

ML-Based Fracture and Fault Identification

WP 2C.2 Leads: David Alumbaugh (LBNL) and Joe Morris (LLNL)

Youzuo Lin (Formerly Affiliated with LANL)
School of Data Science and Society
University of North Carolina at Chapel Hill

U.S. DEPARTMENT OF ENERGY

Task Objectives, Preliminary Study, and Challenges

Objectives:

1. To improve the imaging description of fractures & faults at IBDP
2. To obtain a site-agnostic ML-assisted toolset and workflow
3. To demonstrate the benefits of using ML methods

Significance:

- Provide fast fluid pathway and flow barrier (input for update reservoir model in task 5)
- Indicate and monitor potential induced seismicity (input for ORION in task 6)

Study Site & Preliminary Study:

- Study Site: IBDP
- Preliminary Study: Dando et al., (2021) deployed a modified double-difference method to identify microseismicity, and further delineated linear clustering of events with uncertainty.

Challenges:

- Expensive computational cost
- Limited imaging resolution
- Lack of training dataset

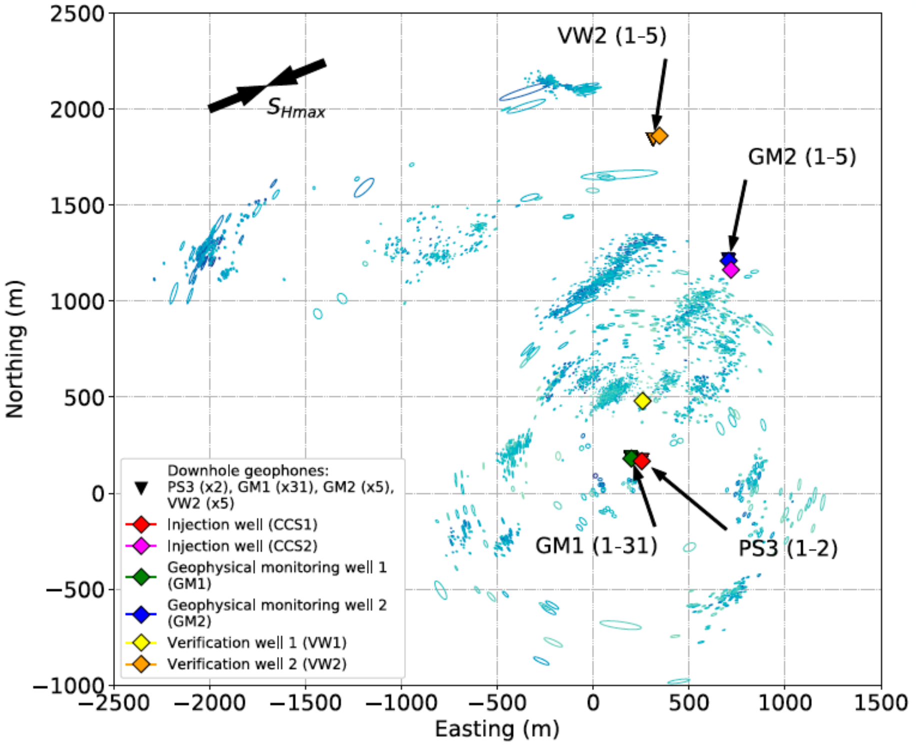
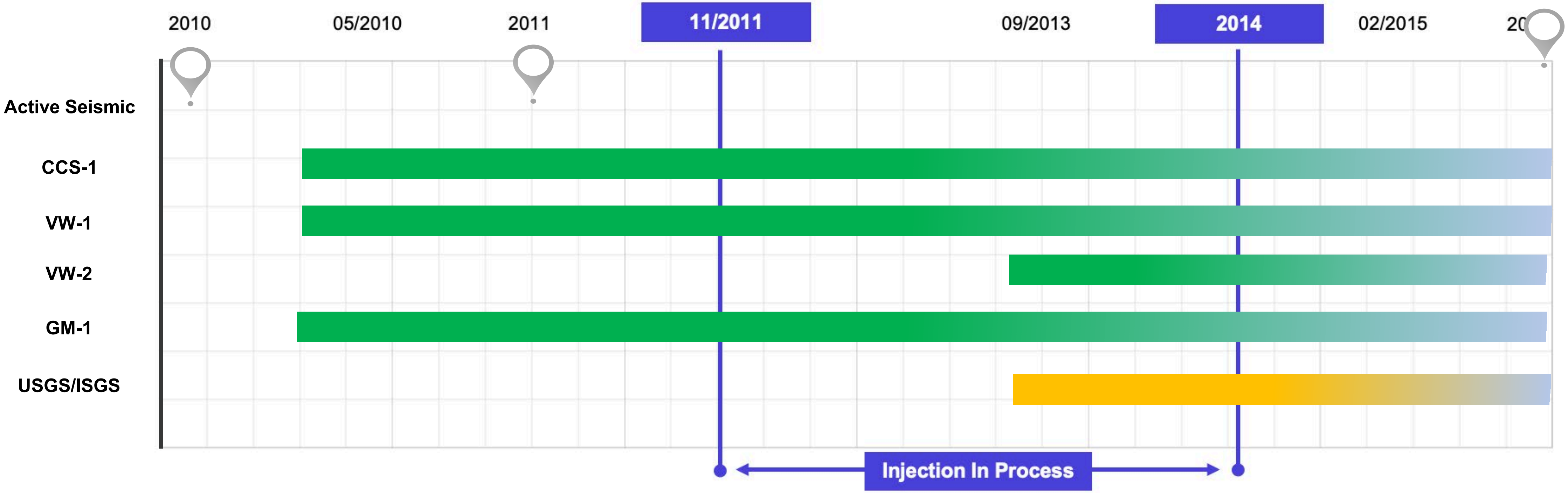


Figure Courtesy of Dando et al., 2021

Dando et al., "Relocating microseismicity from downhole monitoring of the Decatur CCS site using a modified double-difference algorithm" GJI , 2021.

Data Availability – A Case Study at IBDP

- Passive Seismic Acquisition
  - IBDP Installation
    - Borehole arrays located at **CCS-1**, **VW-1**, **VW-2**, **GM-1**
    - Total: 31 stations (z-component: 2/4 CCS-1 + 29/31 GM-1)
  - USGS/ISGS Installation
    - 20 surface seismometers (15 **USGS** + 5 **ISGS**)



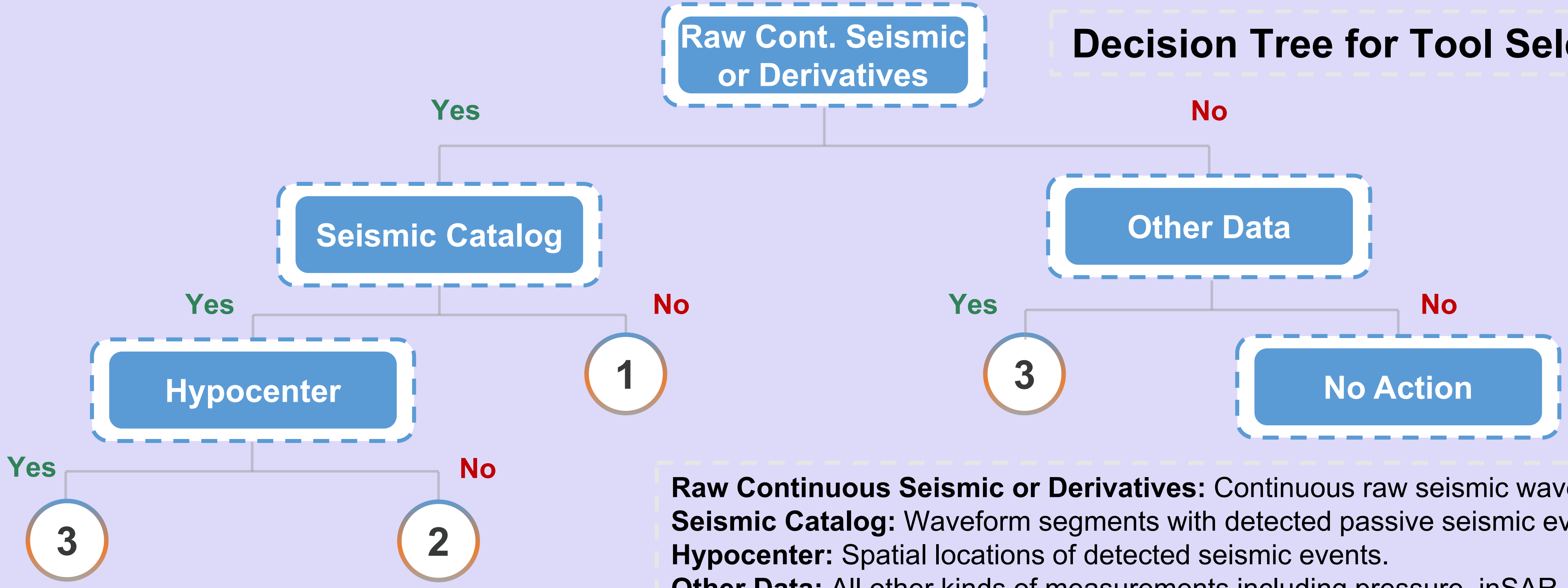
3

Proposed Fracture Imaging Workflow – An Overview

01	• **Event Detection**	Extract useful microseicmic events from continuous waveform measurements
02	• **Velocity Inversion**	Produce 3D velocity model from active/passive seismic data (SubTask 4.4.1)
03	**Hypocenter Locating**	Obtain microseismic source parameters (location, moment tensors, amplitude, etc)
04	• **Fracture Estimation**	Deploy spatio-temporal clustering analysis to obtain fracture lines
05	• **Uncertainty Quantification**	Analyze the uncertainty of the fracture and fault zones
06	• **Fracture Visualization**	Display final fault/fracture representation to field operators

Decision Tree for Tool Selections

Data Availability



Raw Continuous Seismic or Derivatives: Continuous raw seismic waveform data.
Seismic Catalog: Waveform segments with detected passive seismic events.
Hypocenter: Spatial locations of detected seismic events.
Other Data: All other kinds of measurements including pressure, inSAR, EM, etc.).

Decision

1 Events Detection

1. Traveltime: SNL/ORNL
2. Waveform: ISGS/UIC

2 Hypocenter Locating

1. Traveltime: SNL/ORNL
2. Waveform: LANL
3. Waveform: ISGS/UIC

3 Joint Interpretation

1. Seismic Pressure: NETL
2. Seismic, Well logs: ISGS/UIC
3. Seismic: LBL/LLNL
4. Seismic: FACT

4 Fracture Estimation

1. K-Mean: LANL
2. NMF: SNL
3. Clustering: NETL
4. HypoDD+Focal Mech: ISGS/UIC
5. OpenDtect: FACT

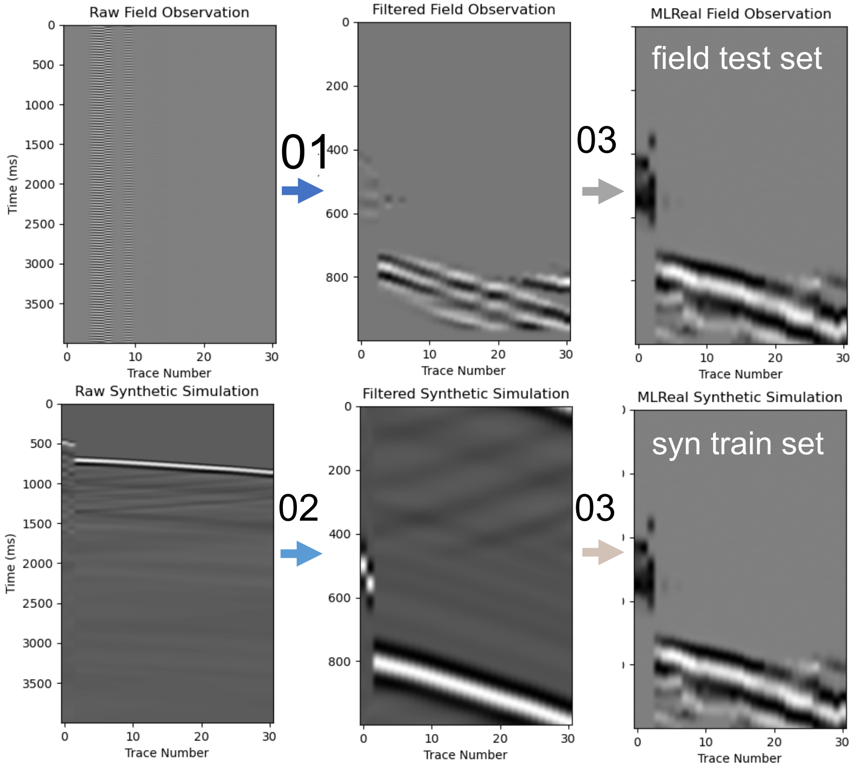
5 Fracture Visualization

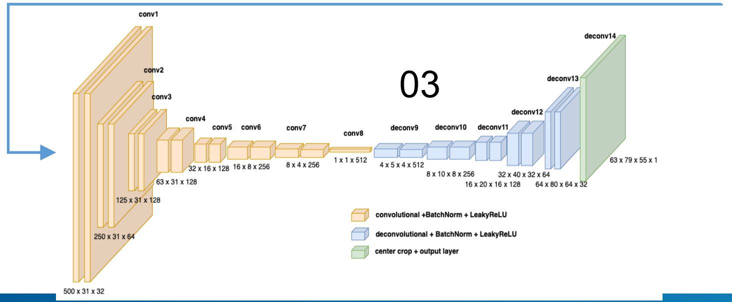
1. Petrel: EERC

Technique

Org	Data	Key Technique Descriptions	Event Detection	Hypocenter Locating	Joint Interpretation	Fracture Estimation	Fracture Visualization
NETL	Catalog, Pressure	**1.** Spatial-temporal analysis on MS **2.** Deploy five clustering methods to characterize fault & fracture.			●	●	
ORNL SNL	Raw Seismic	**1.** Detect events (raw seismic data) **2.** Leverage traveltime information to invert source location.	●	●		●	
LANL	Catalog	**1.** Use borehole seismic waveform. **2.** Deploy full-waveform inversion to invert the source location.		●		●	
ISGS UIC	Catalog, Well logs	**1.** Infer location (PhaseNet,HypoDD) **2.** Obtain focal mech using ANN **3.** Infer fault (location & focal mech)	●	●	●	●	
LBNL LLNL	Catalog	**1.** Infer fault location using microseismicty **2.** Deploy aCNN method			●	●	
FACT	Catalog	**1.** Use OpenDtect to interpret seismic data 2. Identify useful seismic features			●	●	
EERC		**1.** Use Petrel for visualizing fault & fracture maps.					●

LANL – DeFault: Deep-learning-based Fault Delineation (POC: Y. Lin)





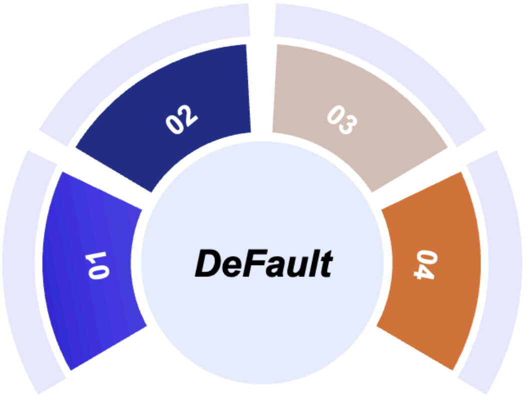
01 – Data Pre-Processing on Raw Seismic Waveform

Enhance waveform data by carrying out bandpass filters, amplitude normalization, F-k dipping filter, time-domain noise removal, averaged F-k envelop filter

02 – Full-waveform data synthesis to build high-fidelity traning set

Leverage 3D velocity model and acoustic wave equation to generate full-physics training data

Gaussian heatmaps centering at true source locations – spatial distribution



DeFault [Wang et al., 2024]

03 – ML-based Full-Waveform Inversion to Relocate Source Parameters

MLReal data domain adaptation, deploy encoder-decoder full-waveform inversion to obtain microseismic event location heatmaps
Heatmap upsampling to remove griding effect, interpolation of first and second maximum values to get coordinates predictions

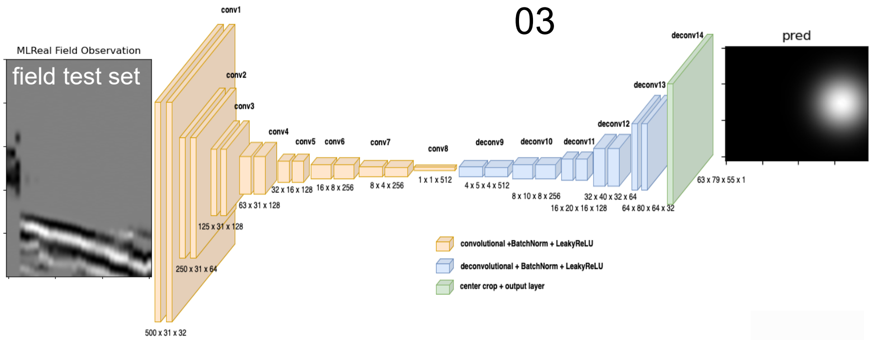
04 – Employ Spatio-temporal clustering analysis to delinate Fracture imaging

Temporal period selection, K-means spatial clustering, outlier removal, least squared distance fault plane estimation

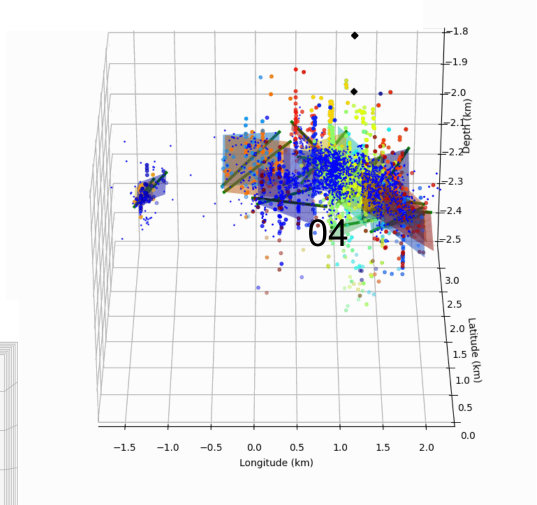
Hanchen Wang, et al, “DeFault: Deep-learning-based Delineation Using Domain Adaptation Training and Automatic Clustering”, ESS 2024 (Under Review).

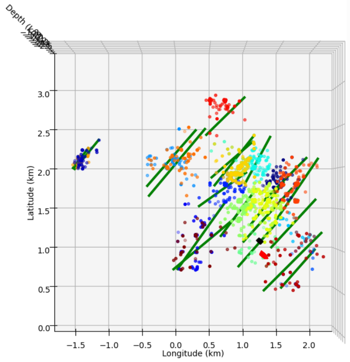
LANL – DeFault: Deep-learning-based Fault Delineation (POC: Y. Lin)

03



04





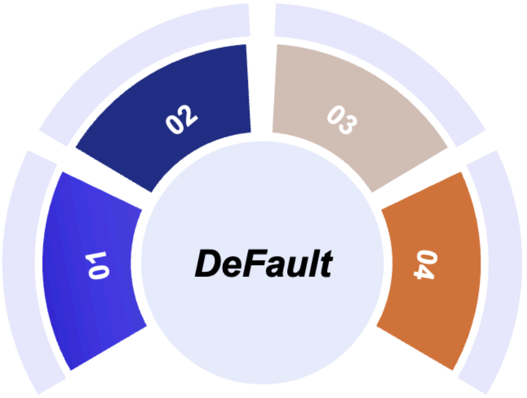
01 – Data Pre-Processing on Raw Seismic Waveform

Enhance waveform data by carrying out bandpass filters, amplitude normalization, F-k dipping filter, time-domain noise removal, averaged F-k envelop filter

02 – Full-waveform data synthesis to build high-fidelity traning set

Leverage 3D velocity model and acoustic wave equation to generate full-physics training data

Gaussian heatmaps centering at true source locations – spatial distribution



DeFault [Wang et al., 2024]

03 – ML-based Full-Waveform Inversion to Relocate Source Parameters

MLReal data domain adaptation, deploy encoder-decoder full-waveform inversion to obtain microseismic event location heatmaps

Heatmap upsampling to remove griding effect, interpolation of first and second maximum values to get coordinates predictions

04 – Employ Spatio-temporal clustering analysis to delinate Fracture imaging

Temporal period selection, K-means spatial clustering, outlier removal, least squared distance fault plane estimation

Hanchen Wang, et al, "DeFault: Deep-learning-based Delineation Using Domain Adaptation Training and Automatic Clustering", ESS 2024 (Under Review).

SNL – Fault Imaging via Event Detection & Source estimation (POC: J. Harding & H. Yoon)

- Integrated ML approaches of event detection and source location estimation
- Data pre-processing of raw continuous microseismic data & event detection
- Data augmentation using WGAN (Wasserstein Generative Adversarial Network)
- PhaseNet used to downselect generated event data with high quality
- CNN model with multi-modal input for source location estimation of events

Data processing of raw waveform continuous data

- 10s raw continuous waveform data (e.g., 4 to 3 channels)

Event detection & arrival time: CNN & U-Net models

- Train a CNN model for event detection with 3 channel & energy feature as input and retrain PhaseNet for arrival time

Synthetic data generation: SeismoML (WGAN)

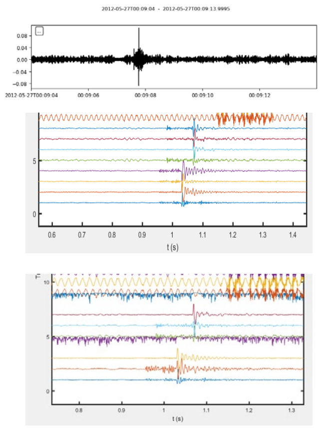
- WGAN model Input: source locations and distance of ~400 events from catalog and output of waveform
- Apply trained seismoML model to generate synthetic waveforms of each channel (H1,H2,V)
- Screen generated waveform data by phase arrival times (PhaseNet)

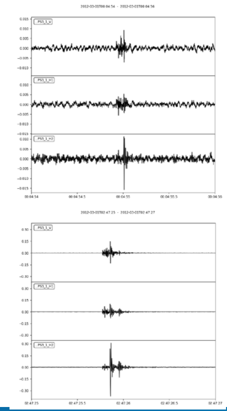
Multi modal CNN for source locations of newly detected events

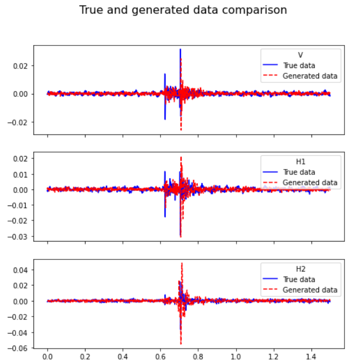
- Construct synthetic waveform data over a range of source locations & distance
- Train a multi modal CNN with spectrograms and P&S arrival (binary) of each channel data

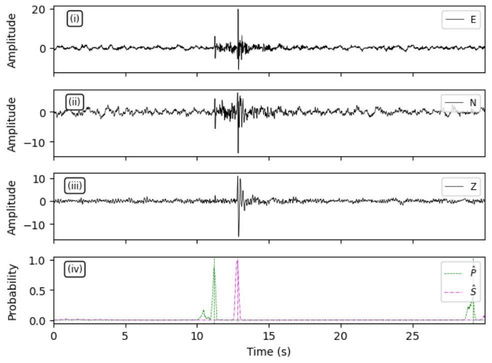
Event clustering & construct faults

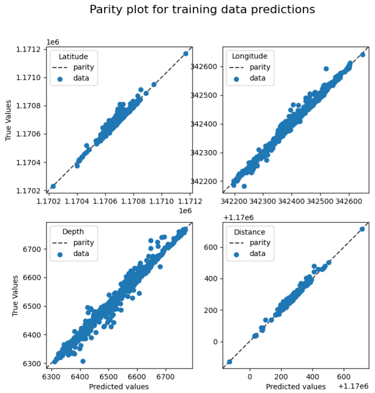
- Event clustering using NMF-HMM (Non-negative matrix factorization Hidden Markov Model) to construct planar faults

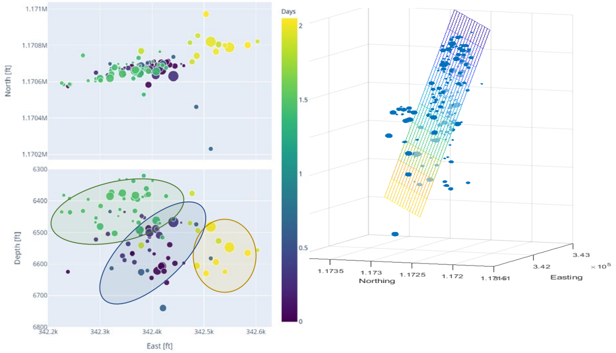












U.S. DEPARTMENT OF ENERGY

SMART

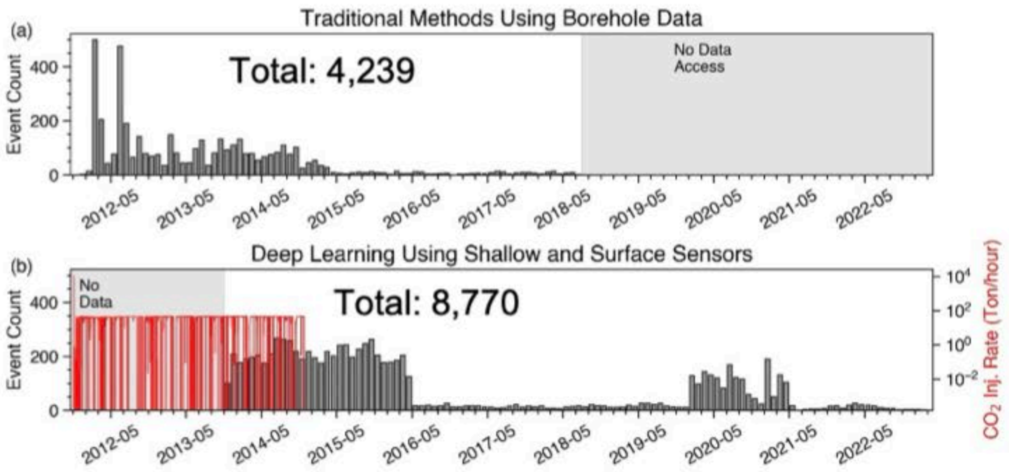
ORNL – Fault & Fracture Identification (POC: C. Chai & M. Maceira)

USGS Dataset (surface)

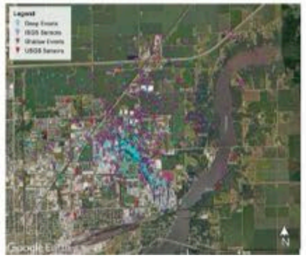
ML Picker → ML Models

Continuous Seismic Data → Signal Arrival Times → Seismic Event Catalog

Physics-based Algorithm



We detected and located 8,770 seismic events using 10 years of continuous seismic data from 17 USGS seismic stations.

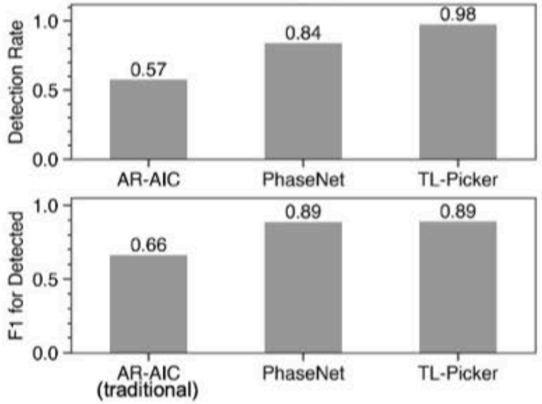


IBDP Dataset (borehole)

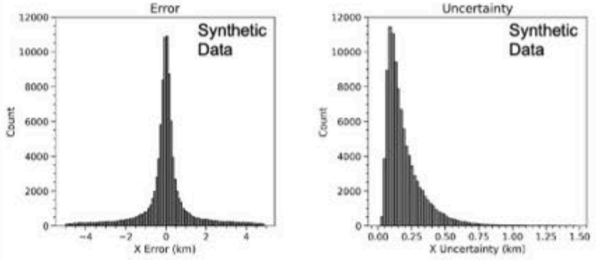
TL-Picker → DL-TT-Locator → ML-Fault-Mapper

Passive Seismic Data → Signal Arrival Times → Seismic Event Catalog → Fault & Fracture

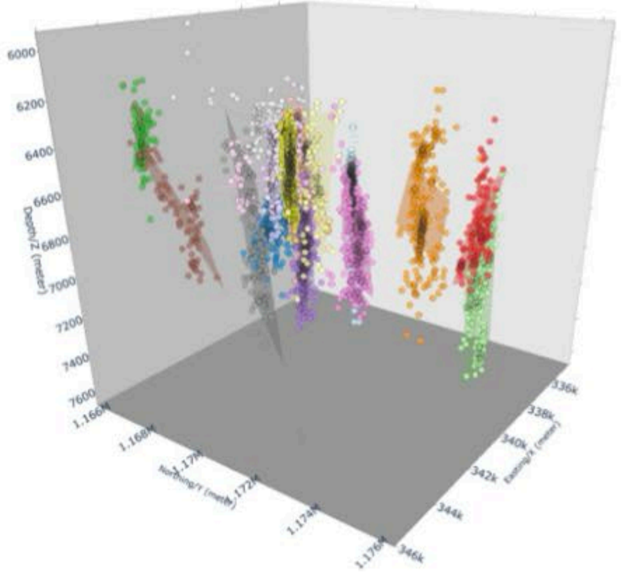
TL-Picker



DL-TT-Locator

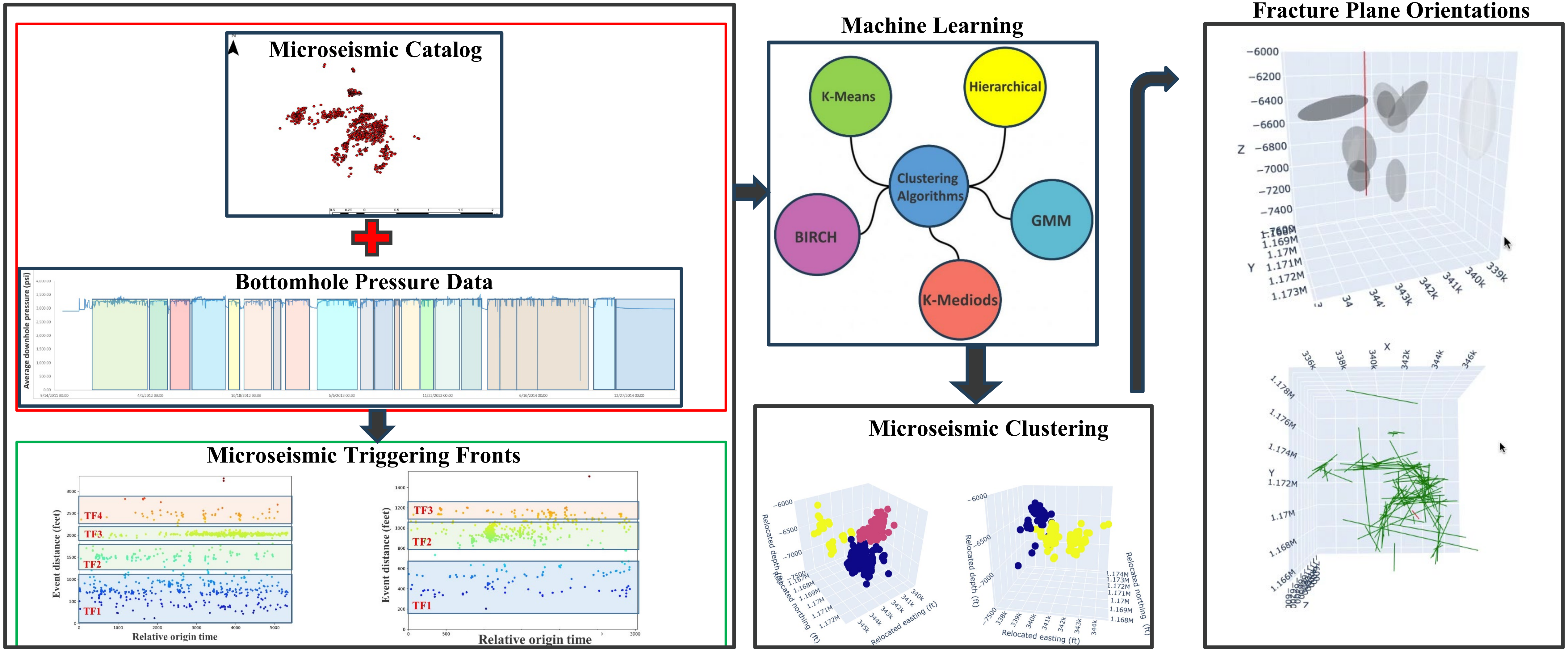


ML-Fault-Mapper



Three ML modules have been developed to process passive seismic data precisely and efficiently.

NETL – ML-Based Fracture Network Quantification (POC: A. Kumar)



Org	Data	Key Technique Descriptions	Event Detection	Hypocenter Locating	Joint Interpretation	Fracture Estimation	Fracture Visualization
NETL	Catalog, Pressure	**1.** Spatial-temporal analysis on MS **2.** Deploy five clustering methods to characterize fault & fracture.			●	●	
ORNL SNL	Raw Seismic	**1.** Detect events (raw seismic data) **2.** Leverage traveltime information to invert source location.	●	●		●	
LANL	Catalog	**1.** Use borehole seismic waveform. **2.** Deploy full-waveform inversion to invert the source location.		●		●	
ISGS UIC	Catalog, Well logs	**1.** Infer location (PhaseNet,HypoDD) **2.** Obtain focal mech using ANN **3.** Infer fault (location & focal mech)	●	●	●	●	
LBNL LLNL	Catalog	**1.** Infer fault location using microseismicty **2.** Deploy aCNN method			●	●	
FACT	Catalog	**1.** Use OpenDtect to interpret seismic data 2. Identify useful seismic features			●	●	
EERC		**1.** Use Petrel for visualizing fault & fracture maps.					●

Fracture Imaging Workflow – Lessons Learned

		Accuracy	Efficiency	Generalization	Data Demanding
01	Event Detection				
02	Velocity Inversion	SubTask 4.4.1 (LBL/LANL)			
03	Source Inversion				
04	Fracture Analysis				
05	Uncertainty Quantification				
06	Visualization				

Physics Wins

Pure ML Wins

Fracture Imaging Workflow – Lessons Learned

- Pure ML methods suffers from **weak generalization ability**, **high training cost**, and **require a large volume of training data**
- One solution is to incorporate underlying physics, geology knowledge – **Physics-guided ML**

Questions?