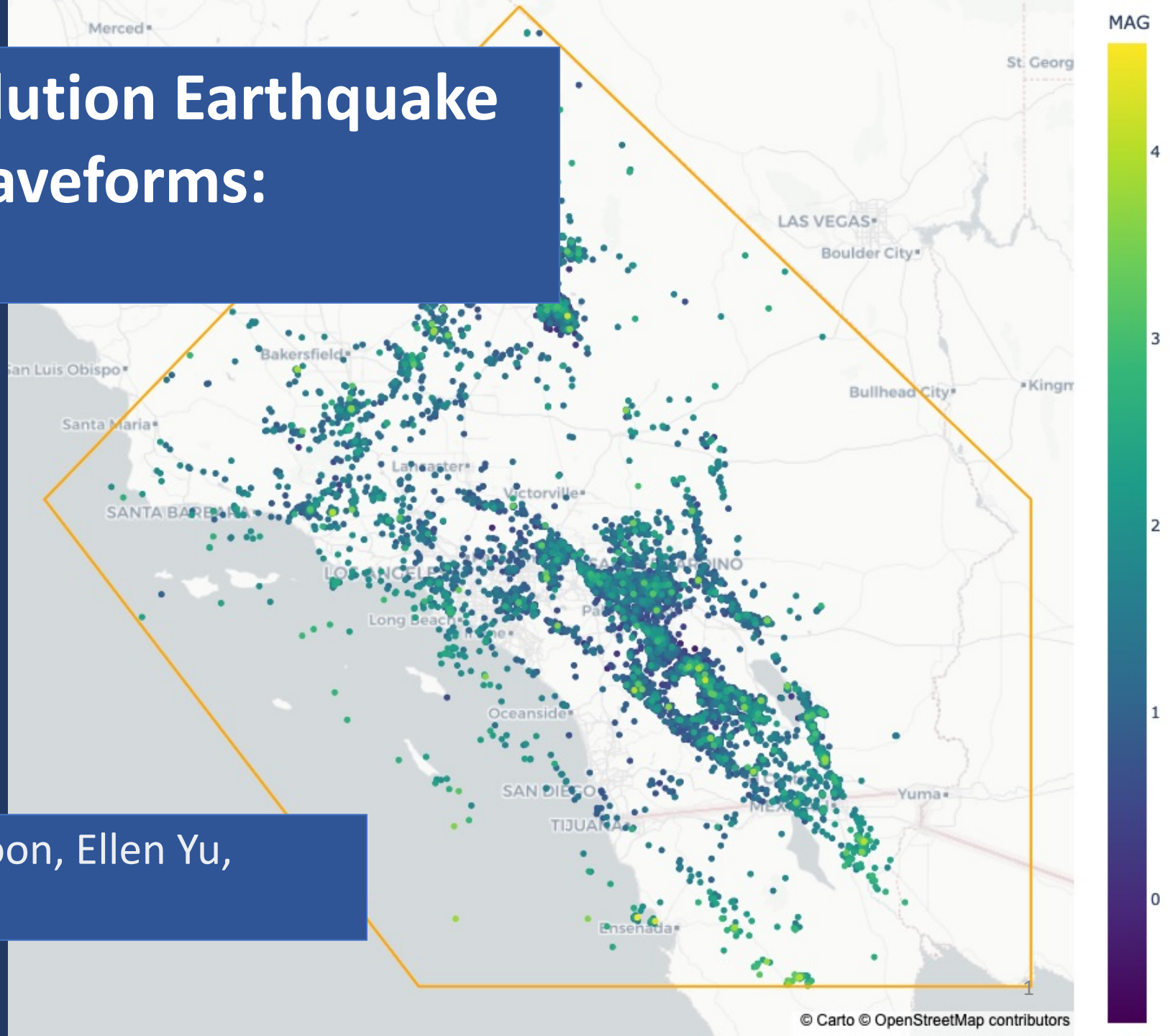
Building a High-Resolution Earthquake Catalog from Raw Waveforms:

A Step-by Step Guide

SSA Workshop
Baltimore, MD

April 14, 2025 (1 – 5 pm)

Presenters: Eric Beaucé, Clara Yoon, Ellen Yu, Weiqiang Zhu, Gabrielle Tepp



Agenda

20 min (1:00 – 1:20 pm) - Introductions & Overview [Eric, All]

45 min (1:20 – 2:05 pm) - Regional Seismic Networks: Official Catalogs and Data Access [Gabrielle, Ellen]

10 min (2:05 - 2:15 pm) - Break

105 min (2:15 – 4:00 pm) - Building Custom Catalogs with Modern Tools
Deep-learning (2:15 – 3:15 pm) [Weiqiang, Clara]
Template-Matching (3:15 – 4:00 pm) [Eric]

10 min (4:00 – 4:10 pm) - Break

35 min (4:10 – 4:45 pm) - Evaluating Catalog Quality, Completeness [Clara]

15 min (4:45 – 5:00 pm) - Conclusions [Clara, All]

Introductions

Eric Beaucé

Columbia/LDEO

enhanced catalogs, statistical seismology, earthquake science

Ellen Yu

Data Center Manager
Caltech/SCSN

data archival, distribution

Weiqiang Zhu

Assistant Professor
UC Berkeley

machine learning, earthquake catalogs

Gabrielle Tepp

Staff Seismologist
Caltech/SCSN

network seismology, hydroacoustics, volcano seismology

Clara Yoon

Research Geophysicist
USGS Pasadena

enhanced earthquake catalogs from a user perspective

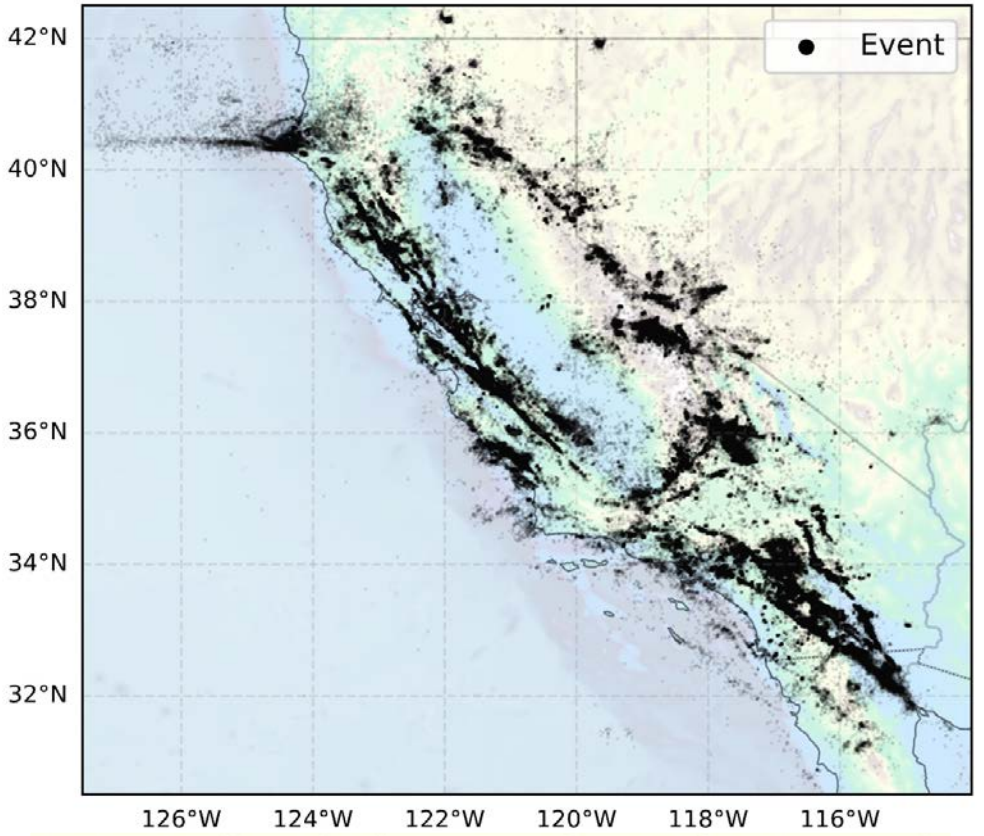
What is an earthquake catalog?

Database of known earthquakes typically including:

- **origin time**
- **location**
- **magnitude**

and maybe additional information:

- seismic phase picks
- source characteristics (e.g., focal mechanism)
- event type (e.g., natural eq, volcanic explosion, quarry blast)
- etc

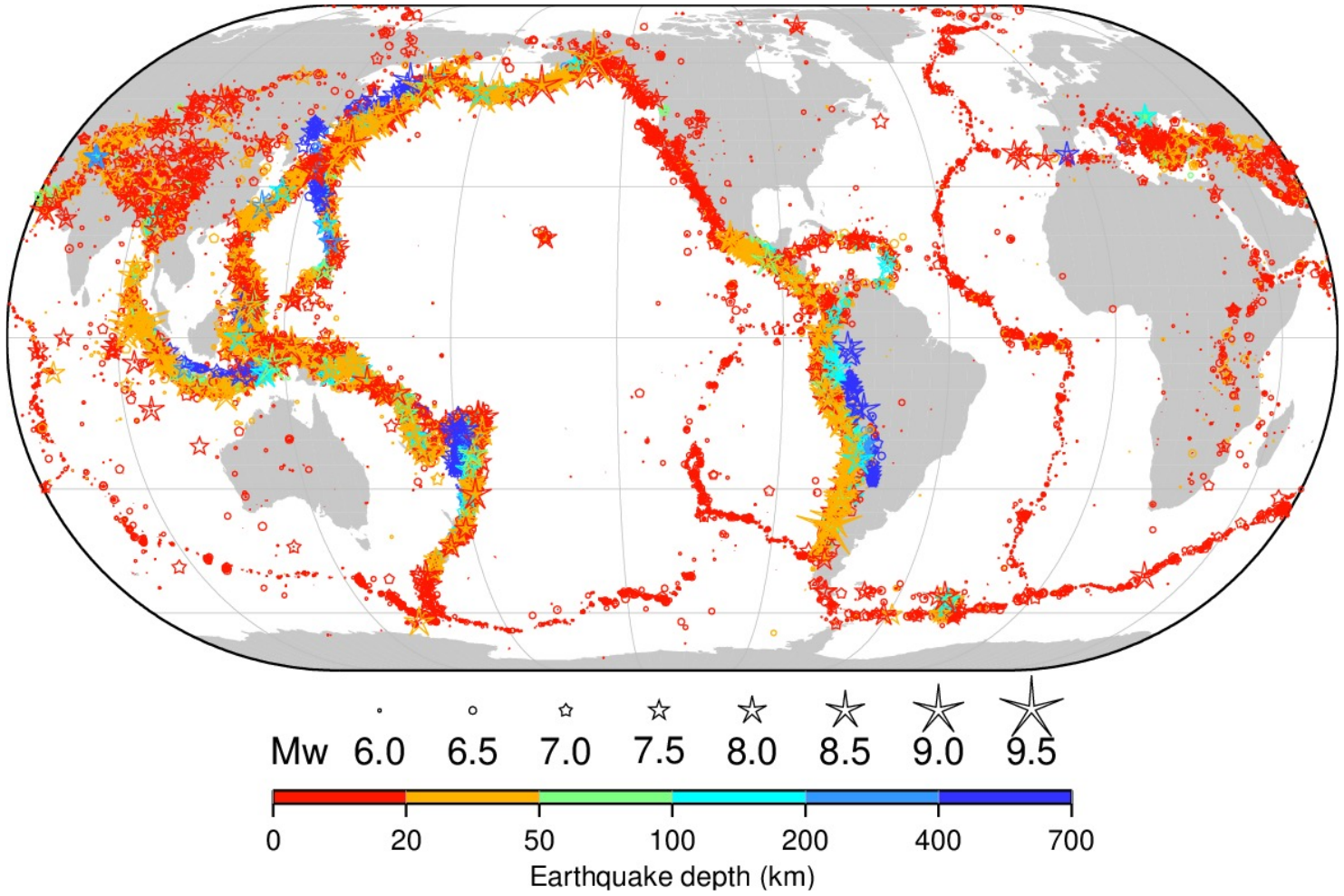



```
#YYY/MM/DD HH:mm:SS.ss ET GT MAG  M     LAT     LON     DEPTH Q  EVID       REMARKS
2023/03/06 04:37:29.14  eq l 0.35 l    33.486 -116.400  12.5 A 40421392
2023/03/06 04:38:43.63  eq l 1.95 l    32.643 -115.724   9.9 A 40421400
2023/03/06 04:50:21.47  eq l 0.51 l    33.765 -116.929  10.7 A 40421408
2023/03/06 06:11:39.95  eq l 0.64 l    33.468 -116.553   8.6 A 40421416
2023/03/06 08:25:38.38  eq l 1.36 l    34.045 -118.937  10.7 B 40421424
2023/03/06 09:17:14.03  eq l 1.41 l    34.324 -118.500   7.3 A 40421432
2023/03/06 09:27:07.06  eq l 1.07 l    33.399 -116.386   2.4 A 40421440
2023/03/06 09:34:43.69  eq l 1.25 l    33.991 -118.339   9.2 A 40421448
2023/03/06 09:48:56.80  eq l 1.17 l    34.457 -117.954   8.9 A 40421456
2023/03/06 12:05:54.76  eq l 1.51 l    32.983 -115.525   6.0 A 40421480
2023/03/06 12:46:15.72  eq l 1.36 l    34.023 -116.740  15.9 A 40421488
2023/03/06 13:01:37.18  eq l 1.26 l    35.506 -118.391   8.6 A 40421496
2023/03/06 17:29:52.47  eq l 1.29 l    35.804 -117.627   6.1 A 40421512
2023/03/06 17:40:07.96  eq l 0.31 l    33.671 -116.686  16.1 A 40421520
2023/03/06 18:11:35.09  eq l 1.43 l    35.644 -117.385  10.3 A 40421528
2023/03/06 18:20:11.62  eq l 2.72 l    35.617 -117.537   7.6 B 40421536
2023/03/06 19:34:27.21  eq l 0.63 l    33.502 -116.779   4.3 A 40421552
2023/03/06 19:34:50.65  eq l 1.20 l    33.930 -116.788  18.6 A 40421560
2023/03/06 22:45:51.61  eq l 0.66 l    33.528 -116.727   4.8 A 40421576

# Number of events: 19
```

Why do we need earthquake catalogs?

- **Seismic hazard**
  - long term hazard and risk
  - short term forecasting
  - ...
- **Understanding the earth**
  - plate tectonics
  - earthquake physics
  - deformation processes
  - rheology
  - ...



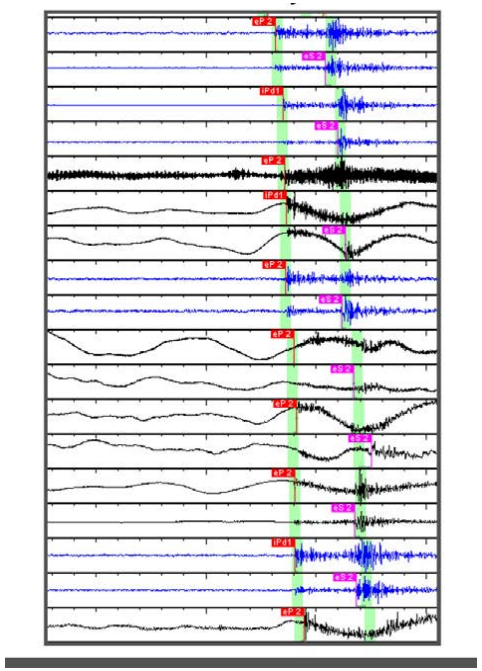
Source: https://www.isc.ac.uk/iscgem/overview.php

How do we build earthquake catalogs?

There are as many types of catalogs as there are techniques.

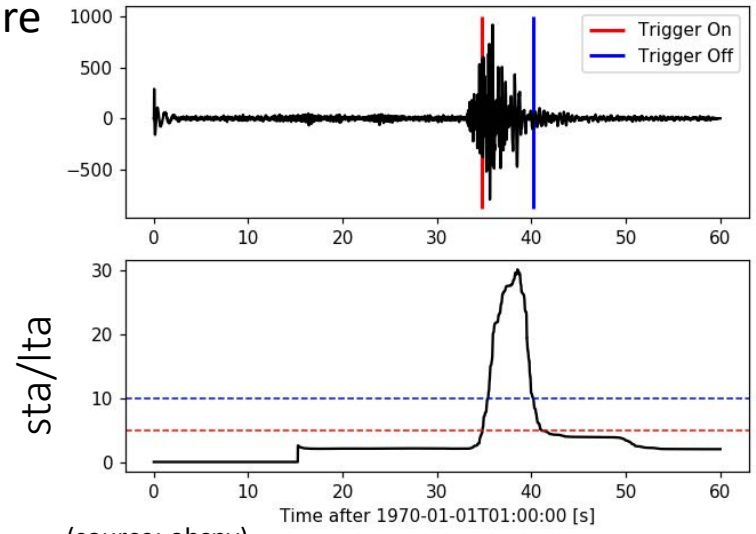
Historical Developments

Traditionally, earthquakes were all hand-picked!



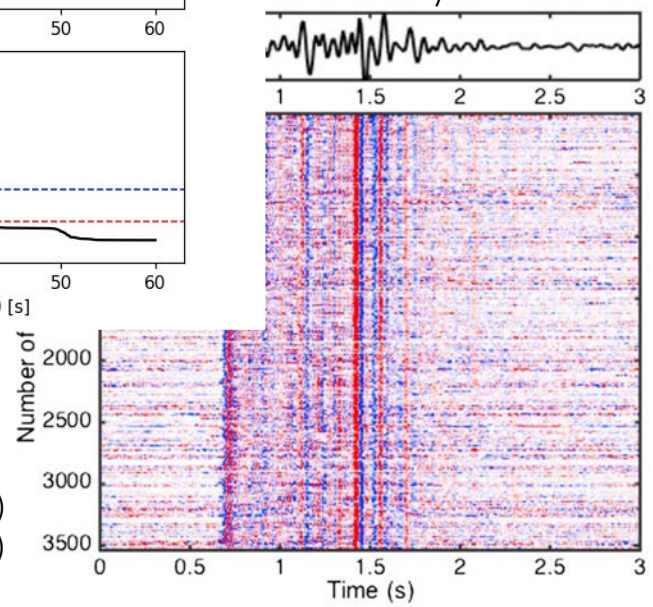
https://www.scsn.org/index.php/seismologists-tools/eqprocessing/index.html

Allen (1978)



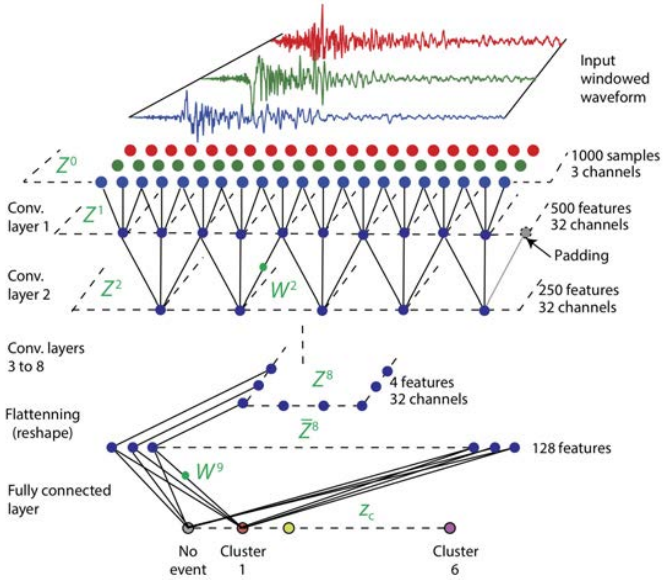
(source: obspy)

similarity-based



Huang & Beroza (2015)
Yoon et al. (2015)

Rapid advances boosted by the deep learning revolution.



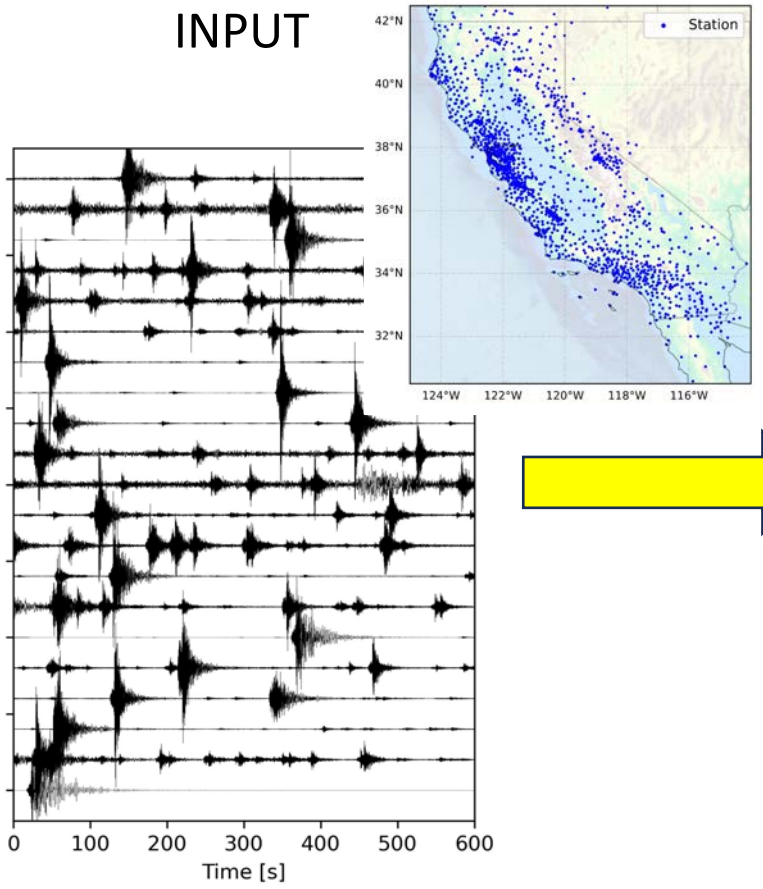
Perol, Gharbi & Denolle (2018)

The rise of automated techniques.

Generic workflow for catalog production

Workflows follow a similar structure despite differences in the techniques

INPUT



Seismic Data

Phase Detection & Picking

Phase Association

Earthquake Location

Earthquake Magnitude

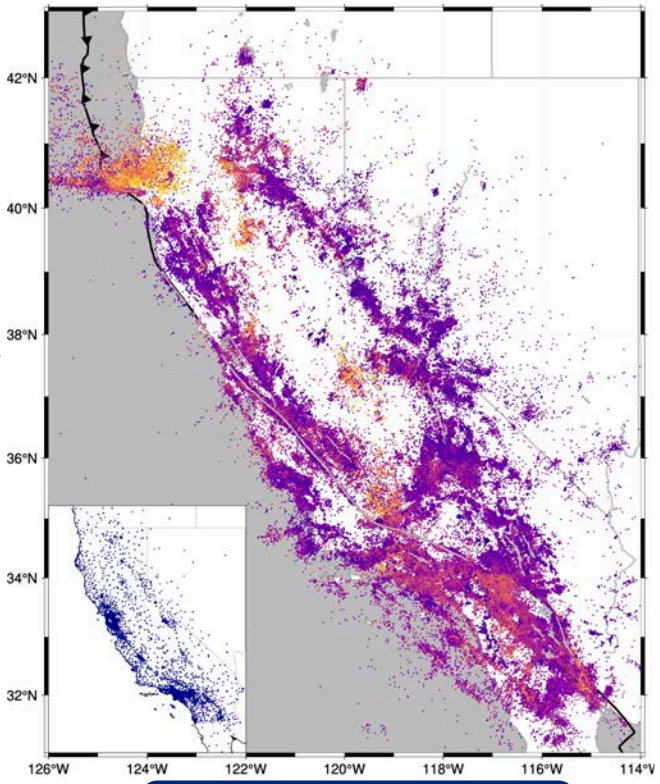
Focal Mechanism / Moment Tensor

First motion polarity

S/P amplitude ratio

Full-waveform detection & location

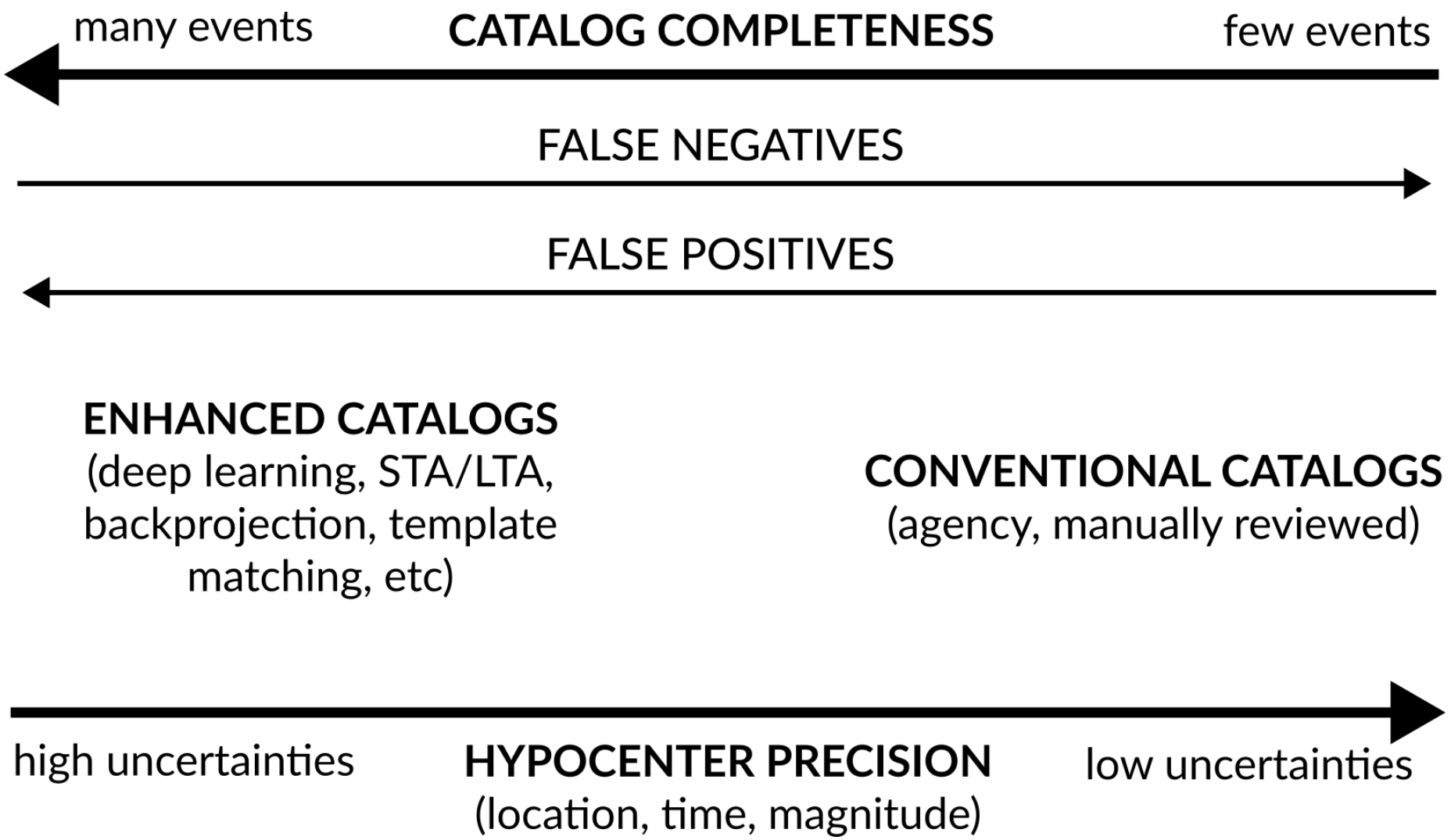
OUTPUT



Earthquake Catalog

Different catalogs, different strengths and weaknesses

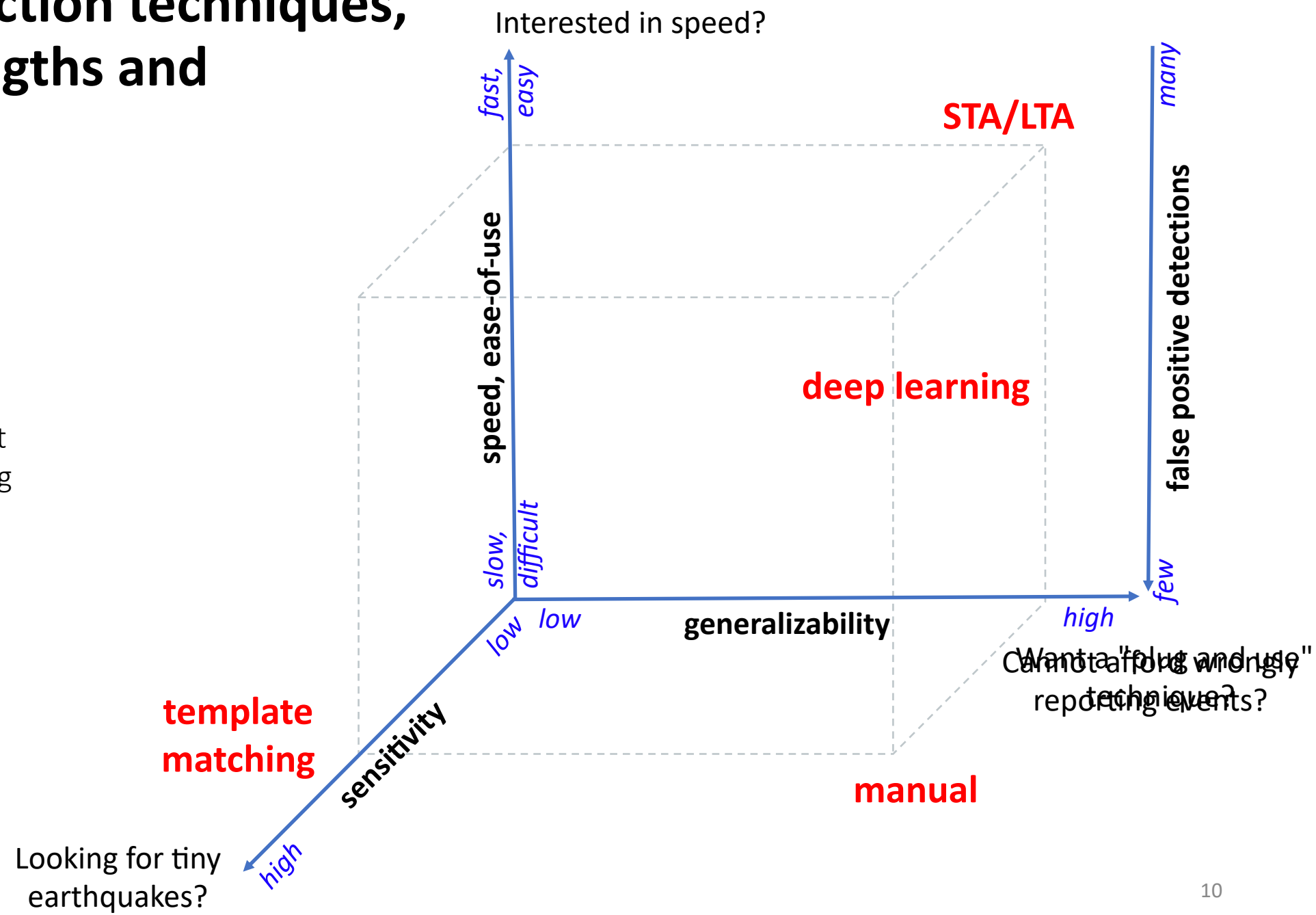
The completeness vs accuracy trade-off



Different detection techniques, different strengths and weaknesses

Purpose and resources guide your choice.

You will learn more about these techniques all along the workshop!



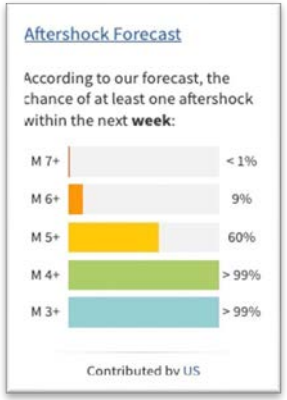
What are some **use-cases** of earthquake catalogs?

Less complete, more accurate

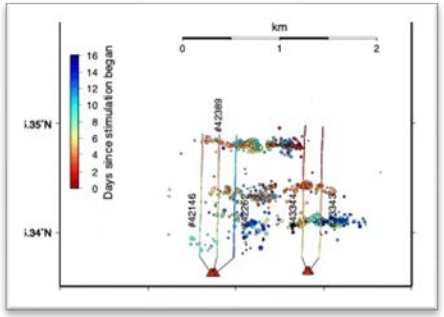
More complete, less accurate

Real-time earthquake monitoring

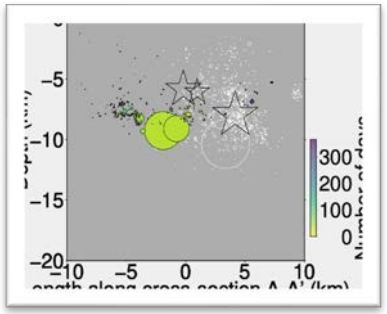
Operational aftershock forecasting



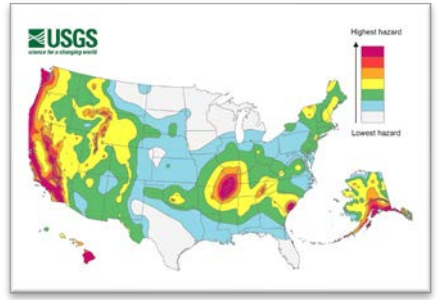
Understand earthquake processes



Outline active fault structure at depth



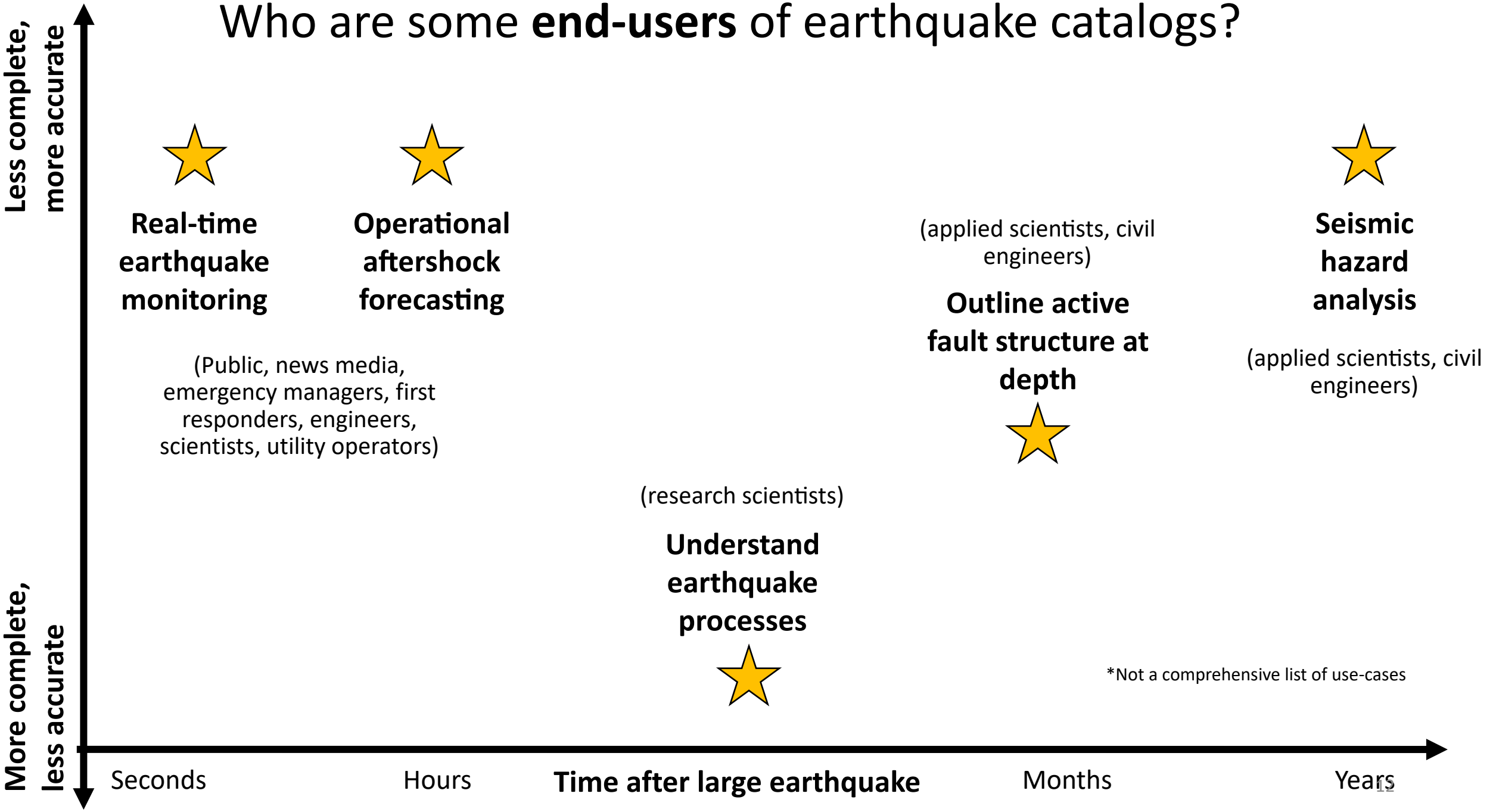
Seismic hazard analysis



*Not a comprehensive list of use-cases

Seconds | Hours | **Time after large earthquake** | Months | Years

Who are some **end-users** of earthquake catalogs?

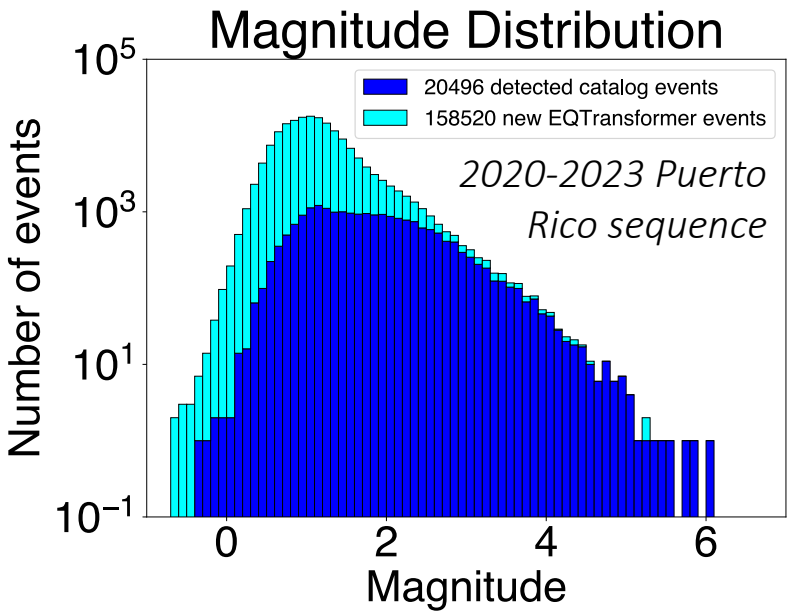


*Not a comprehensive list of use-cases

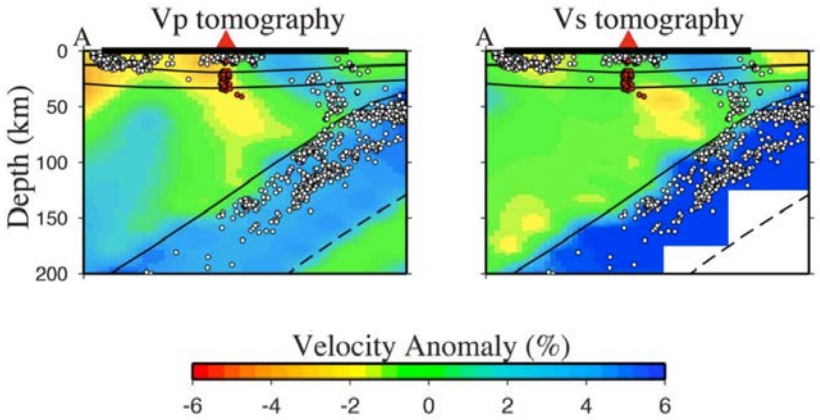
What do we learn from modern, enhanced earthquake catalogs?

Enhanced catalogs are advancing our understanding of the earth.

More complete earthquake catalogs:
e.g., fault structure, monitoring

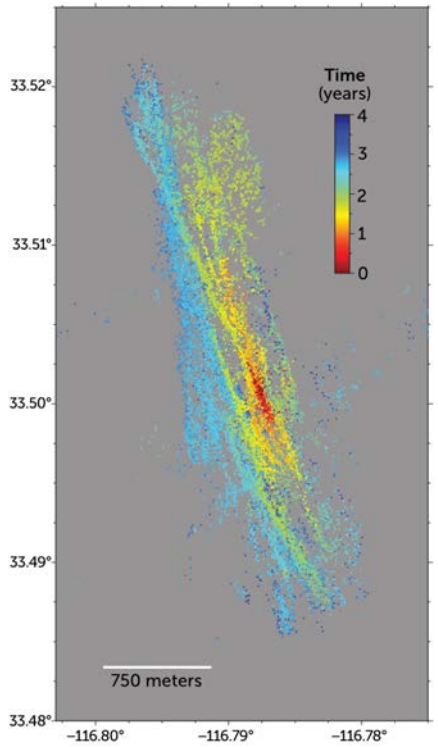


More input to infer subsurface conditions:
e.g., tomography, velocity models



Wang et al. (2019) - Japan

More detail about earthquake processes:
e.g., space-time evolution, fluid migration, stress transfer



Ross et al. (2020) - California

Learning goals

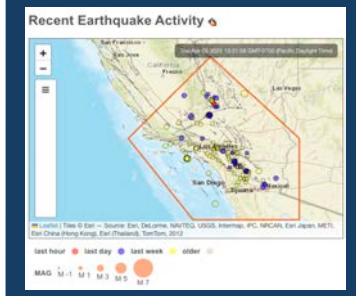
- Access and manipulate earthquake catalogs
- Different types of catalogs
- Different techniques used to build them
- Quality control and critical thinking
- Which catalog for which purposes (~~best catalog~~)

What are some **use-cases** of earthquake catalogs?

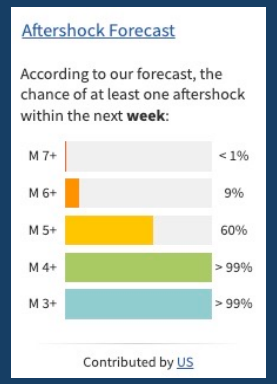
Less complete, more accurate

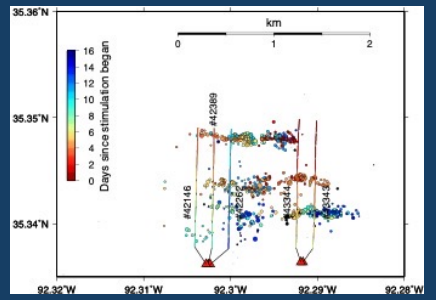
More complete, less accurate

Real-time earthquake monitoring

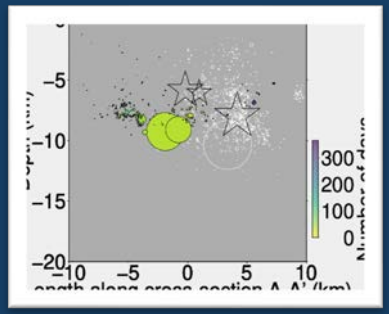


Operational aftershock forecasting



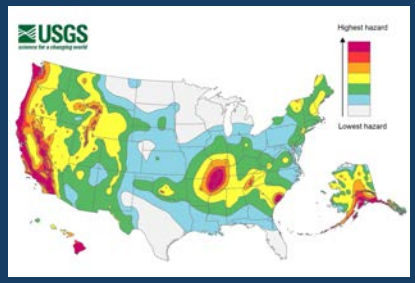


Understand earthquake processes



Outline active fault structure at depth

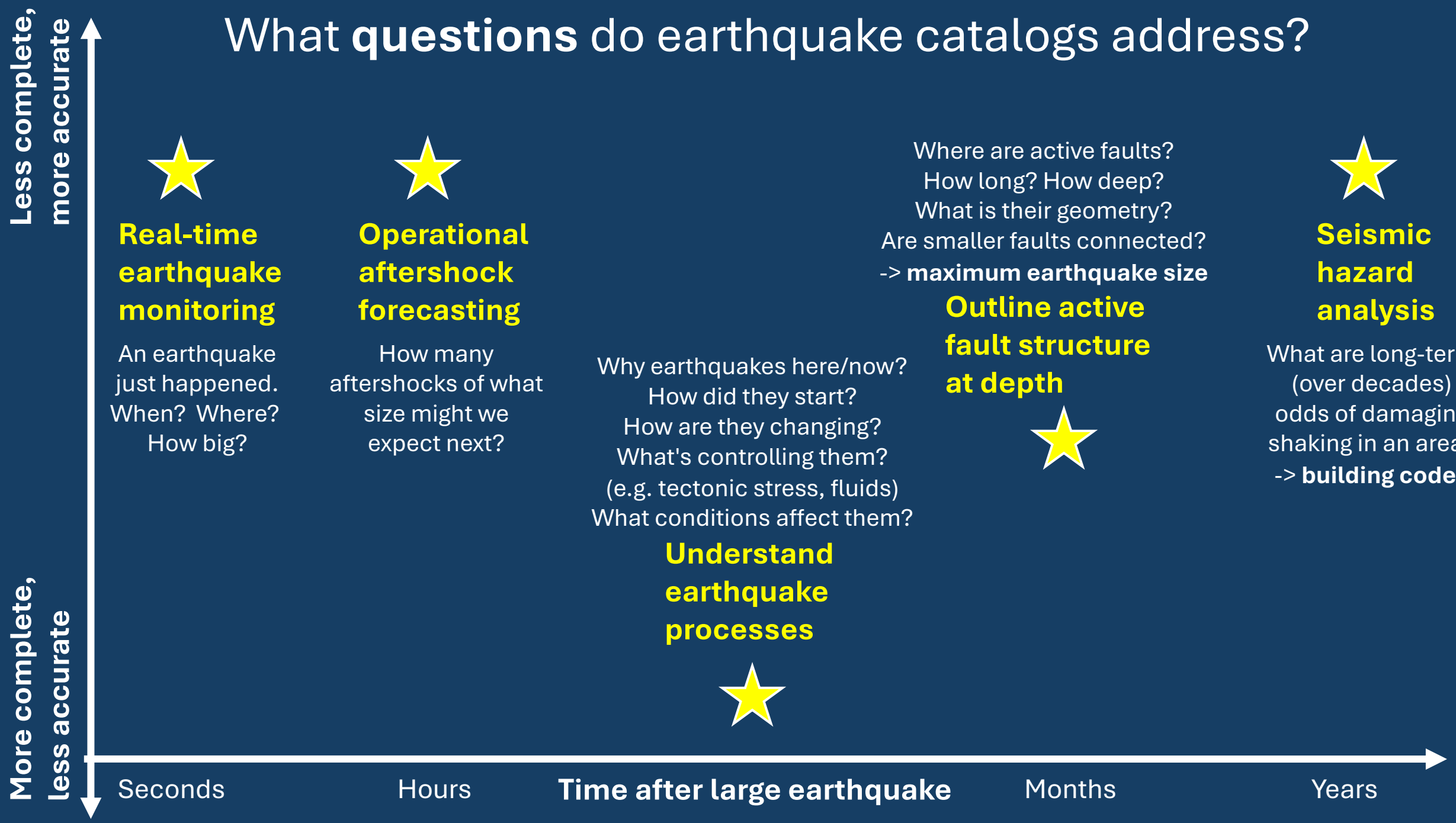
Seismic hazard analysis



*Not a comprehensive list of use-cases

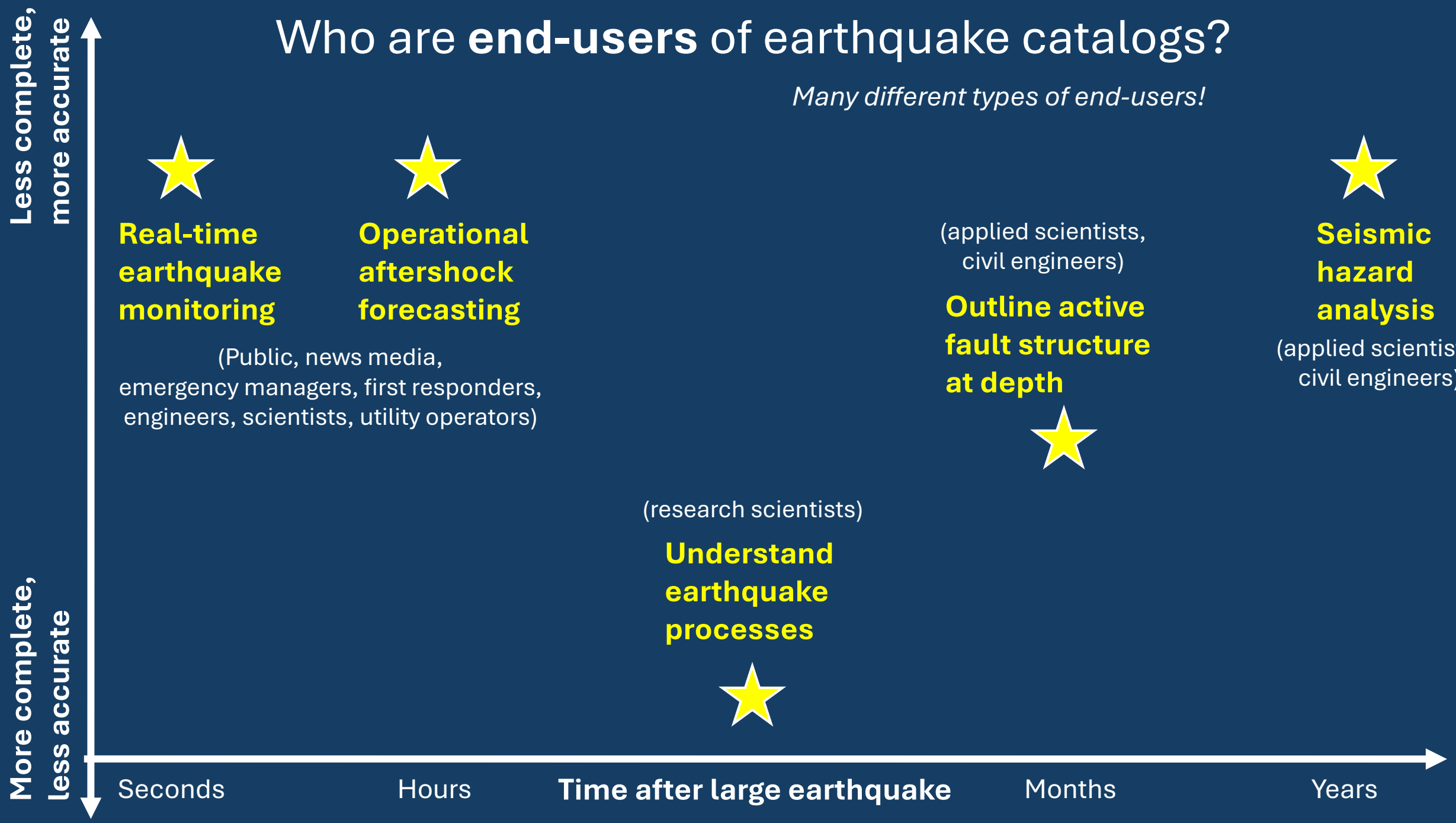
Seconds | Hours | **Time after large earthquake** | Months | Years

What **questions** do earthquake catalogs address?



Who are **end-users** of earthquake catalogs?

Many different types of end-users!



What do **end-users need** from earthquake catalogs?

It varies!

Less complete, more accurate

More complete, less accurate

Real-time earthquake monitoring

Speed, reliability, accuracy of earthquake catalog, for situational awareness.

Need only events

Operational aftershock forecasting

Speed, quality. Accurate **magnitudes**: essential

Mc ~ 3 ok, but need accurate Mc

Need only events

Understand earthquake processes

More complete catalogs (low Mc): essential

Precise locations: useful
Need phase picks & event waveforms too

Outline active fault structure at depth

Precise locations: essential
Need phase picks & event waveforms too

Focal mechanisms / moment tensors: useful

More complete catalogs (low Mc): useful, for sparse network

Seismic hazard analysis

Accurate, consistent magnitude (Mw) & seismicity rates: essential

Mc ~ 5 ok, but need accurate Mc

Catalog continuity over decades: essential

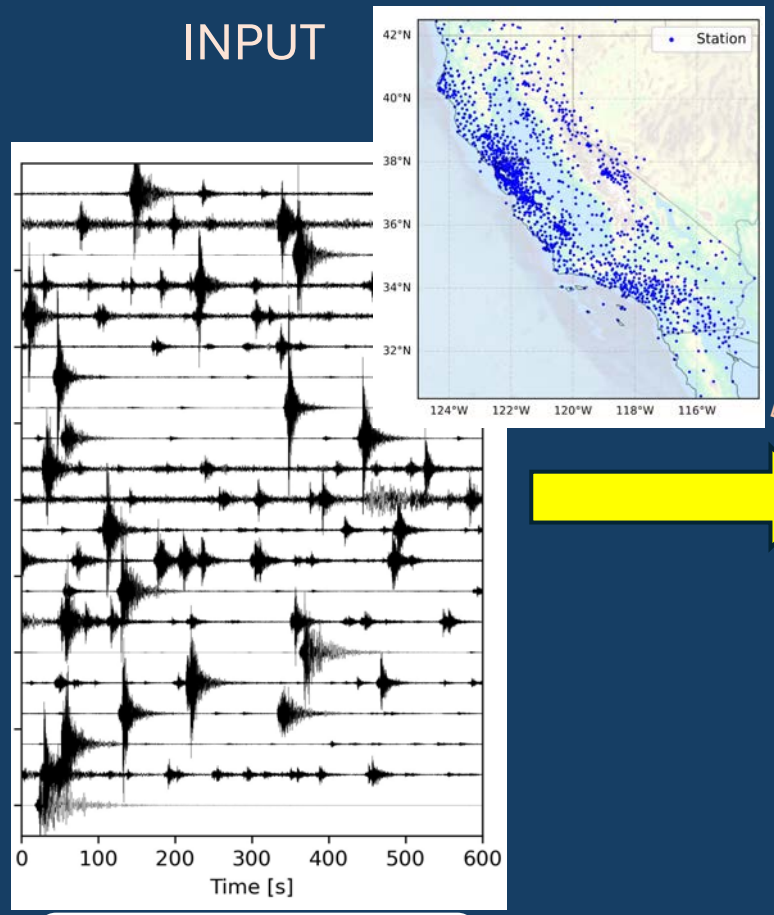
Need only events

*Mc: magnitude of completeness (lower is better)

Seconds | Hours | **Time after large earthquake** | Months | Years

How are earthquake catalogs produced?

INPUT



Seismic Data

First motion polarity

S/P amplitude ratio

Phase Detection & Picking

Phase Association

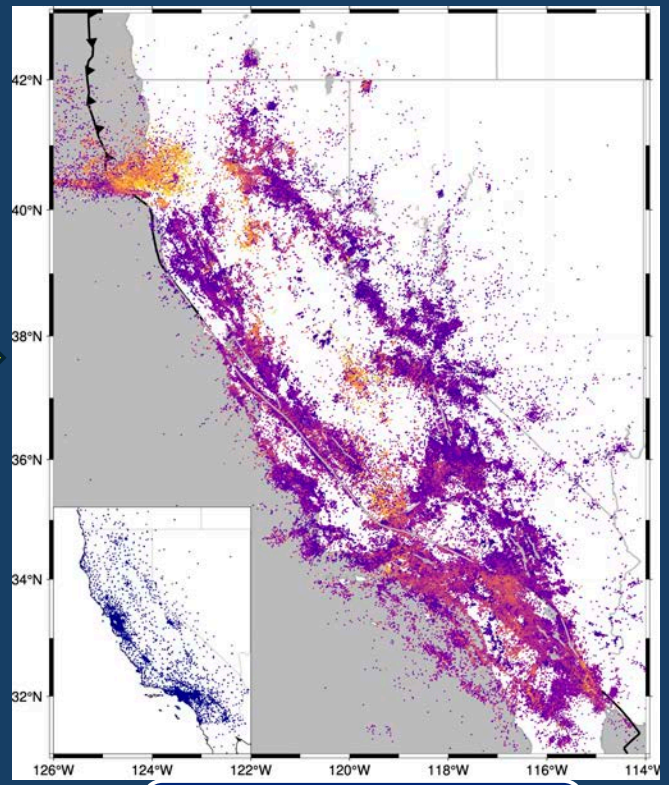
Earthquake Location

Earthquake Magnitude

Focal Mechanism / Moment Tensor

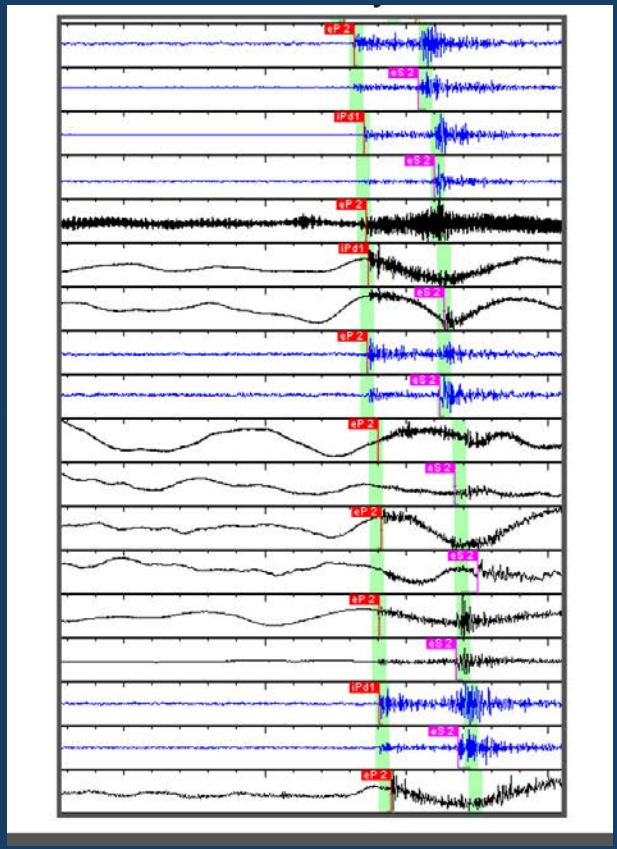
Full-waveform detection & location

OUTPUT



Earthquake Catalog

What methods exist for creating custom earthquake catalogs?



Manual
- Less common now, but still used if event quality is important

Simple/traditional automatic
- Event detection, phase-picking: STA/LTA
- Association: grid-search, back-projection
- Location: grid-search, linearized-inversion

High-Resolution: Machine Learning
- Event detection, phase-picking: deep-learning/neural networks
- Association: Bayesian Gaussian Mixture Model (GaMMA)
- Location: graph neural networks

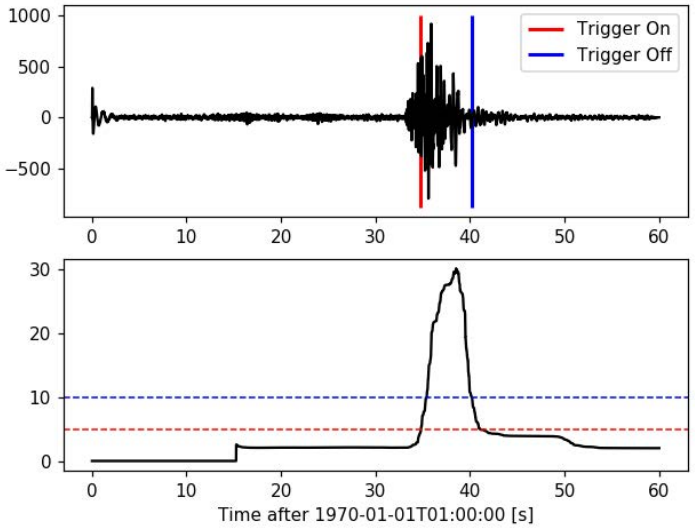
High-Resolution: Template Matching
- Event detection + Association: waveform similarity (cross-correlation) with template waveform (at multiple stations) of known earthquake

High-Resolution: Relocation
- Location: relative (double-difference)

How do different event detection & phase picking methods compare?

1. Characteristic-Function Based

- (e.g., STA/LTA)



(source: obspy)

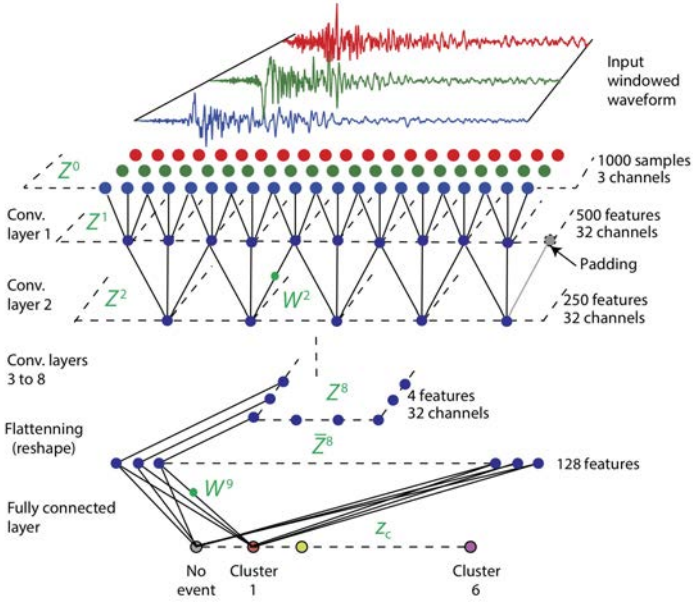
Allen (1978)

Pros:
Fast
Generalized

Cons:
Sensitive to Noise
High False Positives

2. Neural-Network Based

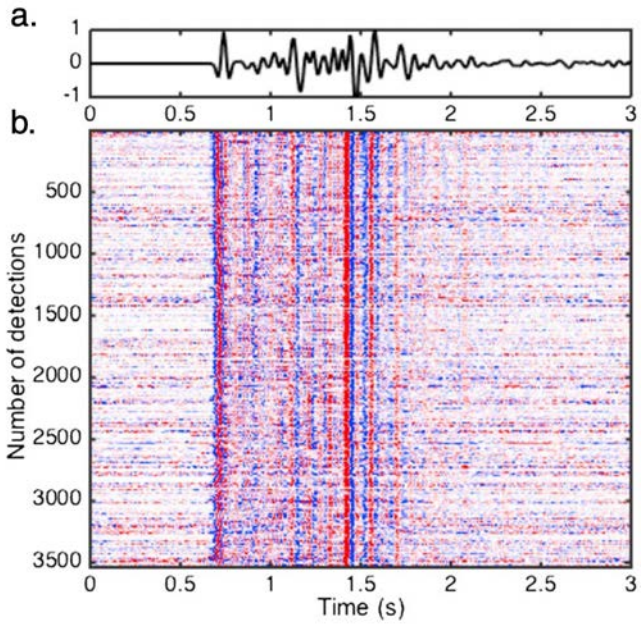
(e.g., CNN)



Perol, Gharbi & Denolle (2018)

3. Similarity-Search Based

(e.g., Template Matching, FAST)



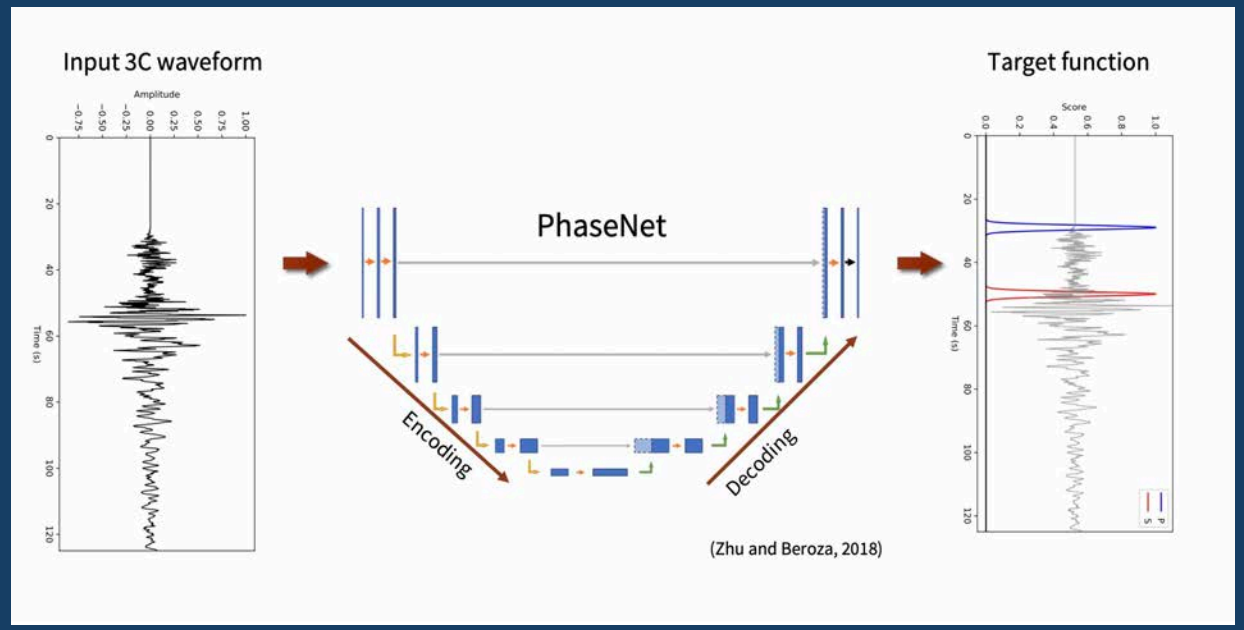
Huang & Beroza (2015)
Yoon et al. (2015)

Pros:
Robust
Low False Positive

Cons:
Limited to Similar Events
High Computational Time

High-resolution methods for event detection & phase-picking

Machine (deep) learning



Template-matching

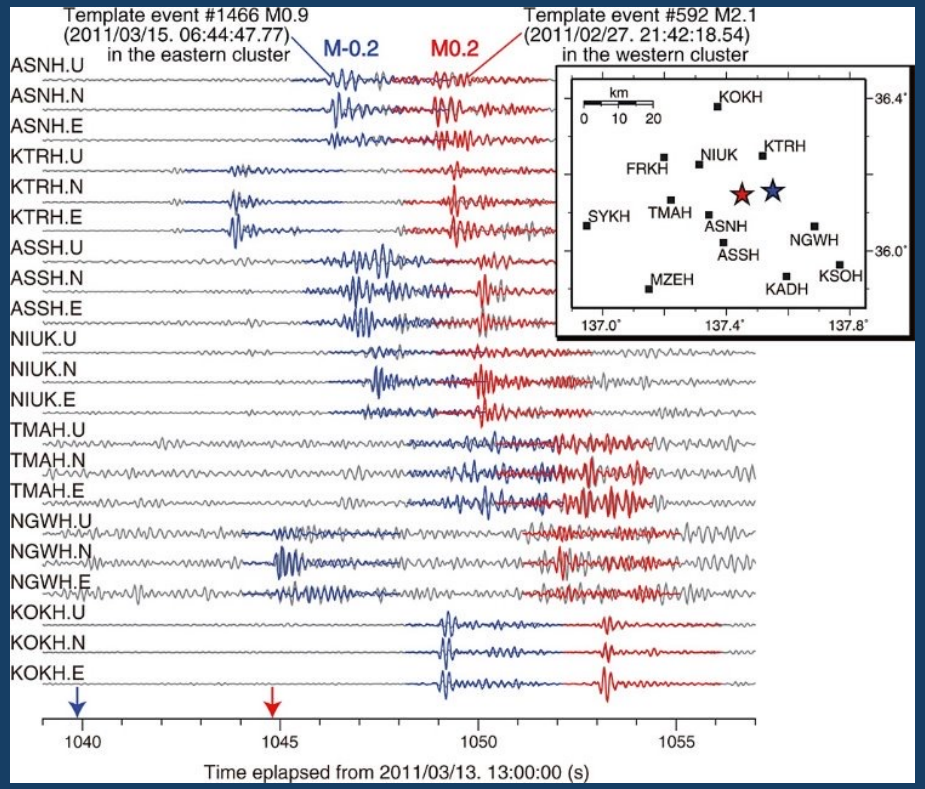
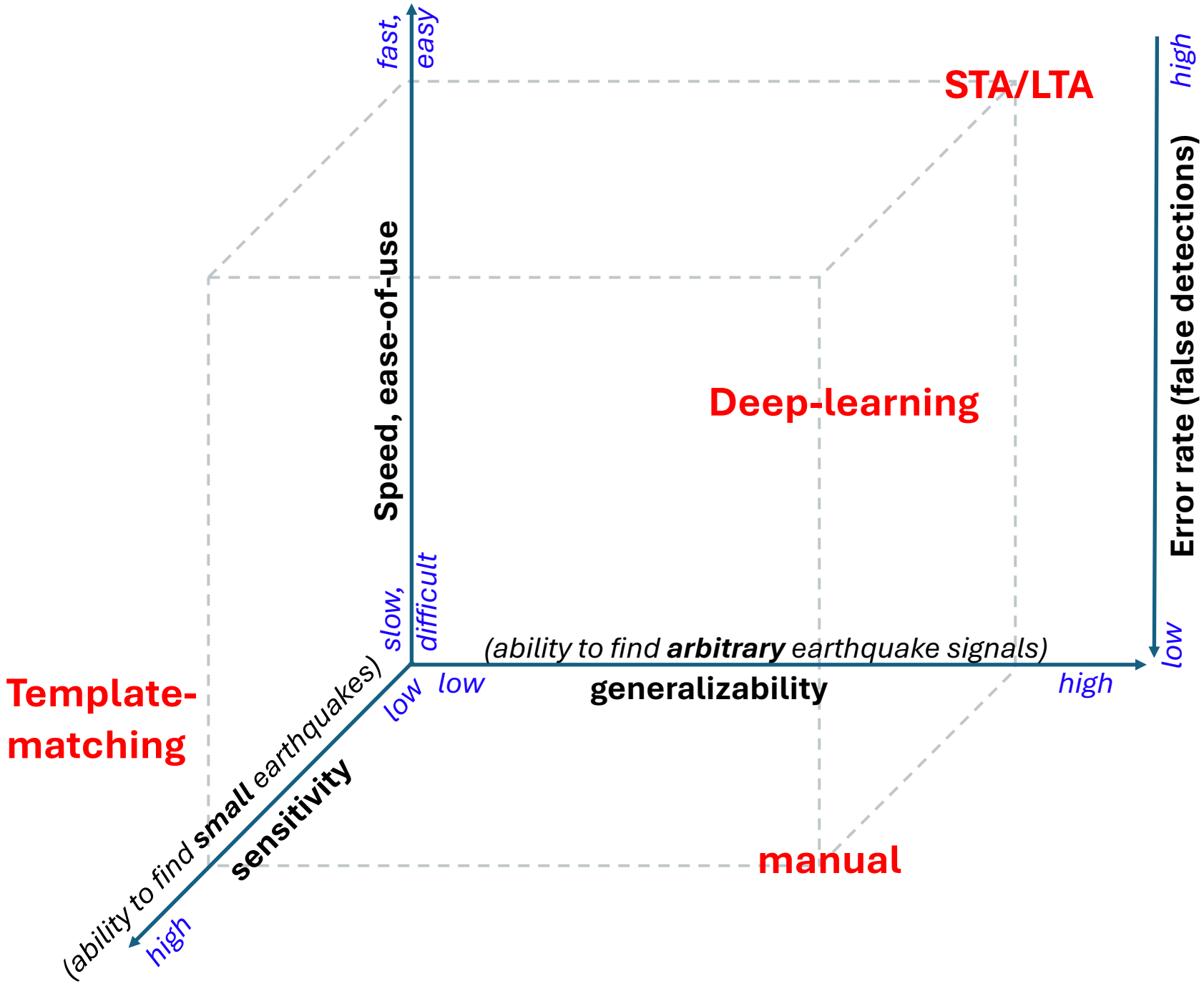


Figure: Kato et al. (2013)

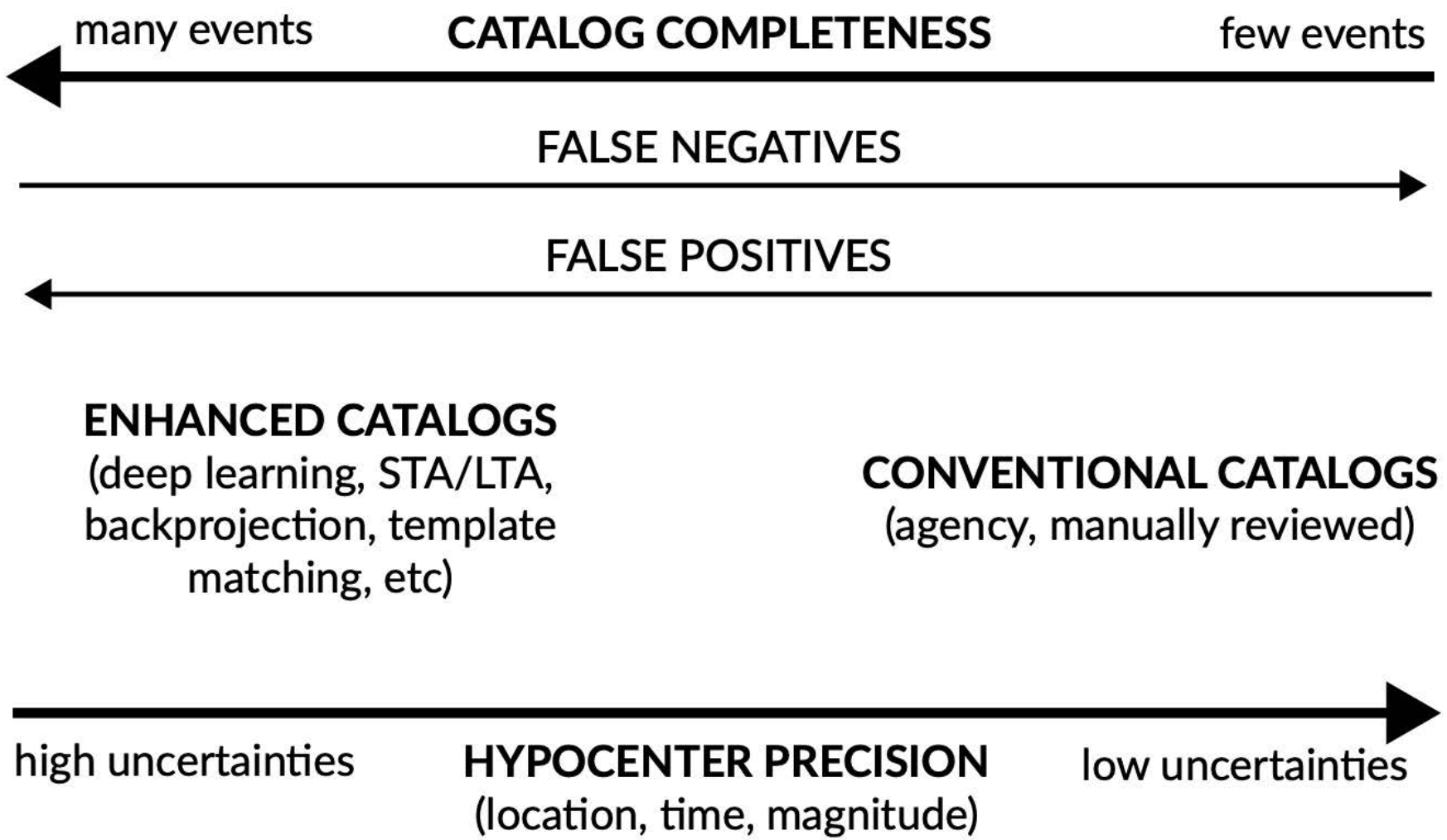
How do different event detection & phase picking methods compare?

Each method has their strengths & weaknesses

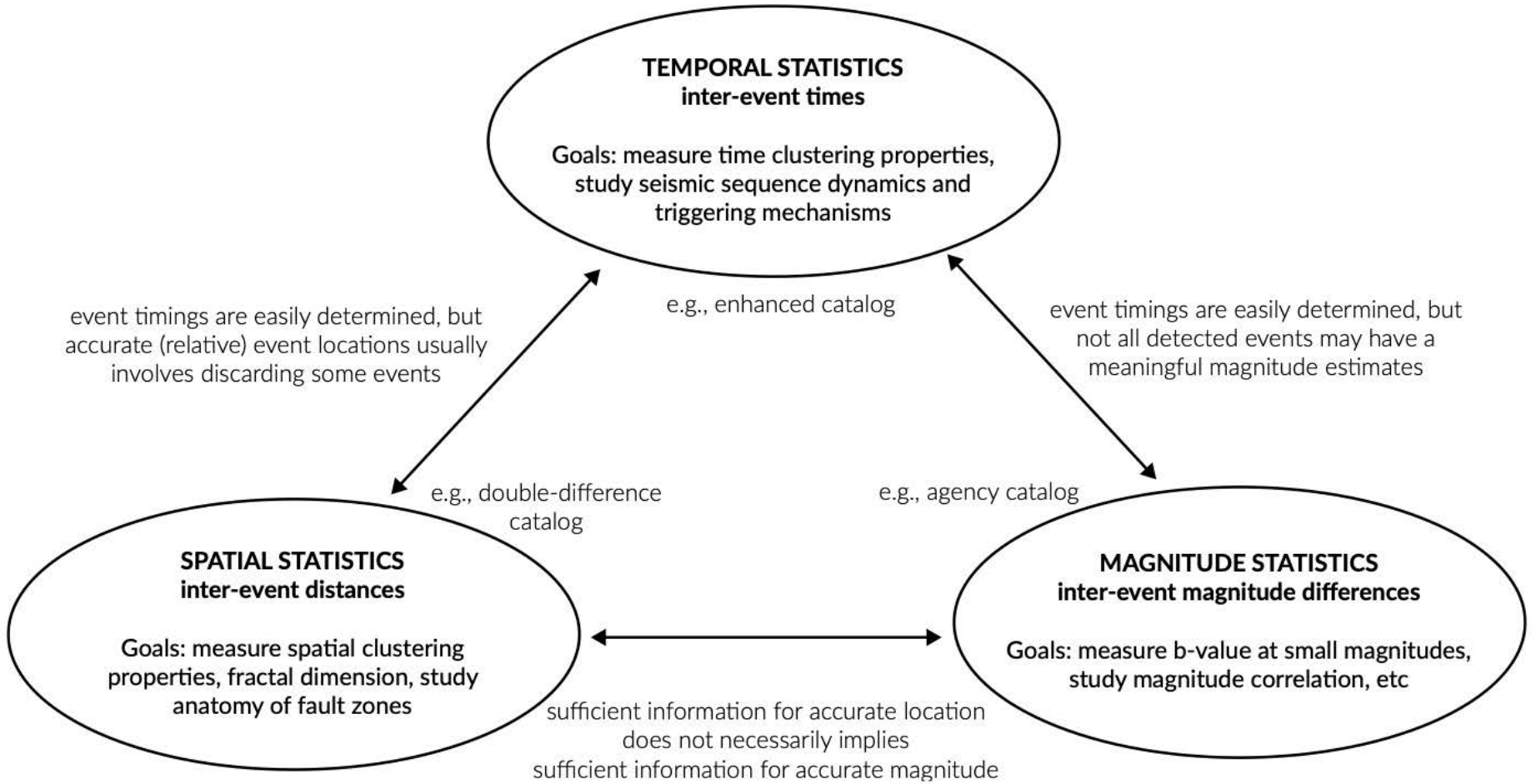
Note **tradeoff**:
speed/ease-of-use & error rate



Completeness vs precision trade-off

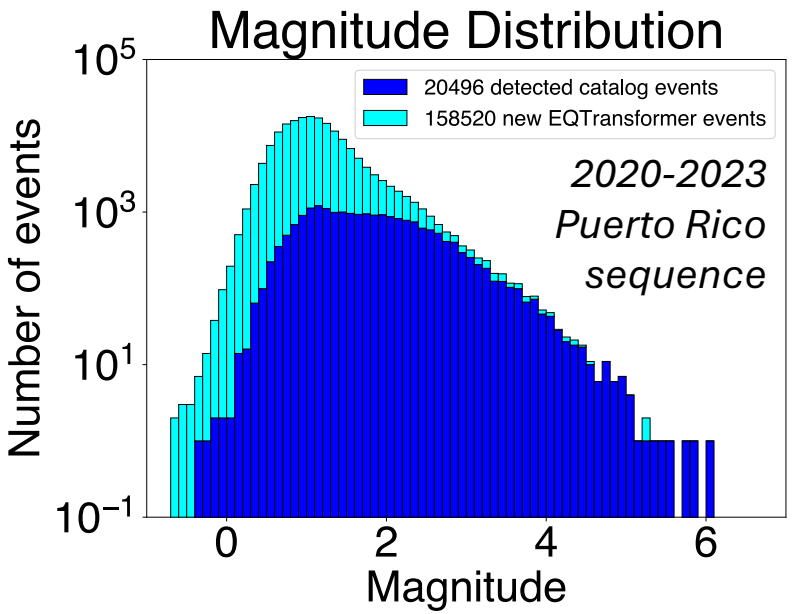


Hypocentral parameter precision trade-off

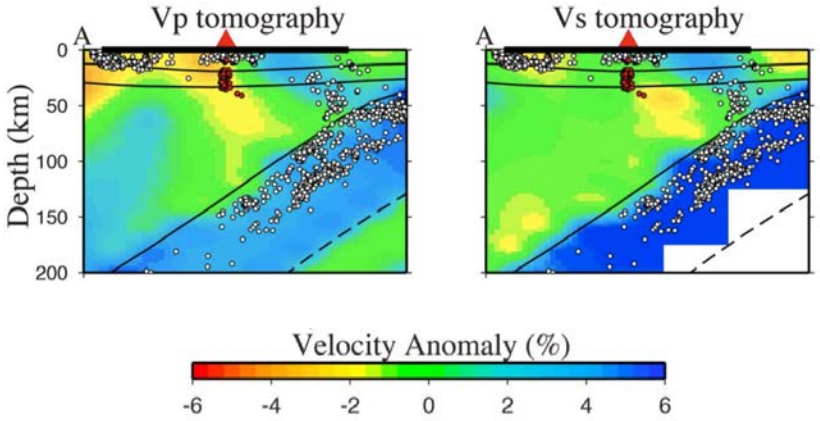


Modern algorithms excel at finding small earthquakes, with many scientific benefits

More complete earthquake catalogs:
fault structure, monitoring

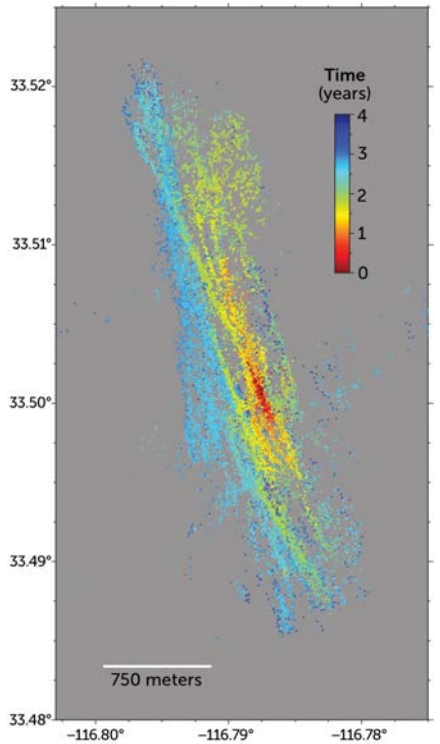


More input to infer subsurface conditions:
tomography, velocity models



Wang et al. (2019) - Japan

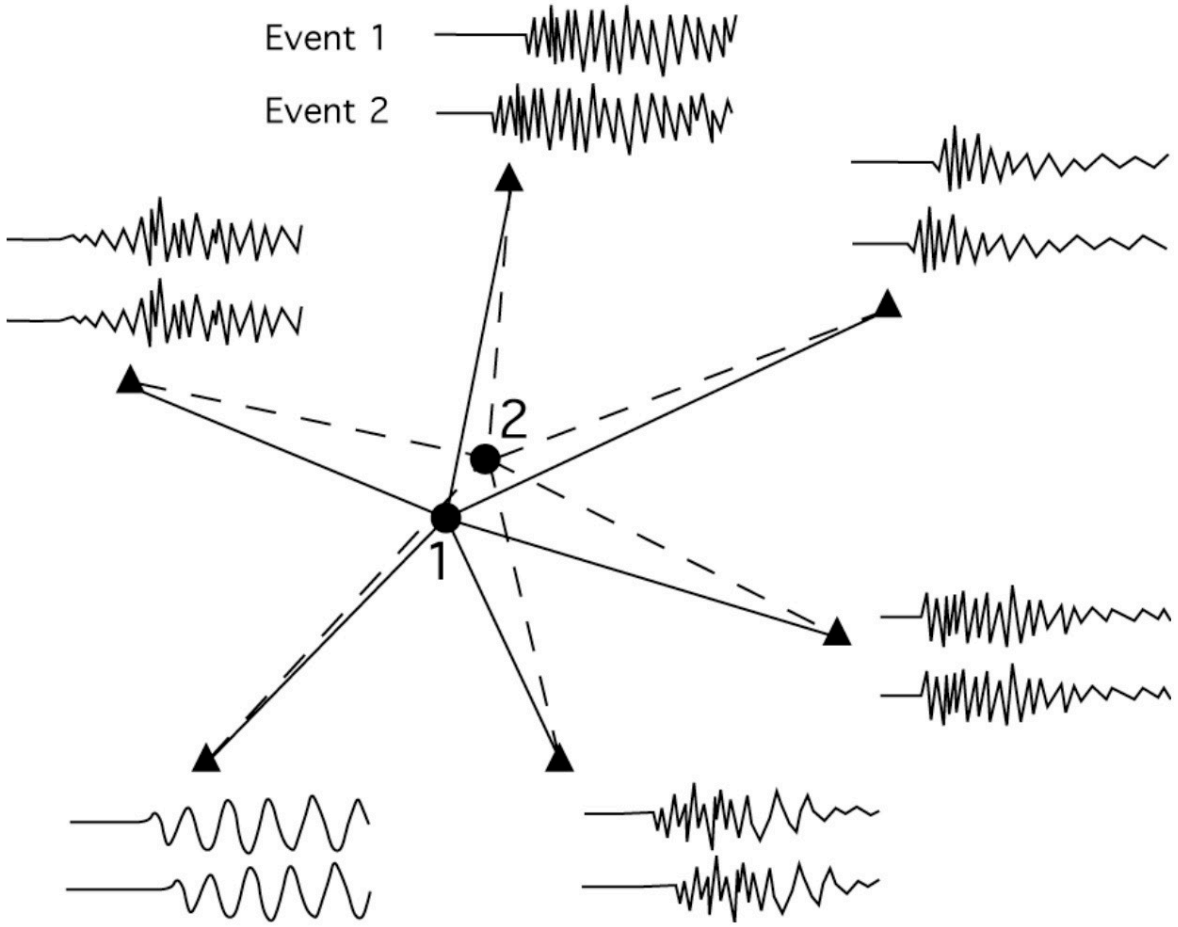
More detail about earthquake processes:
space-time evolution, fluid migration, stress transfer



Ross et al. (2020) - California

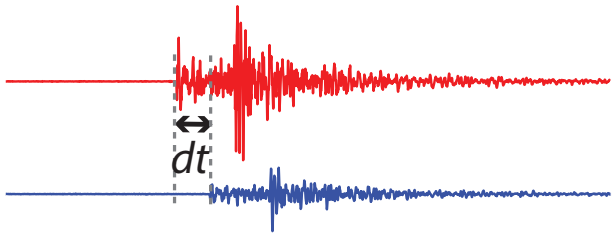
Relative event location using waveform cross-correlation

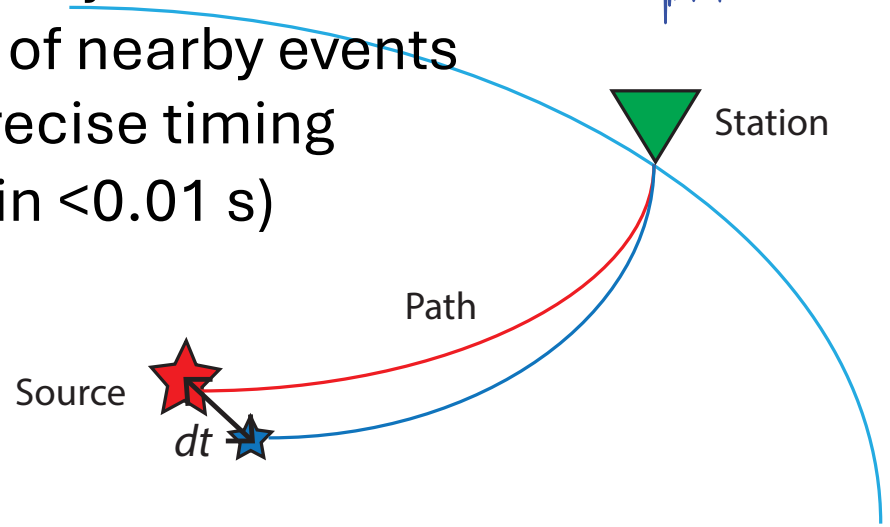
"Relocation"



https://igppweb.ucsd.edu/~shearer/ERI/3_location.pdf

Cross-correlation:
exploit waveform similarity between pairs of nearby events for precise timing (within <0.01 s)





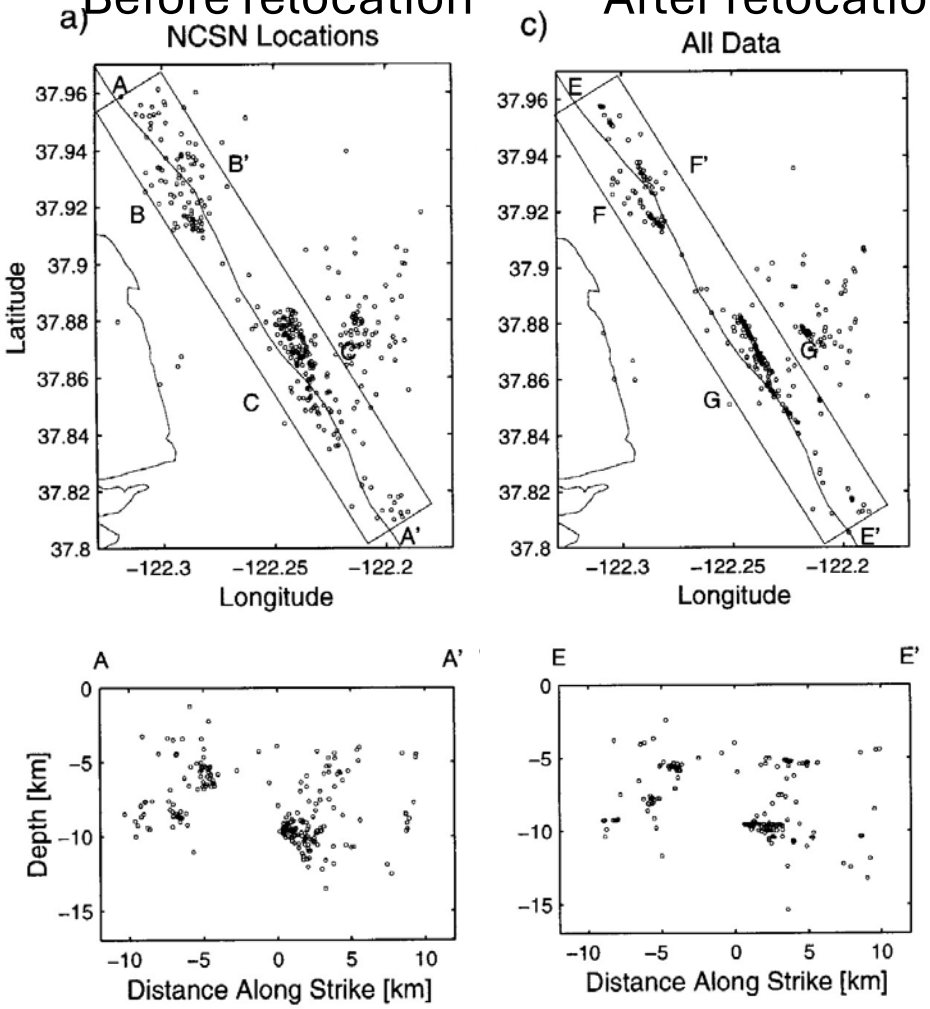
😀 Very low (meter-scale) relative location uncertainties → resolve fault structure

☹️ Absolute location still uncertain; not all events (especially larger / isolated earthquakes) are relocated

HypoDD – Double Difference

Before relocation

After relocation



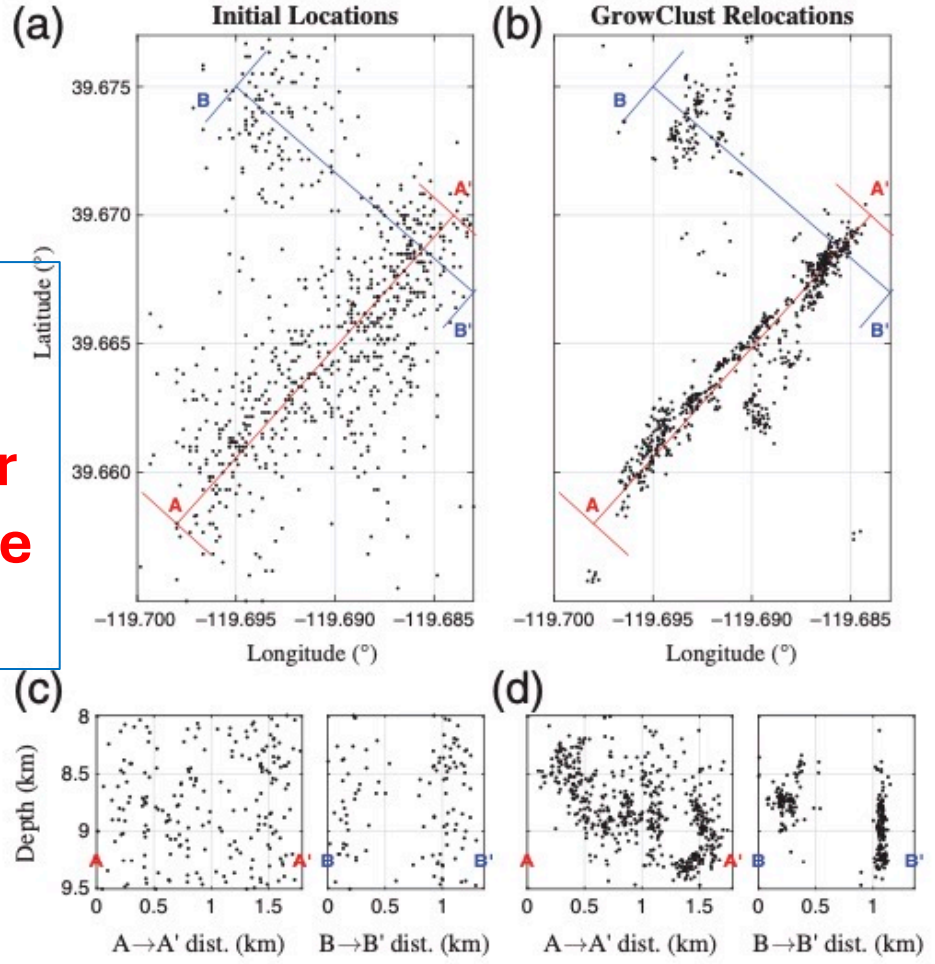
Waldhauser and Ellsworth (2000)

Relocation results are overall similar between these 2 methods

GrowClust

Before relocation

After relocation



Trugman and Shearer (2017); Trugman et al. (2022)

What methods can help these earthquake catalog use-cases?