

Earthquake Catalog: Quality Control (QC)

- I created a custom (deep-learning) enhanced earthquake catalog automatically. How good is it?
 - No ground truth available
- Before using or interpreting your enhanced catalog, do this first:
 - Compare with a **reference catalog**
 - **Data visualization** (make plots!) to check for quality

What is a reference catalog?

- *Authoritative, high-quality* earthquake catalog for a region
 - Good baseline for comparison: manually reviewed, meets performance standards
 - **United States:** Comprehensive Catalog (**ComCat**), complete to M2.5-3.0, <https://earthquake.usgs.gov/earthquakes/search/> , <https://earthquake.usgs.gov/data/comcat/>
 - **International:** ComCat is complete to M4.5 globally. For more complete catalogs, refer to the country/region's authoritative monitoring website.
 - Turkey: <https://deprem.afad.gov.tr/event-catalog>
 - New Zealand: <https://quakesearch.geonet.org.nz/>
 - Italy: <https://terremoti.ingv.it/bsi>

Event comparison: reference catalog vs. enhanced catalog

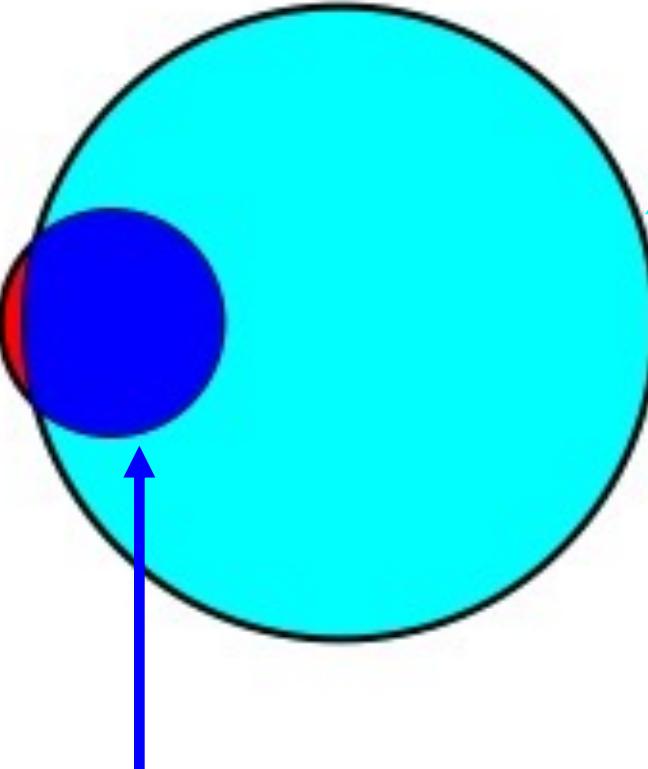
(Need locations)

MISSED events
(only in reference catalog)

- *Manually detected and/or picked by analysts* →
- *Check why they were missed (esp. larger events?)*

(small circle)

(big circle)



MATCH events (common to both catalogs)

- *Origin times match within X~5 seconds*
- *Hypocenters match within Y~25 km*

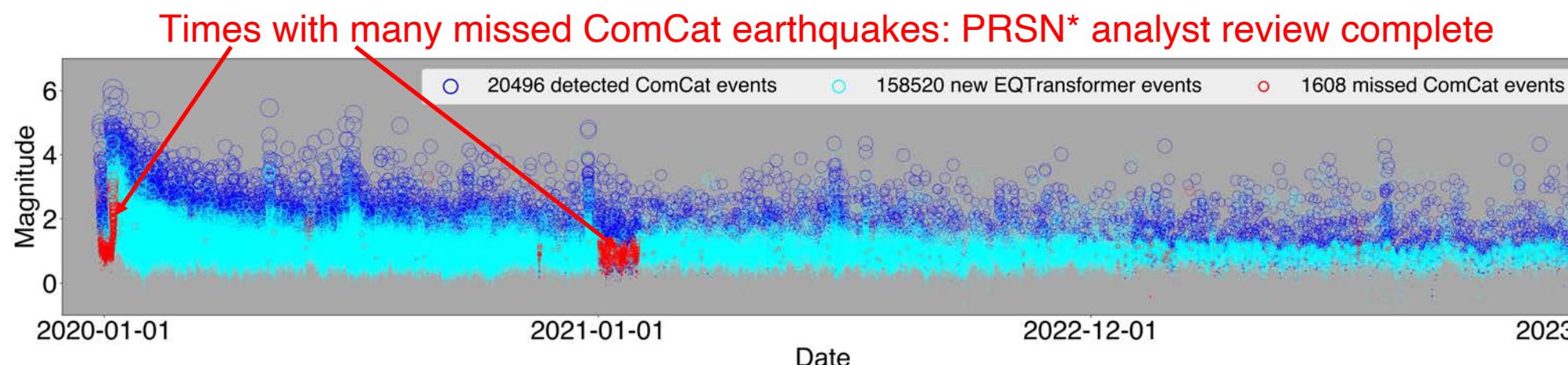
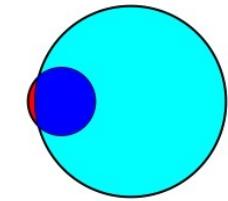
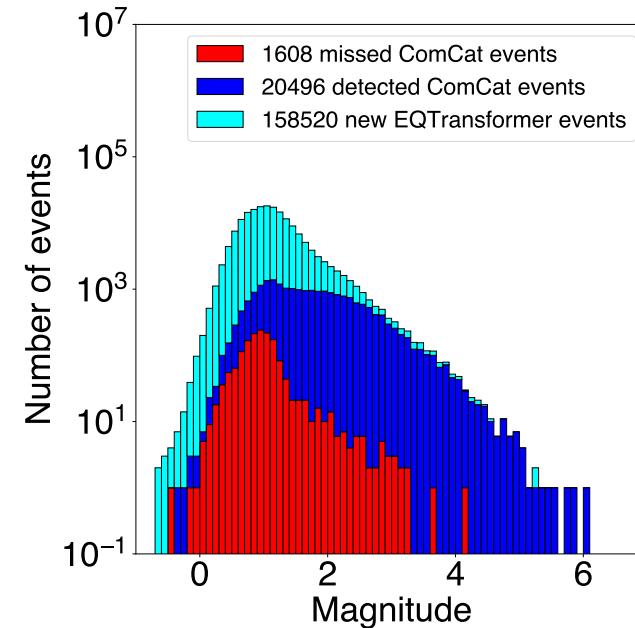
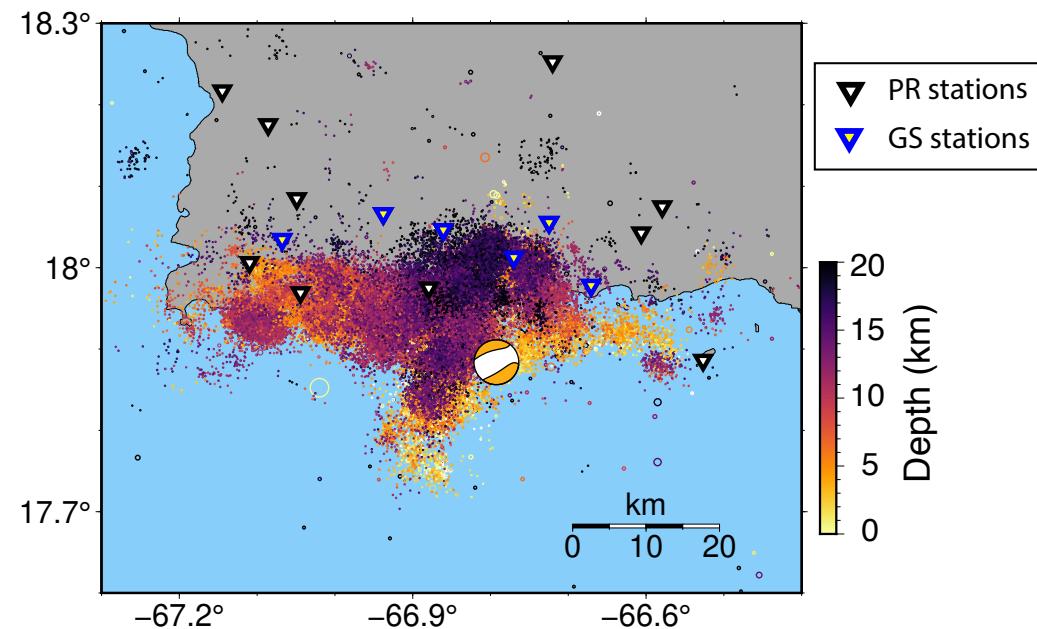
NEW events

(only in enhanced catalog)

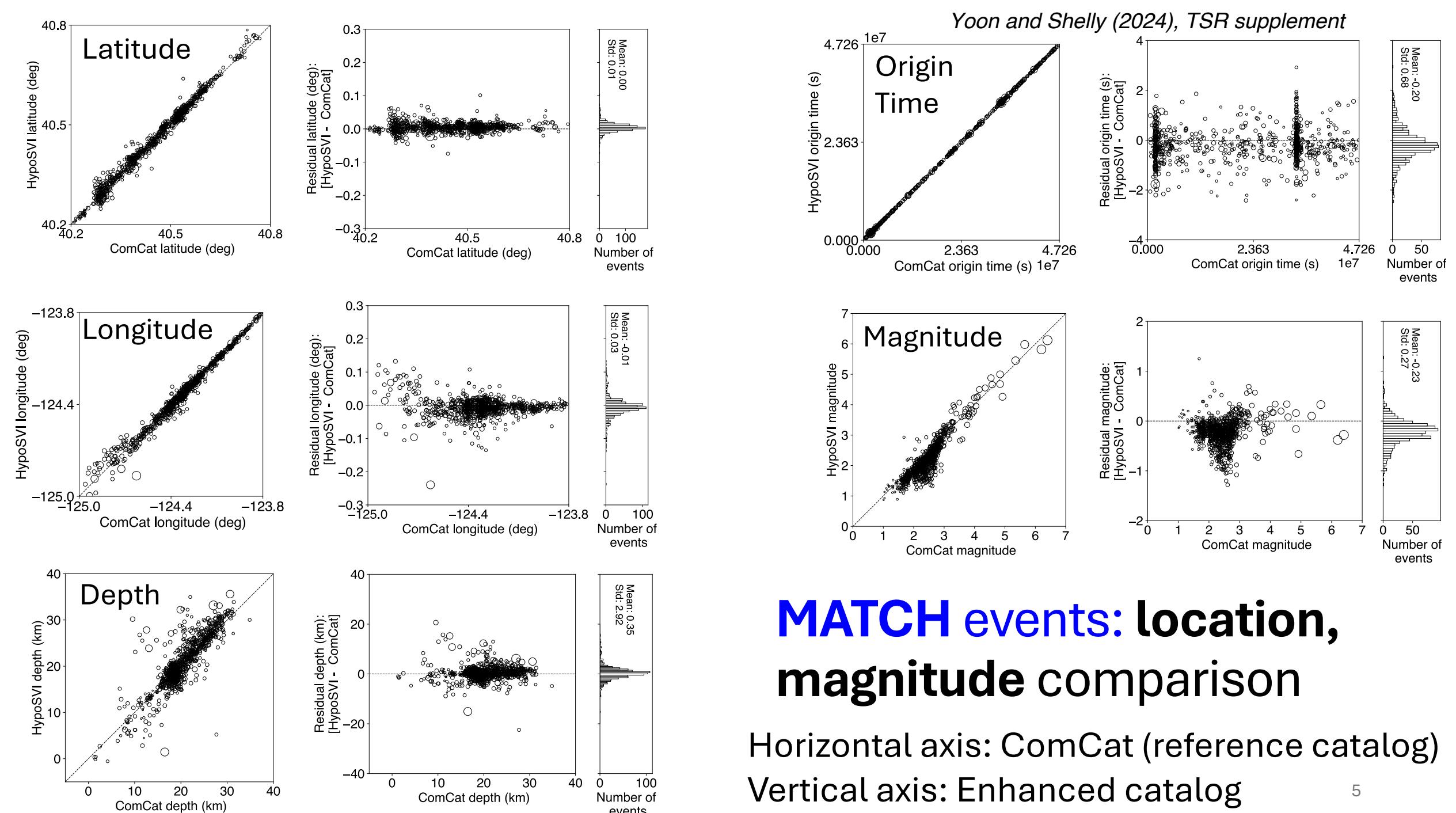
- *Newly detected small local earthquakes? 😊*
- *False detections from noise? 😞*
- *Other signals (quarry blasts, sonic booms, explosions)? 🤔*
- *Regional or teleseismic earthquakes? 😞*

Event comparison example: SW Puerto Rico sequence

Yoon et al. (2023), BSSA



*PRSN:
Puerto Rico
Seismic
Network



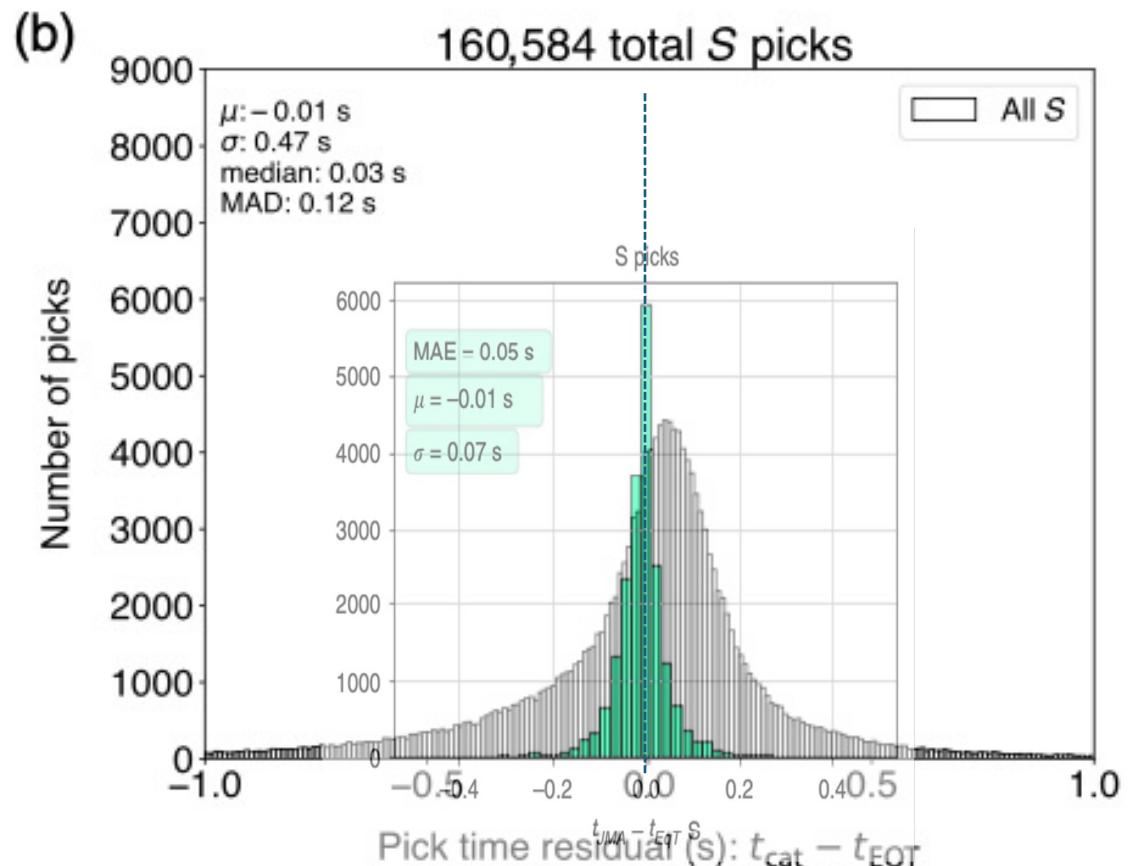
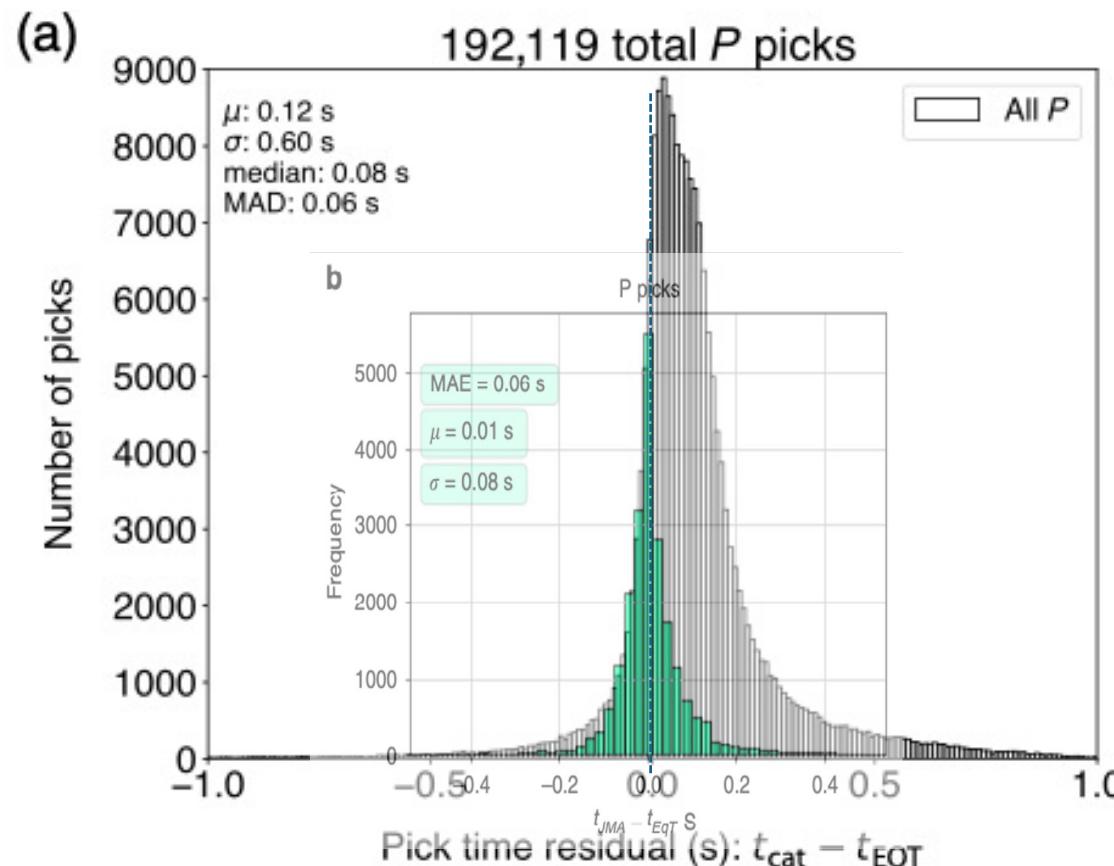
MATCH events: location, magnitude comparison

Horizontal axis: ComCat (reference catalog)
Vertical axis: Enhanced catalog

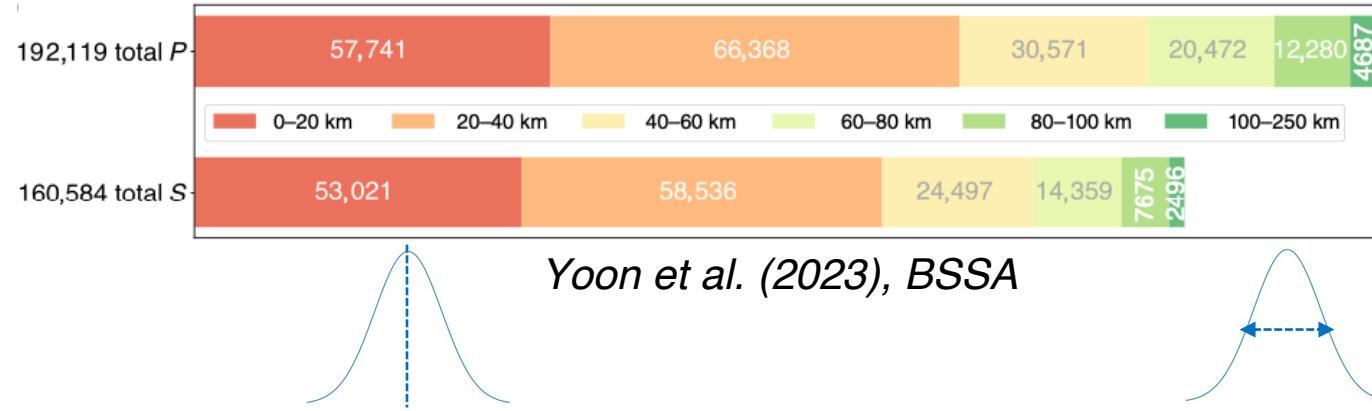
MATCH events: Pick time comparison for each station

→ residual histograms

EQTransformer on Puerto Rico seismic data:
less accurate pick times than
Mousavi et al. (2020) on STEAD test dataset
... but good enough for automatic catalog workflow

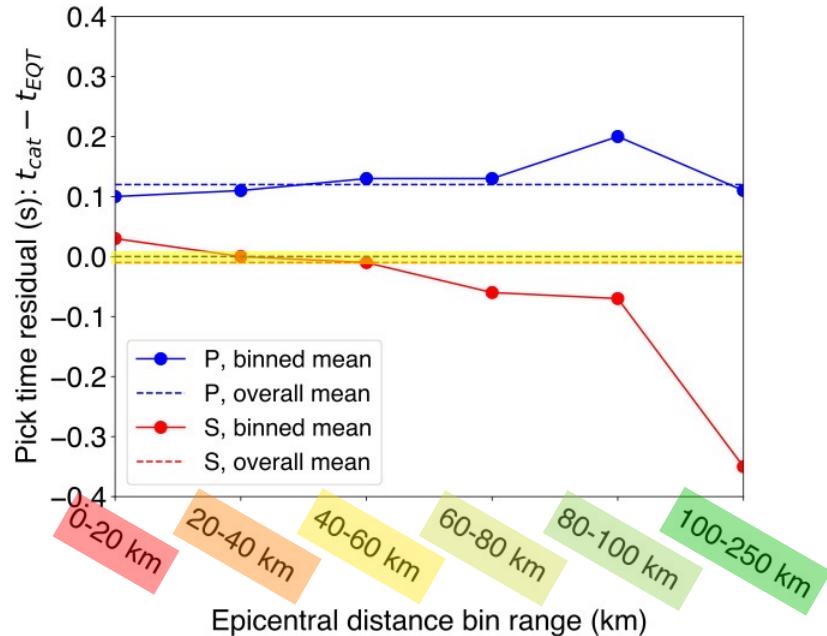


EQT pick quality (number of picks, pick accuracy) degrades with increasing event-station distance

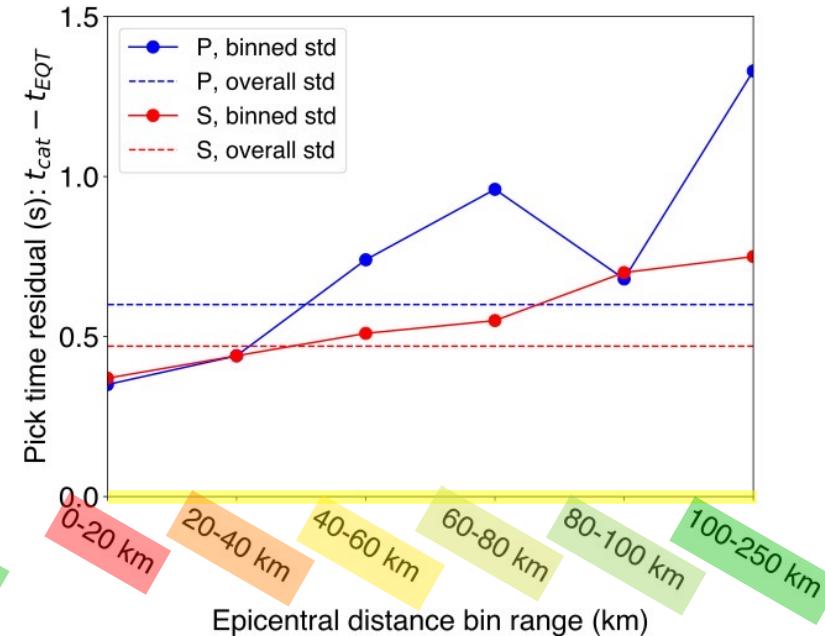


In practice: use only picks out to a maximum distance (~100 km?) for location

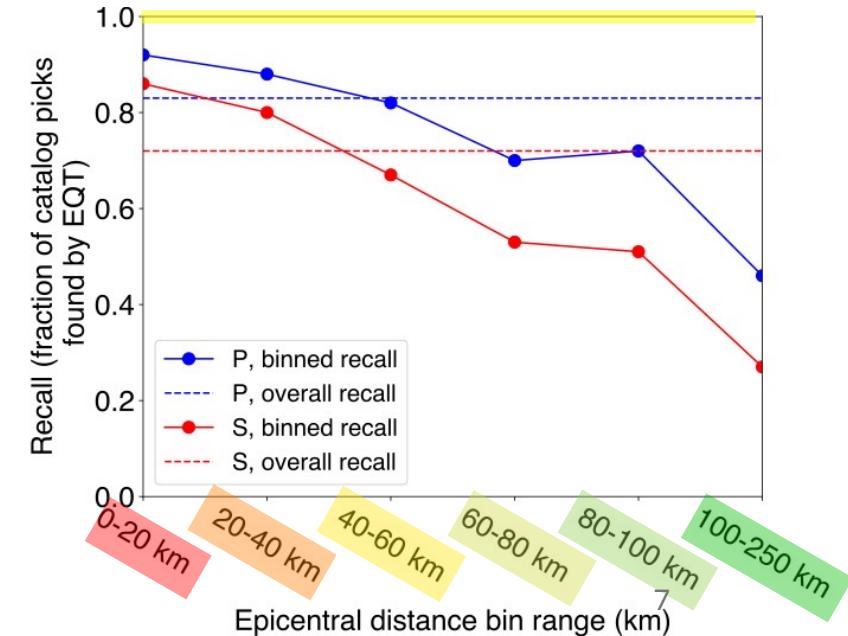
Mean $\mu[t_{\text{cat}} - t_{\text{EQT}}]$
of pick time residual



Stdev $\sigma[t_{\text{cat}} - t_{\text{EQT}}]$
of pick time residual



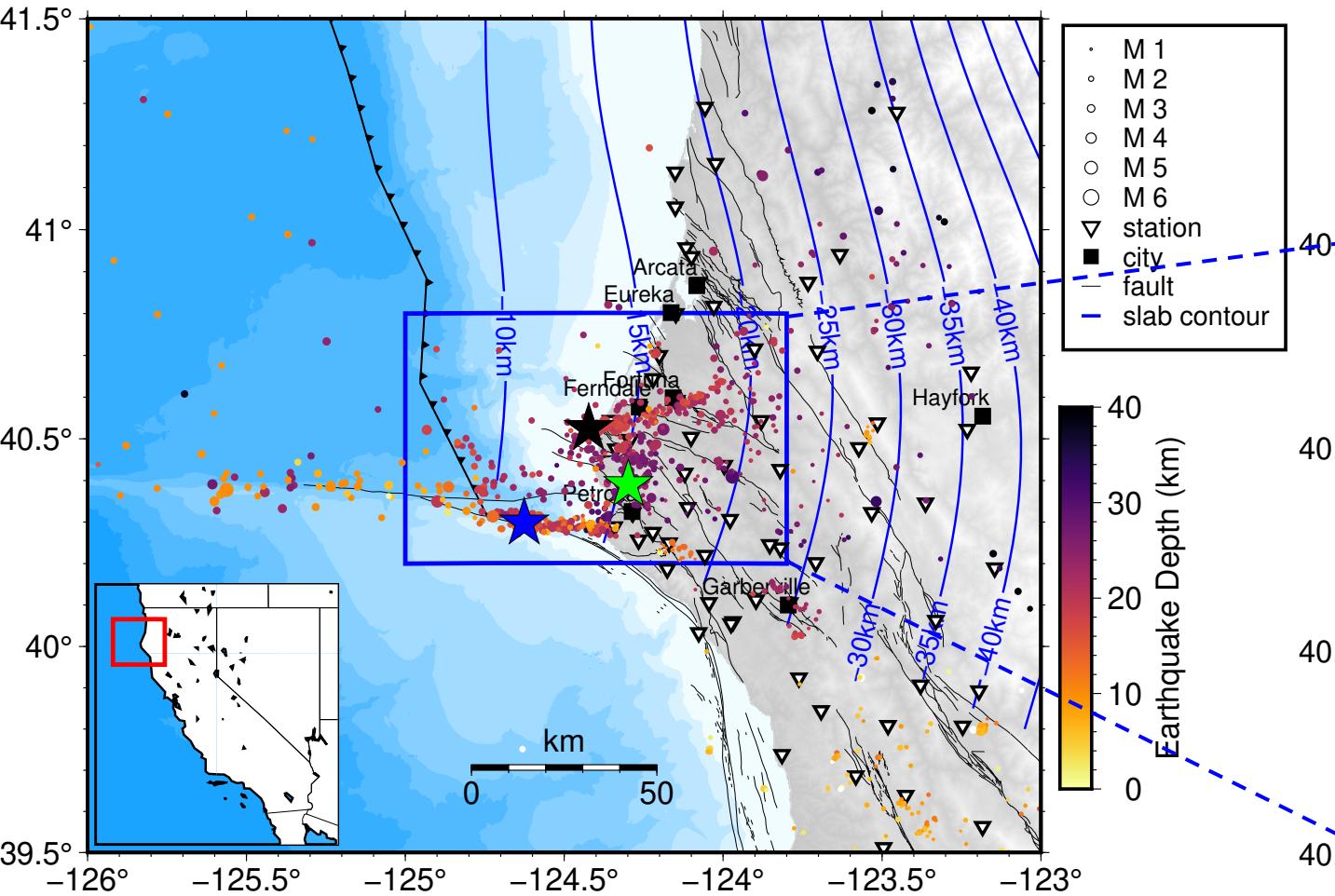
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{N_{\text{cat,EQT}}}{N_{\text{cat,EQT}} + N_{\text{cat,only}}},$$



Reference Catalog Tip: Select boundaries carefully

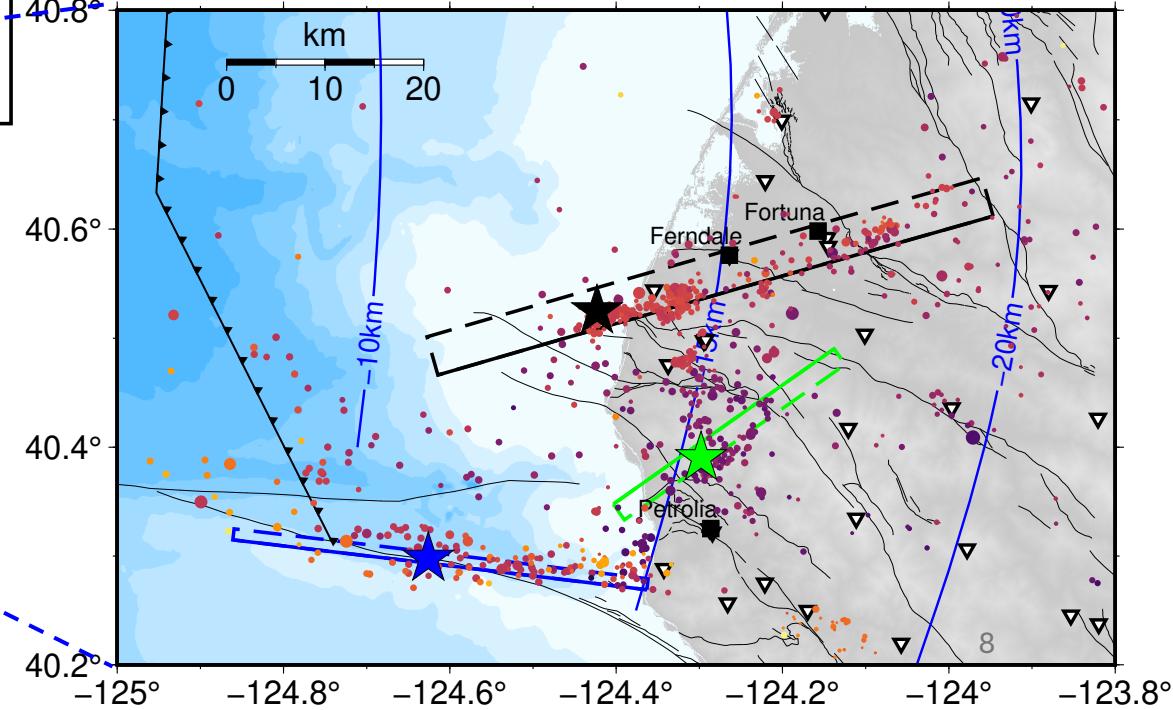
Example: Mendocino Triple Junction, CA

Region big enough: includes most seismicity and all stations used



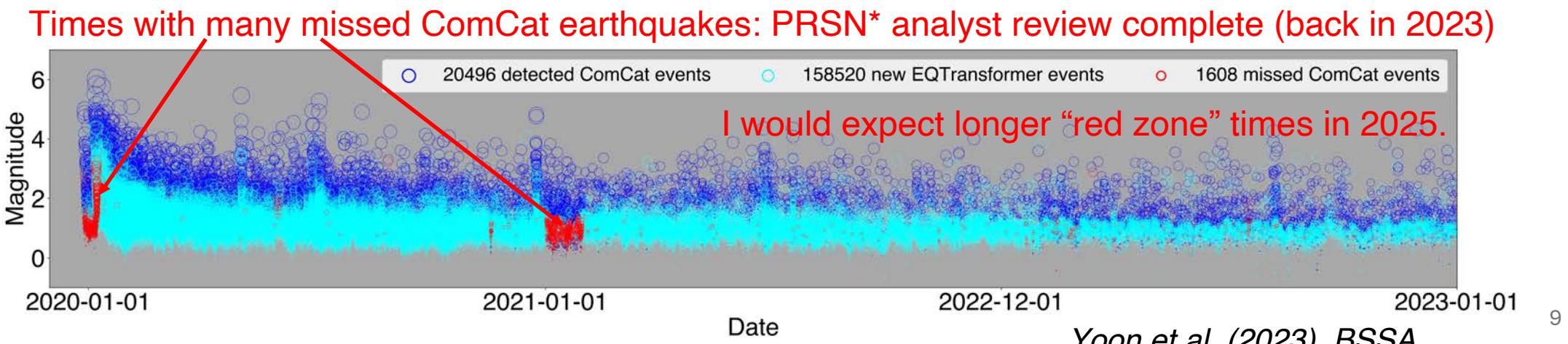
Yoon and Shelly (2024), TSR

Region too small: missing some regions with seismicity; not all stations



Reference Catalog Tip: It changes over time

- Reference catalog changes over time due to analyst review
 - New or deleted events; updated location/magnitude for existing events
 - Expect frequent changes to catalog: **hours to days** after large earthquake
 - Complete analyst review for active aftershock sequence: **months to years**
 - July 2019 M7.1 Ridgecrest CA sequence
 - 2019-07-07: 45% complete, 2019-07-08 to 2019-07-31: <10% complete each day
 - Catalog processing status (updated daily):
https://service.scedc.caltech.edu/ftp/catalogs/catalog_status/



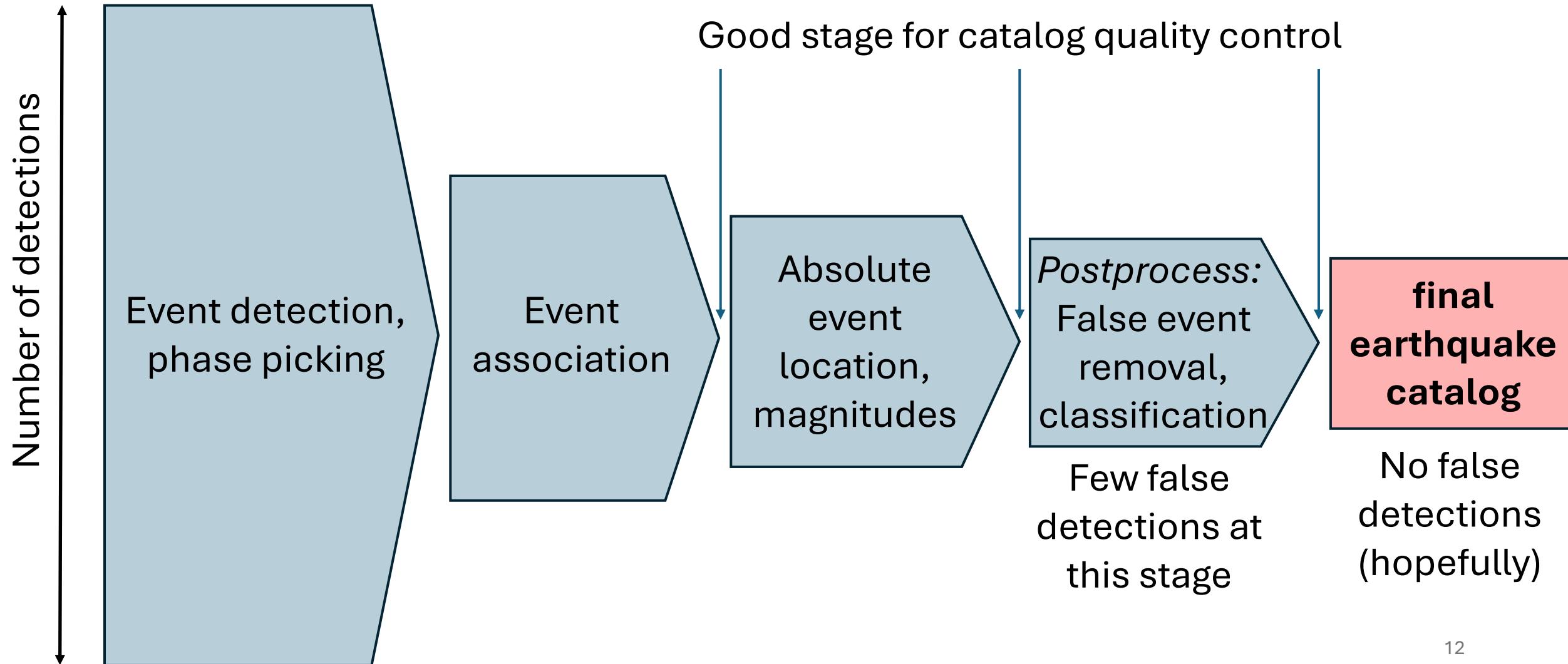
Reference Catalog Tip: It changes over time

- **Save the reference catalog file itself, with filename including the date you downloaded it.**
 - When reference catalog changes in the future, you still have the original version, so you can easily reproduce your results.

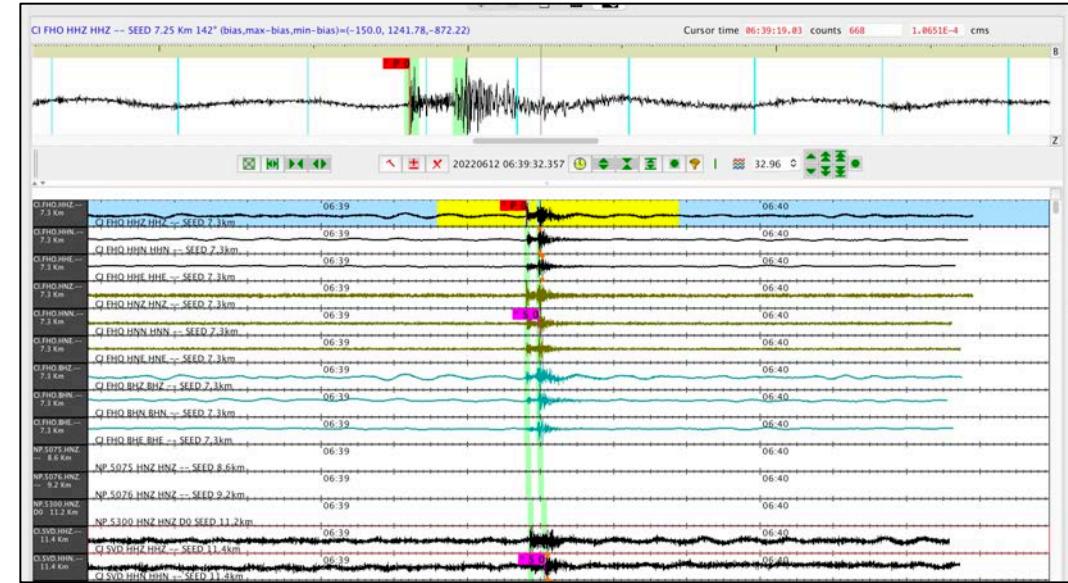
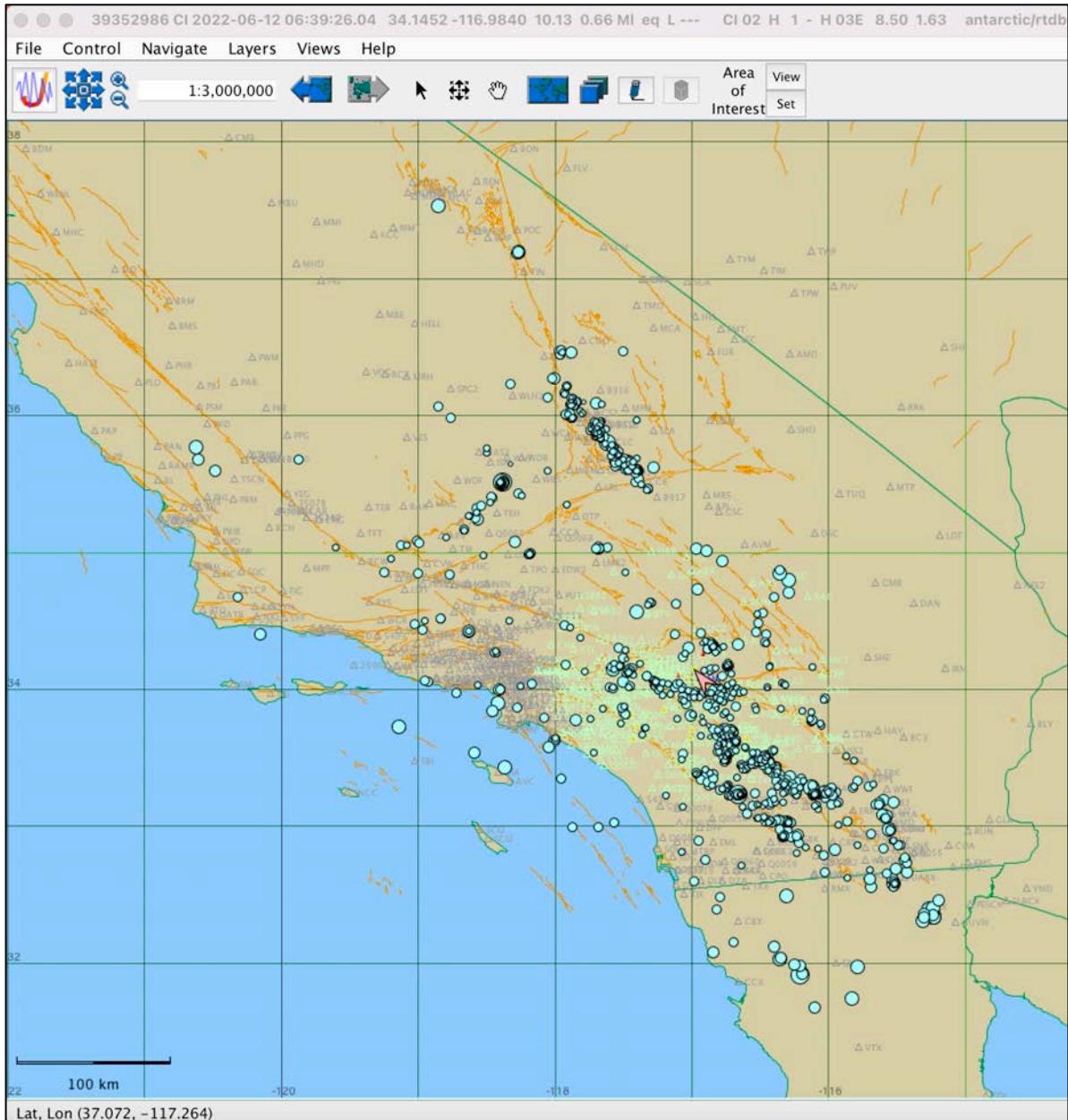
Earthquake Catalog: Quality Control (QC)

- I created a custom (deep-learning) enhanced earthquake catalog automatically. How good is it?
 - No ground truth available
- Before using or interpreting your enhanced catalog, do this first:
 - Compare with a reference catalog
 - **Data visualization** (make plots!) to check for quality

Catalog QC: Most false detections drop out at each step of automatic workflow, but not all.



Earthquake Catalog QC: Think like an analyst



AQMS Jiggle manual review and picking interface at SCSN

ID	RHO	DATETIME	LAT	LONG	REF	Z	PAC	MTP	AUTH	ETYPE	GT	SEC	ESPC	CAP	DEST	PER	ERD	ERL	ERR_S	ERR_T	OBJS	USED	B	STL	PE	LBSATE	SOUTH	COMMENT
39352938		2022-06-12 05:12:40.000,600	33,734	-116,457	2,93	2,46	0,95	ML	CI	earthquake	L	RTE	136,0	9,00	0,24	4,45	8,01	27	27	0,00	A	25,1	0	2022-06-12	45:28:41	CI		
39352946		2022-06-12 05:15:39,600	33,391	-116,872	5,89	4,79	1,19	ML	CI	earthquake	L	RTE	182,0	4,89	0,12	2,27	8,63	27	27	0,00	A	26,3	0	2022-06-12	46:03:35	CI		
39352954		2022-06-12 06:21:39,572	36,462	-117,879	9,42	7,73	2,08	ML	CI	earthquake	L	RTE	92,0	12,00	0,21	0,46	1,19	38	26	0,00	A	38,4	0	2022-06-12	46:27:41	CI		
39352962		2022-06-12 06:25:06,976	33,663	-116,931	16,54	15,79	0,92	ML	CI	earthquake	L	RTE	59,0	7,08	0,17	0,43	0,82	36	36	0,00	A	25,0	0	2022-06-12	46:29:45	CI		
39352978		2022-06-12 06:34:02,938	33,239	-116,998	11,36	18,75	1,31	ML	CI	earthquake	L	RTE	161,0	15,00	0,19	0,51	1,84	26	23	0,00	A	26,9	0	2022-06-12	46:58:36	CI		
39352986		2022-06-12 06:41:42,938	33,239	-116,998	11,36	18,75	1,31	ML	CI	earthquake	L	RTE	161,0	15,00	0,19	0,51	1,84	26	23	0,00	A	26,9	0	2022-06-12	46:58:36	CI		
39352994		2022-06-12 07:05:19,140	33,248	-116,182	10,80	18,63	1,14	ML	CI	earthquake	L	RTE	95,0	15,00	0,23	0,49	1,84	37	35	0,00	A	26,1	0	2022-06-12	07:09:50	CI		
39353002		2022-06-12 08:24:11,946	32,121	-116,396	19,46	18,63	0,85	ML	CI	earthquake	L	RTE	274,0	33,00	0,18	0,78	0,76	19	17	0,00	A	26,9	0	2022-06-12	08:29:19	CI		
39353018		2022-06-12 09:08:01,444	33,237	-116,998	9,39	9,03	1,13	ML	CI	earthquake	L	RTE	96,0	15,00	0,21	0,52	2,83	27	22	0,00	A	26,9	0	2022-06-12	09:13:17	CI		
39353026		2022-06-12 09:28:24,439	33,238	-116,998	9,71	7,35	1,44	ML	CI	earthquake	L	RTE	96,0	15,00	0,20	0,49	1,88	23	22	0,00	A	27,6	0	2022-06-12	09:42:47	CI		
39353034		2022-06-12 10:08:01,444	33,238	-116,998	9,39	9,03	1,13	ML	CI	earthquake	L	RTE	64,0	15,00	0,24	0,48	0,94	48	44	0,00	A	24,4	0	2022-06-12	10:08:51	CI		
39353054		2022-06-12 10:55:28,999	33,237	-116,998	9,71	7,35	1,44	ML	CI	earthquake	L	RTE	136,0	15,00	0,23	0,39	1,37	37	35	0,00	A	26,9	0	2022-06-12	10:58:07	CI		
39353058		2022-06-12 10:01:05,960	33,436	-116,382	4,88	2,89	1,28	ML	CI	earthquake	L	RTE	42,0	18,00	0,28	0,28	1,28	57	54	0,00	A	26,8	0	2022-06-12	10:18:52	CI		
39353074		2022-06-12 10:46:45,538	33,449	-116,476	10,85	15,58	1,28	ML	CI	earthquake	L	RTE	137,0	12,00	0,11	0,49	1,12	28	28	0,00	A	26,4	0	2022-06-12	10:46:11	CI		
39353090		2022-06-12 10:35:12,766	35,560	-118,395	8,85	7,79	1,72	ML	CI	earthquake	L	RTE	102,0	19,00	0,15	0,36	1,17	35	33	0,00	A	27,0	0	2022-06-12	10:39:41	CI		
39353106		2022-06-12 10:56:45,600	33,927	-117,876	17,27	15,52	1,22	ML	CI	earthquake	L	RTE	176,0	12,00	0,29	0,26	0,88	59	55	0,00	A	26,5	0	2022-06-12	11:01:02	CI		
39353107		2022-06-12 10:56:45,600	33,927	-117,876	17,27	15,52	1,22	ML	CI	earthquake	L	RTE	66,0	15,00	0,21	0,49	0,99	49	49	0,00	A	26,5	0	2022-06-12	11:01:02	CI		
39353120		2022-06-12 07:07:37,420	33,223	-115,974	5,11	3,49	1,23	ML	CI	earthquake	L	RTE	135,0	26,00	0,25	0,36	1,38	64	53	0,00	A	26,9	0	2022-06-12	07:12:02	CI		
39353130		2022-06-12 08:10:40,515	33,219	-115,969	11,79	11,67	1,87	ML	CI	earthquake	L	RTE	33,0	18,00	0,29	0,36	0,93	84	64	0,00	A	29,0	0	2022-06-12	12:13:52	CI		
39353162		2022-06-12 11:11:22,640	32,889	-115,538	5,11	5,11	1,85	ML	CI	earthquake	L	RTE	171,0	3,00	0,18	0,59	0,68	18	18	0,00	A	27,5	0	2022-06-12	15:25:27	CI		
39353170		2022-06-12 11:32:11,554	32,889	-115,529	4,65	4,63	0,87	ML	CI	earthquake	L	RTE	168,0	3,00	0,18	1,14	0,75	12	8	0,00	A	25,3	0	2022-06-12	15:21:57	CI		
39353178		2022-06-12 13:02:27,539	35,179	-118,161	6,12	7,35	1,57	ML	CI	earthquake	L	RTE	68,0	15,00	0,18	0,29	0,88	51	49	0,00	A	25,0	0	2022-06-12	13:07:03	CI		
39353210		2022-06-12 13:25:22,439	33,289	-116,382	4,88	2,89	1,28	ML	CI	earthquake	L	RTE	162,0	18,00	0,23	0,29	0,55	25	25	0,00	A	26,5	0	2022-06-12	13:25:27	CI		
39353218		2022-06-12 13:25:23,446	33,449	-116,965	4,05	2,66	1,52	ML	CI	earthquake	L	RTE	162,0	18,00	0,23	1,42	7,41	18	16	0,00	A	28,0	0	2022-06-12	13:39:27	CI		
39353226		2022-06-12 13:49:27,680	33,298	-117,377	14,09	8,04	0,91	ML	CI	earthquake	L	RTE	74,0	4,00	0,22	0,45	1,86	33	32	0,00	A	25,0	0	2022-06-12	13:52:52	CI		
39353234		2022-06-12 14:51:48,600	33,927	-116,195	23,44	22,94	1,99	ML	CI	earthquake	L	RTE	235,0	45,00	0,17	0,51	2,31	38	15	0,00	A	30,4	0	2022-06-12	14:57:22	CI		
39353259		2022-06-12 15:45:17,990	33,267	-119,793	13,59	12,67	1,46	ML	CI	earthquake	L	RTE	56,0	5,00	0,19	0,32	0,71	48	39	0,00	A	27,7	0	2022-06-12	15:49:58	CI		
39353274		2022-06-12 16:19:52,320	33,214	-116,240	10,85	11,55	1,11	ML	CI	earthquake	L	RTE	286,0	11,00	0,17	0,44	1,24	14	14	0,00	A	26,9	0	2022-06-12	16:24:28	CI		
39353280		2022-06-12 16:24:07,390	34,438	-115,255	15,54	14,88	1,23	ML	CI	earthquake	L	RTE	127,0	12,00	0,28	0,62	8,87	26	24	0,00	A	26,6	0	2022-06-12	16:29:13	CI		
39353299		2022-06-12 16:28,333,648	35,828	-116,967	5,09	4,13	2,16	ML	CI	earthquake	L	RTE	58,0	17,00	0,17	0,16	0,67	103	44	0,00	A	31,2	0	2022-06-12	16:33:33	CI		
39353314		2022-06-12 16:33:31,747	35,661	-117,528	14,68	13,95	1,34	ML	CI	earthquake	L	RTE	159,0	18,00	0,17	0,45	1,18	28	21	0,00	A	27,1	0	2022-06-12	16:38:33	CI		
39353322		2022-06-12 16:38:26,600	34,269	-118,448	16,79	16,37	1,32	ML	CI	earthquake	L	RTE	107,0	4,00	0,21	0,59	0,96	38	38	0,00	A	27,0	0	2022-06-12	16:43:13	CI		
39353346		2022-06-12 16:40:51,700	33,277	-117,437	8,18	8,63	1,49	ML	CI	earthquake	L	RTE	134,0	11,00	0,17	0,46	1,63	22	22	0,00	A	26,8	0	2022-06-12	16:45:25	CI		
39353362		2022-06-12 17:08:05,660	33,147	-117,878	5,11	19,24	1,89	ML	CI	earthquake	L	RTE	63,0	11,00	0,17	0,49	0,94	23	22	0,00	A	25,4	0	2022-06-12	17:12:19	CI		
39353366		2022-06-12 17:08:55,390	32,852	-118,388	7,48	6,44	0,94	ML	CI	earthquake	L	RTE	118,0	17,00	0,18	0,47	1,48	21	18	0,00	A	25,2	0	2022-06-12	18:03:19	CI		
39353394		2022-06-12 18:25:16,976	35,597	-118,424	28,36	27,81	1,78	ML	CI	earthquake	L	RTE	264,0	27,00	0,26	0,74	1,96	32	11	0,00	A	29,4	0	2022-06-12	18:33:34	CI		
39353426		2022-06-12 19:36:51,600	36,172	-119,729	3,72	32,2	1,53	ML	CI	earthquake	L	RTE	108,0	9,00	0,32	0,64	1,12	22	21	0,00	A	28,1	0	2022-06-12	19:41:14	CI		
39353436		2022-06-12 19:40:51,700	33,277	-117,437	8,18	8,63	1,49	ML	CI	earthquake	L	RTE	134,0	11,00	0,18	0,46	1,63	19	17	0,00	A	25,6	0	2022-06-12	19:45:25	CI		

Earthquake Catalog QC Tips: Check Parameters

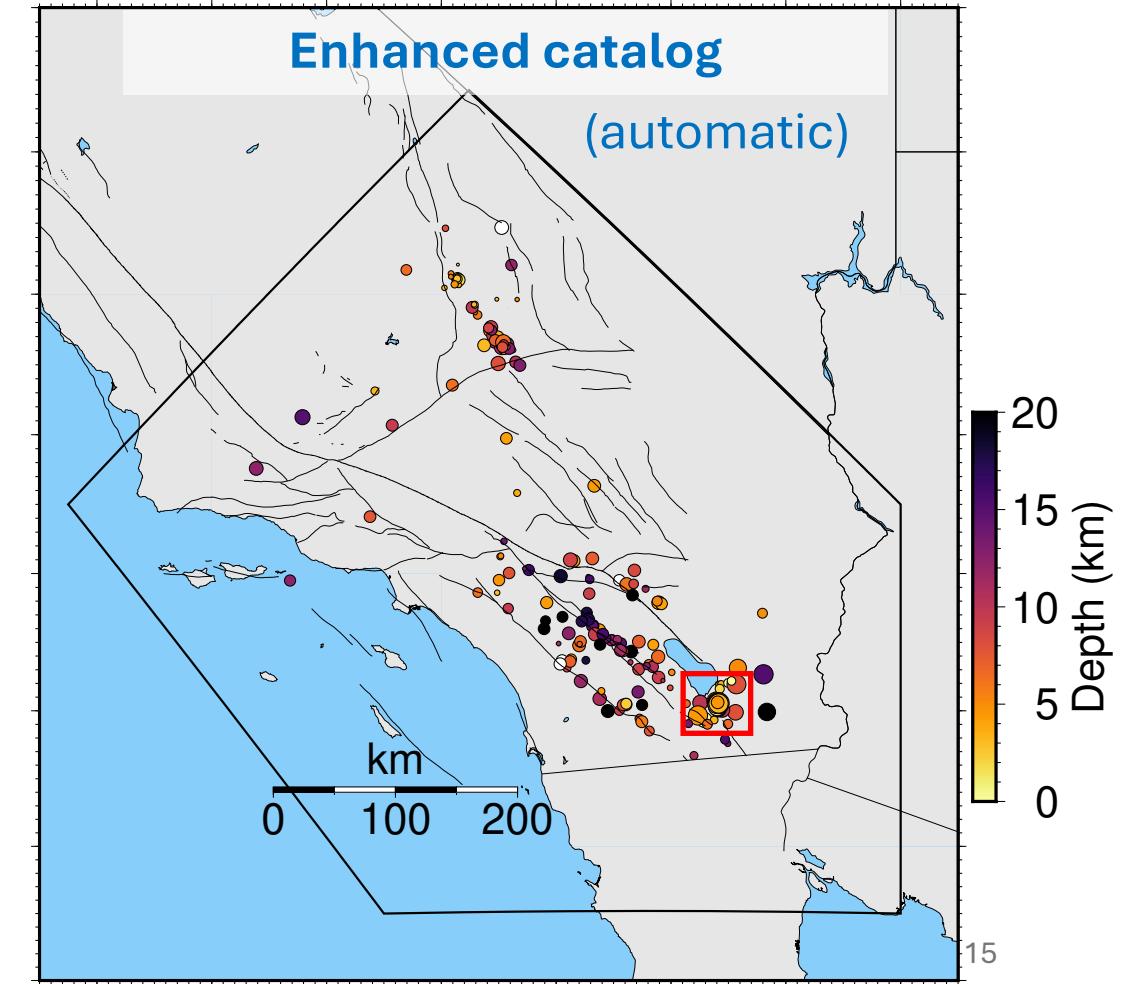
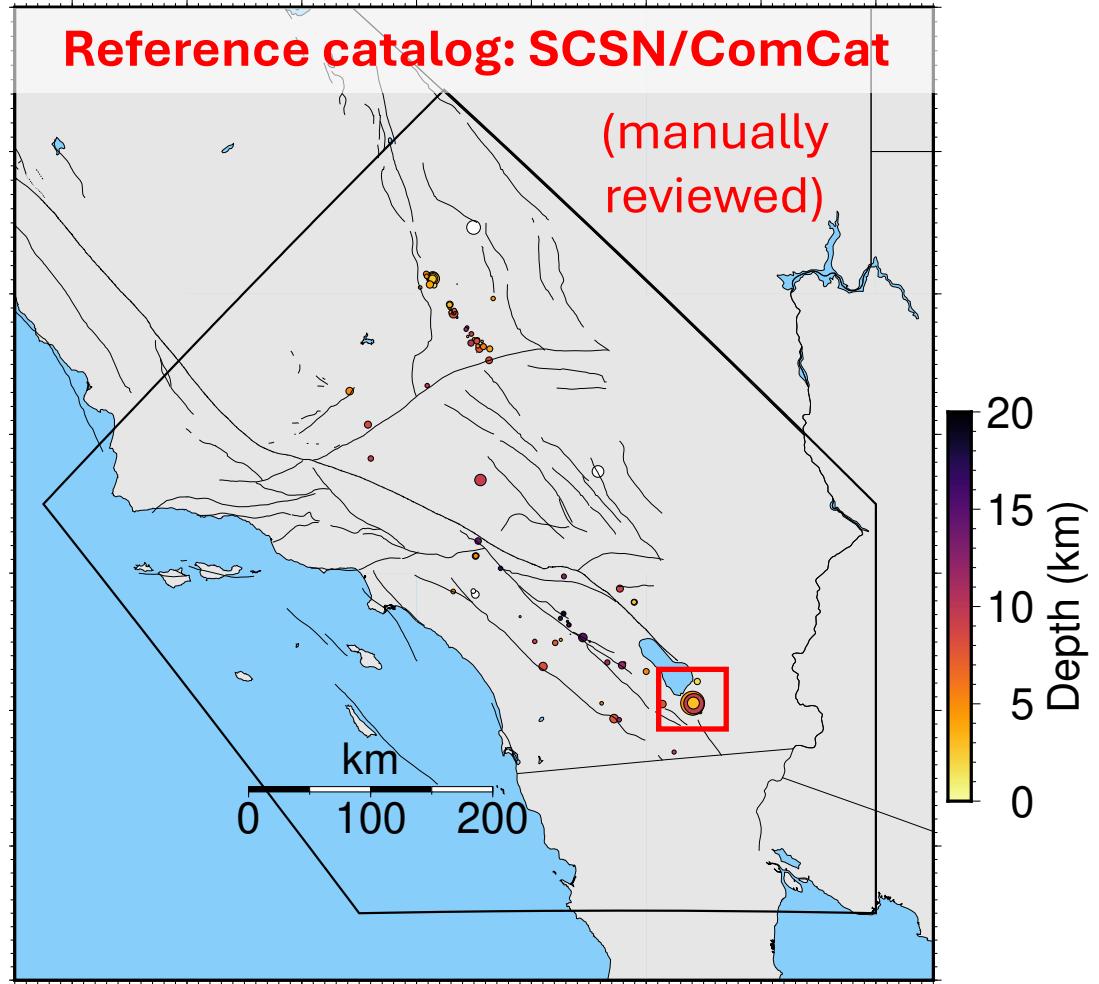
- Compare enhanced catalog with reference catalog on plots
 - Plot reference & enhanced catalogs side-by-side
 - Separate plots for MATCH, NEW, MISSED category events
- Visualize enhanced catalog → are results reasonable for earthquakes?
 - Locations (epicenters) in map view
 - Are events near known seismogenic areas, past seismicity, faults? Known quarries?
 - Depth cross-sections
 - Are events at expected seismogenic depths (0-30 km, unless in subduction zone)?
 - Magnitude-frequency distribution
 - Do event magnitudes follow Gutenberg-Richter distribution (lots of small events)?
 - Magnitude vs. time, Seismicity rate (number of events vs. time)
 - If aftershock sequence, does number of events decrease as 1/time after mainshock (Omori decay)?

Catalog QC, locations in map view: SCSN



2020-09-30 00:00:00 UTC to 2020-10-01 00:00:00 UTC (24 hours)

Enhanced catalog: events are on known faults with active seismicity

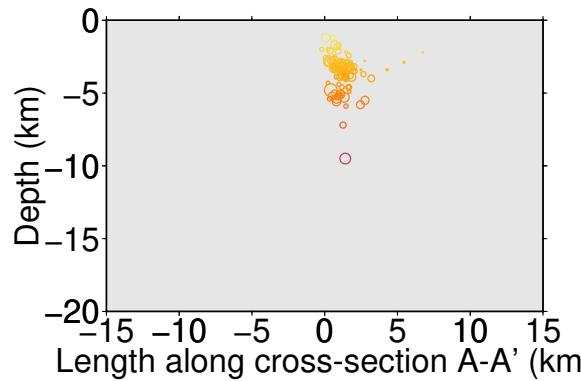
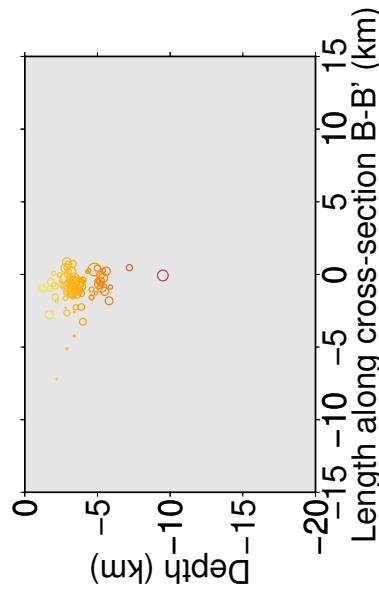
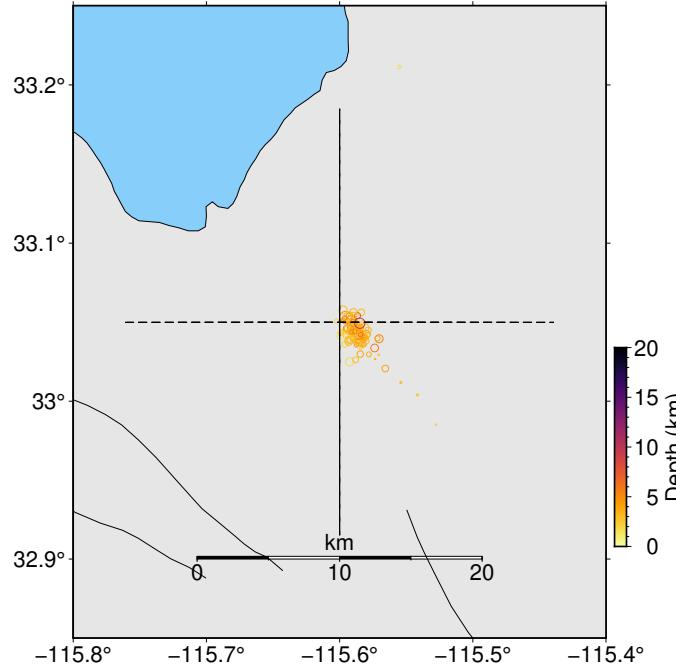


Catalog QC, zoomed locations in map view & depth cross-section: SCSN

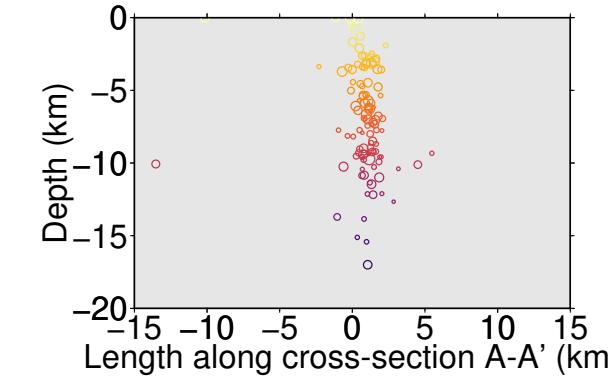
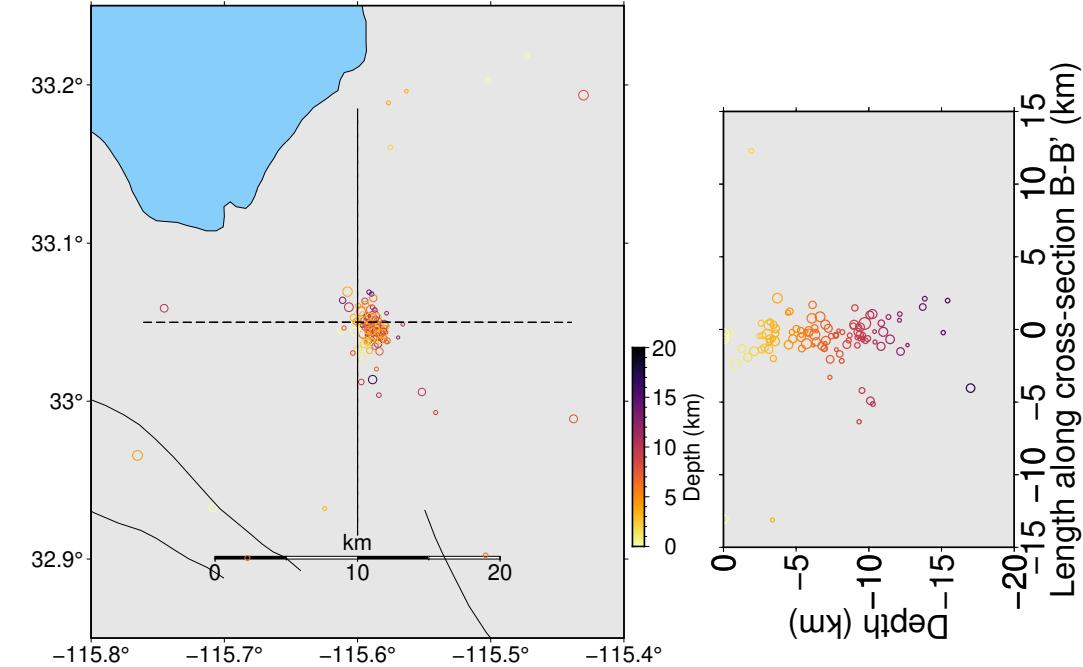
2020-09-30 00:00:00 UTC to 2020-10-01 00:00:00 UTC (24 hours)



Enhanced catalog: Similar epicenters. Larger range of depths, but still reasonable



Reference catalog:
SCSN/ComCat
(manually reviewed)



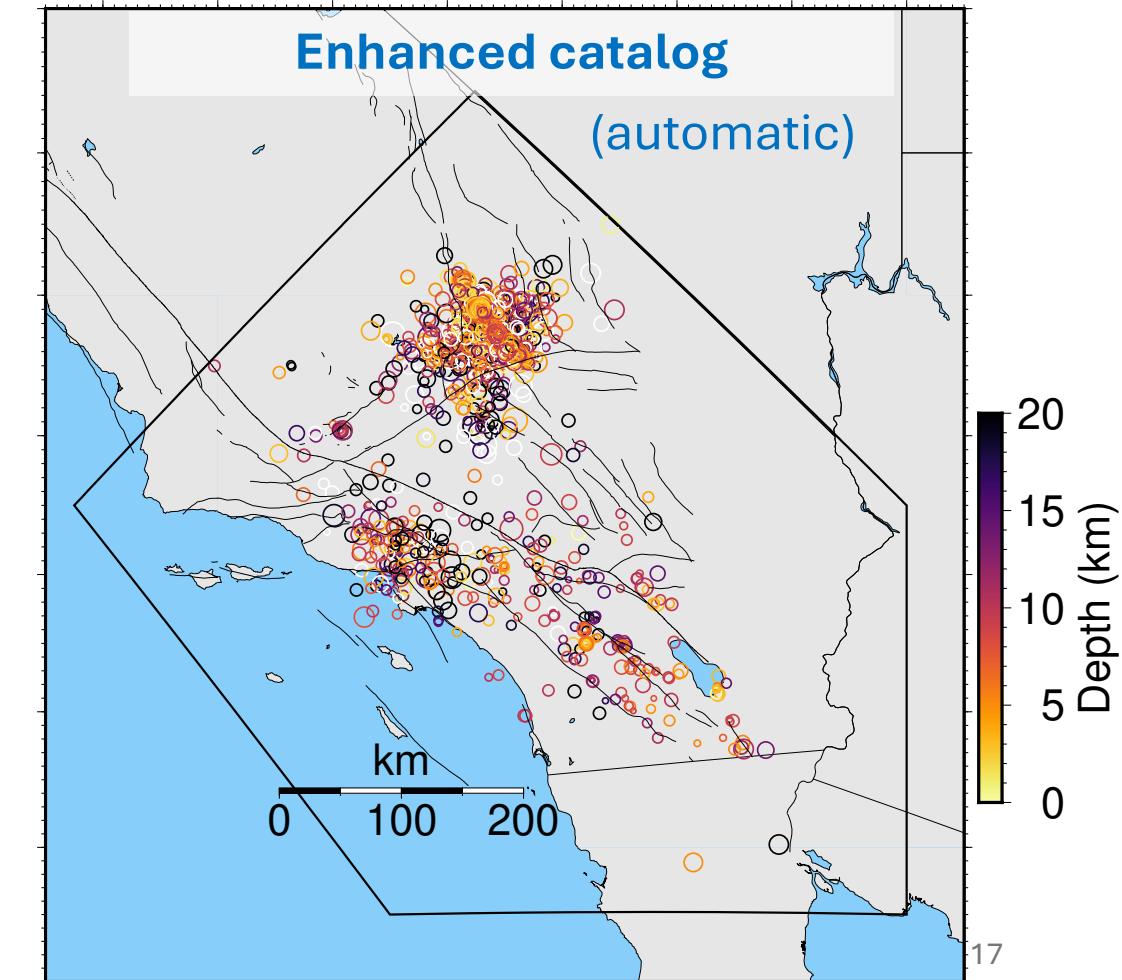
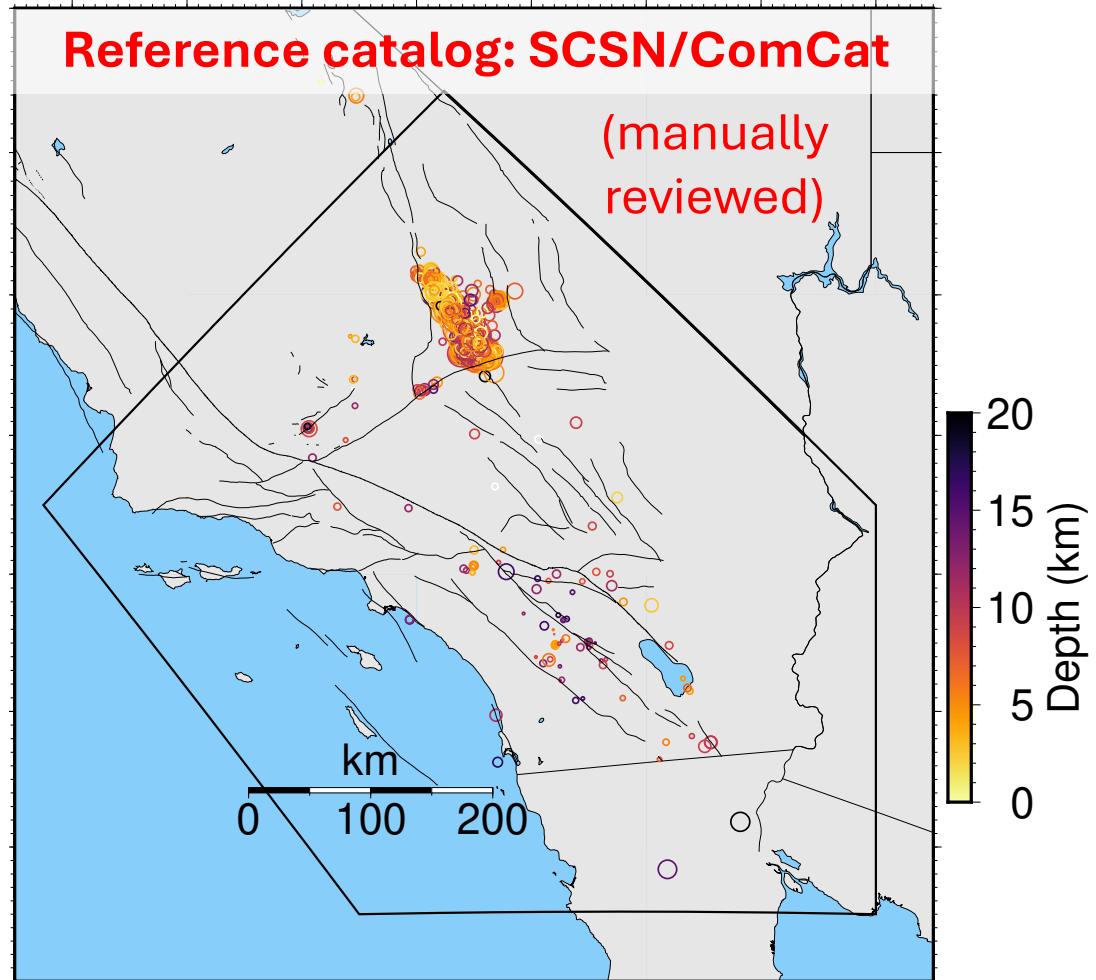
Enhanced catalog
(automatic)

Catalog QC, locations in map view: SCSN



2019-07-04 00:00:00 UTC to 2019-07-07 01:14:00 UTC (73 hours)

Ridgecrest first 3 days (M_w 6.4 & 7.1, aftershocks), enhanced catalog:
association errors → false detections with bad locations/magnitudes



Earthquake Catalog QC Tips: Check Parameters

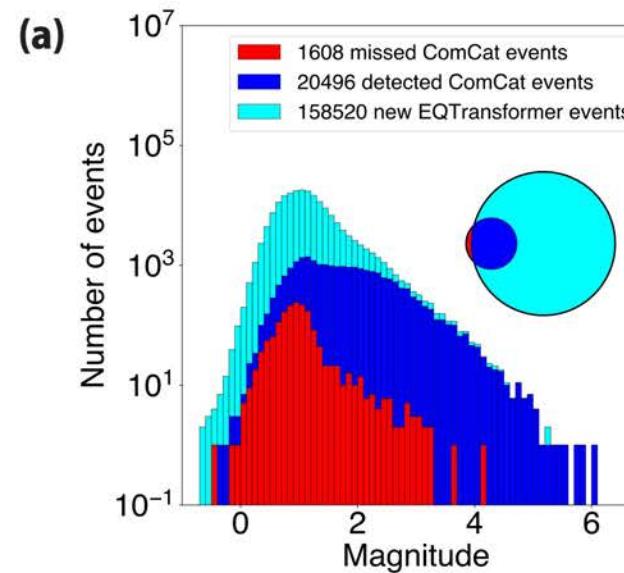
- Compare enhanced catalog with reference catalog on plots
 - Plot reference & enhanced catalogs side-by-side
 - Separate plots for MATCH, NEW, MISSED category events
- Visualize enhanced catalog → are results reasonable for earthquakes?
 - Locations (epicenters) in map view
 - Are events near known seismogenic areas, past seismicity, faults? Known quarries?
 - Depth cross-sections
 - Are events at expected seismogenic depths (0-30 km, unless in subduction zone)?
 - **Magnitude-frequency distribution**
 - Do event magnitudes follow Gutenberg-Richter distribution (lots of small events)?
 - **Magnitude vs. time, Seismicity rate (number of events vs. time)**
 - If aftershock sequence, does number of events decrease as $1/\text{time}$ after mainshock (Omori decay)?

Catalog QC: Puerto Rico



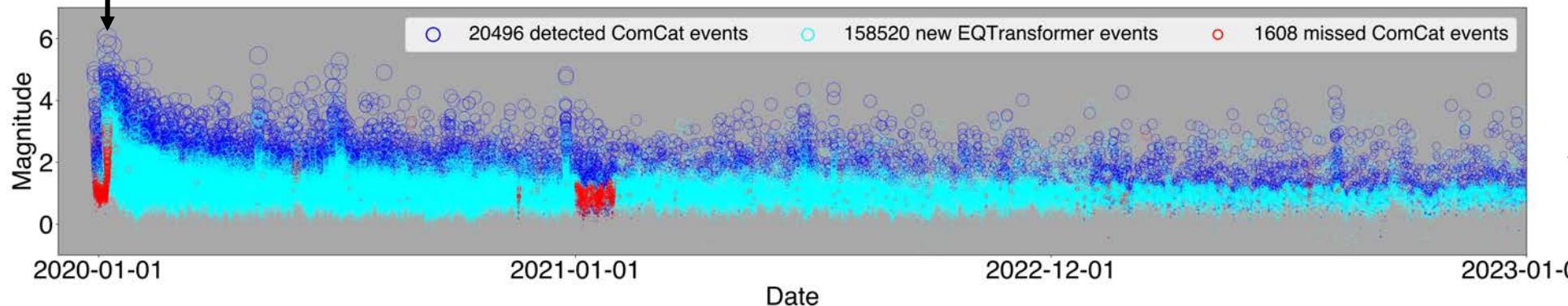
MATCH,
MISSED,
NEW

M_w 6.4 mainshock



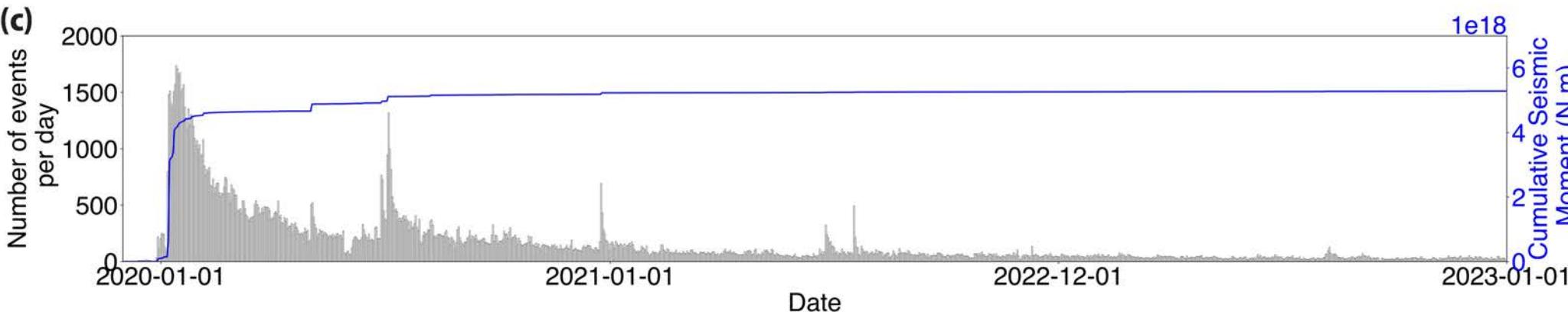
Magnitude-frequency distribution (MFD)

(b)



Magnitude
vs. time

(c)

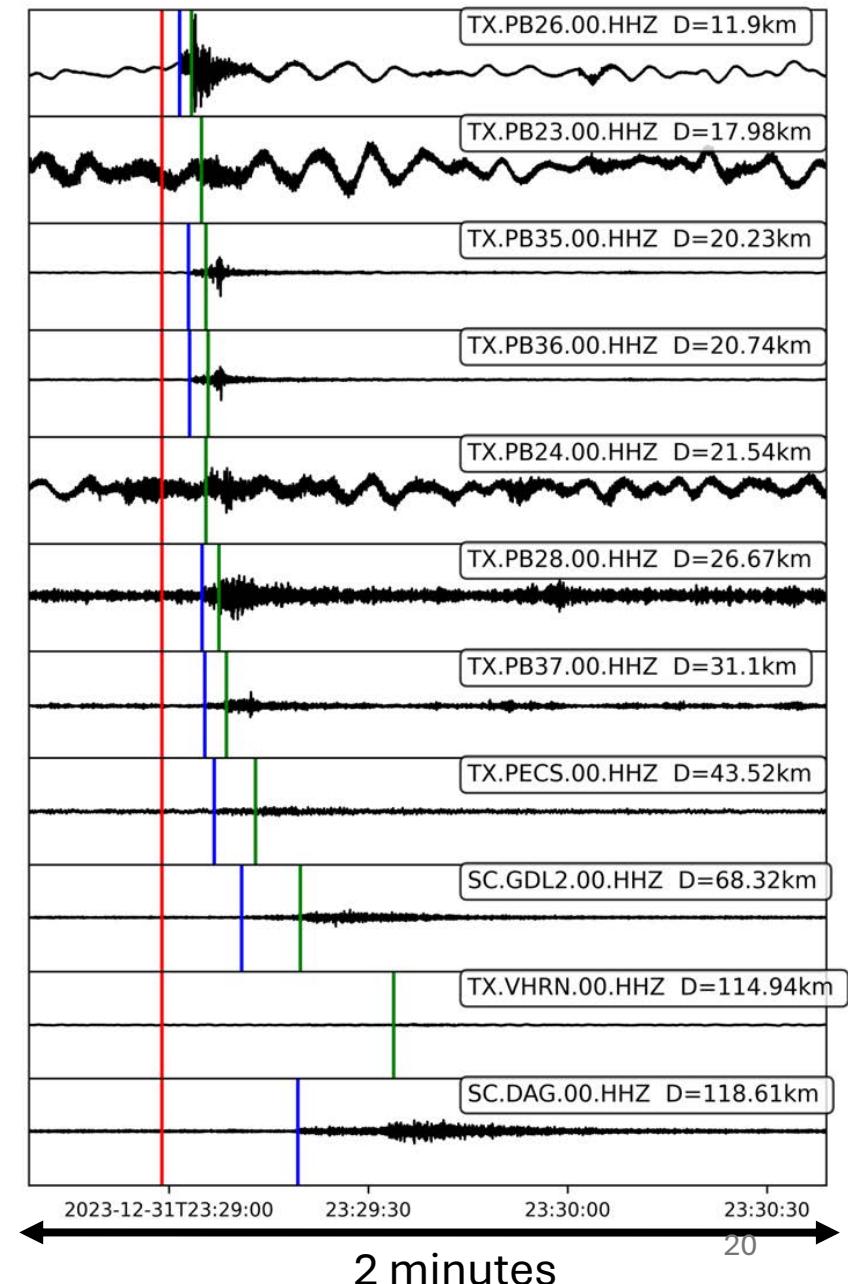


Number of events/day,
Cumulative moment
vs. time

Earthquake Catalog QC Tips: Check Waveforms

- When in doubt, **plot event waveforms**
 - Overlay picks on event waveforms, ordered by event-station distance, especially for “**NEW**” events not in reference catalog
 - **Earthquake or noise?**
 - **Distance from event to nearest station?**
 - **Frequency content & time duration** – local, regional, or teleseismic earthquake? Quarry blast?
 - **Moveout** (from P, S pick times) & **attenuation** (from amplitudes) **with distance?**
 - Do actual picks match predicted arrival times (from ray-tracing through velocity model)?

PyOcto Event #340 : 5/5/10 No Match to TexNet, Match to GaMMA, NSD<75.0
Red: Event, Blue: P Pick, Green: S Pick



Check “**NEW**” events in enhanced catalog, not in reference catalog

Sorted by **magnitude**
(Southwest Puerto Rico)

Origin time (UTC)	OT (seconds)	Latitude	Longitude	Depth	Magnitude	EventID
2020-01-07T08:35:15.030000	30915.030000	17.871833	-66.721667	6.420000	4.320000	1000335
2020-01-07T11:21:01.620000	40861.620000	18.034167	-66.800000	13.780000	4.360000	1000569
2020-01-07T08:29:18.320000	30558.320000	18.139667	-66.810000	7.660000	4.440000	1000328
2020-01-07T08:29:36.930000	30576.930000	17.963500	-66.752333	29.160000	4.480000	1000330
2020-01-07T08:51:09.400000	31869.400000	17.676333	-66.773667	27.610000	4.760000	1000356

Catalog QC: False detection in coda of larger earthquake (also, too deep?)

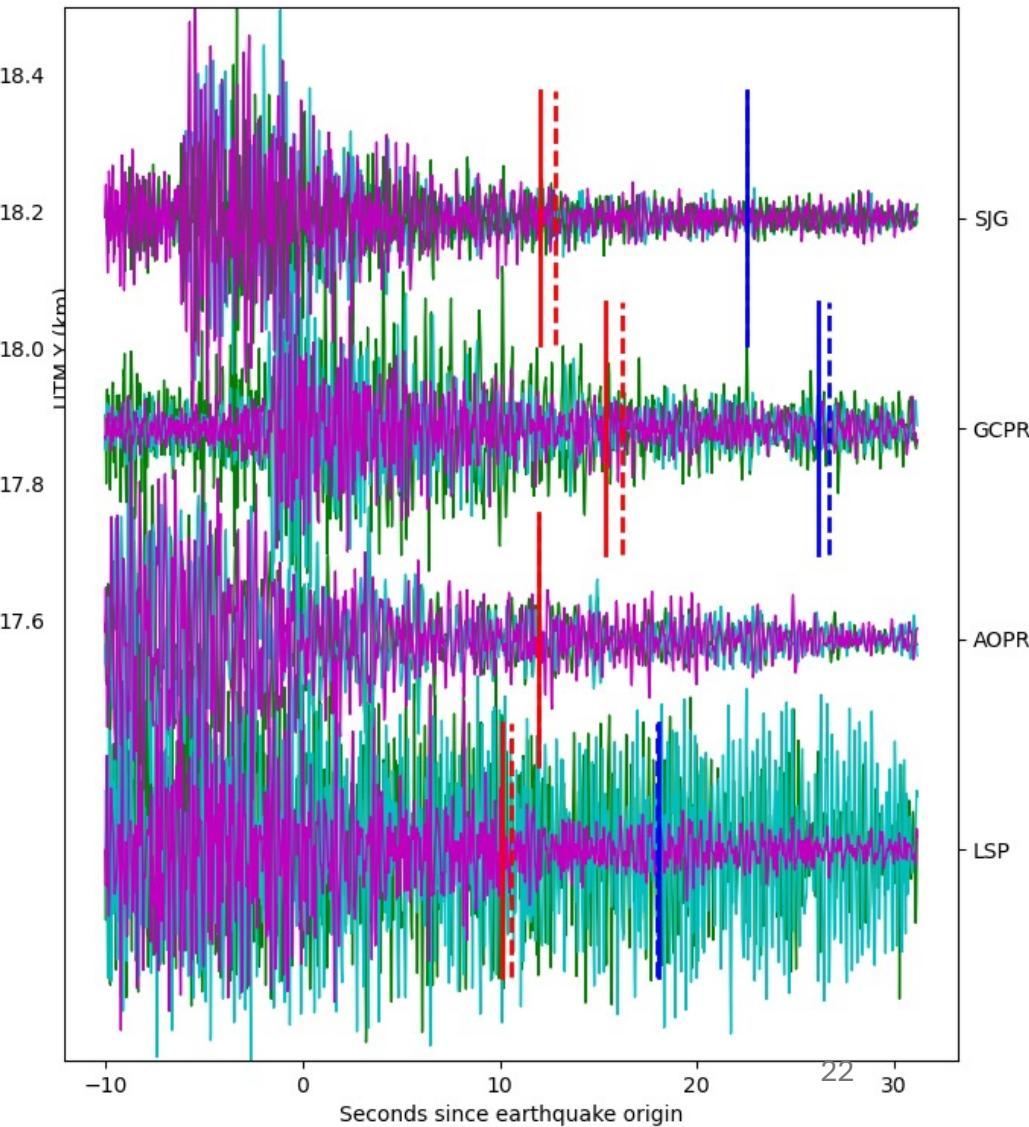
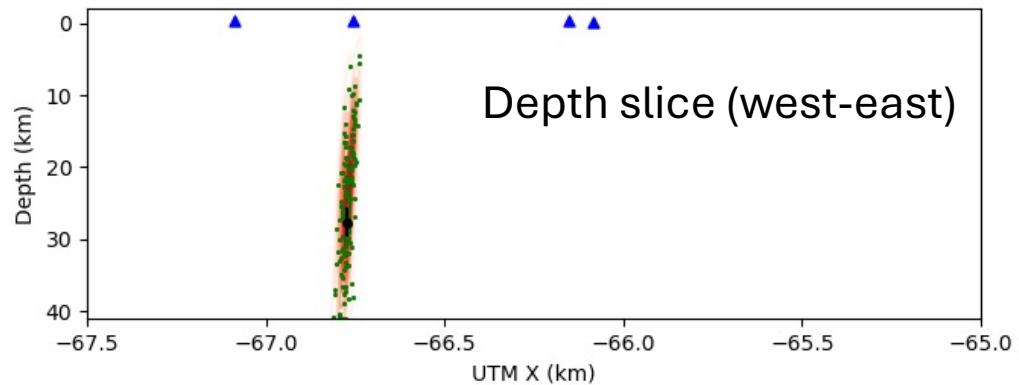
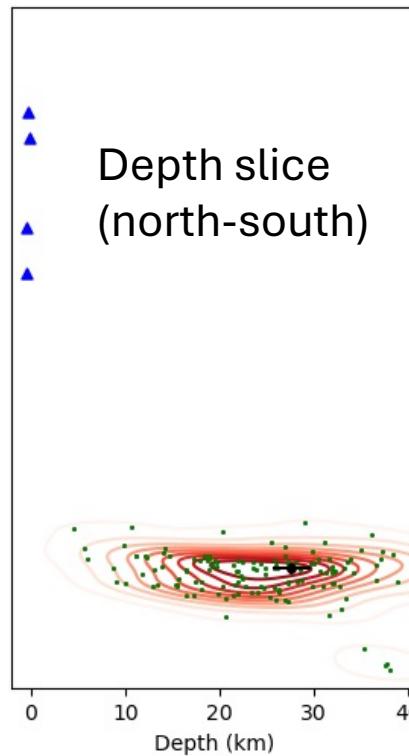
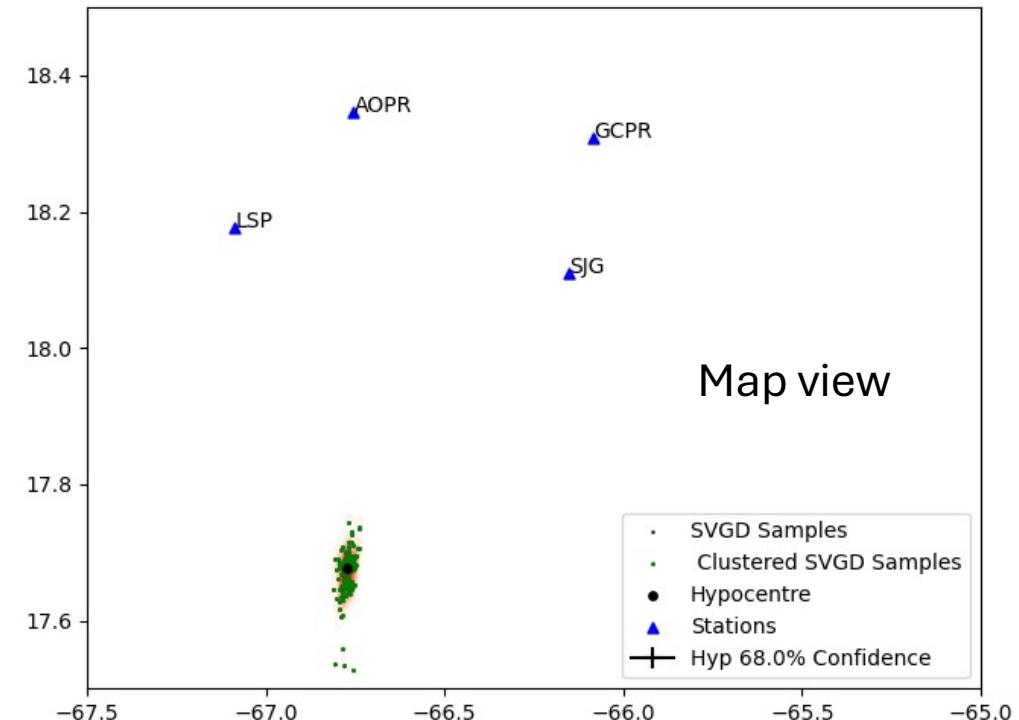


2020-01-07T08:51:09.400000 31869.400000 17.676333 -66.773667 27.610000 4.760000 1000356

Earthquake 2020-01-07 08:51:09.396302640 +/- 0.50s
Hyp=[-66.77,17.68,27.61] - Hyp Uncertainty (km) +/- [0.35,0.84,1.97]

Event waveforms with P, S picks

Solid: auto-picks, Dashed: predicted arrivals



Catalog QC: Real (newly detected) earthquake (reasonable depth)

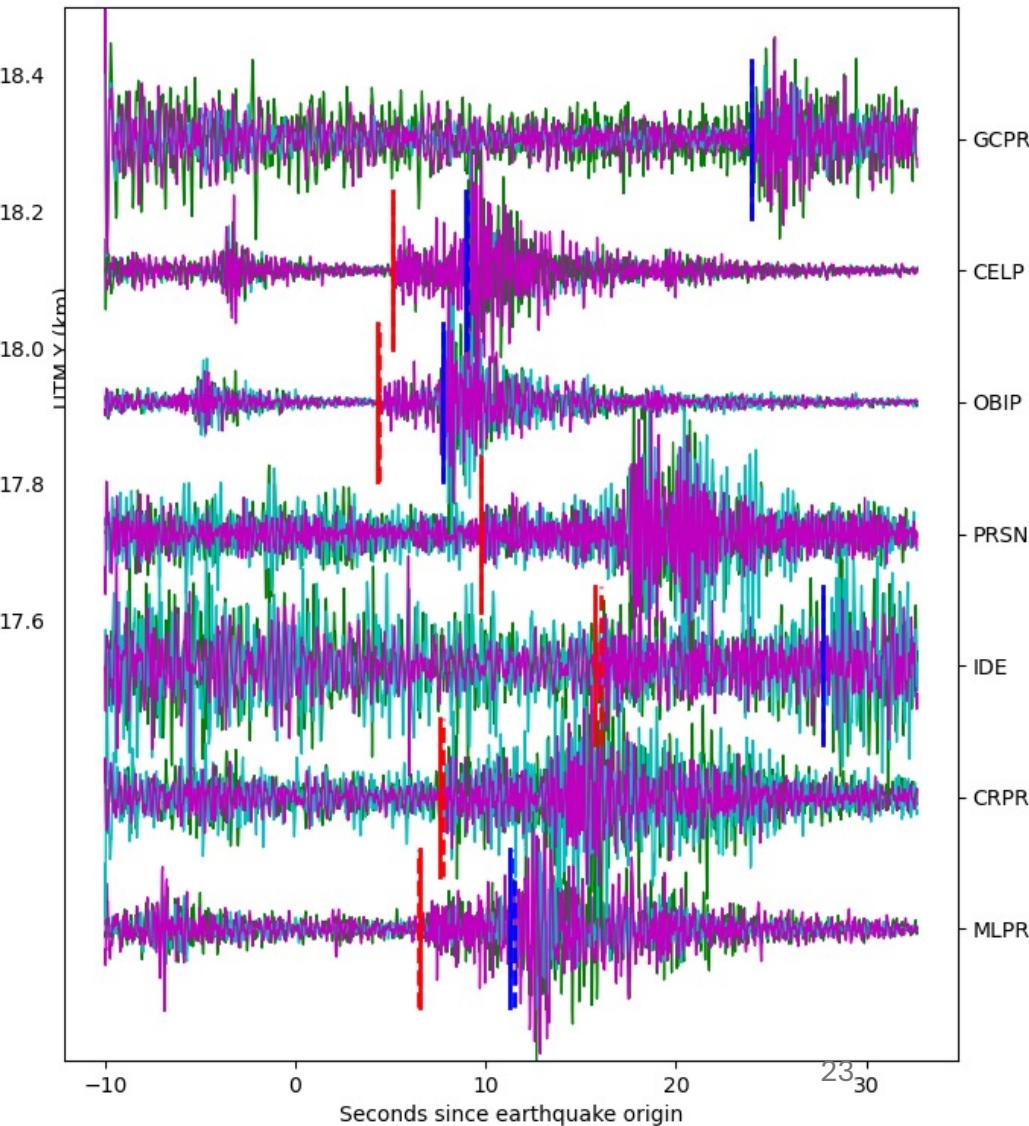
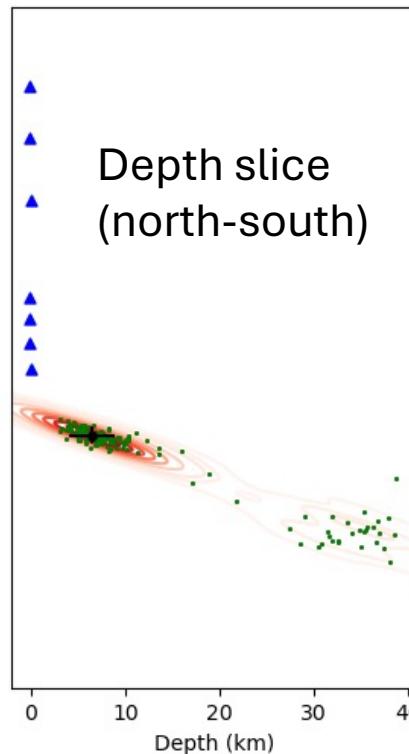
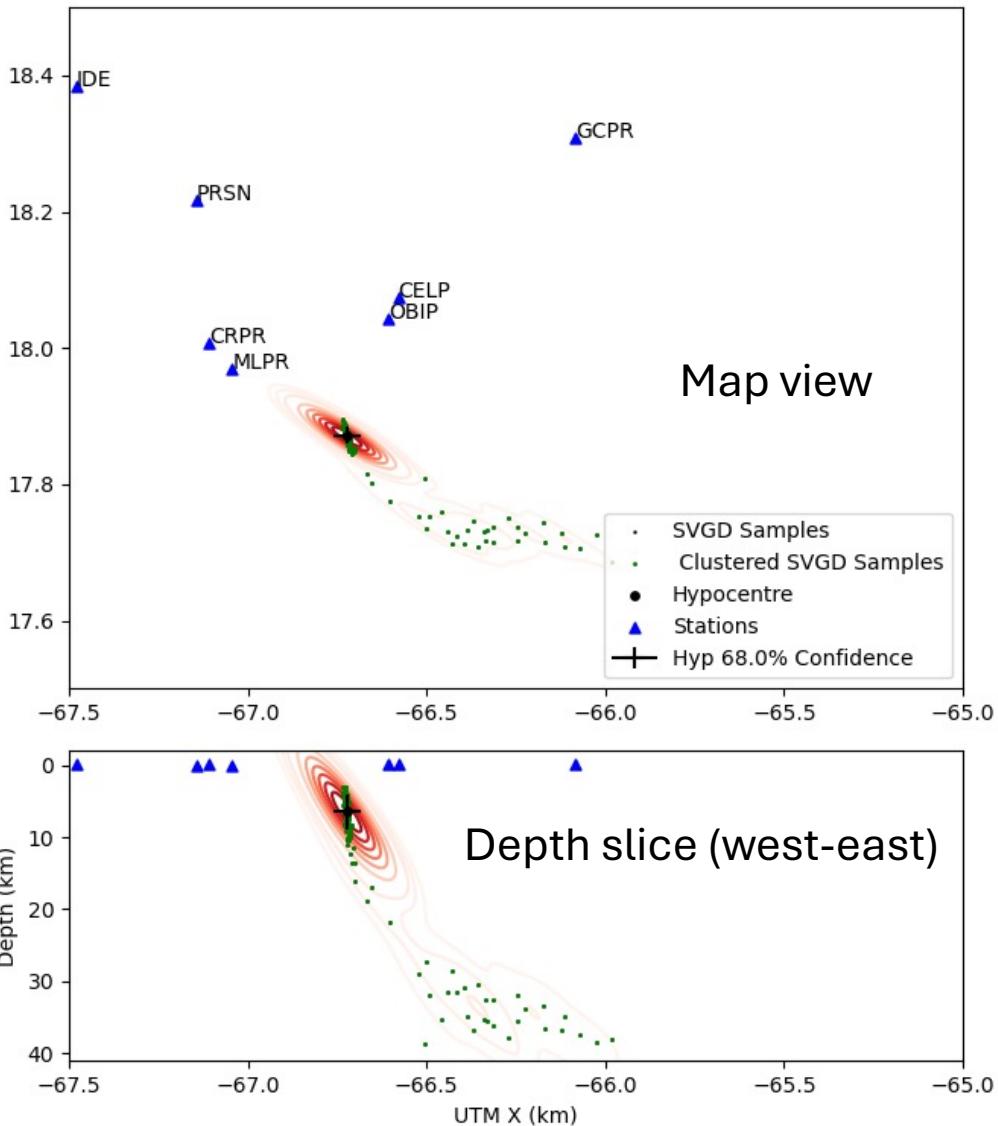


2020-01-07T08:35:15.030000 30915.030000 17.871833 -66.721667 6.420000 4.320000 1000335

Earthquake 2020-01-07 08:35:15.029341138 +/- 0.05s
Hyp=[-66.72,17.87,6.42] - Hyp Uncertainty (km) +/- [4.13,1.33,2.41]

Event waveforms with P, S picks

Solid: auto-picks, Dashed: predicted arrivals



Check “**NEW**” events in enhanced catalog, not in reference catalog

Sorted by **depth**
(Southwest Puerto Rico)

Origin time (UTC)	OT (seconds)	Latitude	Longitude	Depth	Magnitude	EventID
2020-01-07T16:19:28.490000	58768.490000	18.146667	-66.835500	37.840000	2.980000	1001012
2020-01-07T20:43:26.700000	74606.700000	17.933500	-66.923333	38.810000	1.570000	1001342
2020-01-07T15:58:52.370000	57532.370000	18.353667	-67.028000	39.160000	1.310000	1000980
2020-01-07T20:27:22.970000	73642.970000	17.931000	-66.846500	39.800000	1.560000	1001320
2020-01-07T22:21:53.870000	80513.870000	18.466000	-66.986833	40.580000	2.170000	1001466



Unrealistic depths
for the region?
(SW Puerto Rico)

Catalog QC: False detection in coda of larger earthquake (also, too deep?)

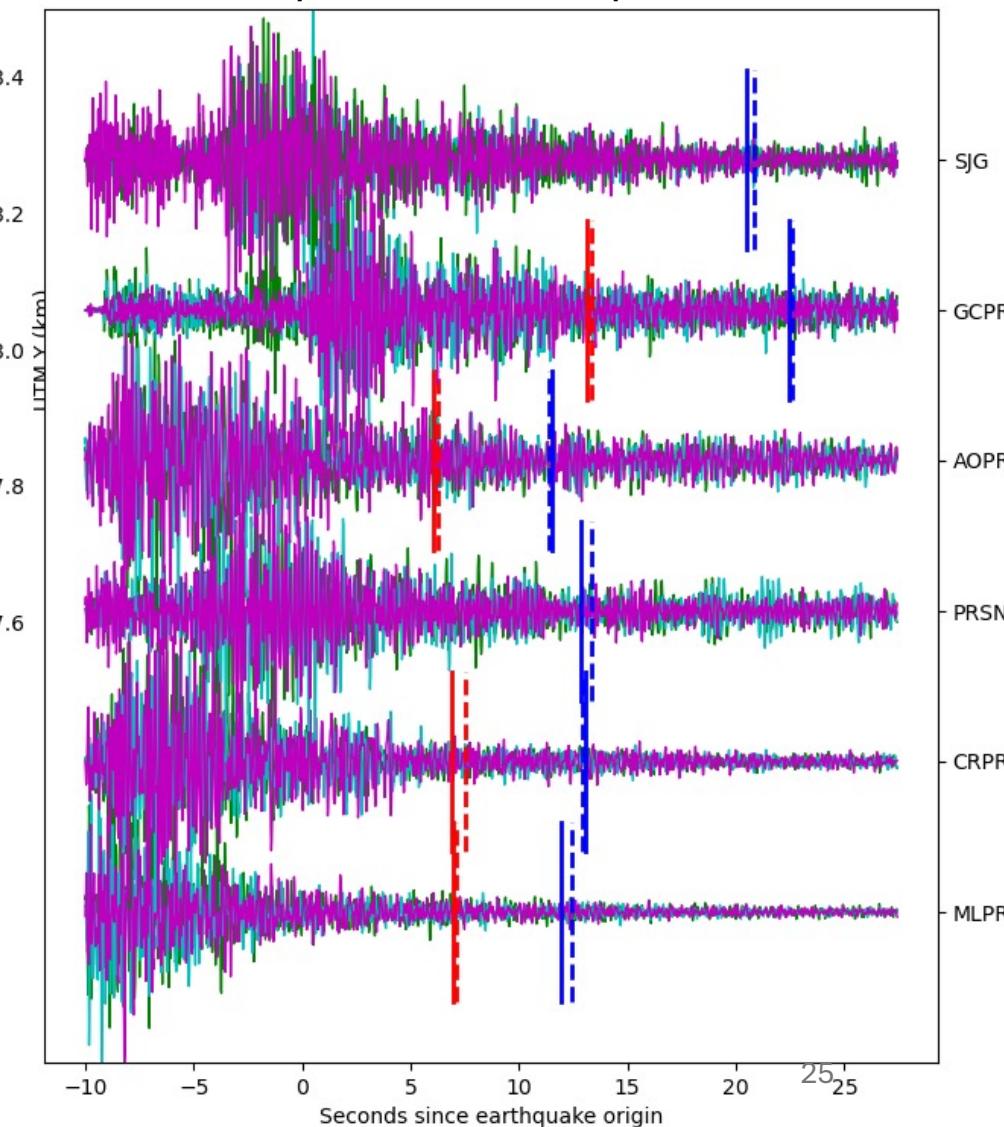
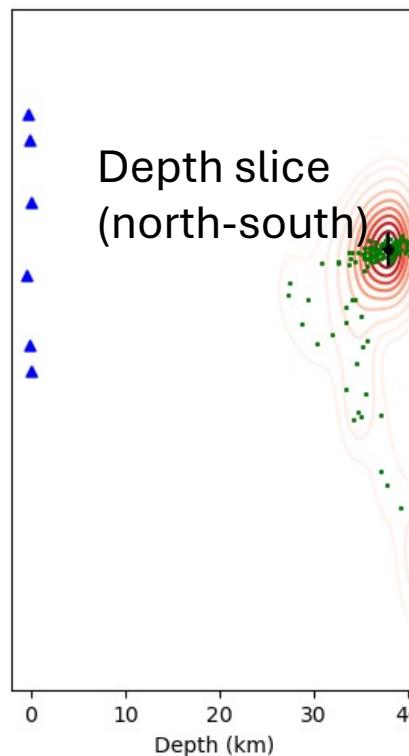
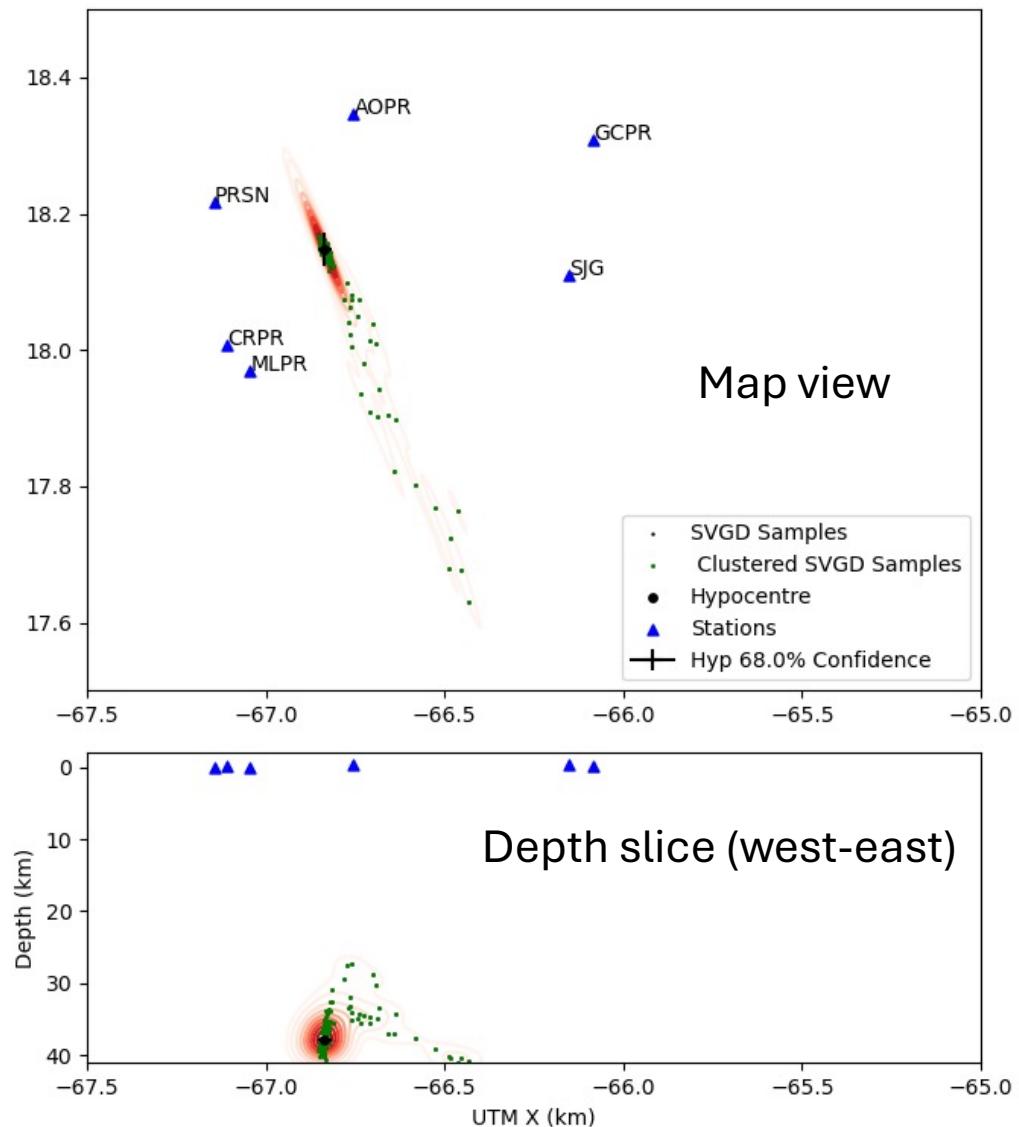
2020-01-07T16:19:28.490000 58768.490000 18.146667 -66.835500 37.840000 2.980000 1001012



Earthquake 2020-01-07 16:19:28.491959877 +/- 0.20s
Hyp=[-66.84,18.15,37.84] - Hyp Uncertainty (km) +/- [2.08,2.71,0.58]

Event waveforms with P, S picks

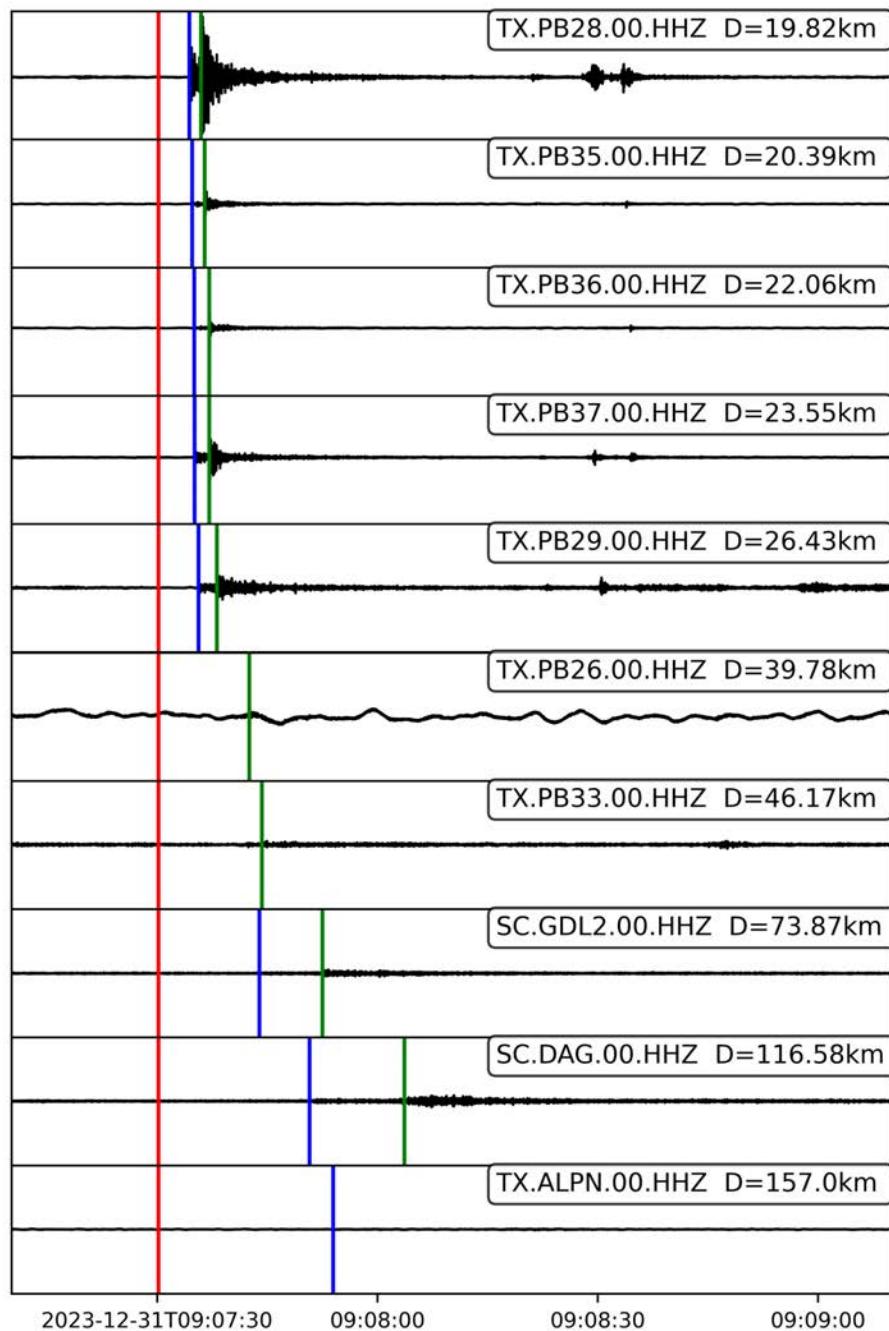
Solid: auto-picks, Dashed: predicted arrivals



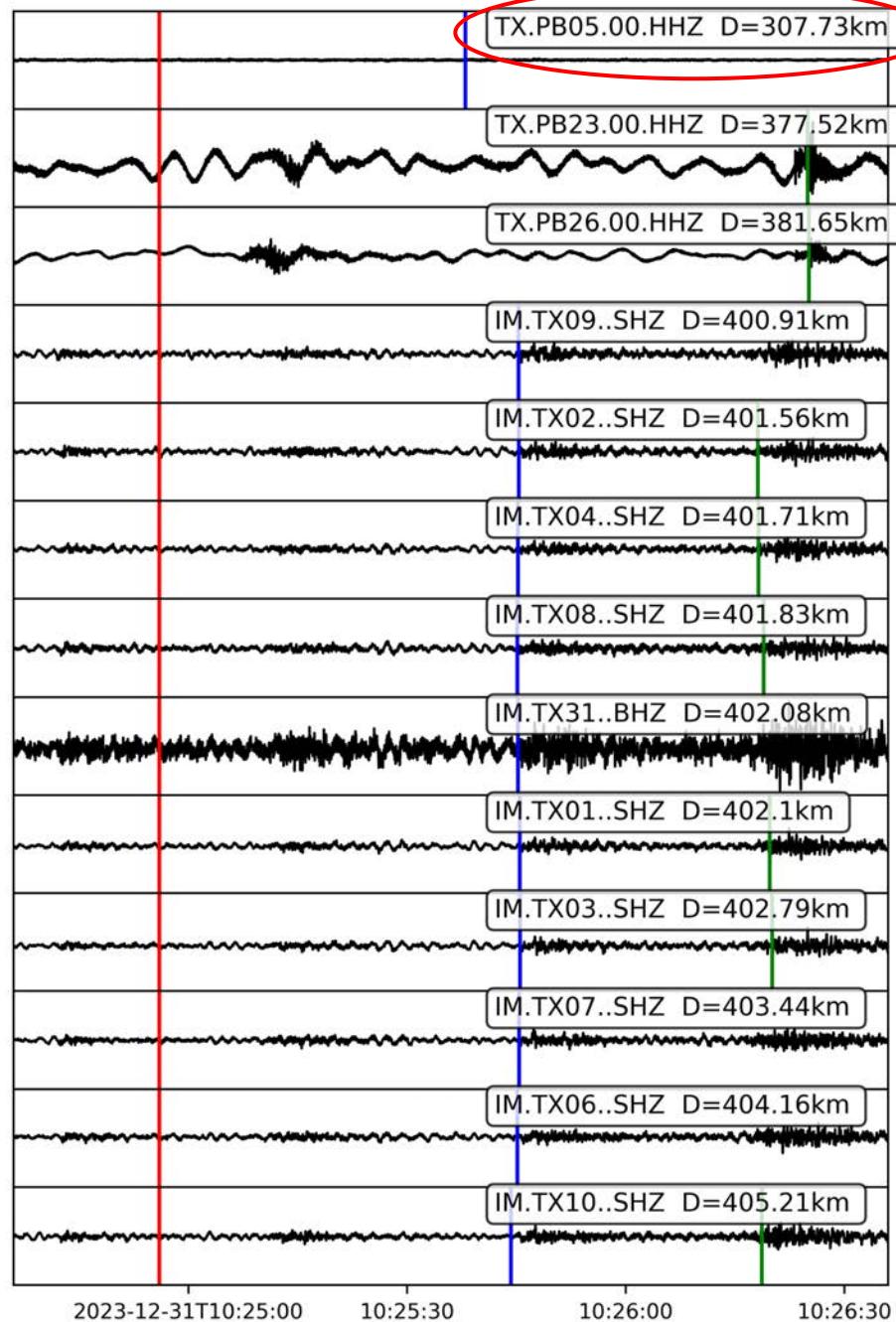
Post-processing: automatic removal of false detections in coda of larger earthquakes

- Visual detection of event waveforms with unrealistic depths (30-40 km): false detections in coda of larger earthquakes
 - EQTransformer (Mousavi et al., 2020) with low thresholds for event detection and P, S picks
- Devised an empirical algorithm to remove these specific false detections (Yoon et al., 2023, BSSA)
 - Apply only to NEW events not in reference catalog, not to MATCH events. Must meet ALL 4 criteria below for removal.
 - 1) current event was within 45 seconds of the previous event in time;
 - 2) current event had 14 or fewer phases;
 - 3) previous event had at least 5 more phases than the current event;
 - 4) current & previous events within 1 local magnitude unit of each other

PyOcto Event #82 : No Match to ComCat or GaMMA
Red: Event, Blue: P Pick, Green: S Pick



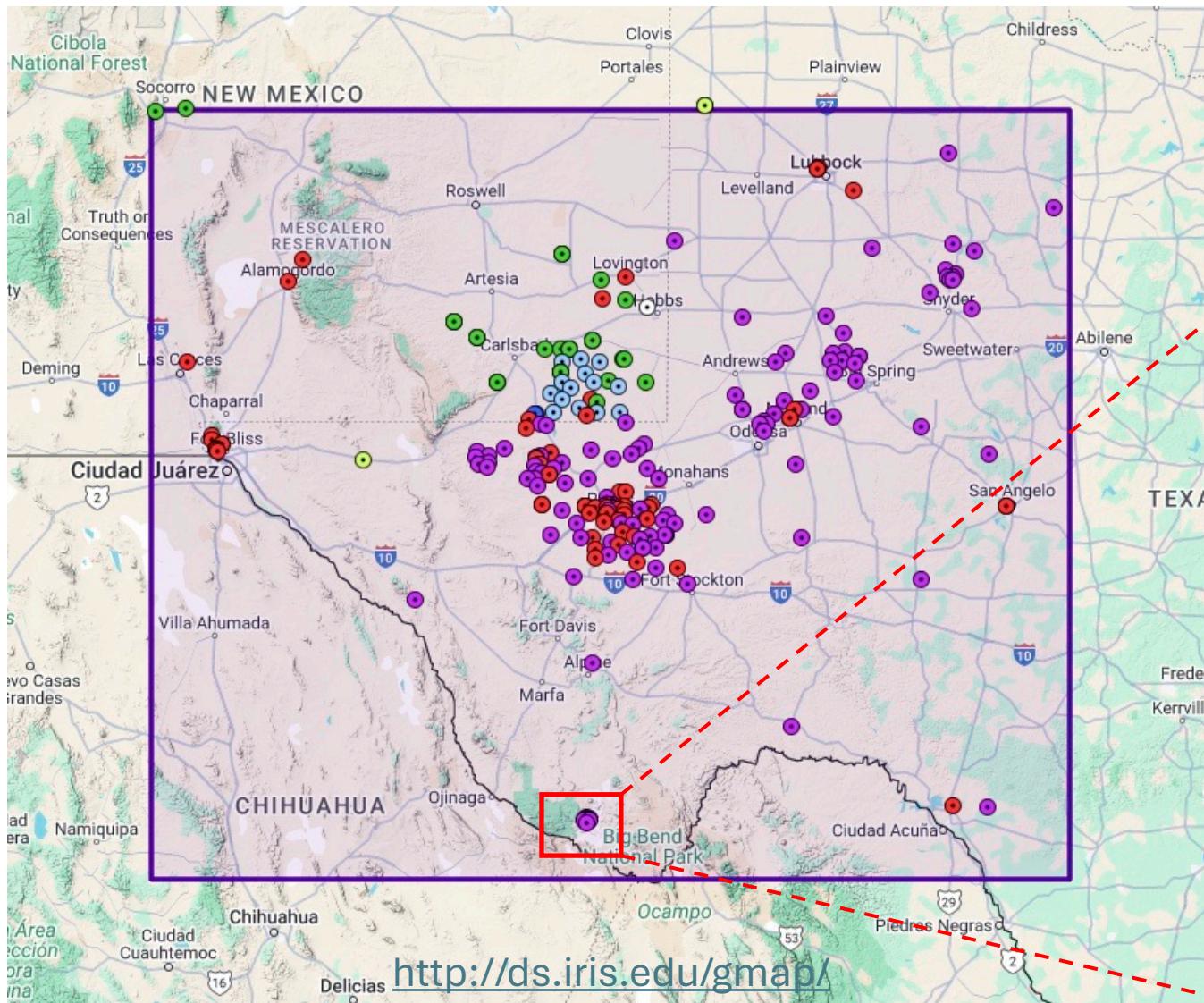
PyOcto Event #90 : No Match to ComCat or GaMMA
Red: Event, Blue: P Pick, Green: S Pick



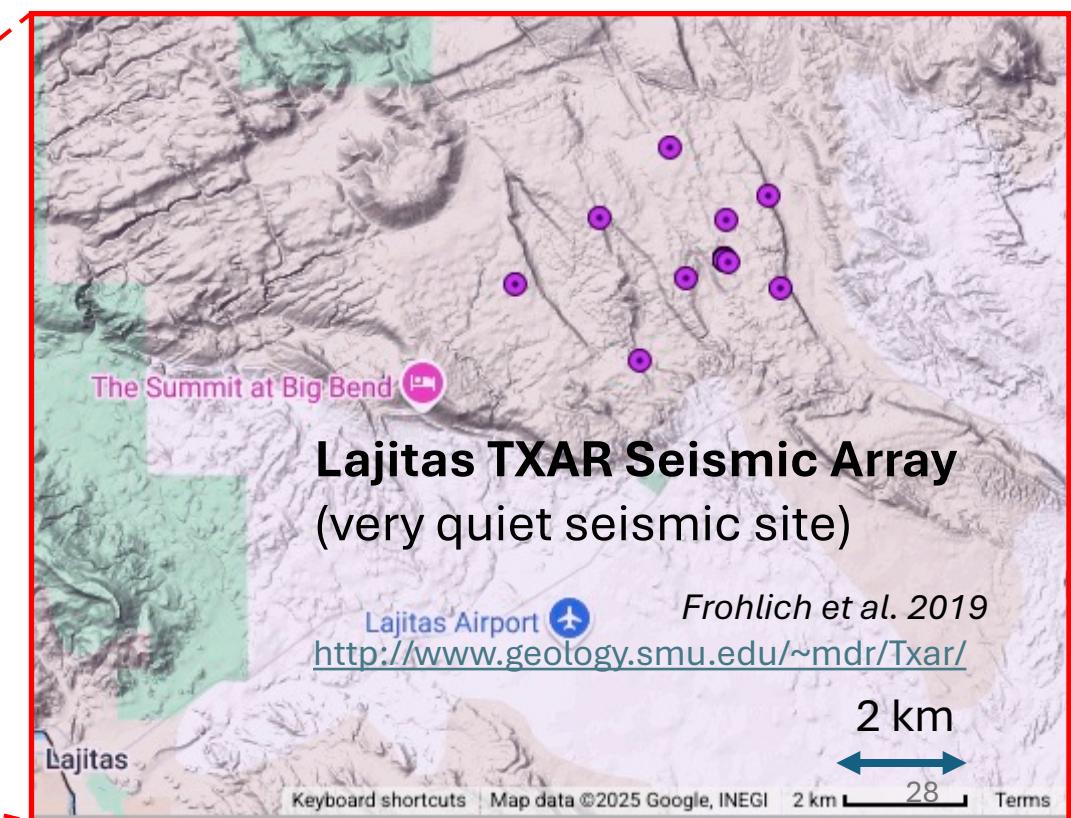
- Catalog QC: West Texas
- After event association
- (Left) new local earthquake 😊
- (Right) false detection 😞
 - large distance (307 km) to nearest station; IM network?

2-minute waveforms
Vertical lines:
Origin time, P, S

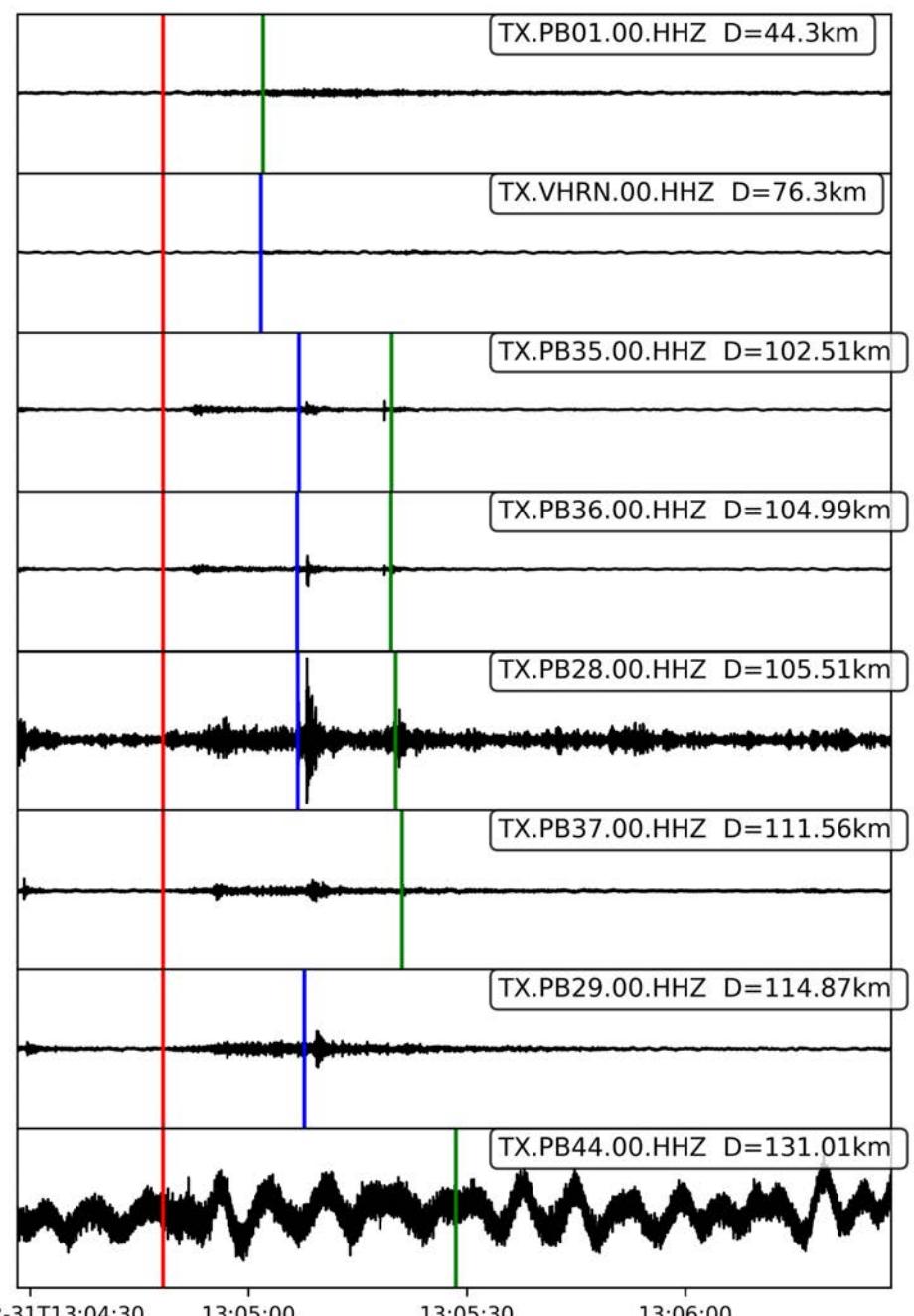
Problem: Incorrect association/location due to **uneven station distribution** (array far from rest of network & seismic sources)



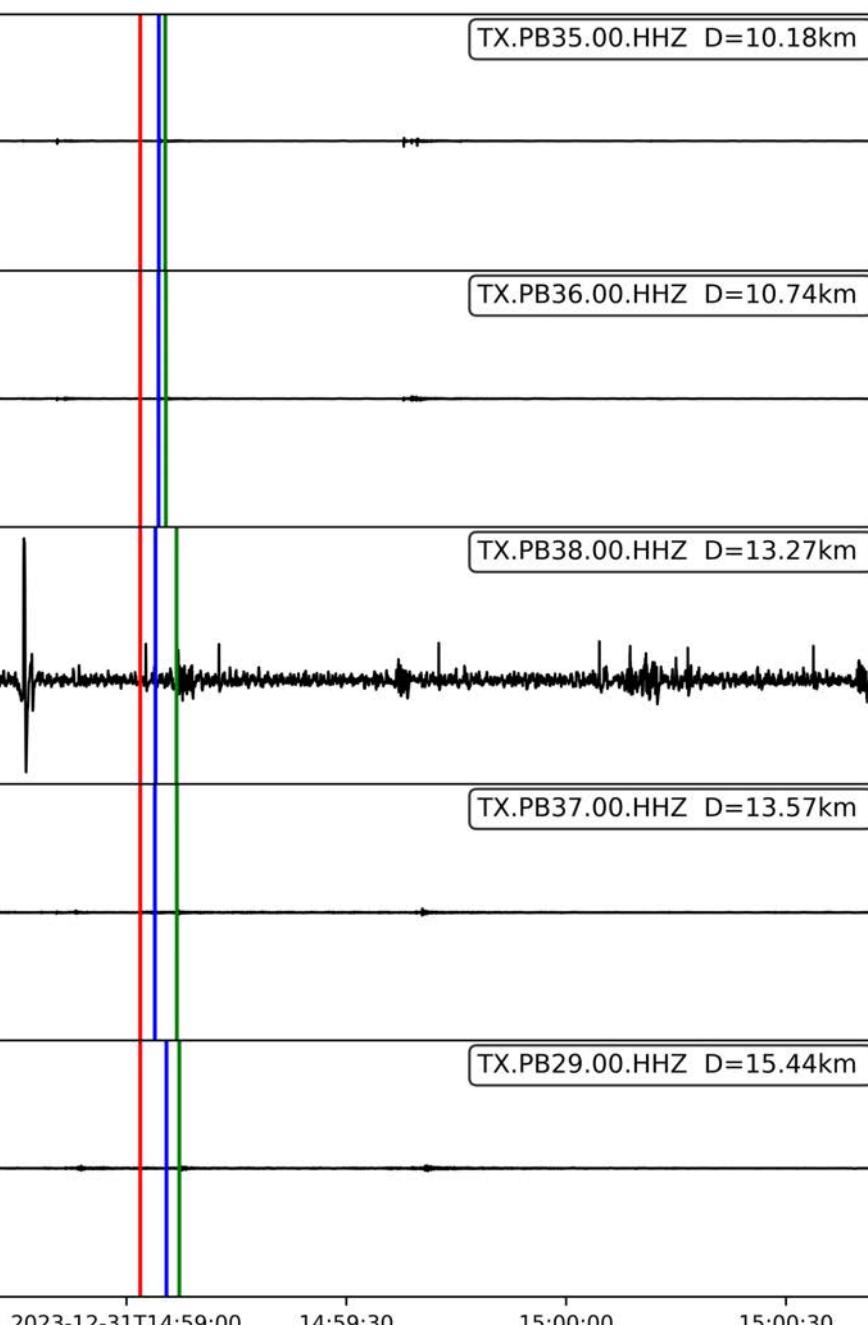
Solution: Use only one station from array (IM.TX01) for event association



PyOcto Event #218 : 5/5/10 No Match to TexNet, Match to GaMMA, NSD<75.0
Red: Event, Blue: P Pick, Green: S Pick



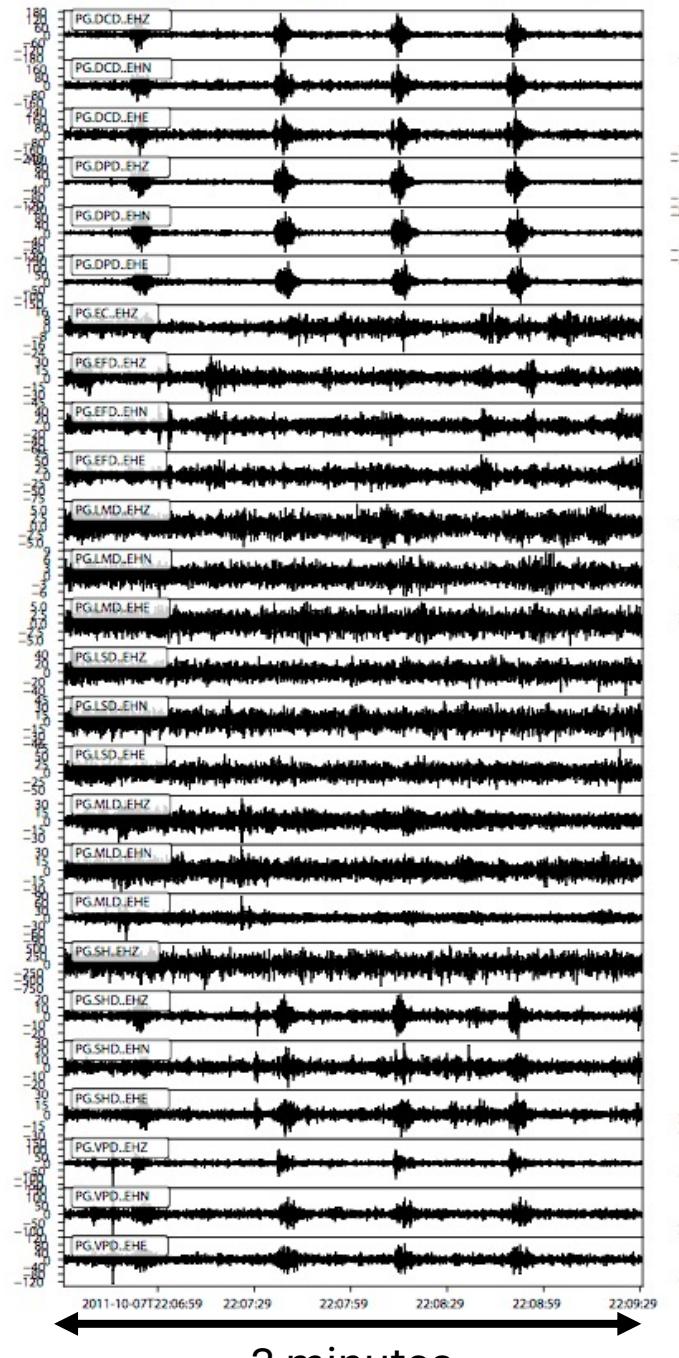
PyOcto Event #250 : 5/5/10 No Match to TexNet, Match to GaMMA, NSD<75.0
Red: Event, Blue: P Pick, Green: S Pick



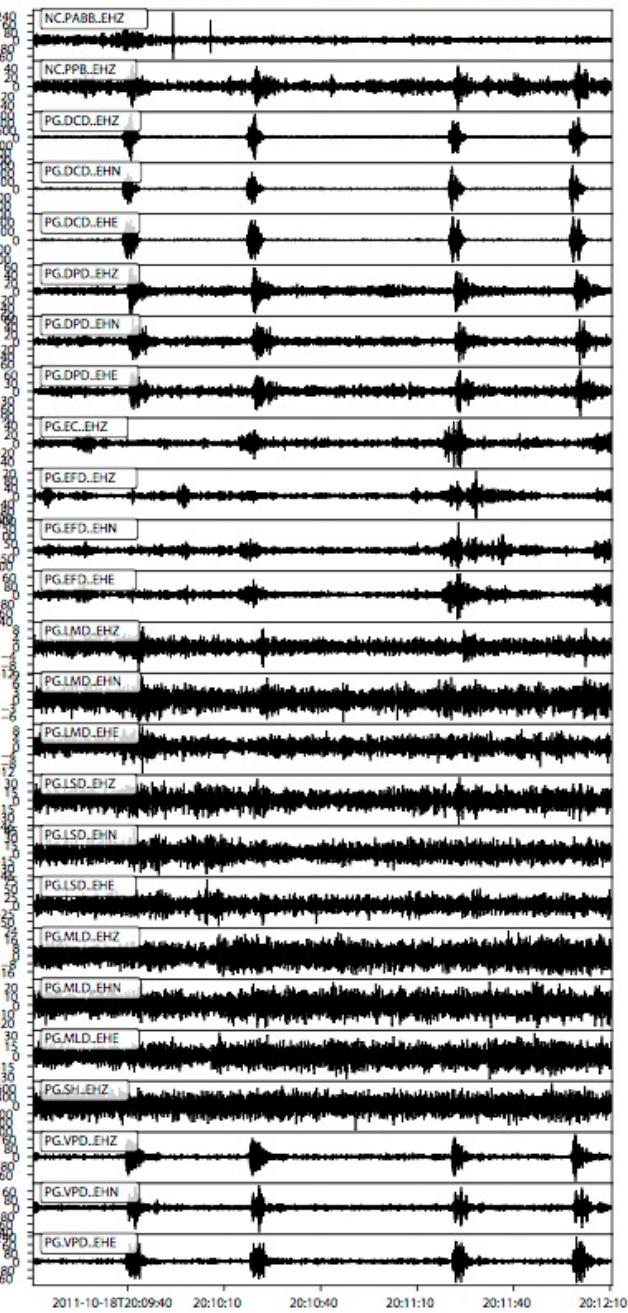
- Catalog QC: West Texas
- Bad event association
- (Left) P, S picks from 2 different earthquakes associated as one event
- (Right) Noise picks associated as event → wrong moveout

2-minute waveforms
Vertical lines:
Origin time, P,₂₅S

2011-10-07T22:06:29.77 - 2011-10-07T22:09:29.77



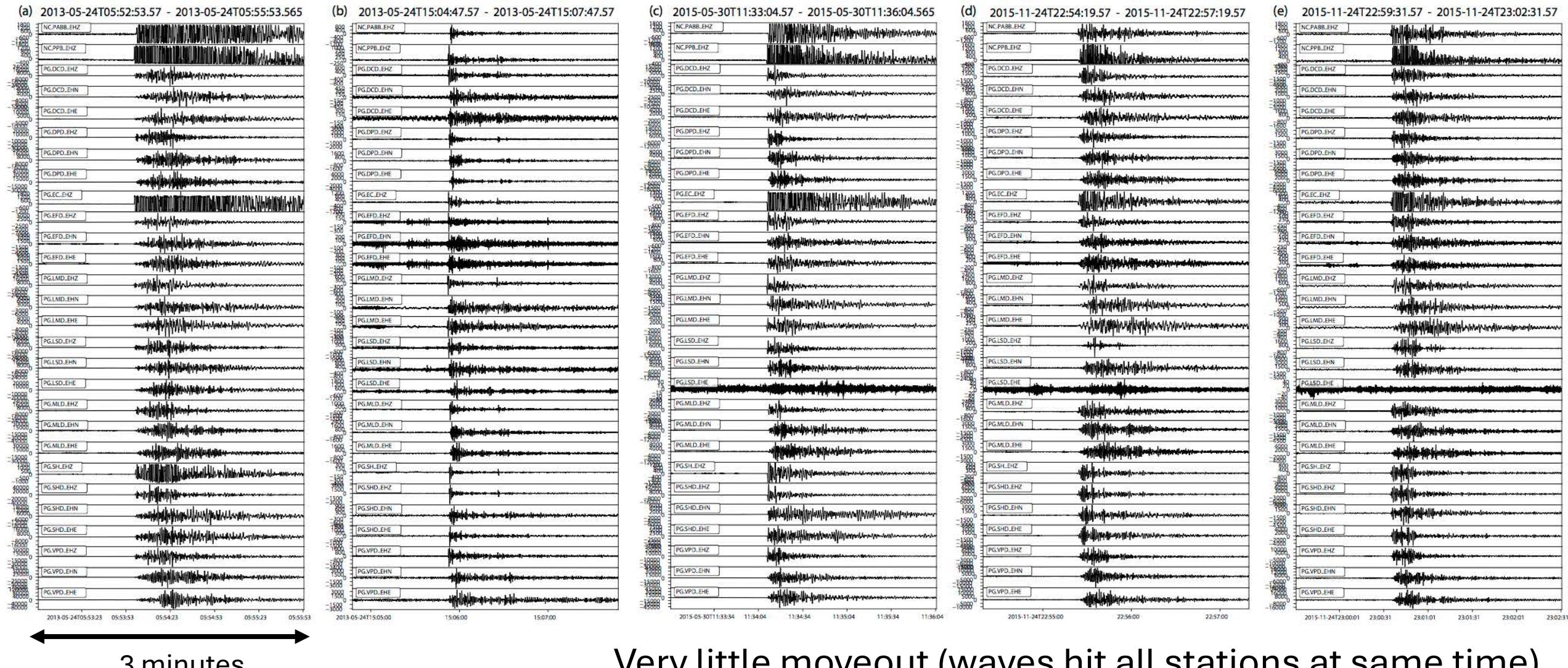
2011-10-18T20:09:10.57 - 2011-10-18T20:12:10.57



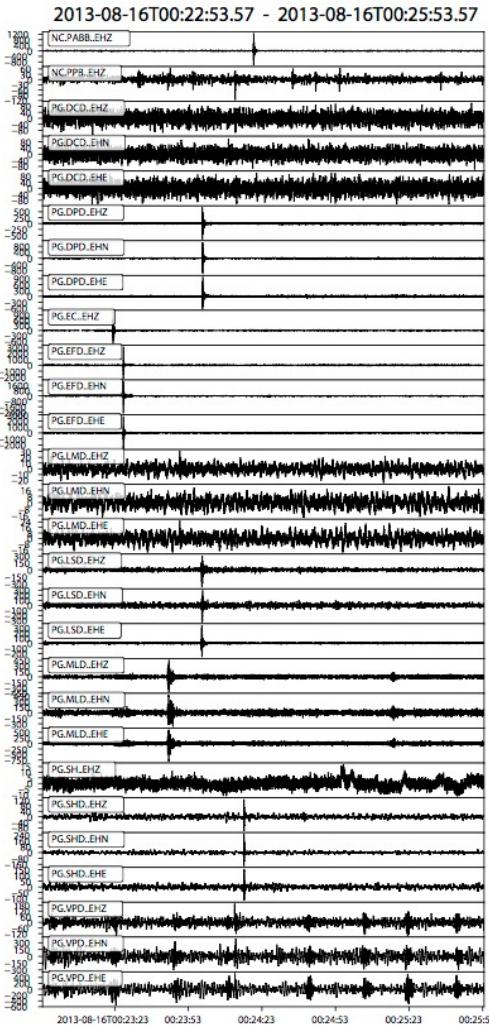
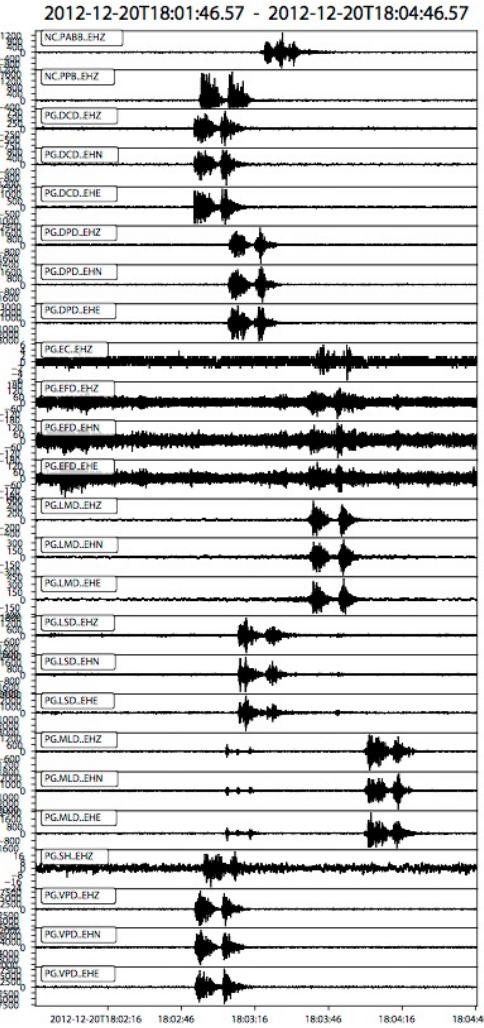
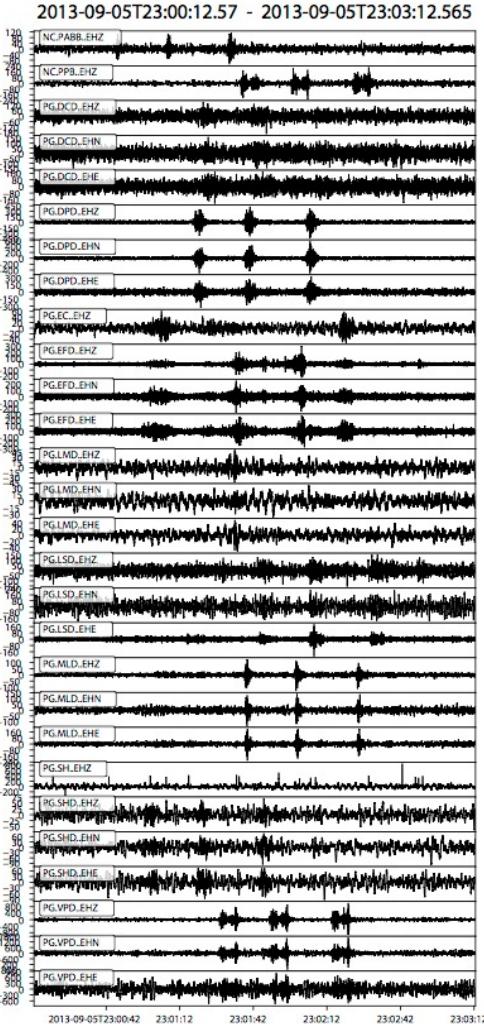
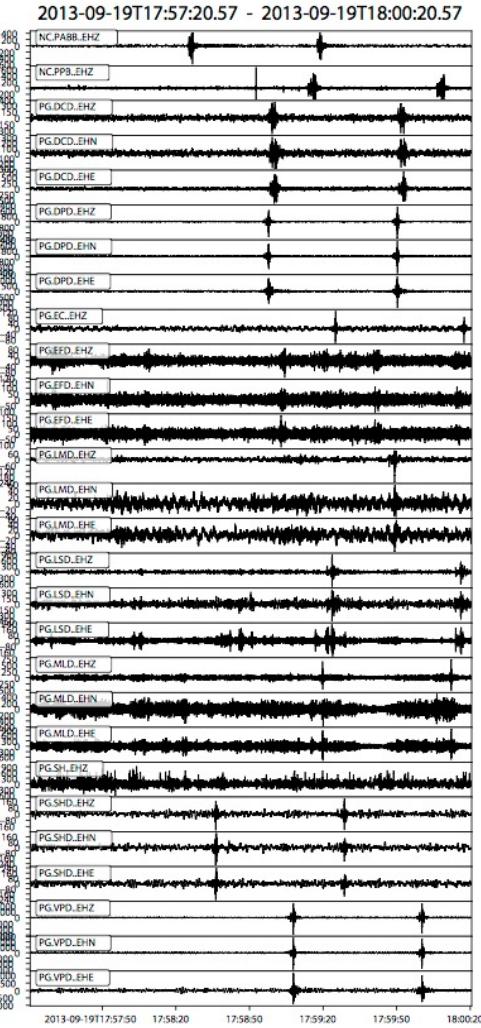
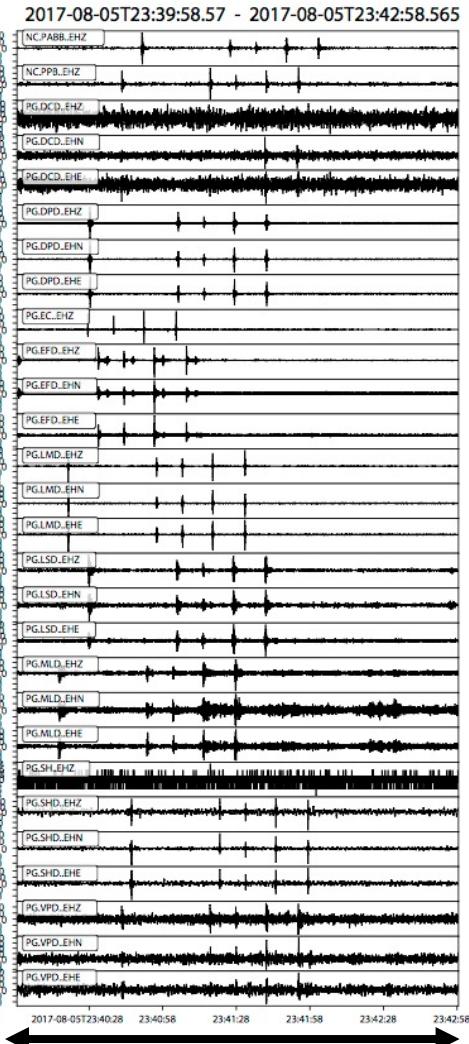
- Catalog QC: Central CA
- Filtered 3-12 Hz
- Non-earthquake signals: active-source survey



Deep teleseismic earthquakes (depth 600+ km) recorded in Central CA



Infrasound signals recorded in Central CA



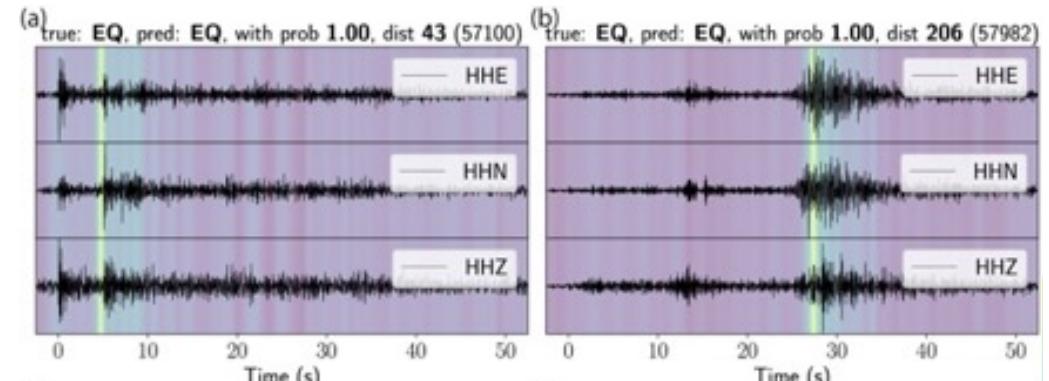
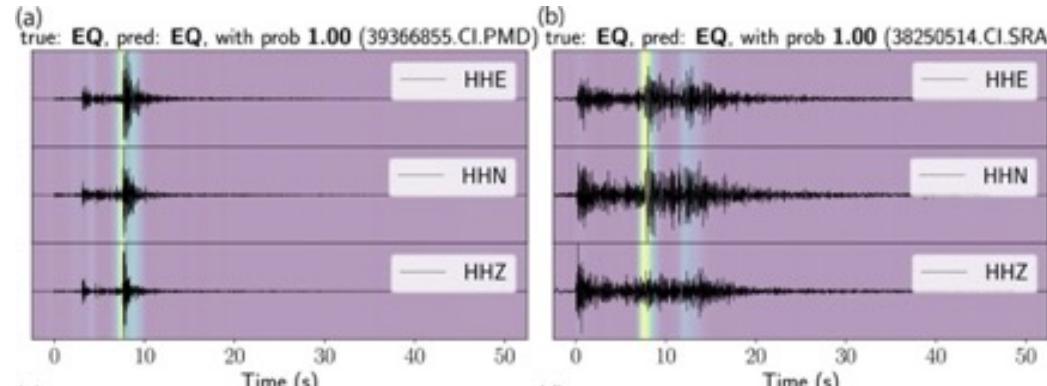
Slower moveout from sound waves

Quarry blasts: often in enhanced catalogs, but not 'interesting' to many seismologists

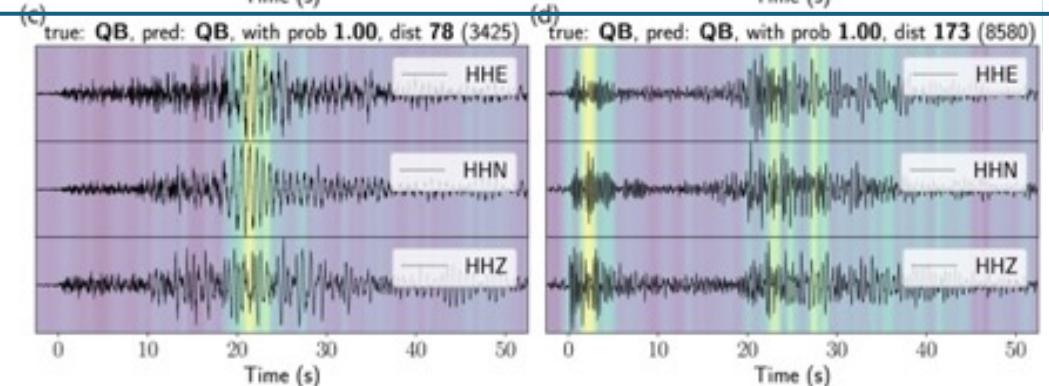
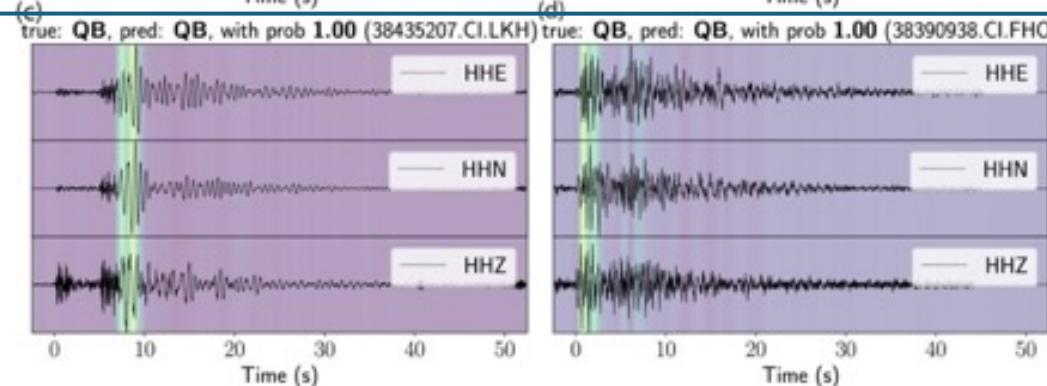


Check ComCat for blast events

earthquake



quarry blast



↔

50 seconds

Check waveform plots:
lower frequency content?

Events located near
known quarry?

Events occur only
in local daytime?

Automated post-processing strategies to remove false detections & unwanted signals from catalog

[Machine-learning]
classifiers: discriminate
different seismic signals

- Teleseismic earthquakes,
quarry blasts, cultural noise
sources, infrasound, ...
- Assemble training data set,
create classifier model



Empirical algorithms: remove
specific false detections after
visual inspection

- e.g. false detections in coda of
larger earthquakes
- may be specific to network
geometry & source distribution

Simple thresholding & filtering to
discard non-earthquake signals

- easy & effective, but be aware of tradeoffs

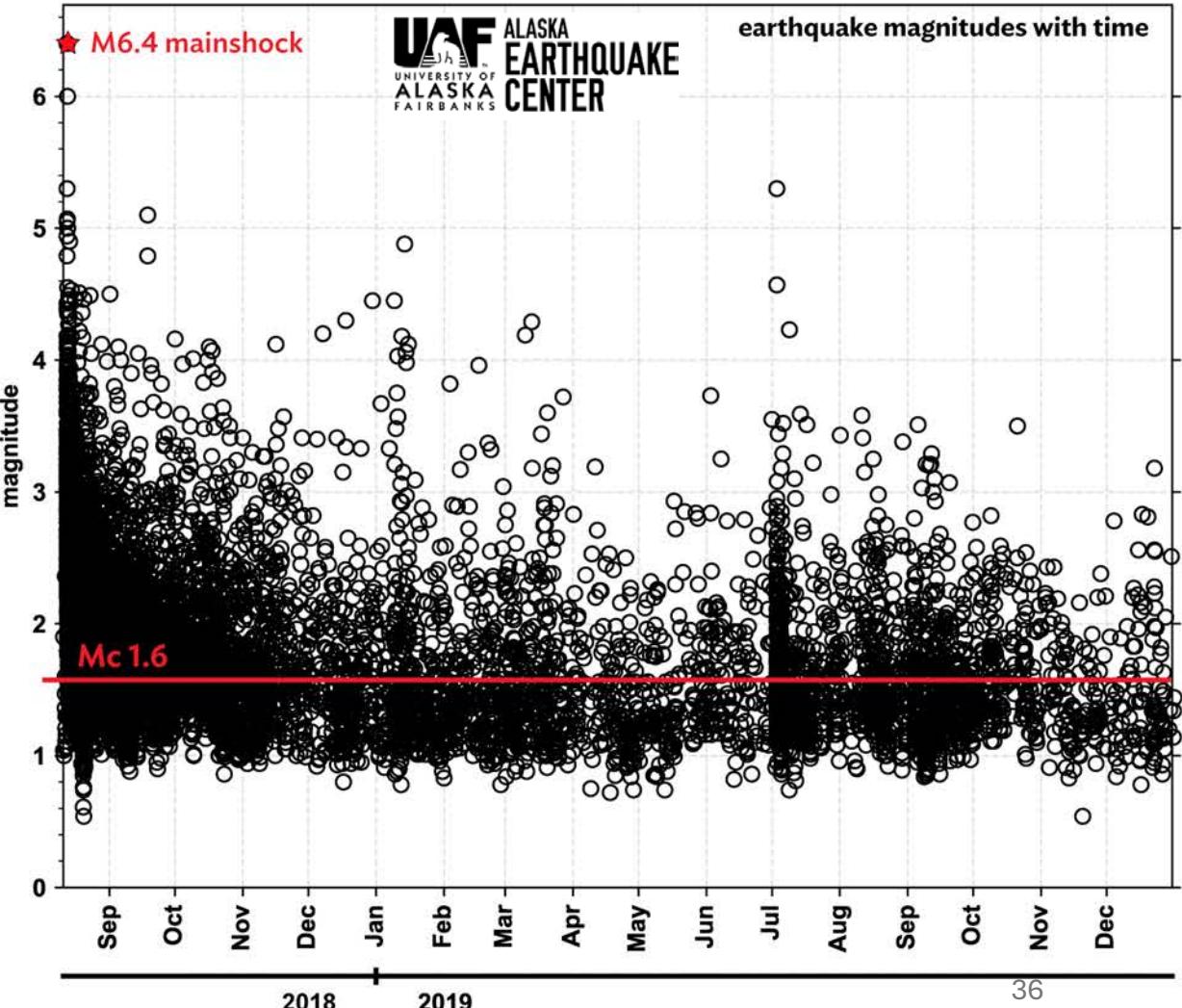
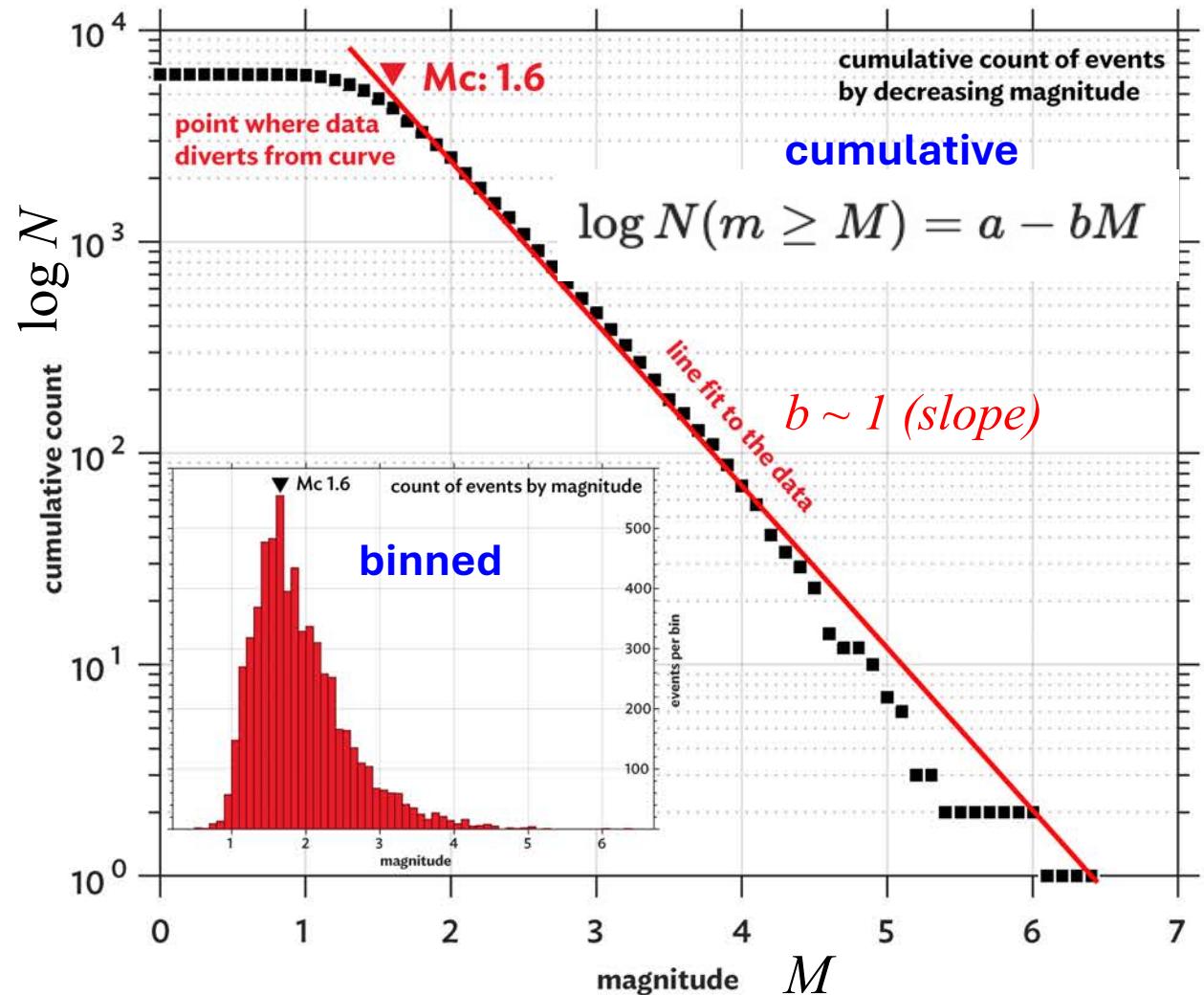
Simple thresholding & filtering to discard non-earthquake signals

- easy & effective, but be aware of tradeoffs

- Restrict locations to a certain region and/or depth range
- Set minimum threshold on:
 - Number of (P, S) picks
 - Number of stations per event
 - Output probabilities for P, S picks
- Set maximum bound on:
 - Distance to nearest station
 - Distance to farthest contributing station with picks
 - Azimuthal gap
 - RMS residual
 - Travel time residual for P, S pick at given station

Magnitude of Completeness (Mc)

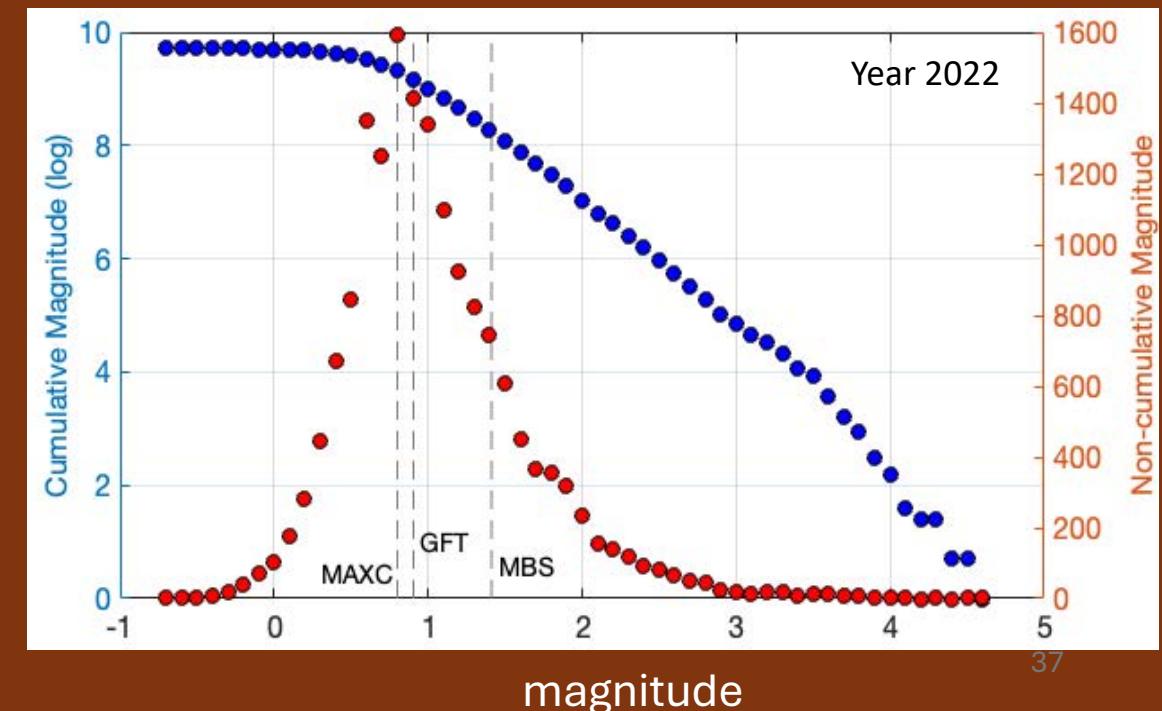
Magnitude-frequency distribution (MFD) (Gutenberg-Richter)



Many different methods to get catalog Mc

- **Maximum Curvature (MAXC)**: max value of 1st derivative of MFD (typically, max of binned MFD)
 - * most reliable for small sample sizes (~50-100 events)
 - * works best with local datasets with fewer heterogeneities
- **Goodness-of-fit Test (GFT)**: lowest magnitude cut-off where Gutenberg-Richter relation holds
- **Mc by b-value Stability (MBS)**: first magnitude increment where $|b_{\text{avg}} - b| <$ uncertainty of b
 - * tends to produce the highest (i.e., most conservative) Mc

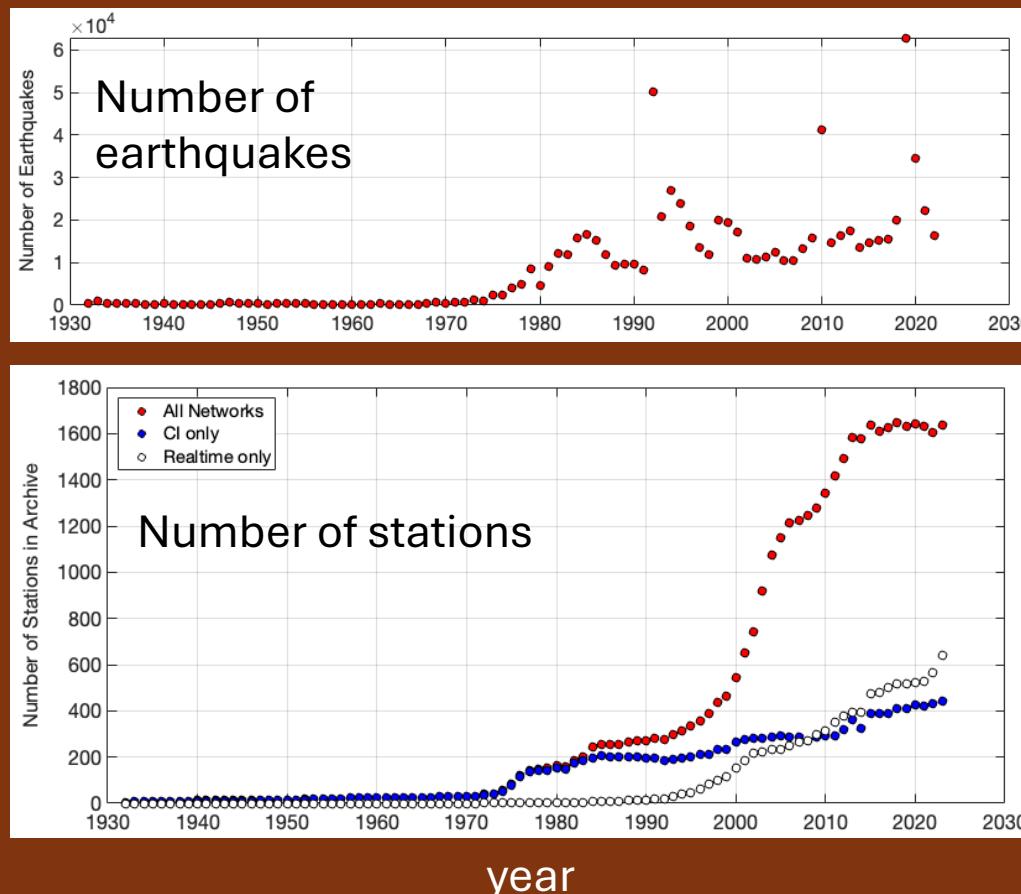
- * More info and original R codes from CORSSA: <https://corssa.org/> <https://doi.org/10.5078/corssa-39071657>
- * Matlab codes available on Github: https://github.com/gtepp/research_codes
- * Other methods to get Mc:
ZMAP (Wiemer, 2001)
b+ positive (van der Elst, 2021)



Temporal Mc changes

Past 90 years: SCSN Mc has decreased
from ~3 to ~1

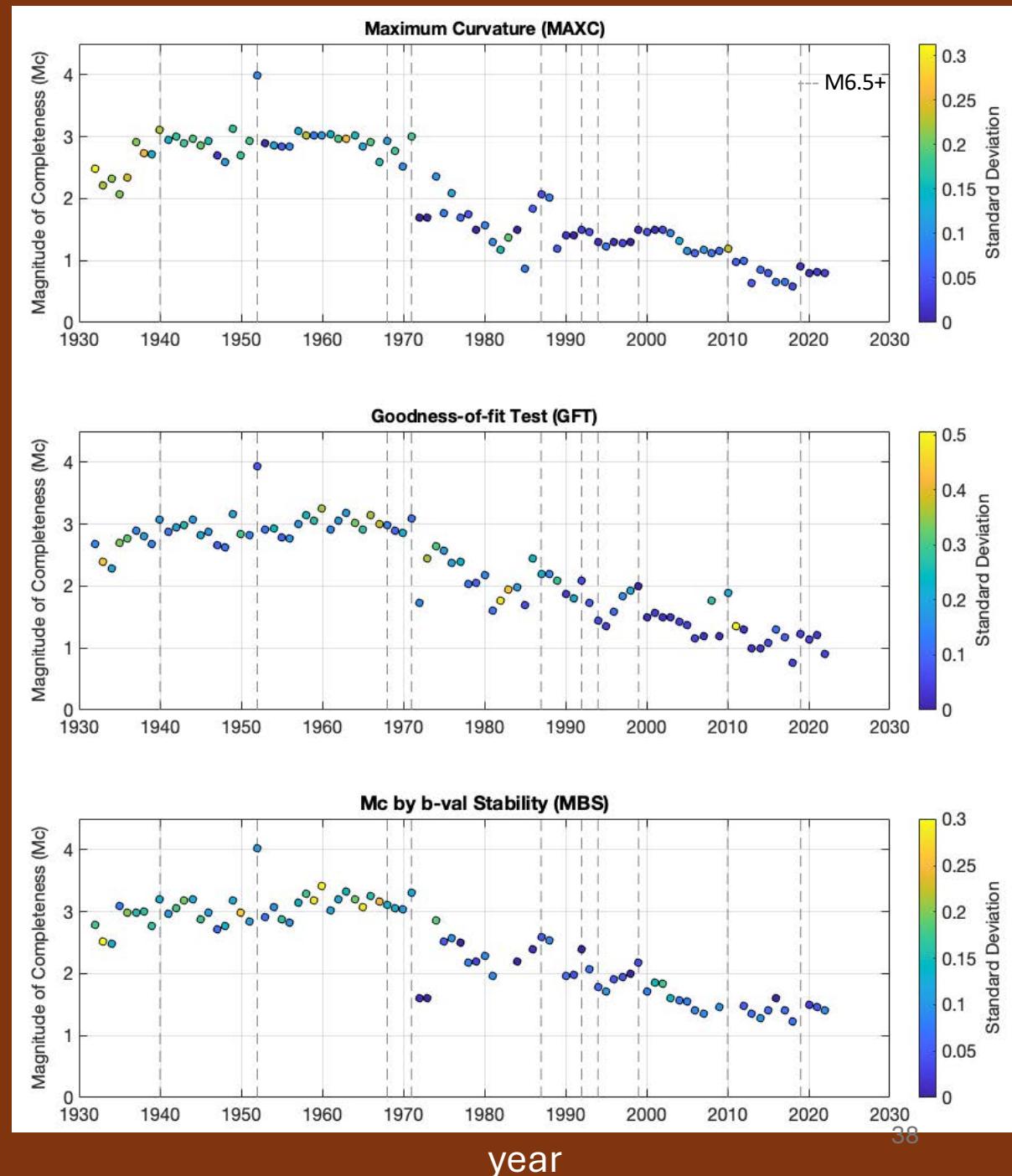
Full SCSN Catalog



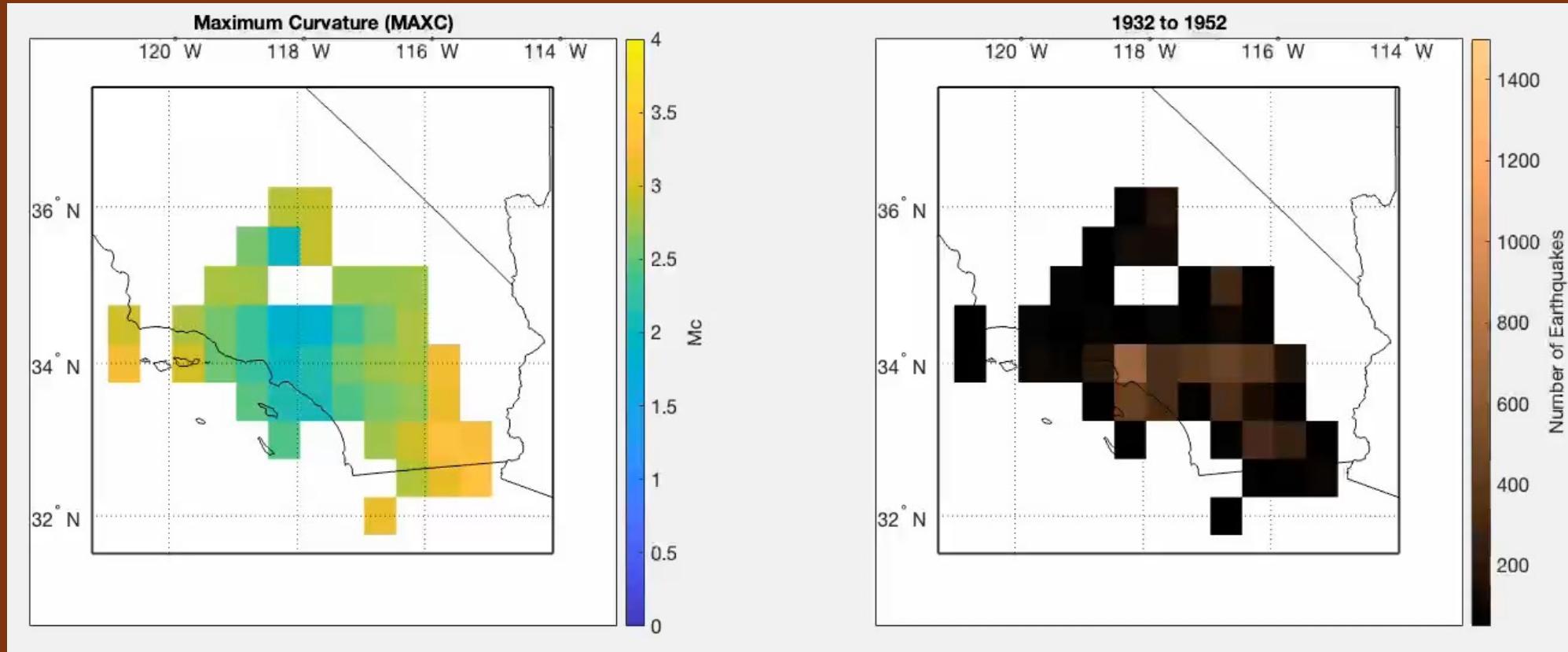
Mc

Mc

Mc

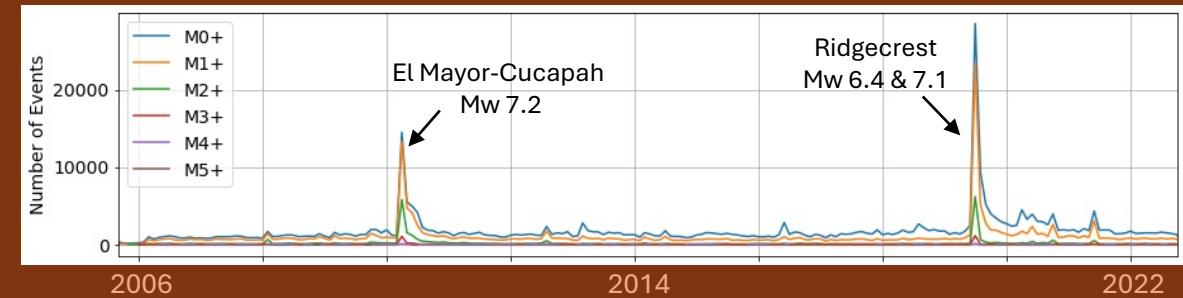
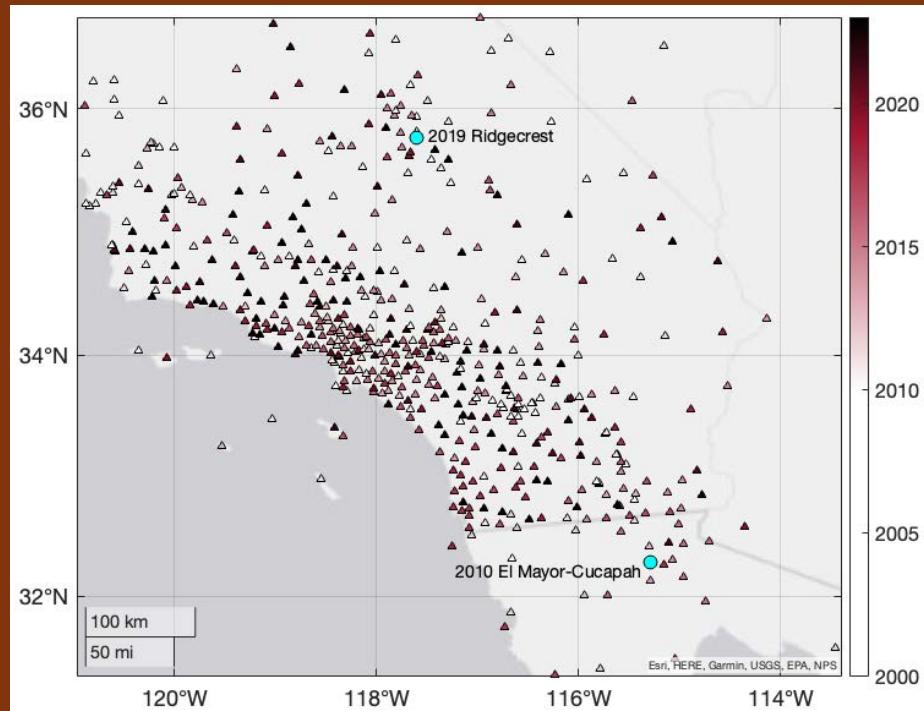


Spatial Mc Changes



- Bins: 0.5° (~ 50 km) squares
- 1932-1972: 20 years
- 1973-2022: 10 years
- Maximum curvature method
- Minimum 30 EQs/bin

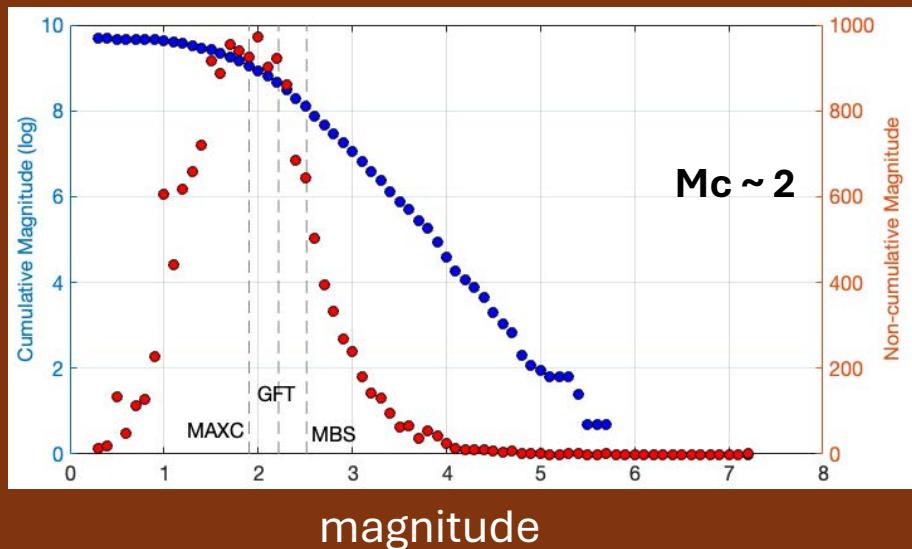
Past 90 years: SCSN Mc has decreased from ~ 3 to ~ 1 , but not uniformly throughout the region



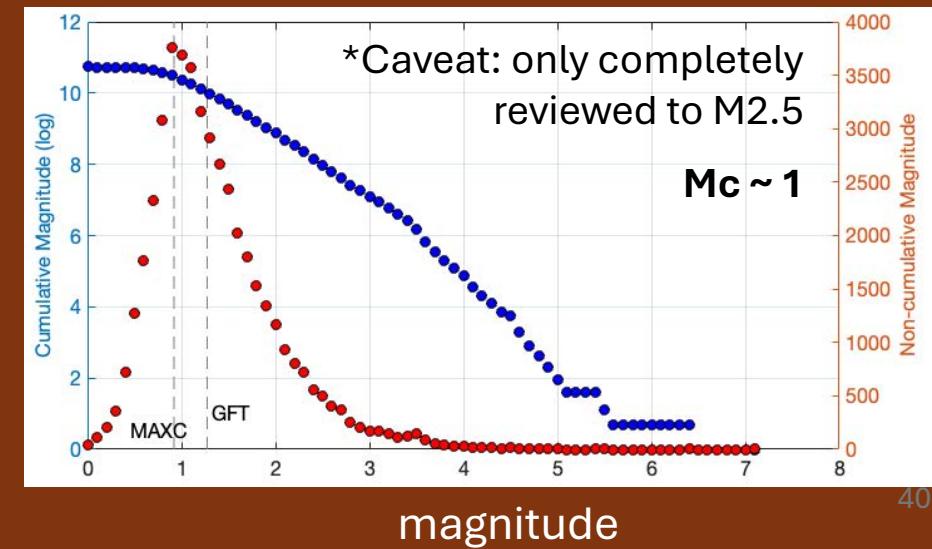
Expect higher Mc earlier in aftershock sequence

Ridgecrest Mc is ~1 mag. unit lower than El Mayor-Cucapah Mc

2010 M_w 7.2 El Mayor-Cucapah sequence

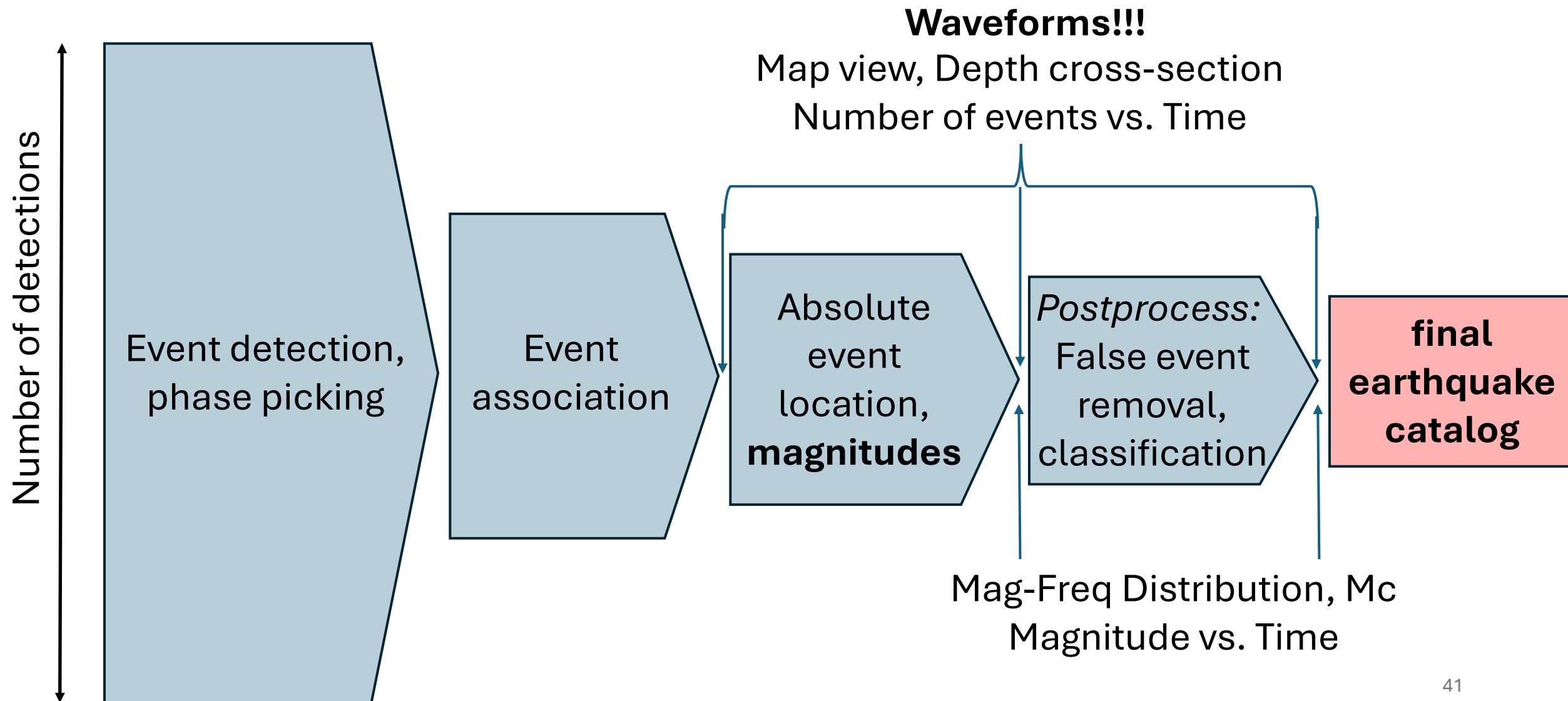


2019 M_w 6.4 & 7.1 Ridgecrest sequence



Catalog QC: Summary of things to check in plots

Compare against a reference catalog!



Extra slides

Earthquake Catalog QC General Tips

- Make plots - think like an analyst – use visual tools. Waveforms!
- Use seismology domain expertise
 - If enhanced catalog events don't look like real earthquakes, figure out why.
- Rule out all sources of noise (background or cultural). Be sure it's not an artifact.
- Check it's not a regional/teleseismic earthquake.
- Check it's not a quarry blast, sonic boom, infrasound, active-source explosion, or other unwanted seismic signal.
- Only then have you discovered new earthquakes.

Reference Catalog: How to download?

- For bulk downloads of catalog events from ComCat
 - ComCat search API: <https://earthquake.usgs.gov/fdsnws/event/1/>
 - (I create the URL string, then call wget with a bash script)
 - Python libcomcat: <https://code.usgs.gov/ghsc/esi/libcomcat-python>
 - ObsPy get_events() with USGS Client (I save QuakeML files with picks):
<https://docs.obspy.org/packages/obspy.clients.fdsn.html>
- For ANSS networks, all finalized local events are sent to ComCat. Can also download reference catalog from regional seismic network, which may contain extra events.
 - SCSN: Quarry blasts, teleseismic/regional events.
 - <https://scedc.caltech.edu/>
 - NCSN: Extra events – not all events are submitted to ComCat; catalogs available in other formats.
 - <https://www.ncedc.org/ncedc/catalog-search.html>

Alternative (non-authoritative) reference catalogs

- Regional seismic network websites (for ANSS networks) also have alternative (non-authoritative) earthquake catalogs and data sets for download, to use as reference catalogs
 - SCSN: template-matching, relocated, deep-learning, focal-mechanism,
<https://scedc.caltech.edu/eq-catalogs/altcatalogs.html>,
<https://scedc.caltech.edu/data/deeplearning.html>
 - NCSN: real-time double-difference.
<https://www.ncedc.org/ncedc/catalog-search.html>

MATCH events: location, magnitude comparison

- Compare enhanced vs. reference catalog
 - Location (latitude, longitude, depth, origin-time)
 - Magnitude
 - Distribution of residuals (ideally zero) – are they low enough for you?
- It's difficult to exactly reproduce locations & magnitudes from ComCat (or other reference catalog)
 - Differences in monitoring software, velocity model, manual pick times/weights, filtering before amplitudes, ...
 - Just check that these values are close enough for your needs.

Reference Catalog Tip: Select boundaries carefully

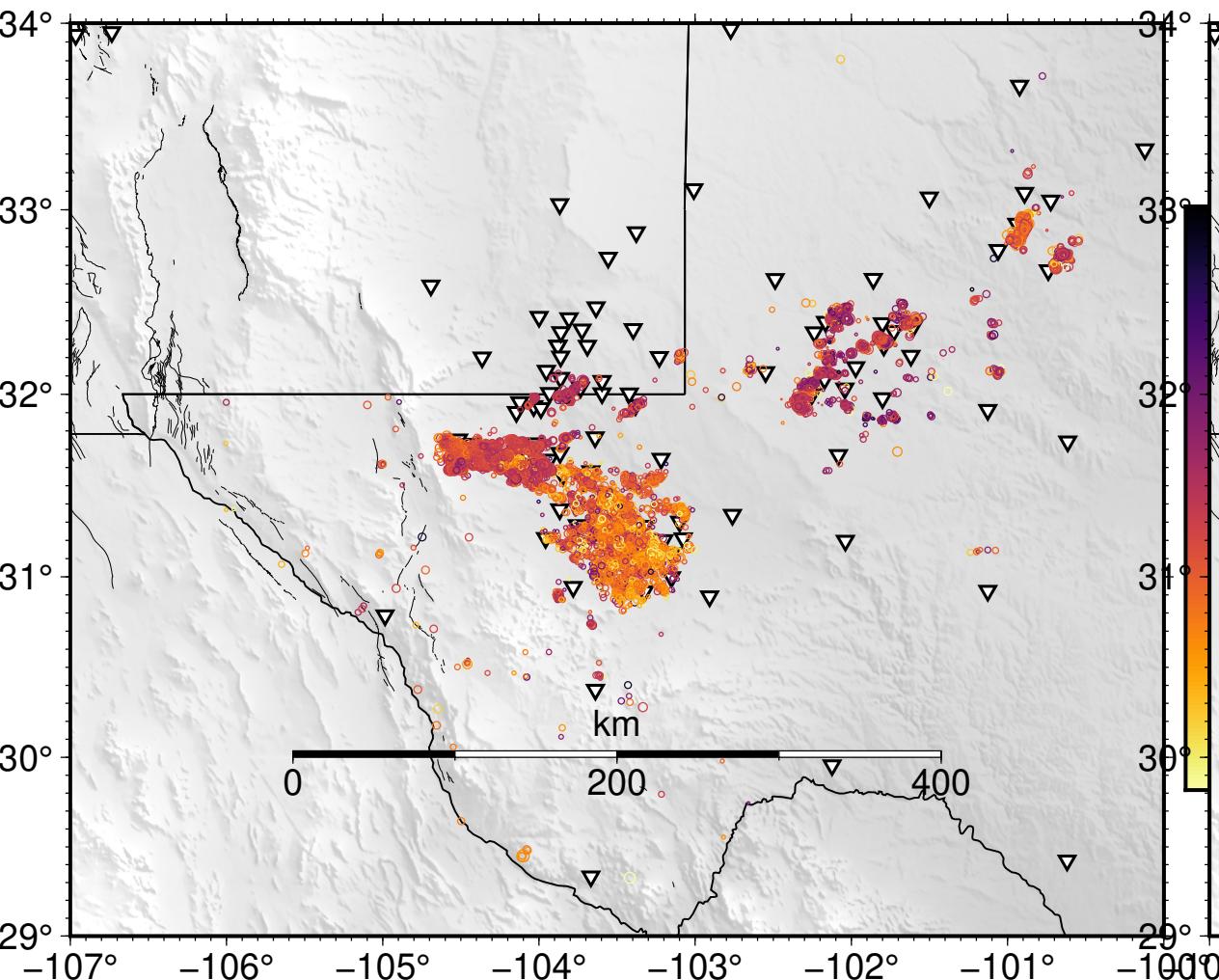
- Carefully select the bounding box or circular radius for downloading your reference catalog, before comparing it to enhanced catalog. It should include:
 - Entire seismogenic region of interest to you
 - Check past seismicity locations on web map: <http://ds.iris.edu/ieb/> or <https://earthquake.usgs.gov/earthquakes/map/>
 - All seismic stations you plan to use for catalog generation
 - Check station locations and availability on web map: <http://ds.iris.edu/gmap/> or <https://www.fdsn.org/networks/>
- **NEW events** in your enhanced catalog might be from reference catalog events outside your selected bounding region
 - Check waveforms: could be a regional distance earthquake if higher duration and amplitude

Catalog QC, locations in map view: TexNet

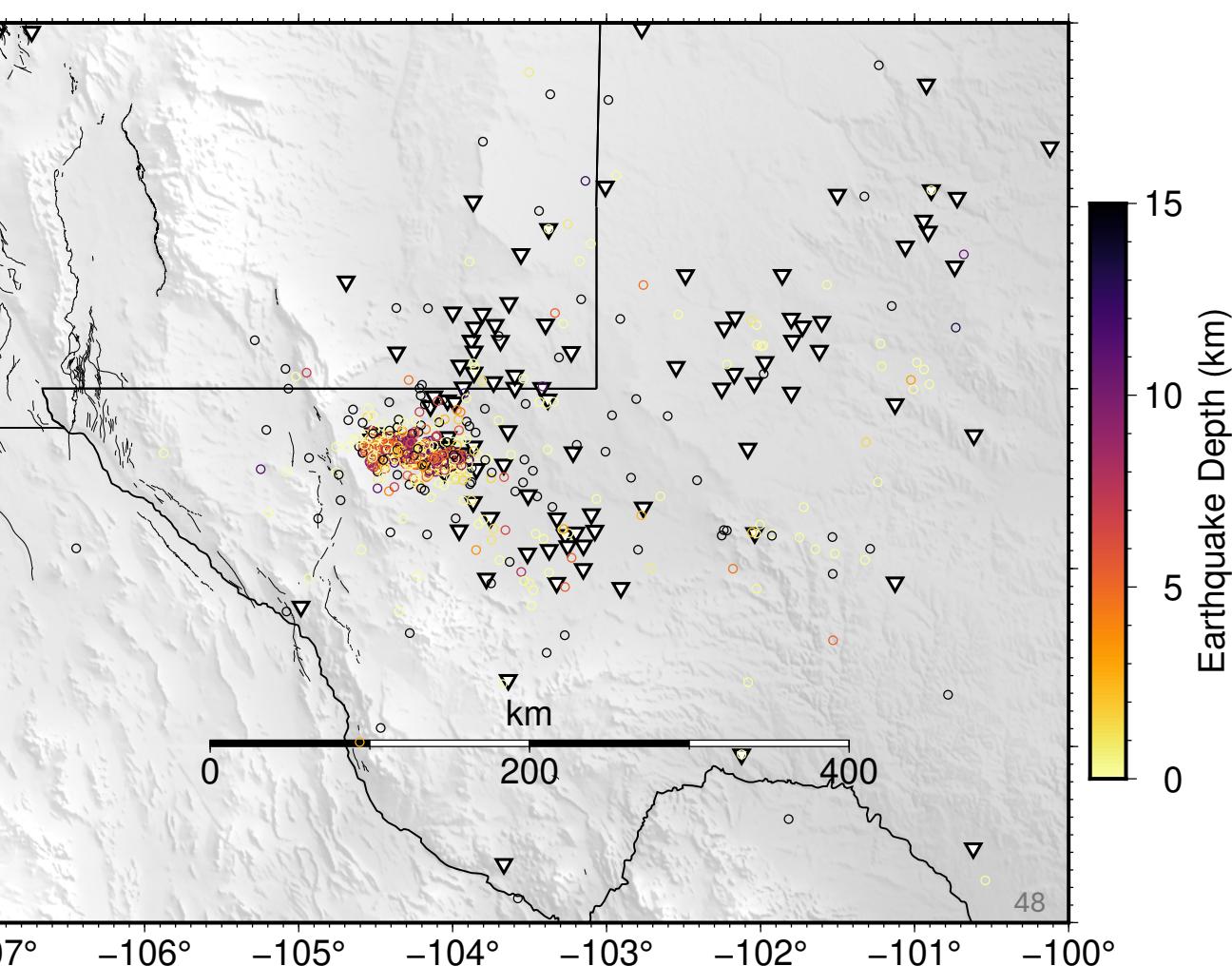


Many scattered event locations –
possible false detections?

Reference catalog:
TexNet (2015-2024: 9 years)



Enhanced catalog: 2023-12-31 (1 day)
GaMMA associator (3 P, 3 S, 6 total picks)

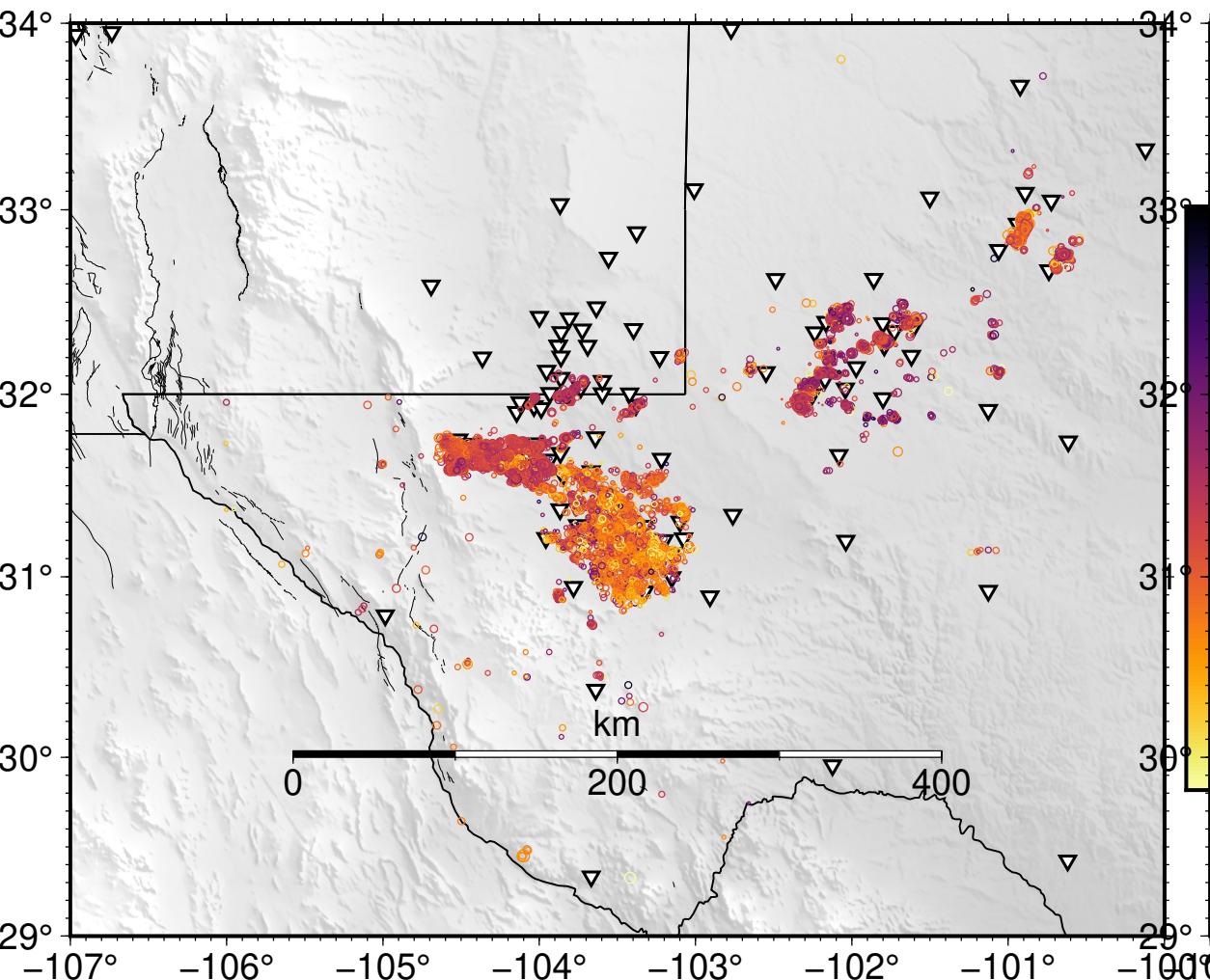


Catalog QC, locations in map view: TexNet

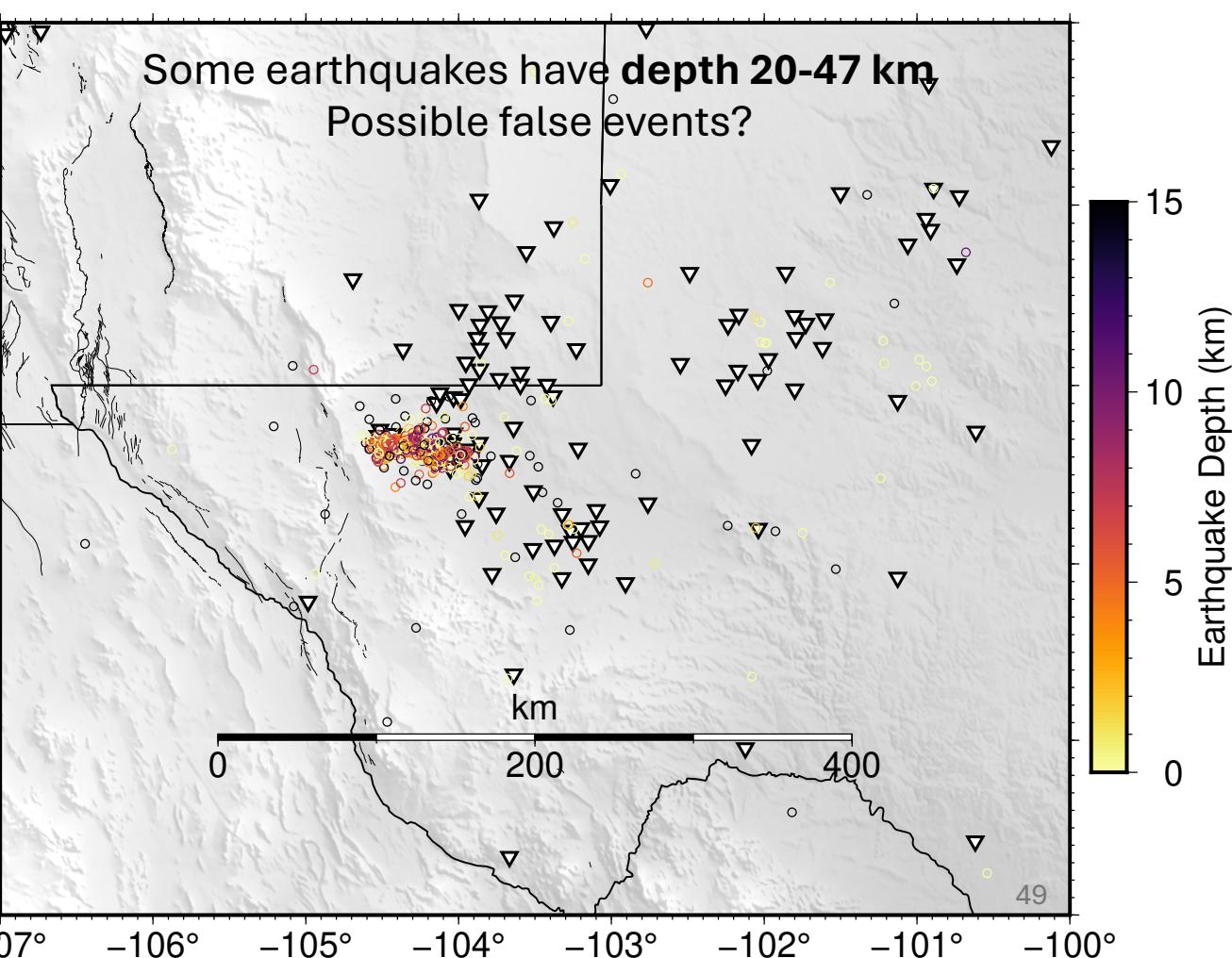


Removed false detections (fewer scattered events)
with stricter association criteria

Reference catalog:
TexNet (2015-2024: 9 years)



Enhanced catalog: 2023-12-31 (1 day)
GaMMA associator (4 P, 4 S, 8 total picks)

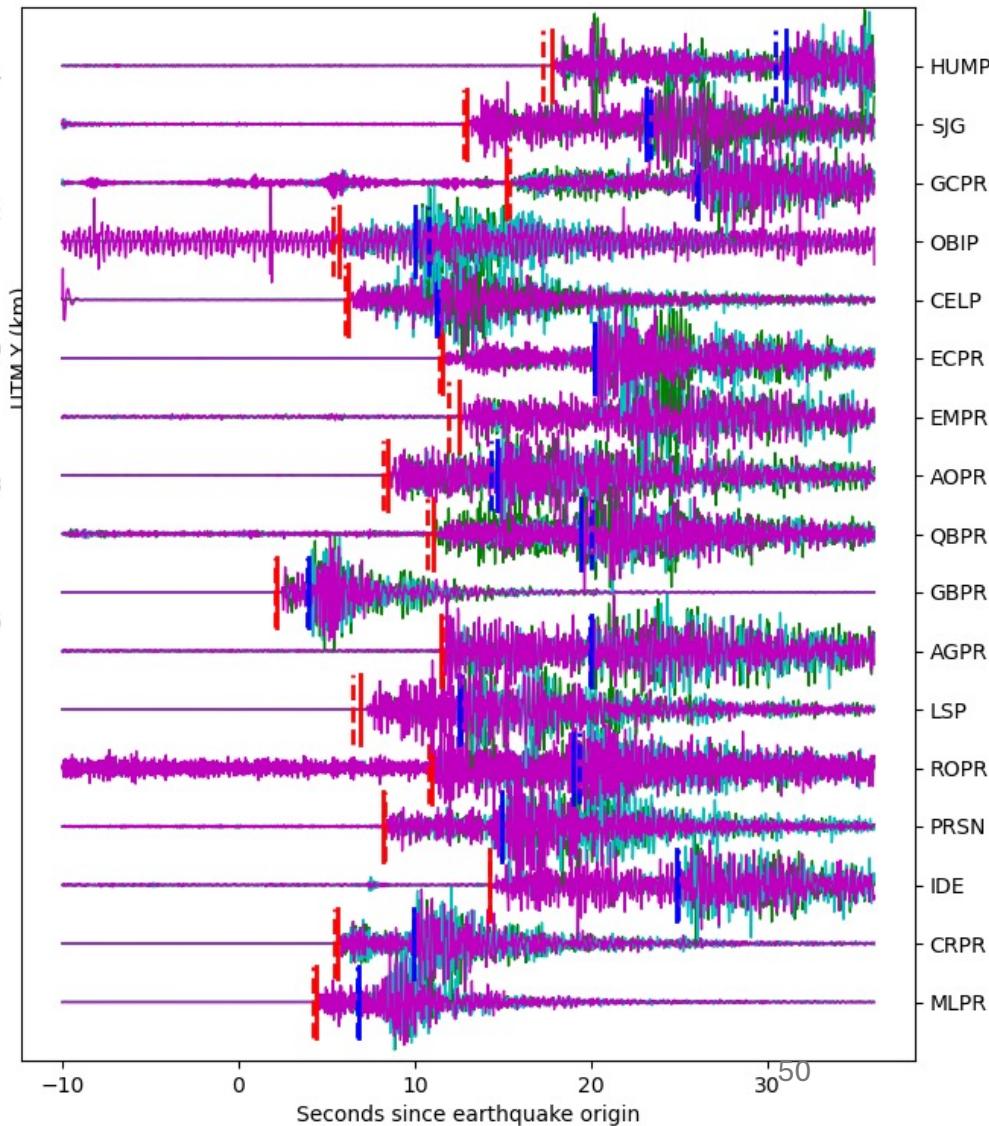
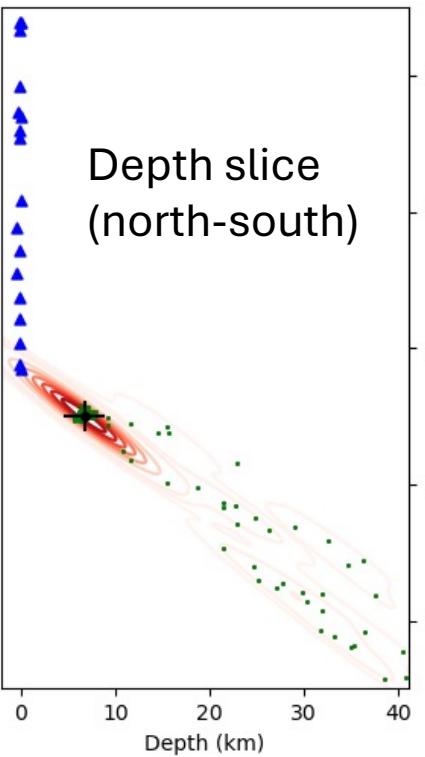
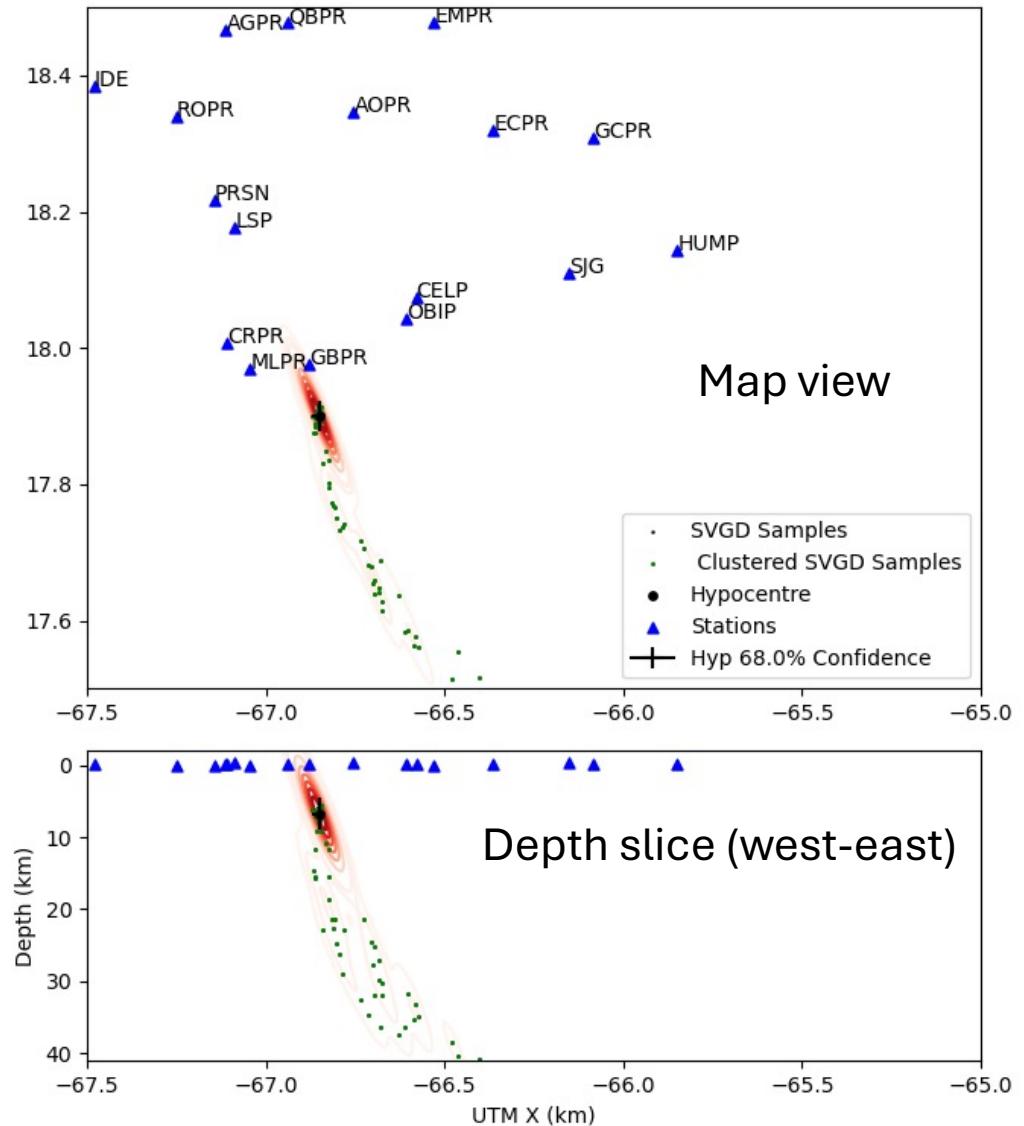


Catalog QC: 2020 M2.8 earthquake, southwest Puerto Rico



Earthquake 2020-01-07 05:59:47.516874763 +/- 0.16s
Hyp=[-66.85,17.90,6.72] - Hyp Uncertainty (km) +/- [1.99,2.48,2.14]

Event waveforms with P, S picks
Solid: auto-picks, Dashed: predicted arrivals

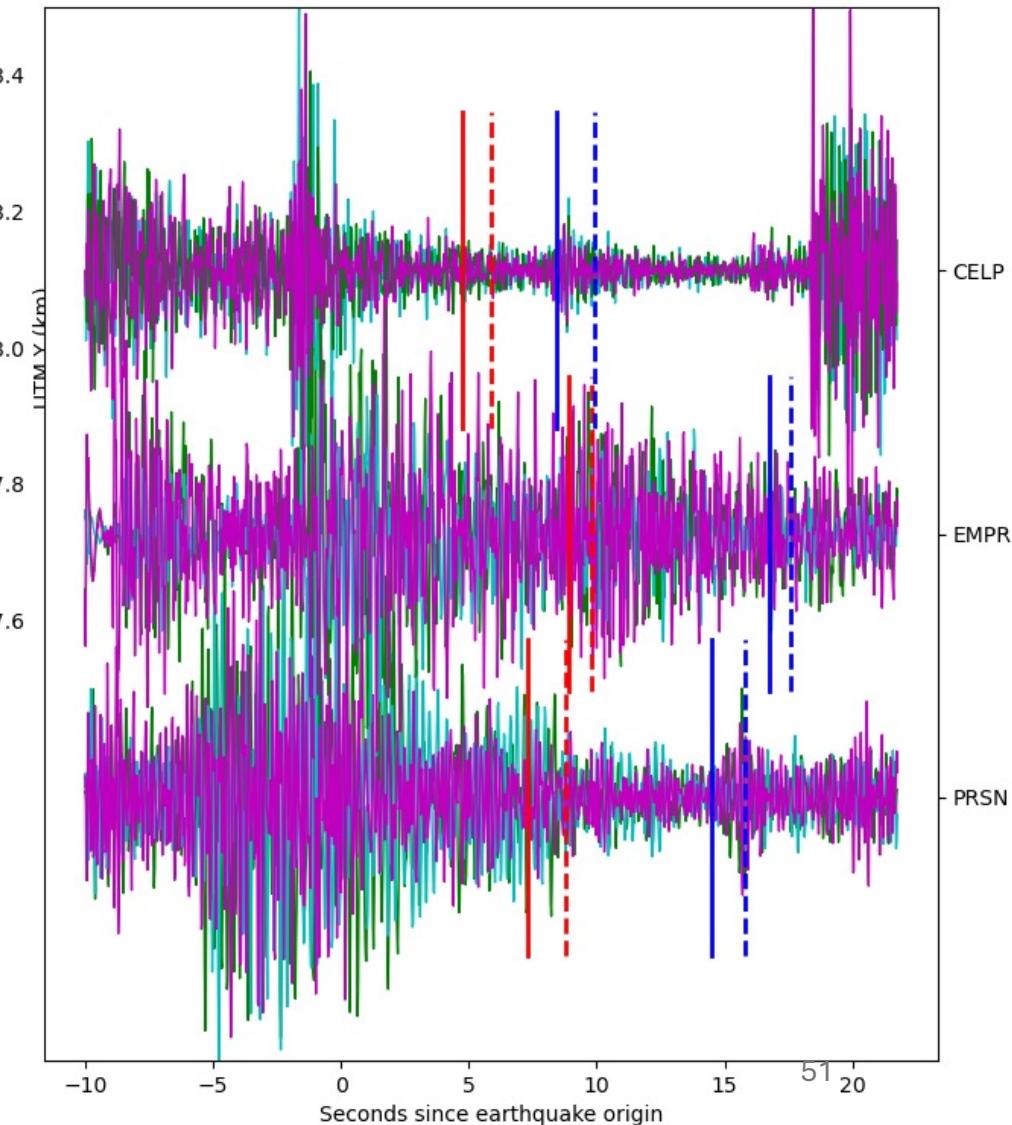
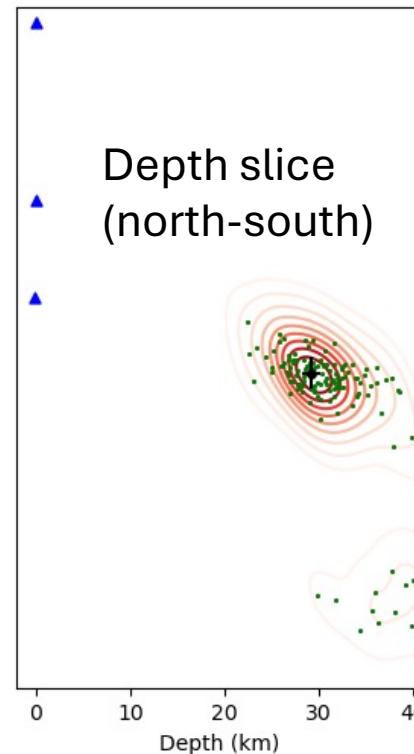
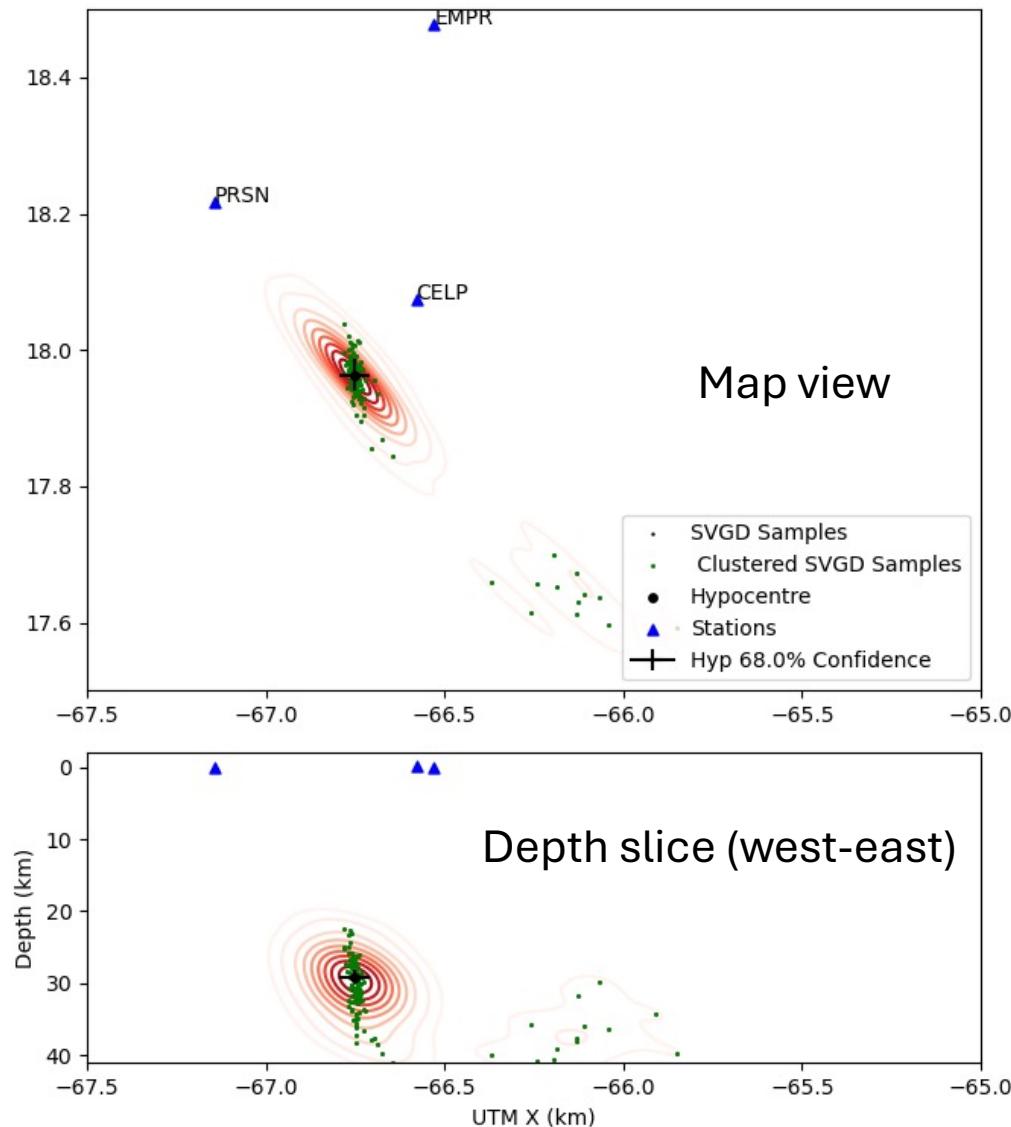


Catalog QC: False detection in coda of larger earthquake (also, too deep?)

2020-01-07T08:29:36.930000 30576.930000 17.963500 -66.752333 29.160000 4.480000 1000330

Earthquake 2020-01-07 08:29:36.925860007 +/- 1.21s
Hyp=[-66.75,17.96,29.16] - Hyp Uncertainty (km) +/- [4.69,2.51,0.93]

Event waveforms with P, S picks
Solid: auto-picks, Dashed: predicted arrivals

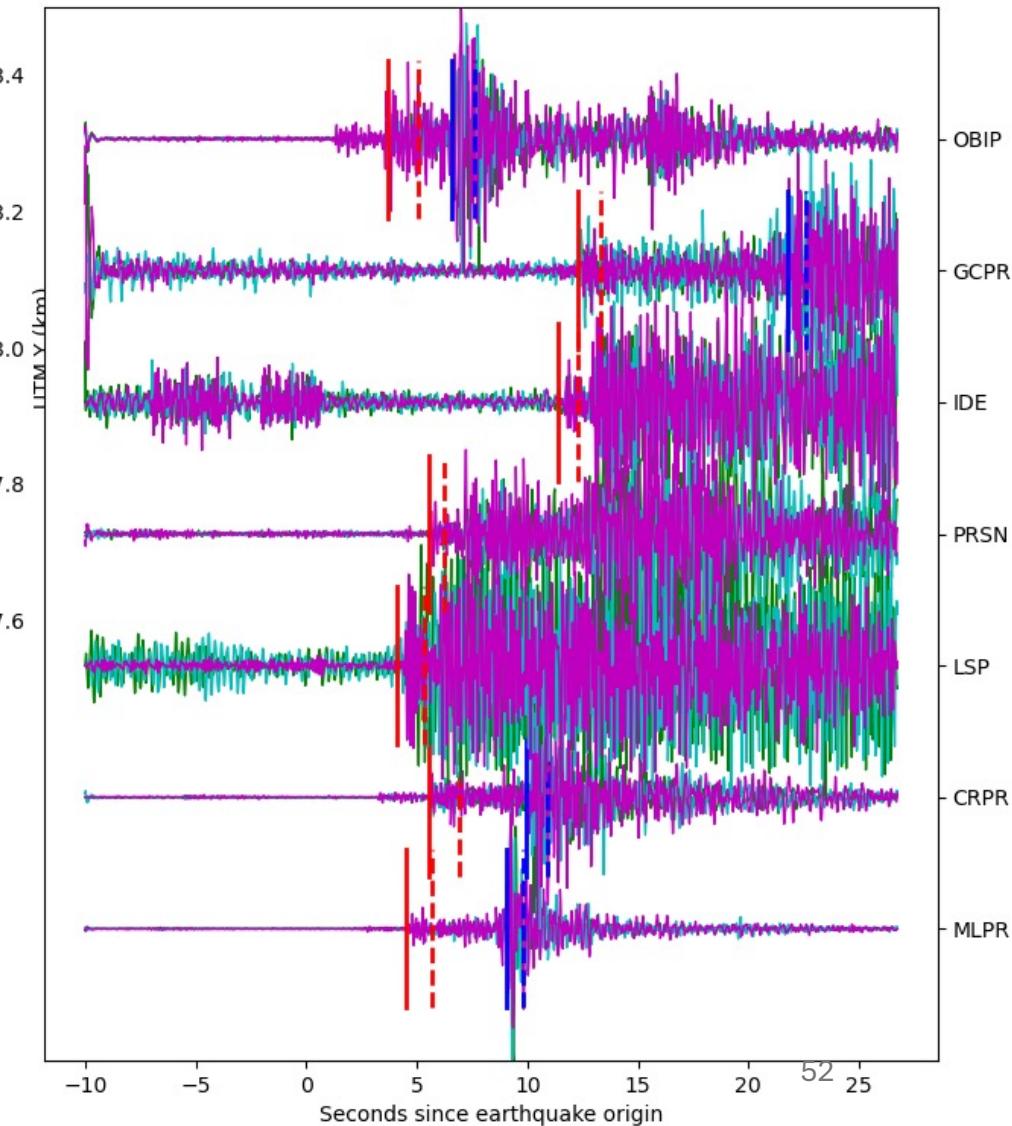
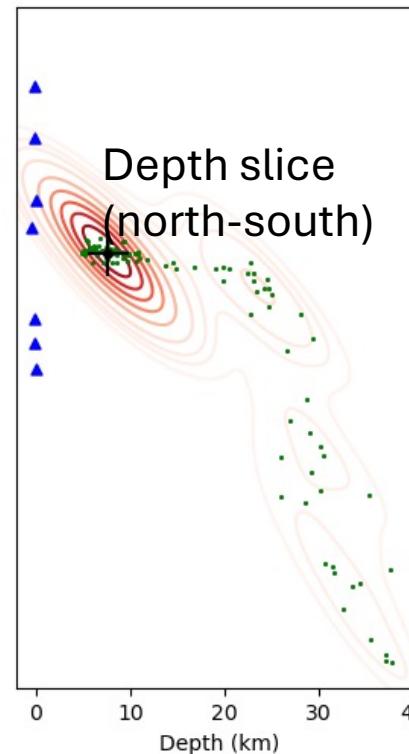
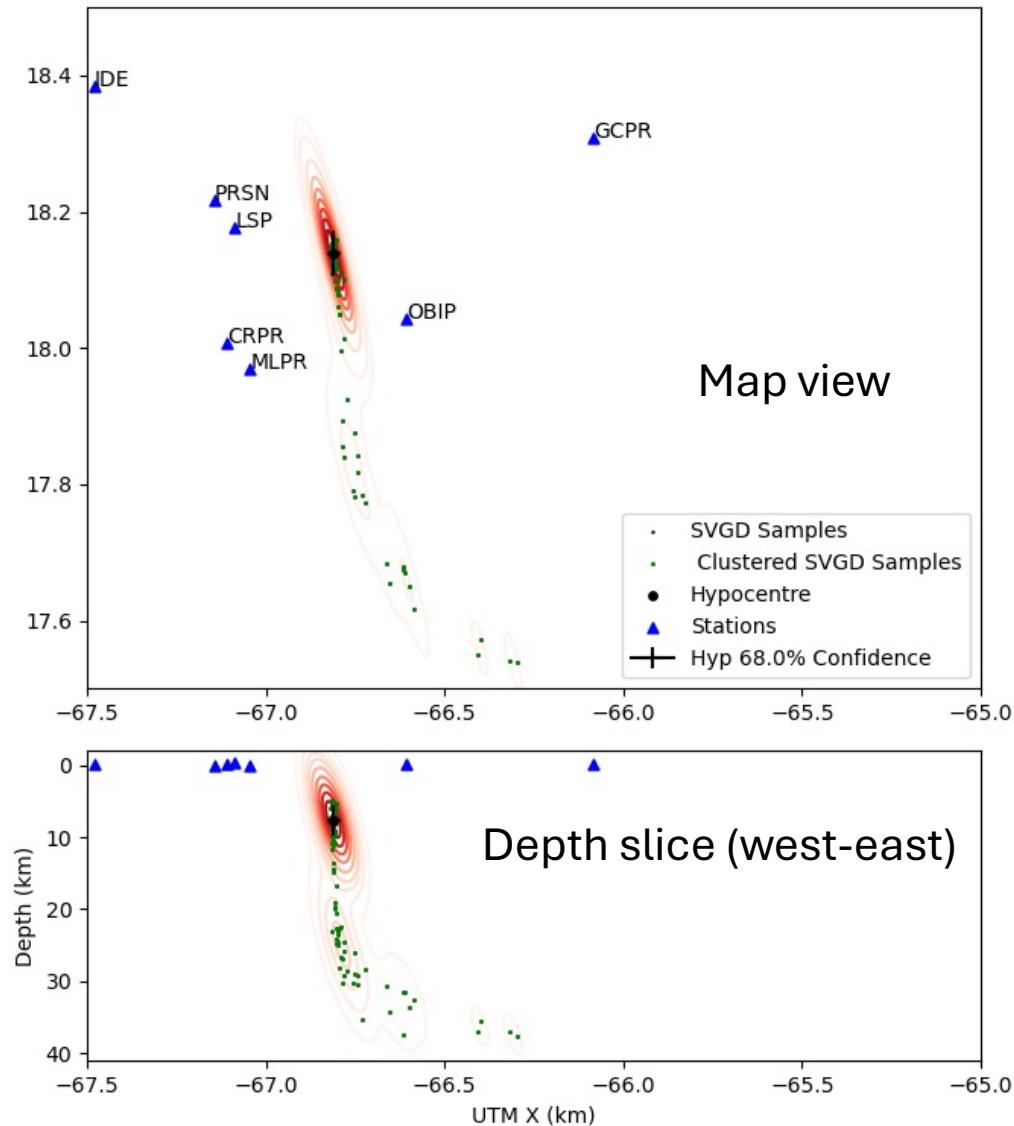


Catalog QC: Real (newly detected) earthquake

2020-01-07T08:29:18.320000 30558.320000 18.139667 -66.810000 **7.660000 4.440000** 1000328

Earthquake 2020-01-07 08:29:18.318609738 +/- 1.01s
Hyp=[-66.81,18.14,7.66] - Hyp Uncertainty (km) +/- [2.00,3.72,2.12]

Event waveforms with P, S picks
Solid: auto-picks, Dashed: predicted arrivals



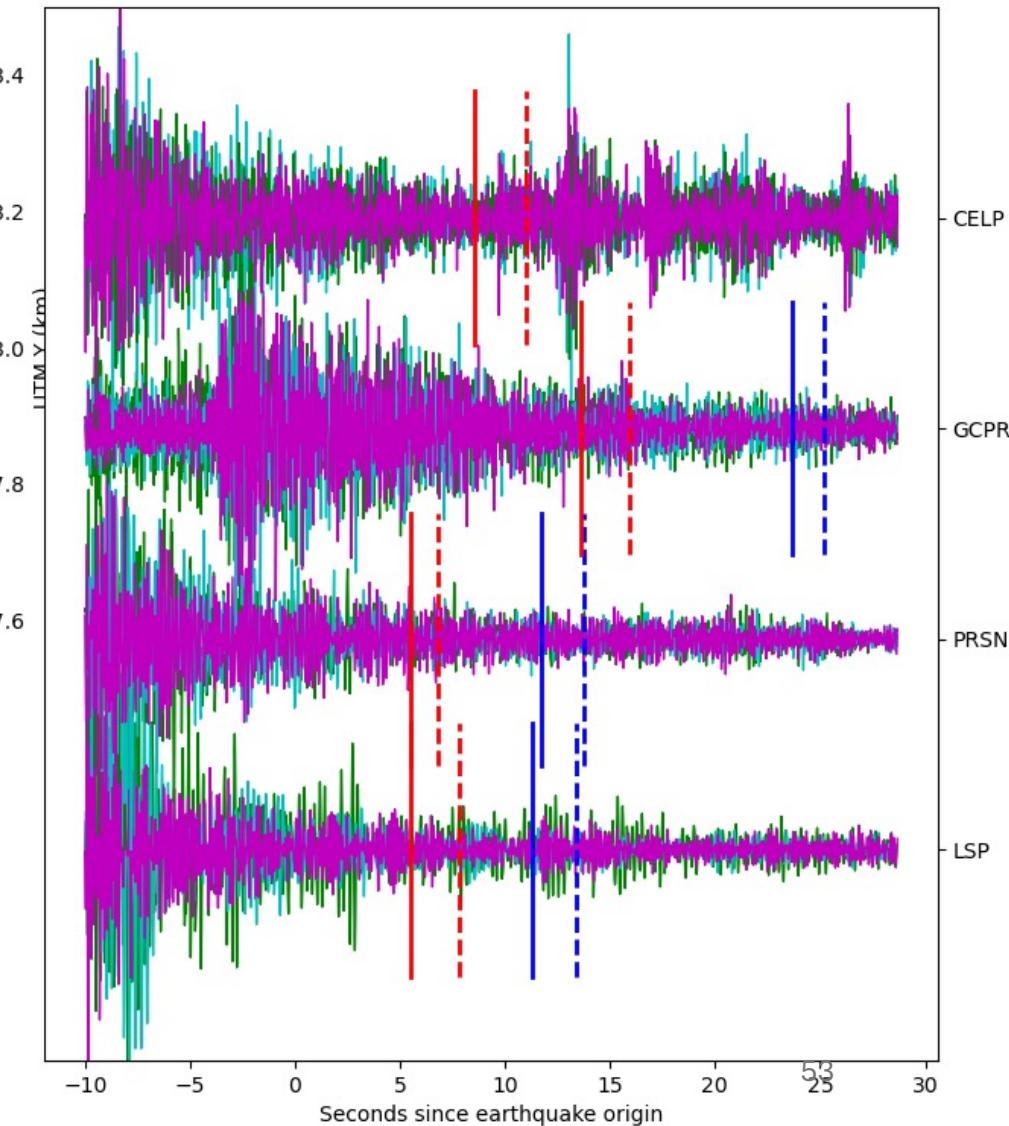
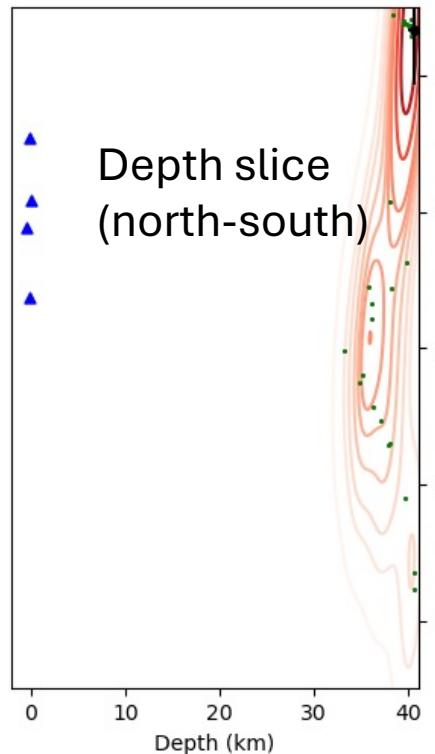
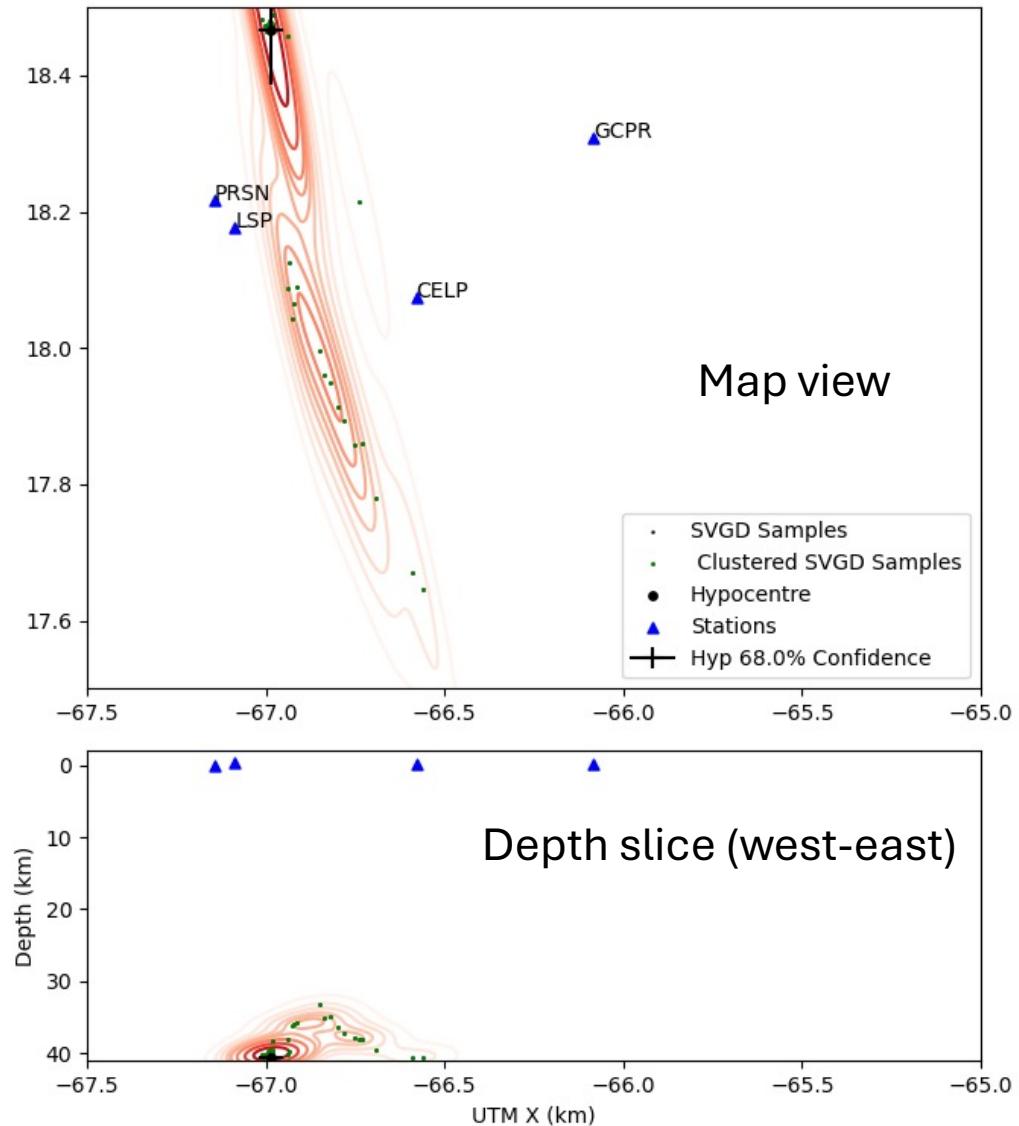
Depth slice (west-east)

Catalog QC: False detection in coda of larger earthquake (also, too deep?)

2020-01-07T22:21:53.870000 80513.870000 18.466000 -66.986833 40.580000 2.170000 1001466

Earthquake 2020-01-07 22:21:53.871442969 +/- 2.08s
Hyp=[-66.99,18.47,40.58] - Hyp Uncertainty (km) +/- [3.84,8.85,0.59]

Event waveforms with P, S picks
Solid: auto-picks, Dashed: predicted arrivals

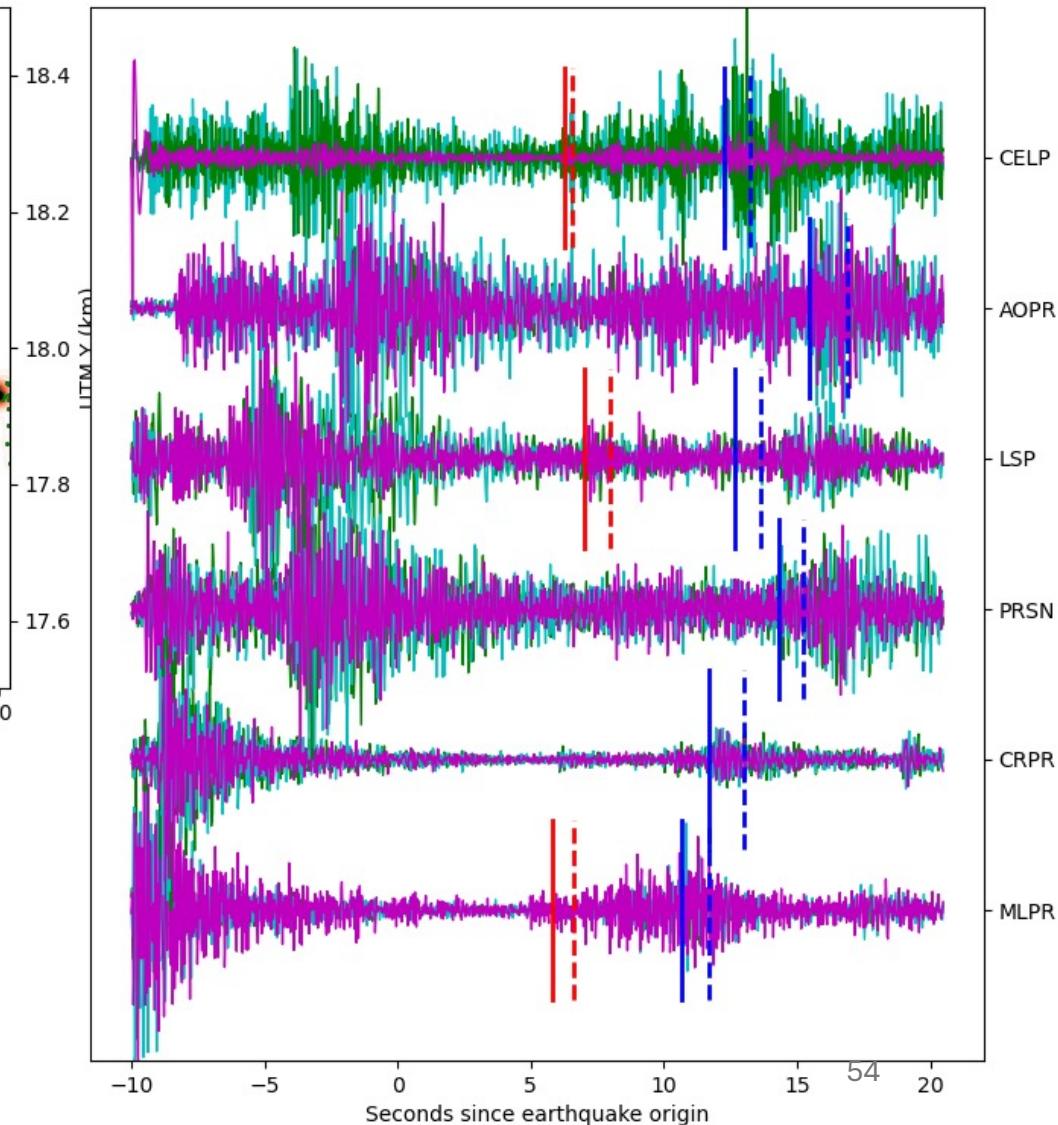
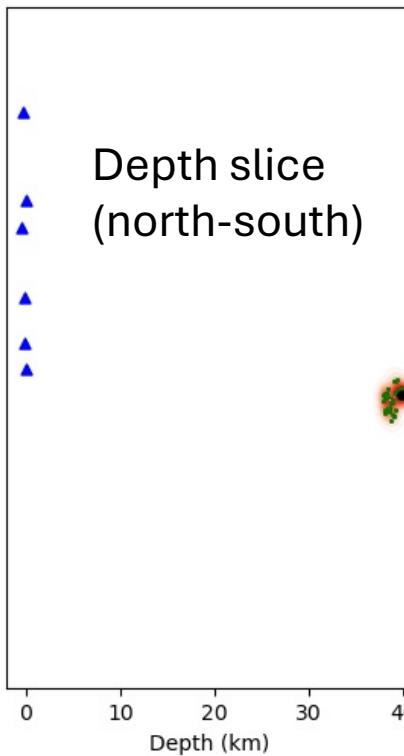
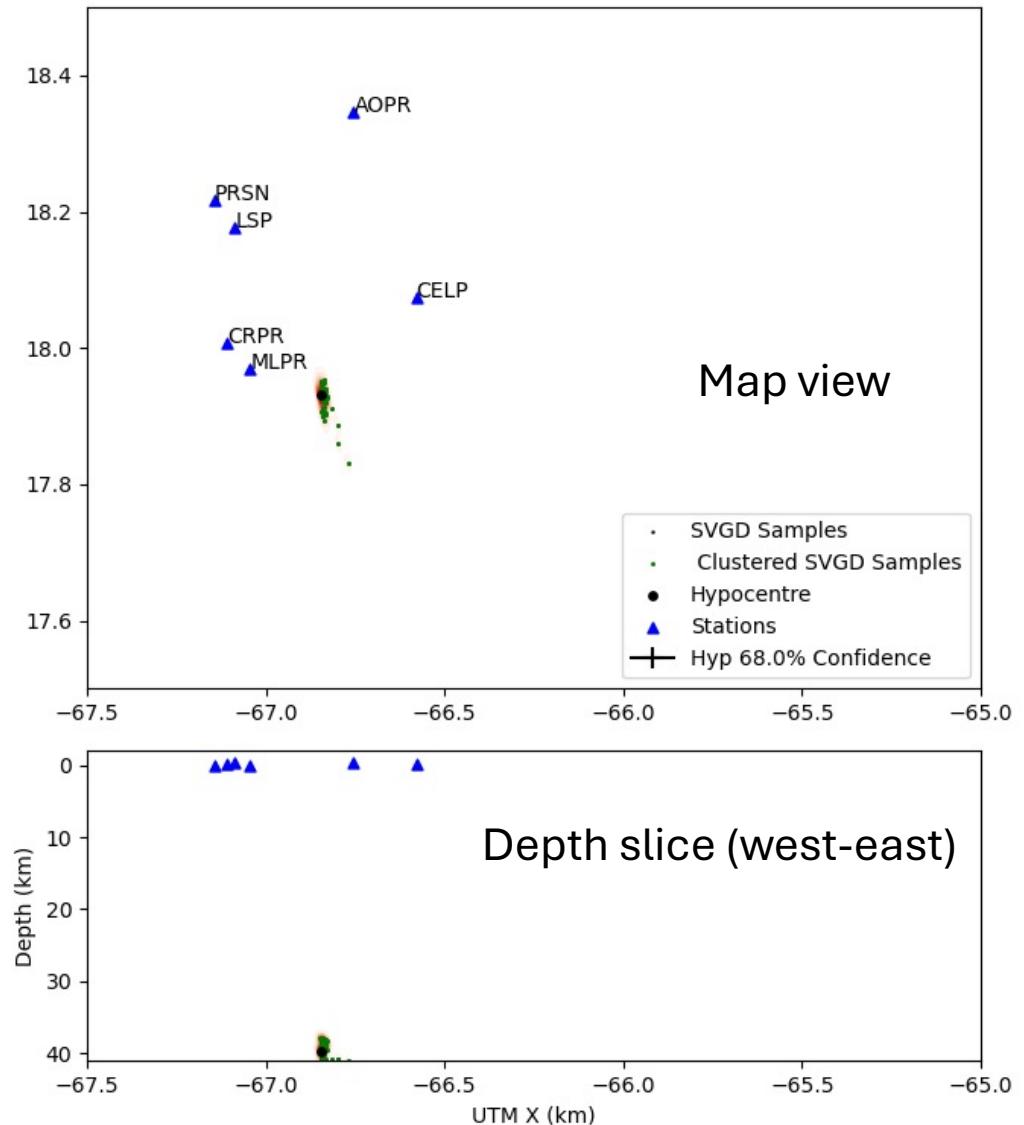


Catalog QC: False detection in coda of larger earthquake (also, too deep?)

2020-01-07T20:27:22.970000 73642.970000 17.931000 -66.846500 39.800000 1.560000 1001320

Earthquake 2020-01-07 20:27:22.965268426 +/- 0.97s
Hyp=[-66.85,17.93,39.80] - Hyp Uncertainty (km) +/- [0.36,0.59,0.24]

Event waveforms with P, S picks
Solid: auto-picks, Dashed: predicted arrivals

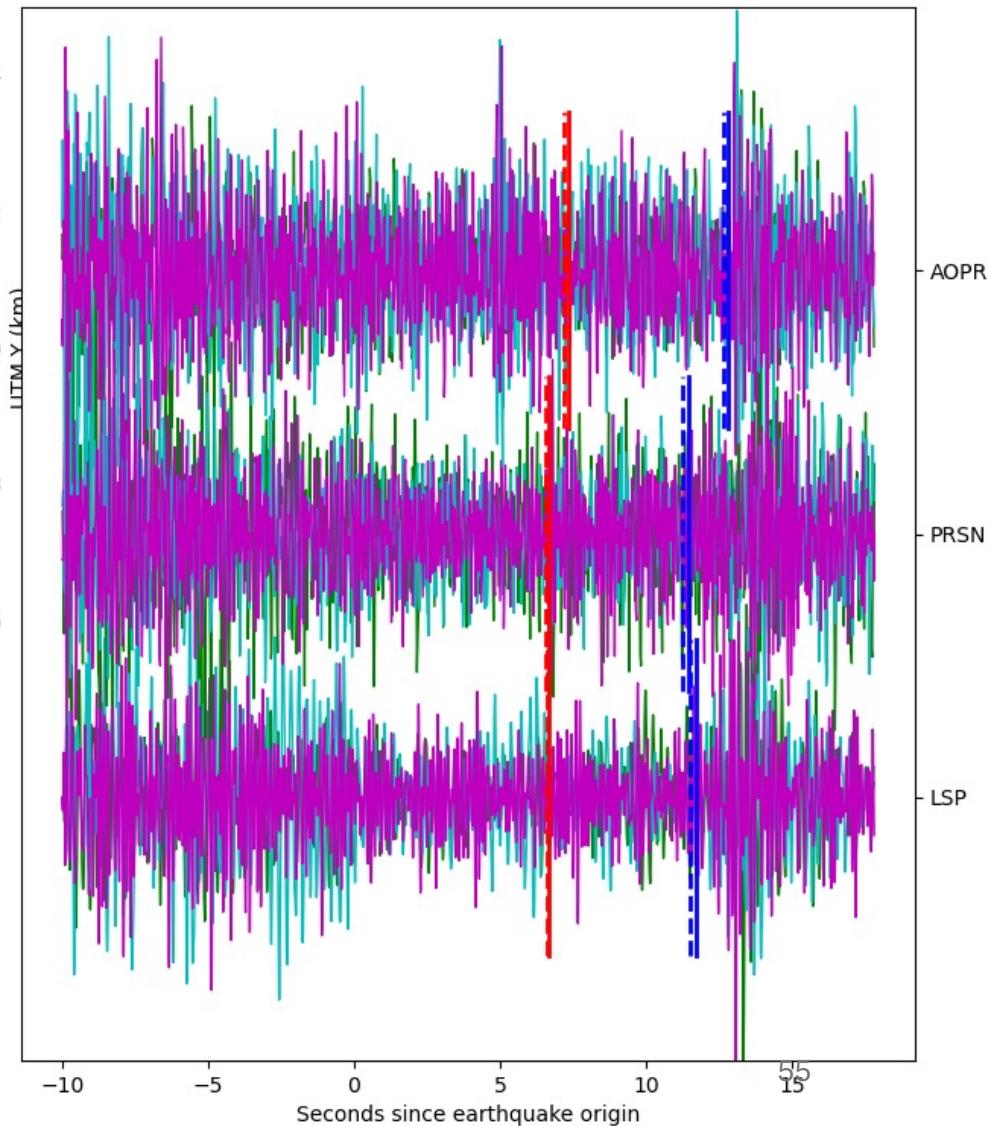
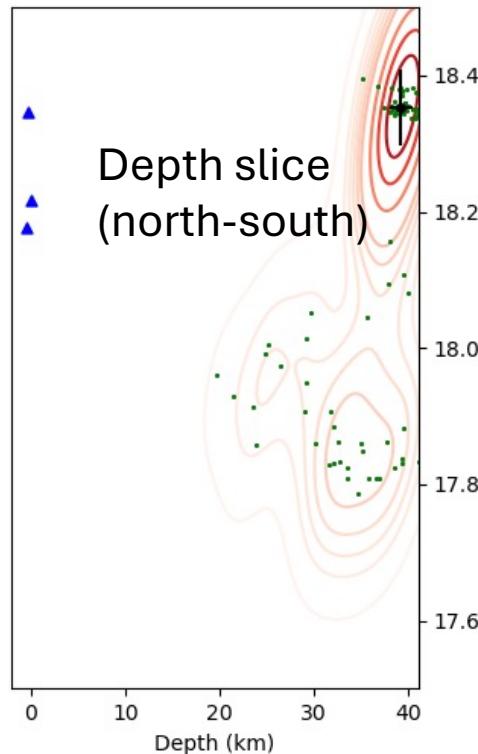
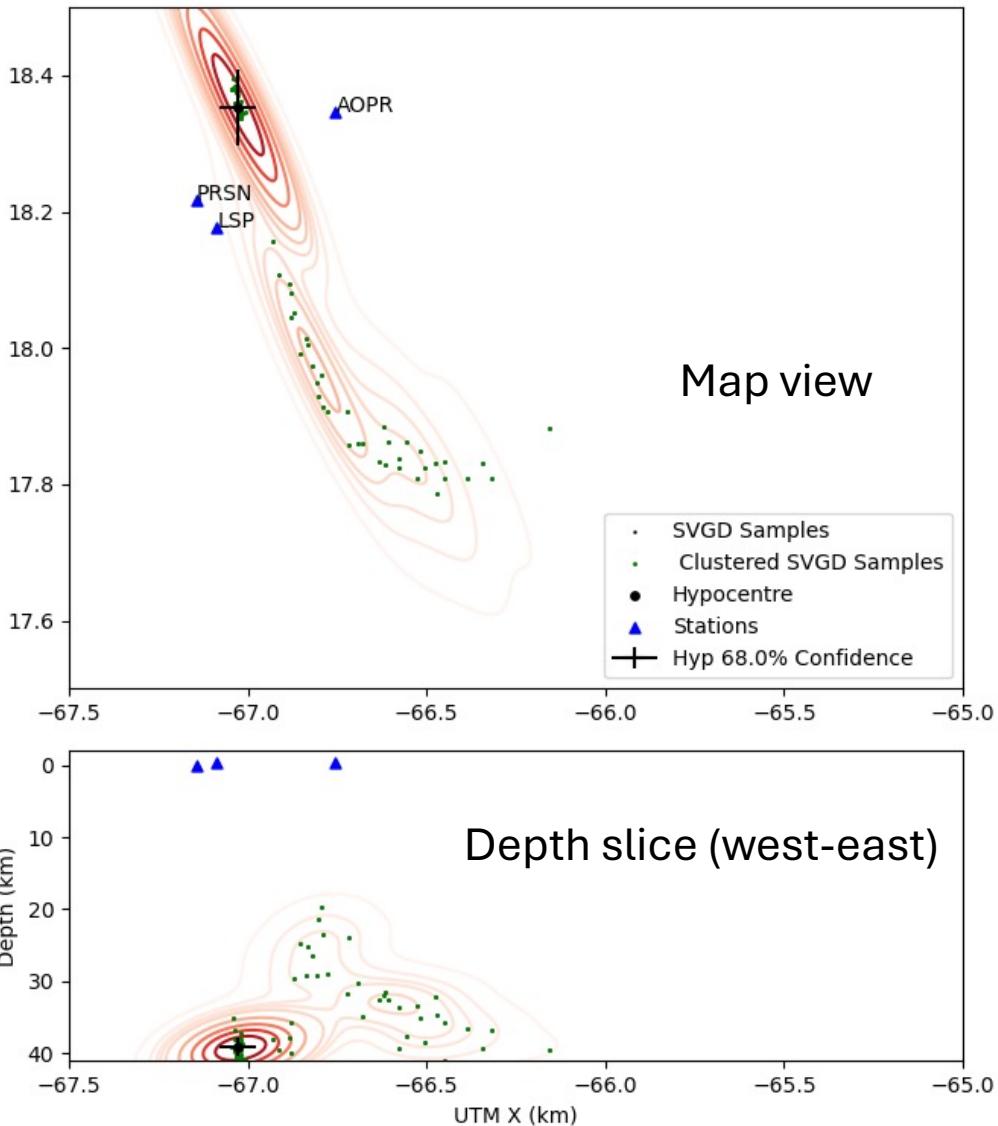


Catalog QC: False detection in noise (also, too deep?)

2020-01-07T15:58:52.370000 57532.370000 18.353667 -67.028000 39.160000 1.310000 1000980

Earthquake 2020-01-07 15:58:52.367385814 +/- 0.17s
Hyp=[-67.03,18.35,39.16] - Hyp Uncertainty (km) +/- [5.86,6.19,1.22]

Event waveforms with P, S picks
Solid: auto-picks, Dashed: predicted arrivals

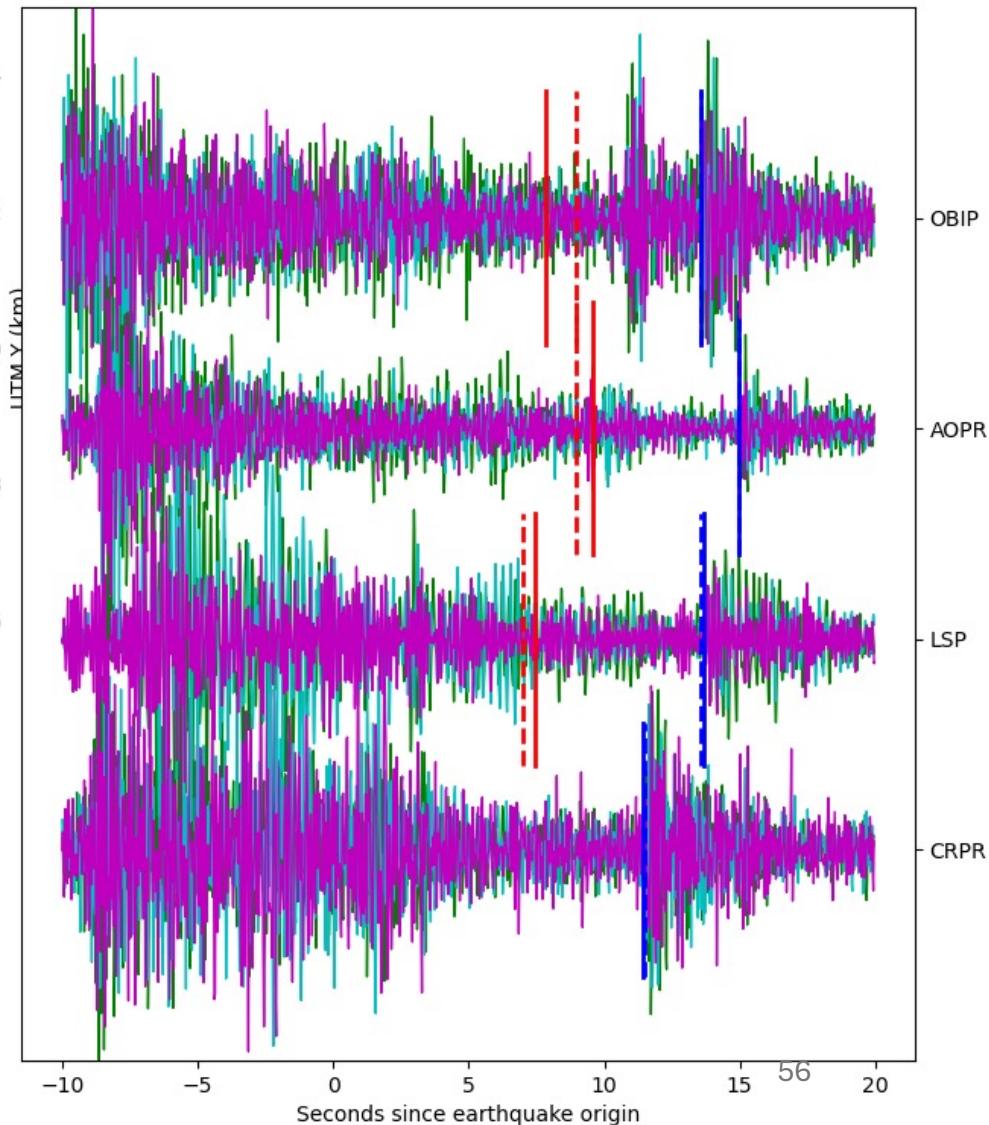
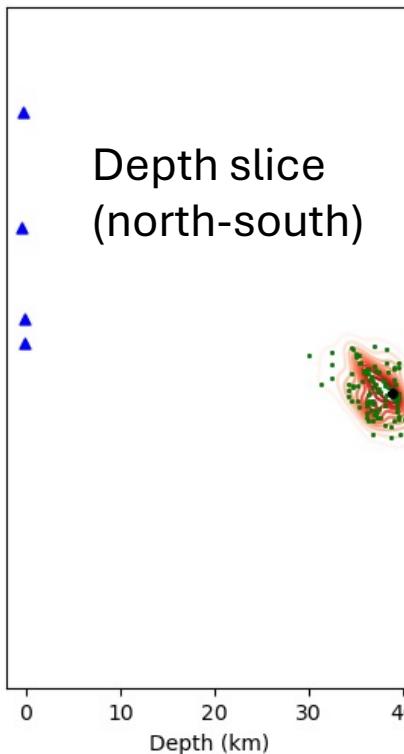
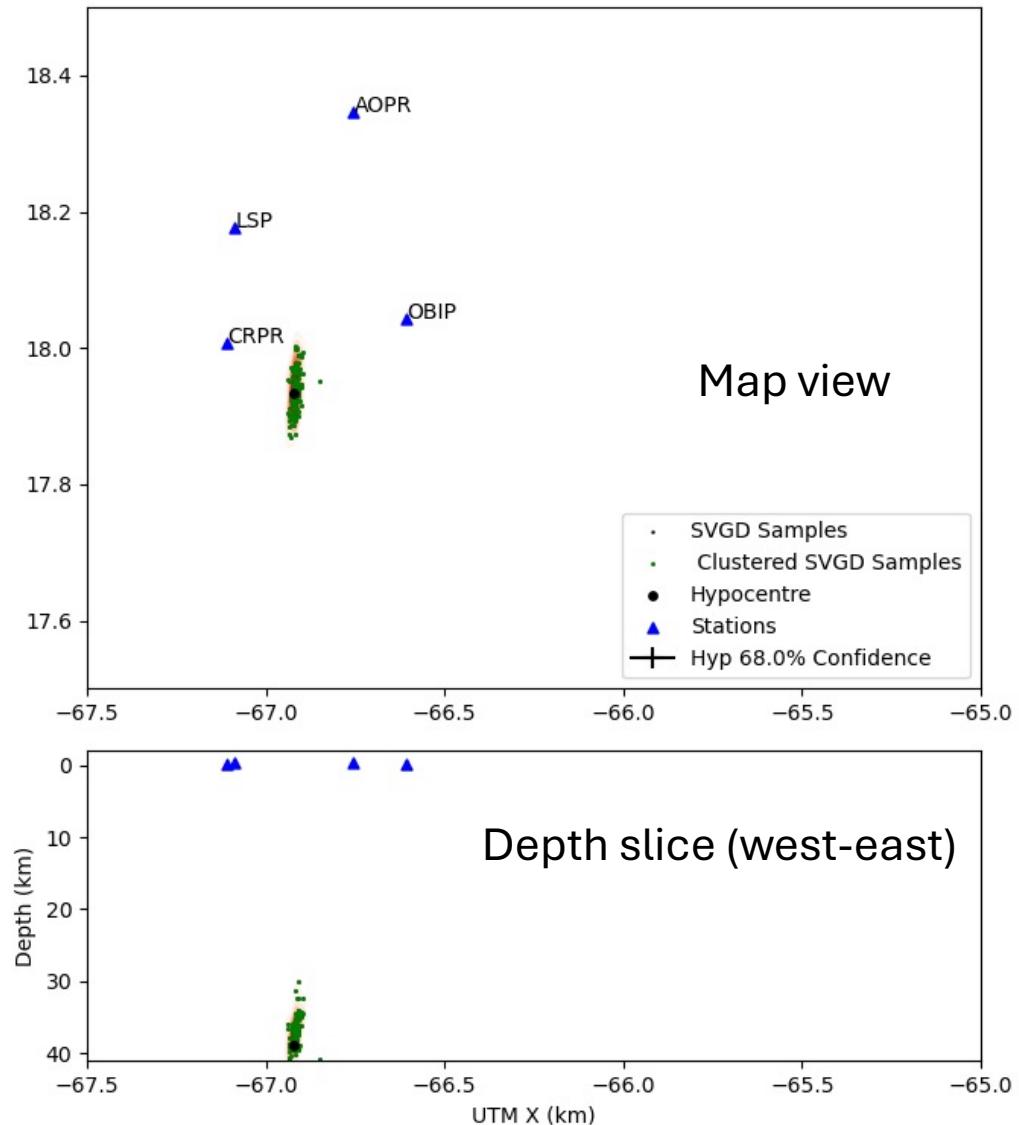


Catalog QC: False detection in coda of larger earthquake (also, too deep?)

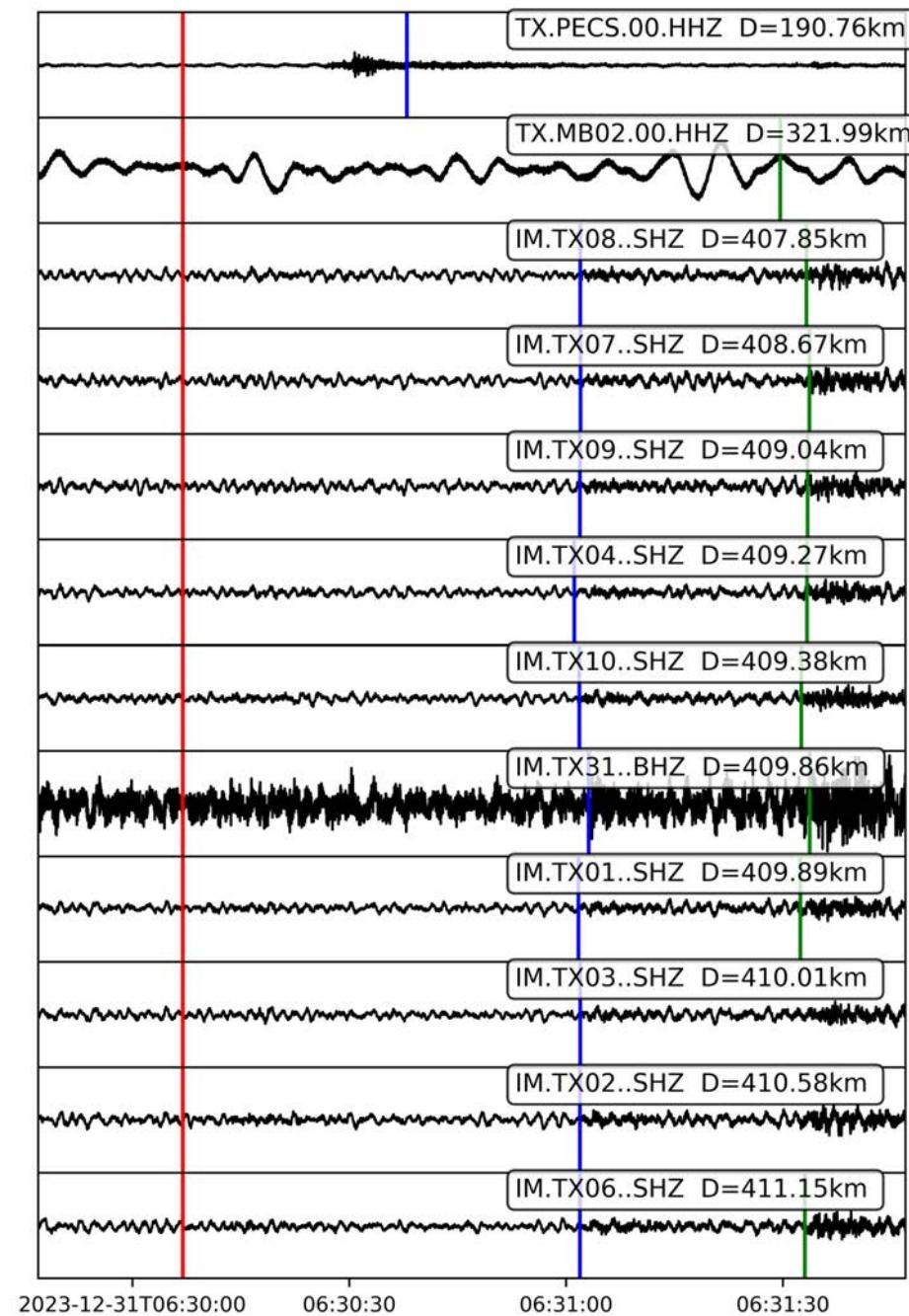
2020-01-07T20:43:26.700000 74606.700000 17.933500 -66.923333 38.810000 1.570000 1001342

Earthquake 2020-01-07 20:43:26.699477316 +/- 0.08s
Hyp=[-66.92,17.93,38.81] - Hyp Uncertainty (km) +/- [0.27,0.82,0.48]

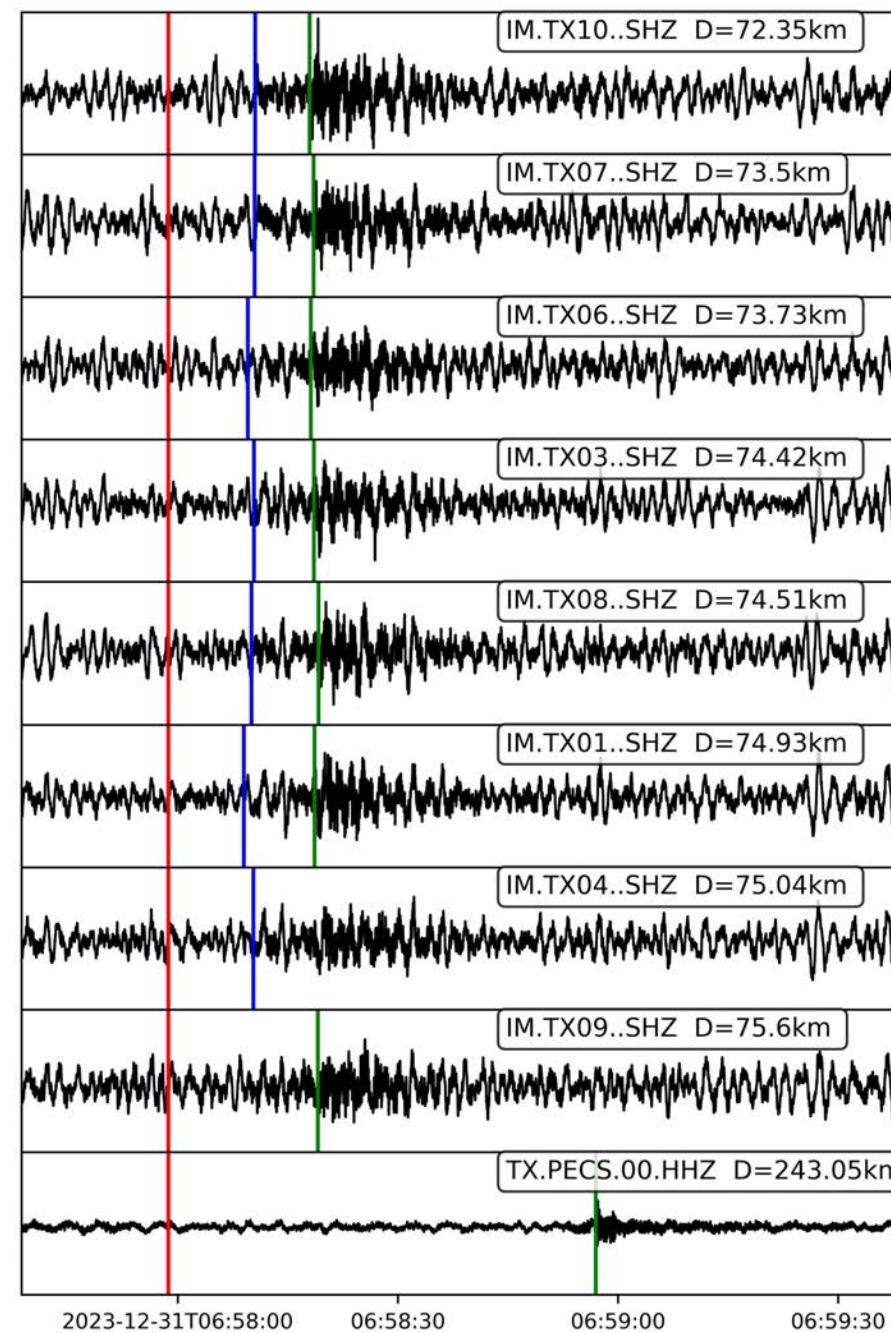
Event waveforms with P, S picks
Solid: auto-picks, Dashed: predicted arrivals



PyOcto Event #79 : 6/6/12 No Match to IRIS or GaMMA
Red: Event, Blue: P Pick, Green: S Pick



PyOcto Event #83 : 6/6/12 No Match to IRIS or GaMMA
Red: Event, Blue: P Pick, Green: S Pick

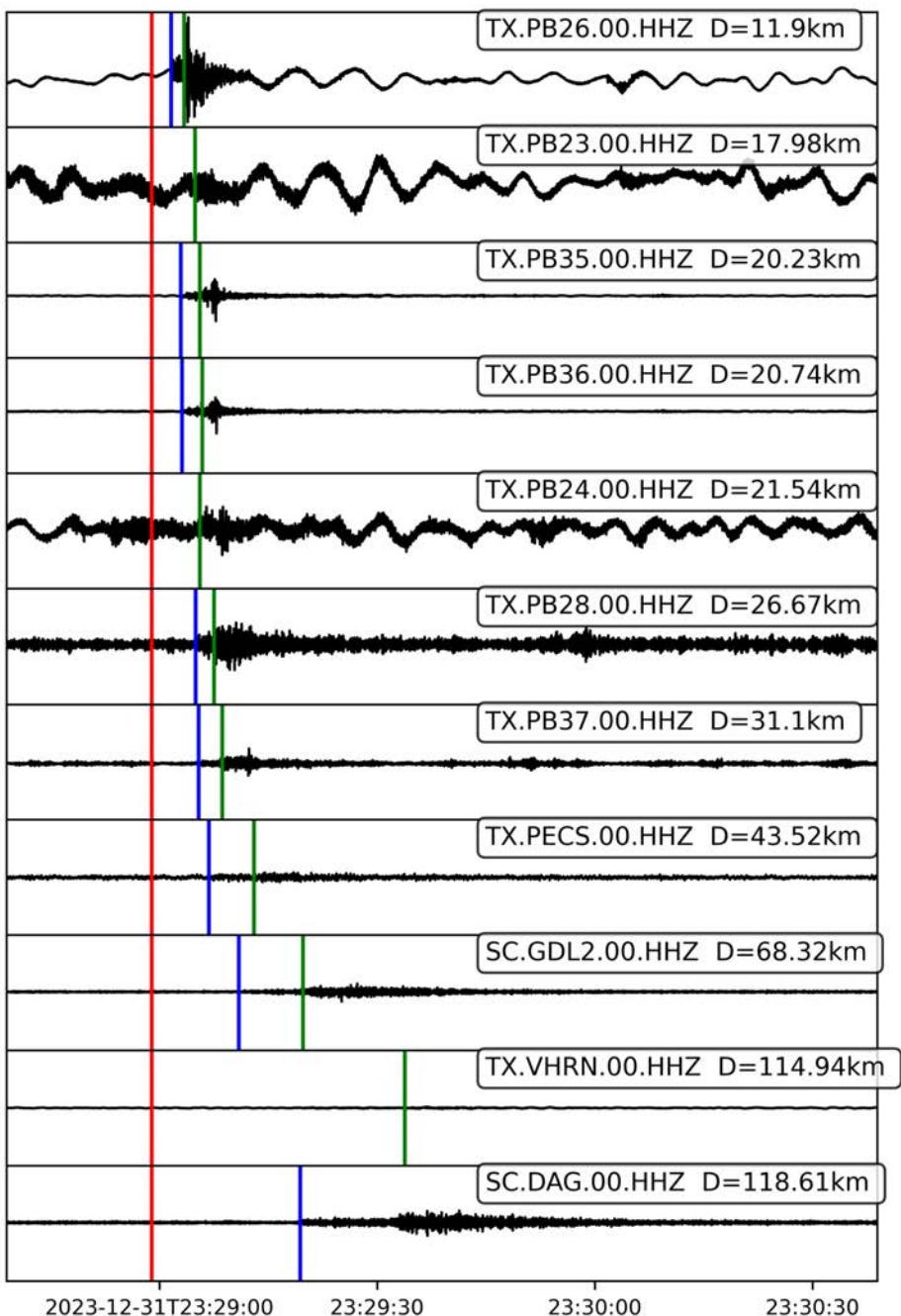


- Catalog QC: West Texas
- Bad event associations – dominated by array
- Notice large distance to nearest station

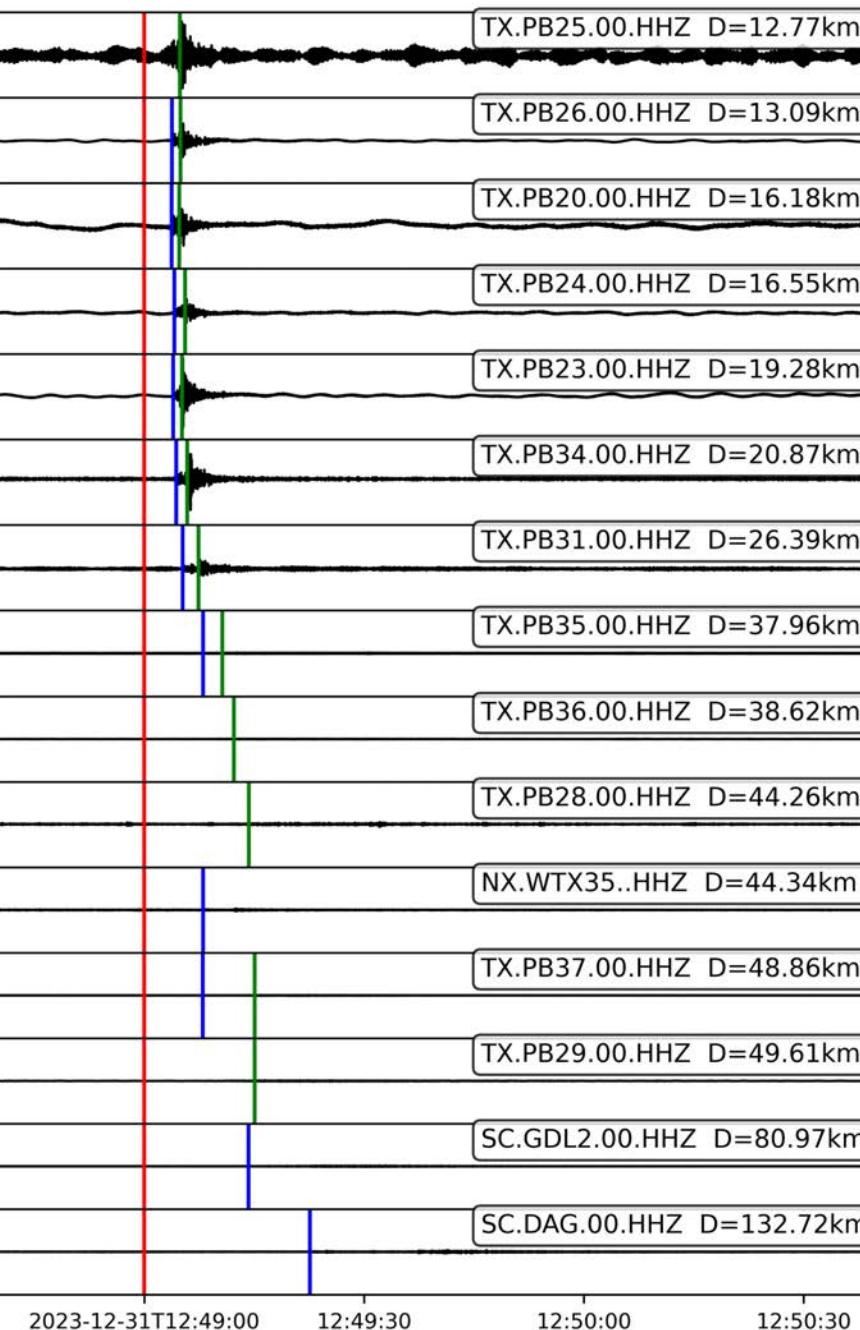
2-minute waveforms

Vertical lines:
Origin time, P, S

PyOcto Event #340 : 5/5/10 No Match to TexNet, Match to GaMMA, NSD<75.0
Red: Event, Blue: P Pick, Green: S Pick



PyOcto Event #213 : 5/5/10 No Match to TexNet, Match to GaMMA, NSD<75.0
Red: Event, Blue: P Pick, Green: S Pick



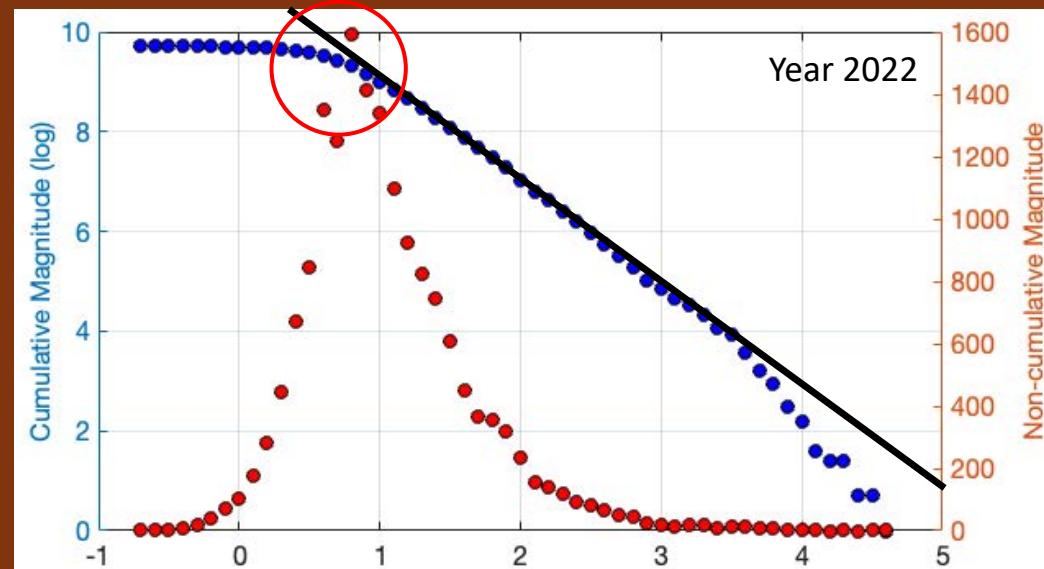
- Catalog QC: West Texas
- Good event association examples
- P, S picks on earthquake signals with expected moveout

2-minute waveforms

Vertical lines:
Origin time, P, S

Magnitude of Completeness

- lowest magnitude at which the catalog is “complete” (based on Gutenberg-Richter relation)
- examine frequency-magnitude distribution obtained by binning earthquake magnitudes
- one approach for evaluating the quality of a catalog



Can consider:

- temporal changes of M_c
- spatial variations of M_c across region

* More info and original R codes from CORSSA: [doi:10.5078/corssa-0018080](https://doi.org/10.5078/corssa-0018080)

* Matlab codes available on Github:
https://github.com/gtepp/research_codes

Calculated with different methods, including

- Maximum Curvature: max value of 1st derivative of FMD (typically, max bin of non-cum FMD)
 - * most reliable for small sample sizes (<~50-100 events)
 - * works best with local datasets with fewer heterogeneities
- Goodness-of-fit Test: lowest magnitude cut-off where Gutenberg-Richter relation holds
- Mc by b-value Stability: first magnitude increment where $|b_{avg} - b| <$ uncertainty of b
 - * tends to produce the highest (i.e., most conservative) Mc

Bootstrapping

- gives sense of variation from sampling
- 200 sets with resampling

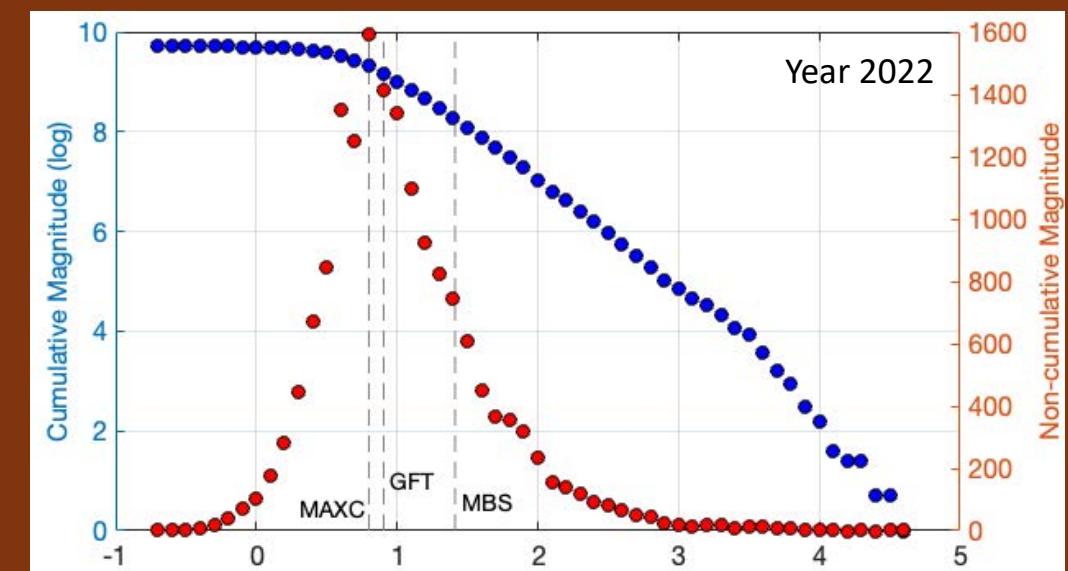
- mean and st. dev.

original: 1, 1, 1, 2, 2, 3

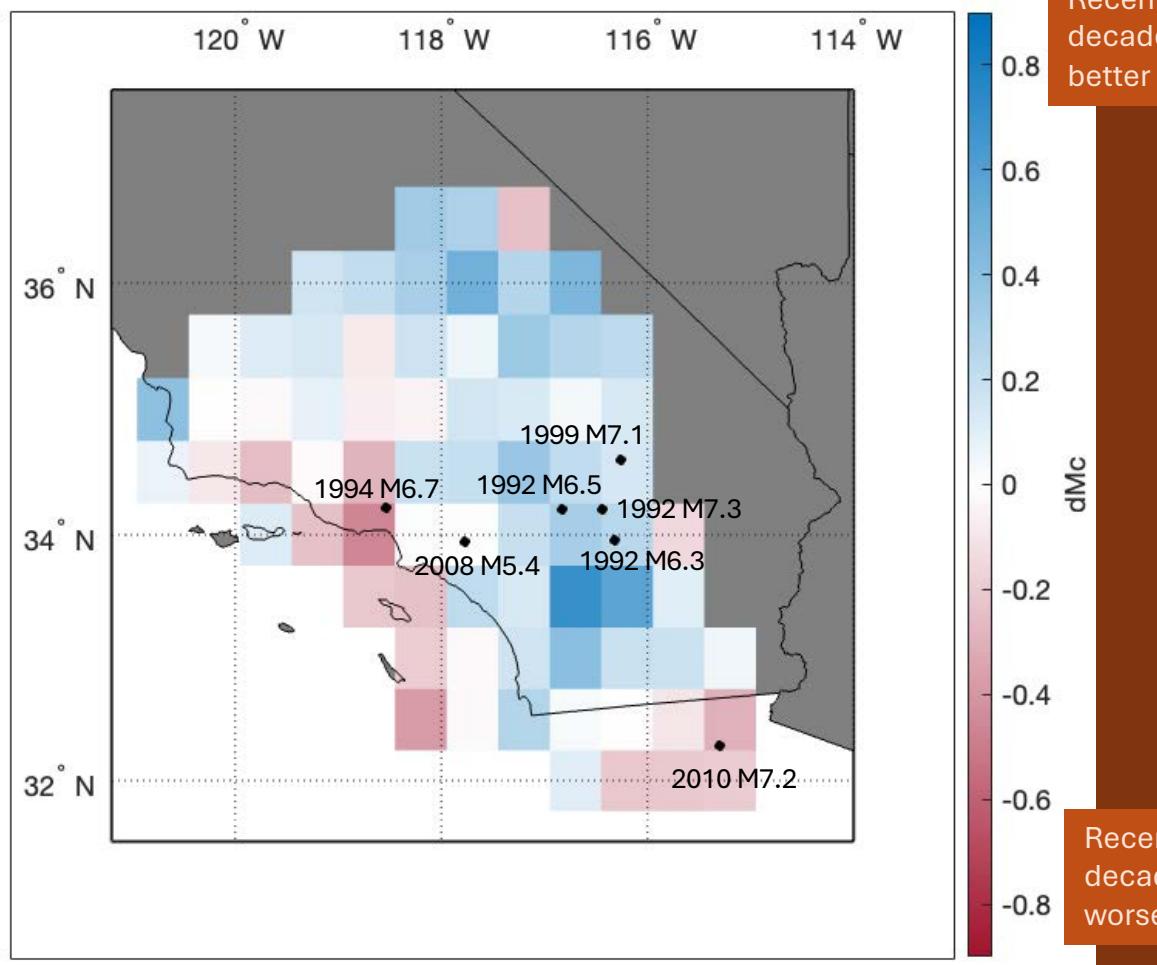
sample A: 1, 2, 2, 2, 3, 3

sample B: 1, 1, 1, 1, 2, 2

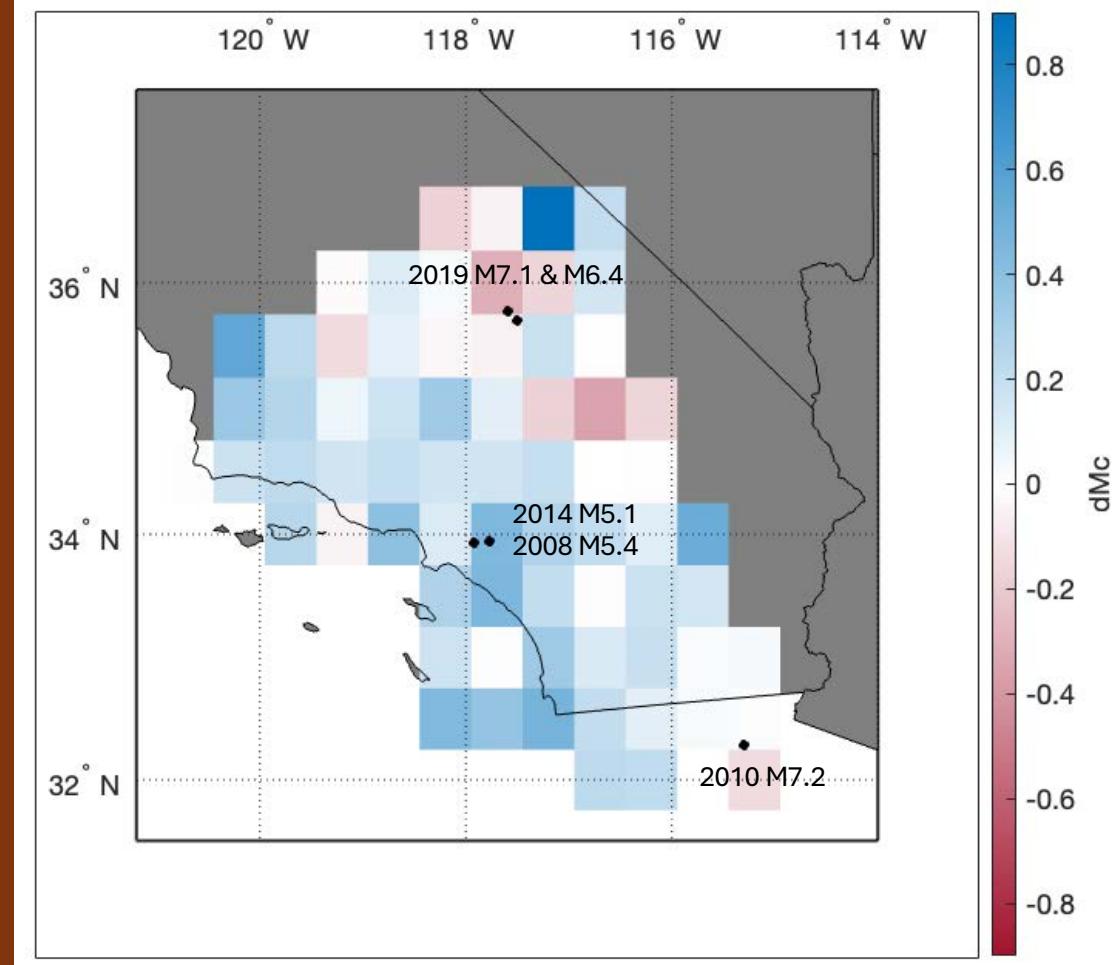
sample C: 1, 1, 2, 2, 2, 3



Difference between 1993-2003 and 2003-2013



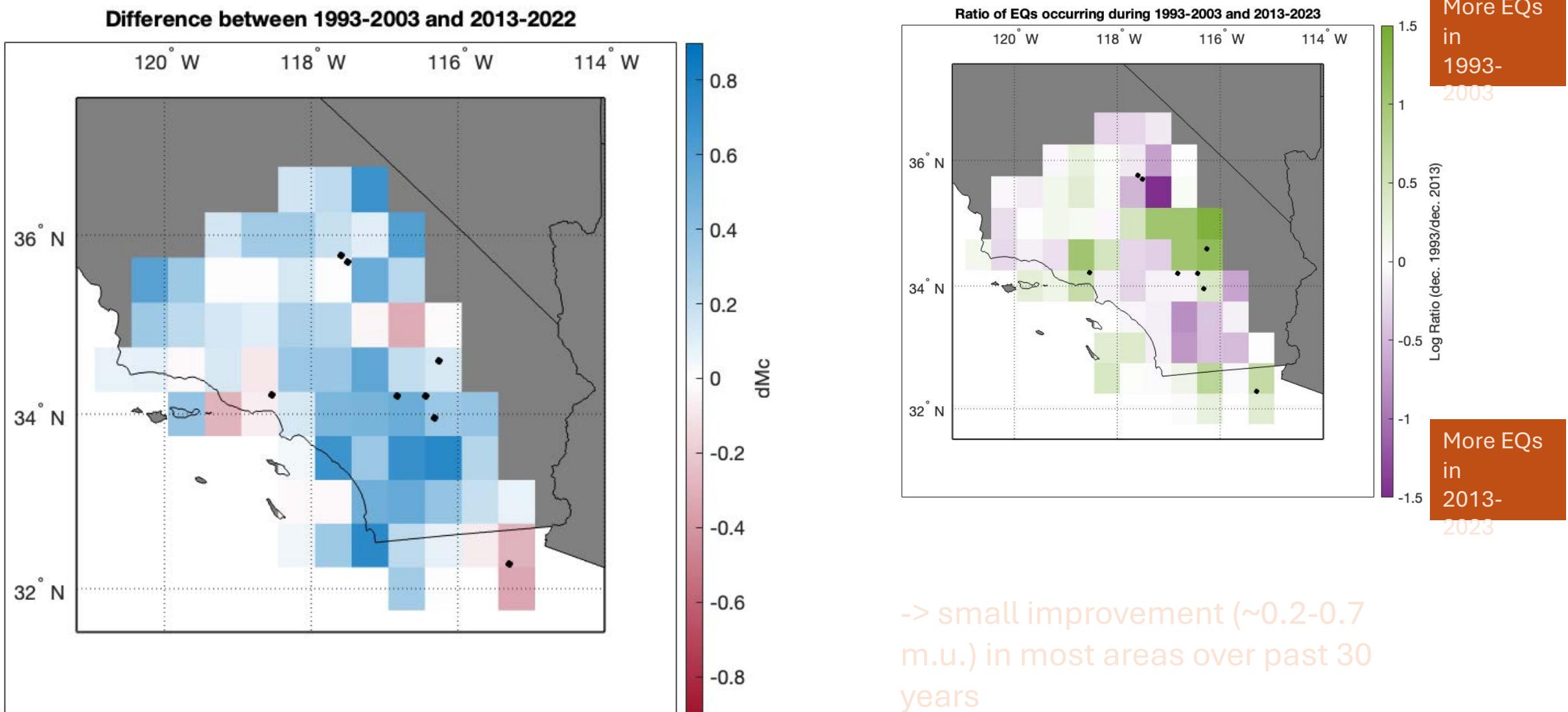
Difference between 2003-2013 and 2013-2023



Better: 1994 Northridge

Worse: 2010 El Mayor, 2019 Ridgecrest*, 1992 Landers-Big Bear & Joshua Tree, 1999 Hector Mine

*incomplete



-> small improvement (~0.2-0.7 m.u.) in most areas over past 30 years

-> regions with worse Mc typically have fewer earthquakes