Stanford SCHOOL OF EARTH, ENERGY & ENVIRONMENTAL SCIENCES | Geophysics

ICME

Efficient blind search for small similar-waveform earthquakes in a decade of continuous seismic data (2007-2017) in coastal central California

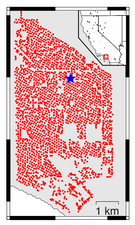
Clara Yoon[1] (ceyoon@stanford.edu), Karianne Bergen[2], Kexin Rong[3], Hashem Elezabi[3], Peter Bailis[3], Philip Levis[3], William Ellsworth[1], Gregory Beroza[1]

[1]*Department of Geophysics,* [2]*Institute for Computational and Mathematical Engineering,* [3]*Department of Computer Science, Stanford University*

Large-T similarity search for earthquake detection, in a decade of continuous seismic data, is now reality: can apply to any seismic network

We find 106 uncataloged earthquakes near Diablo Canyon nuclear power plant that may represent future earthquake sources, though local seismicity rate and magnitude are low

Big data analytics: Comprehensive Large-T Analysis to detect earthquakes with similar waveforms

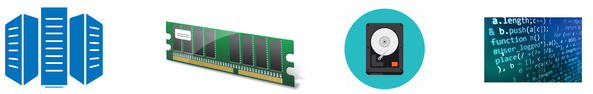


Large-N: dense seismic networks (1000's of stations)

Seismology has big data

Large-T: continuous seismic data over long time periods (>10 years)

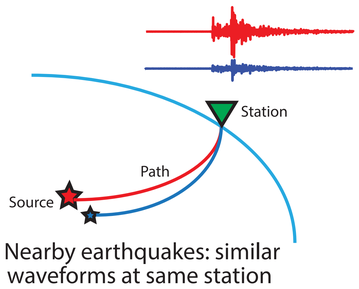
Seismology needs new scalable methods to extract information from massive data volumes



Idea: adapt data mining techniques for search in large databases

Template matching

- Use catalog events to detect smaller unknown events with similar waveforms
- But what if we do not know waveform template in advance?



Nearby earthquakes: similar waveforms at same station

New capability: comprehensive large-T analysis

- Exhaustive search of >10 years continuous seismic data to detect uncataloged low-magnitude earthquakes with similar waveforms

Research question

- Can large-T analysis reveal uncataloged small earthquakes that can illuminate unknown earthquake sources near the Diablo Canyon nuclear power plant (DCNPP) on the central California coast?
- M 6.8 possible on nearby Shoreline Fault *(Hardebeck, 2013, BSSA)*; are there other nearby faults?

Detect earthquakes (1 channel): Fingerprint and Similarity Thresholding (FAST)

Yoon et al. (2015), Sci. Adv.

10 years continuous seismic time series

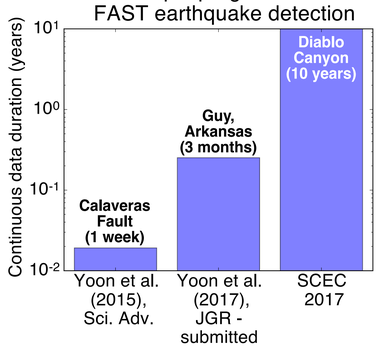
~25 GB/channel, ~500 GB total

3 x 10^8 fingerprints

~120 GB/channel, ~2 TB total

- New, optimized Python/C++ FAST software
- Search for similar fingerprint pairs, but avoid comparing the vast majority of dissimilar times in continuous data
- Runtime ~6 hours/channel (parallel, 48 cores)

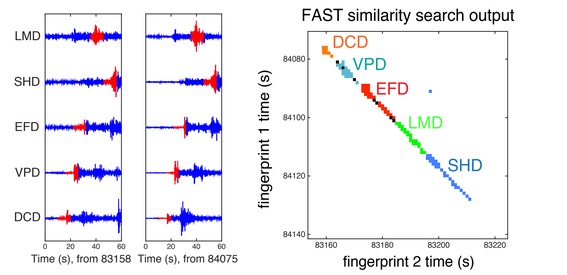
Rapid progress in FAST earthquake detection



Detect earthquakes over a seismic network

Bergen et al. (AGU 2015)

Combine matching earthquake pairs at different stations (consistent inter-event time intervals); reduce false detections



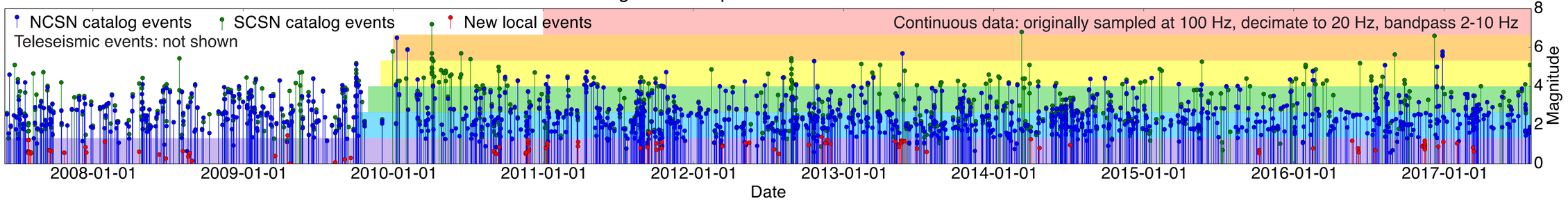
Results: Detected 2,259 earthquakes in 2007-2017 (1,741 catalog events, 412 teleseismic events, 106 new local events)

PG seismic network
6 stations (3 comp. each)

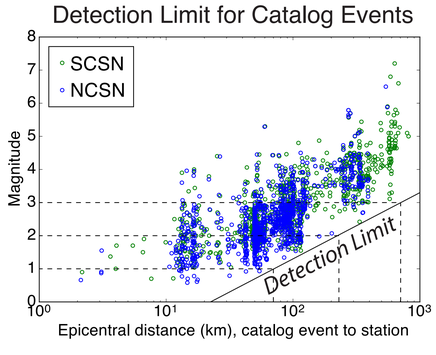
LSD
LMD
SHD
EFD
VPD
DCD

Colors: Data Duration

Large-T earthquake detection results



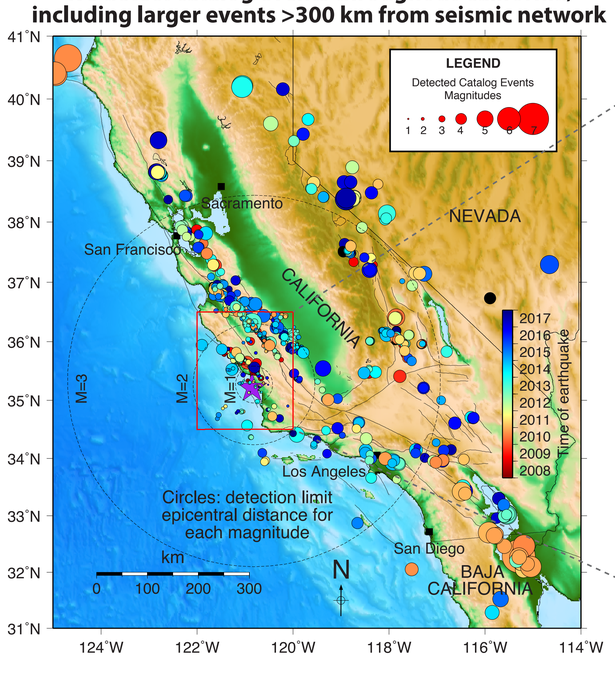
Detection Limit for Catalog Events



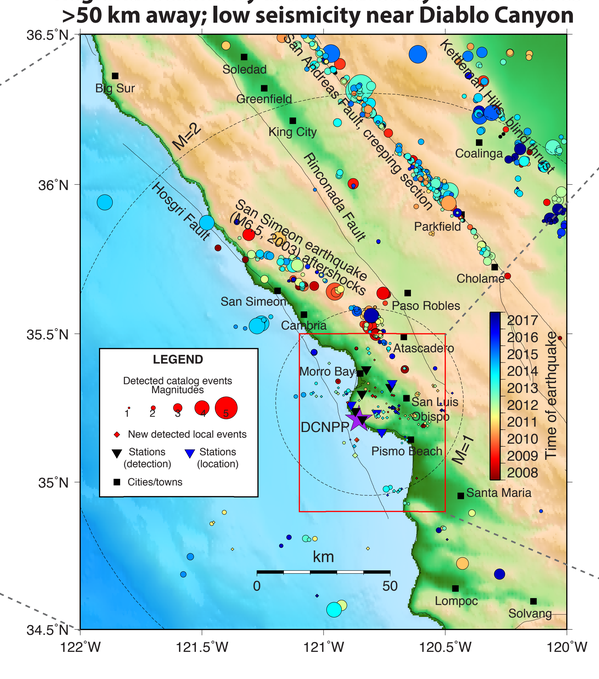
New Detected Event Analysis

- Located 85/106 new detected local events (enough P, S picks)
- VELEST (*Kissling et al., 1994, JGR*): 1D layered velocity model for Vp (*McLaren and Savage, 2001, BSSA*), Vp/Vs = 1.66 (*Hardebeck, 2010, BSSA*)
- Local magnitude ($0.3 < M_L < 1.6$) from Wood-Anderson seismograms

We detect catalog events throughout California, including larger events >300 km from seismic network



Regional seismicity is dominated by other sources >50 km away; low seismicity near Diablo Canyon



Local seismicity: We find 85 new earthquakes (M_L ~ 1) located <20 km from Diablo Canyon nuclear power plant