Supplement to "PyOcto: A high-throughput seismic phase associator"

Jannes Münchmeyer[1]

[1]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ. Gustave Eiffel, ISTerre, Grenoble, France

S1 Distribution of picking stations in synthetics

To determine which stations pick for which event, we use a probabilistic process. Picks are determined independently by station and by event, but correlated between P and S pick within one station. The detection probability p for a pick is defined as:

$$p = (1 - \sigma(\frac{x - r}{w}))(1 - p_{failure}) \tag{1}$$

Here, σ denotes a sigmoid function, x the distance between station and event, r denotes a critical distance, w a decay constant, and $p_{failure}$ the probability that the station is not working or polluted by high noise. This means that in the near field, the probability of a pick will be $1 - p_{failure}$ while it decays to 0 over distance with a sigmoid. The values or r and w are derived from the event magnitude M.

$$r = A * M + B \tag{2}$$

$$w = \max(0.1r, 30\ km) \tag{3}$$

We use $A = 120km$, $B = 80km$, and $p_{failure} = 0.2$.

To account for the fact that detection rates will depend on the current noise conditions at a station, we make the detection of P and S waves correlated. Using the same probability p for both P and S arrival, we generate two correlated Bernoulli variable with a correlation of 0.5.

Table S1: Parameters for the GaMMA associator. All parameters not provided in the table were used at their default values. The parameters for the application to the Iquique sequence are identical to the *Subduction* parameters. Where possible, we use the parameter name from the published implementation. We omit purely functional parameters.

Parameter	Shallow	Subduction
ncpu	15	15
x(km)	[395, 605]	[250, 600]
y(km)	[7560, 7790]	[7200, 8000]
z(km)	[0, 30]	[0, 250]
p velocity	6.3	7.0
s velocity	3.7	4.0
use_dbscan	True	True
use_amplitude	False	False
method	BGMM	BGMM
oversample_factor	4	4
dbscan_eps	25	25
dbscan_min_samples	3	3
min_picks_per_eq	10	10
max_sigma11	2.0	2.0
max_sigma22	1.0	1.0
max_sigma12	1.0	1.0

Table S2: Parameters for the PyOcto associator. All parameters not provided in the table were used at their default values. The parameters for the application to the Iquique sequence are identical to the *Subduction* parameters. Where possible, we use the parameter name from the published implementation. We omit purely functional parameters. If two values are provided, the first is for the homogeneous velocity model, the second for the 1D model.

Parameter	Shallow	Subduction
xlim	[395., 605.]	[250., 600.]
ylim	[7560., 7790.]	[7200., 8000.]
zlim	[0., 30.]	[0., 250.]
time_before	300.	300
n_picks	10	10
n_p_and_s_picks	4	4
pick_match_tolerance	2 / 1.5	2 / 1.5
p_velocity	6.3	7.0
s_velocity	3.7	4.0
velocity_model tolerance	2 / 1.5	2 / 1.5
velocity_model association_cutoff_distance	100	250

Table S3: Parameters for the REAL associator. All parameters not provided in the table were used at their default values. The parameters for the application to the Iquique sequence are identical to the *Subduction* parameters. We omit purely functional parameters. Note that some parameters are only used in the homogeneous (velocities) or in the 1D model (traveltime grid details).

Parameter	Shallow	Subduction
search_range_horizontal_deg	0.2	1.0
search_range_vertical_km	30	250
grid_size_horizontal_deg	0.04	0.1
grid_size_vertical_km	3	10
vp	6.3	7.0
vs	3.7	4.0
shallow_vp	5.3	5.3
shallow_vs	3.1	3.1
elevation_correction	False	True
tt_range_horizontal_deg	3	6
tt_grid_size_horizontal_deg	0.04	0.05
tt_grid_size_vertical_km	3	3

Table S4: Full results of synthetic benchmark. We abbreviate *picks per event* as *ppe*.

Scenario	Events	Noise	Associator	Precision	Recall	F1	Missing ppe	Additional ppe	Run time [s]
shallow	100	0.30	Gamma	1.00	0.99	0.99	0.08	0.04	5.04
shallow	100	0.30	Gamma1D	1.00	0.99	0.99	0.06	0.04	13.02
shallow	100	0.30	PyOcto	1.00	0.96	0.98	0.06	0.03	0.74
shallow	100	0.30	PyOcto1D	1.00	0.96	0.98	0.04	0.03	1.52
shallow	100	0.30	REAL	0.99	0.93	0.96	2.16	0.01	10.11
shallow	100	0.30	REAL1D	1.00	0.93	0.96	5.91	0.00	11.52
shallow	100	1.00	Gamma	1.00	0.97	0.98	0.07	0.21	14.93
shallow	100	1.00	Gamma1D	1.00	0.98	0.99	0.02	0.21	16.13
shallow	100	1.00	PyOcto	1.00	0.96	0.98	0.28	0.21	1.15
shallow	100	1.00	PyOcto1D	1.00	0.96	0.98	0.23	0.16	1.54
shallow	100	1.00	REAL	0.98	0.93	0.95	2.86	0.08	22.86
shallow	100	3.00	Gamma	0.97	0.86	0.91	0.17	0.59	120.52
shallow	100	3.00	Gamma1D	0.98	0.86	0.91	0.06	0.56	49.20
shallow	100	3.00	PyOcto	1.00	0.96	0.98	0.47	0.68	1.69
shallow	100	3.00	PyOcto1D	1.00	0.96	0.98	0.41	0.55	2.11
shallow	100	3.00	REAL	0.99	0.96	0.97	3.20	0.41	59.62
shallow	100	3.00	REAL1D	0.95	0.91	0.93	8.90	0.42	64.44
shallow	500	0.30	Gamma	0.98	0.92	0.95	0.34	0.37	126.18
shallow	500	0.30	Gamma1D	0.98	0.92	0.95	0.32	0.37	48.08
shallow	500	0.30	PyOcto	1.00	0.93	0.96	0.68	0.52	4.47
shallow	500	0.30	PyOcto1D	1.00	0.93	0.97	0.61	0.42	5.37
shallow	500	0.30	REAL	0.99	0.93	0.96	3.07	0.33	64.30
shallow	500	0.30	REAL1D	0.95	0.84	0.89	7.55	0.33	71.81
shallow	500	1.00	Gamma	0.93	0.85	0.89	0.39	1.02	2967.45
shallow	500	1.00	Gamma1D	0.94	0.86	0.90	0.34	1.00	341.33
shallow	500	1.00	PyOcto	1.00	0.95	0.97	0.69	1.07	5.17
shallow	500	1.00	PyOcto1D	1.00	0.95	0.97	0.63	0.76	5.95
shallow	500	1.00	REAL	1.00	0.96	0.98	2.88	0.64	120.98
shallow	500	1.00	REAL1D	0.96	0.87	0.91	7.65	0.62	132.69
shallow	500	3.00	PyOcto	1.00	0.94	0.97	1.04	3.17	7.75
shallow	500	3.00	PyOcto1D	1.00	0.94	0.97	0.93	2.38	8.63
shallow	500	3.00	REAL	0.99	0.92	0.96	3.14	1.79	280.00
shallow	500	3.00	REAL1D	0.95	0.85	0.90	6.94	1.86	304.51
shallow	2000	0.30	PyOcto	0.99	0.88	0.93	1.85	2.31	21.37
shallow	2000	0.30	PyOcto1D	0.99	0.89	0.94	1.53	1.77	24.98
shallow	2000	0.30	REAL	0.99	0.87	0.93	3.96	1.53	293.98
shallow	2000	0.30	REAL1D	0.93	0.78	0.85	8.24	1.68	315.48

Continued on next page

Table S4: Full results of synthetic benchmark. We abbreviate *picks per event* as *ppe*.

Scenario	Events	Noise	Associator	Precision	Recall	F1	Missing ppe	Additional ppe	Run time [s]
shallow	2000	1.00	PyOcto	0.98	0.87	0.92	2.27	4.88	25.85
shallow	2000	1.00	PyOcto1D	0.99	0.89	0.94	1.82	3.69	29.22
shallow	2000	1.00	REAL	0.99	0.88	0.93	4.27	3.03	541.86
shallow	2000	1.00	REAL1D	0.94	0.77	0.85	8.36	3.12	570.99
shallow	2000	3.00	PyOcto	0.65	0.70	0.67	3.81	11.81	118.86
shallow	2000	3.00	PyOcto1D	0.85	0.81	0.83	3.00	9.18	118.70
shallow	2000	3.00	REAL	0.90	0.79	0.84	5.33	7.17	1247.50
shallow	2000	3.00	REAL1D	0.80	0.68	0.74	9.55	7.39	1293.43
subduction	100	0.30	Gamma	0.98	0.97	0.97	3.18	0.02	5.50
subduction	100	0.30	Gamma1D	1.00	0.98	0.99	0.41	0.03	9.73
subduction	100	0.30	PyOcto	1.00	0.99	0.99	3.22	0.07	0.44
subduction	100	0.30	PyOcto1D	1.00	0.98	0.99	0.32	0.05	1.34
subduction	100	0.30	REAL	1.00	0.97	0.98	2.36	0.10	149.07
subduction	100	0.30	REAL1D	1.00	0.97	0.98	0.94	0.08	155.14
subduction	100	1.00	Gamma	0.94	0.95	0.95	3.48	0.06	7.31
subduction	100	1.00	Gamma1D	1.00	0.98	0.99	0.46	0.07	8.19
subduction	100	1.00	PyOcto	0.99	0.98	0.98	3.24	0.14	1.45
subduction	100	1.00	PyOcto1D	1.00	0.98	0.99	0.44	0.12	0.80
subduction	100	1.00	REAL	0.99	0.98	0.98	2.73	0.14	282.84
subduction	100	1.00	REAL1D	1.00	0.98	0.99	1.16	0.32	322.51
subduction	100	3.00	Gamma	0.98	0.94	0.96	3.61	0.28	7.73
subduction	100	3.00	Gamma1D	0.98	0.98	0.98	0.64	0.14	13.13
subduction	100	3.00	PyOcto	1.00	0.99	0.99	3.02	0.33	7.87
subduction	100	3.00	PyOcto1D	1.00	0.98	0.99	0.41	0.23	6.78
subduction	100	3.00	REAL	1.00	0.96	0.98	2.32	0.33	633.18
subduction	100	3.00	REAL1D	1.00	0.99	0.99	1.09	0.41	807.16
subduction	500	0.30	Gamma	0.90	0.83	0.87	3.84	0.31	65.93
subduction	500	0.30	Gamma1D	0.98	0.91	0.94	0.59	0.24	25.80
subduction	500	0.30	PyOcto	0.97	0.94	0.95	3.57	0.35	29.65
subduction	500	0.30	PyOcto1D	1.00	0.97	0.98	0.61	0.27	11.59
subduction	500	0.30	REAL	0.99	0.92	0.95	3.13	0.44	781.53
subduction	500	0.30	REAL1D	0.99	0.93	0.96	1.38	0.55	861.39
subduction	500	1.00	Gamma	0.78	0.70	0.74	3.91	0.64	115.60
subduction	500	1.00	Gamma1D	0.96	0.85	0.90	0.76	0.48	48.98
subduction	500	1.00	PyOcto	0.96	0.92	0.94	3.73	0.70	40.84
subduction	500	1.00	PyOcto1D	0.99	0.96	0.98	0.74	0.56	39.70
subduction	500	1.00	REAL	0.98	0.90	0.93	3.19	0.79	1480.71

Continued on next page

Table S4: Full results of synthetic benchmark. We abbreviate *picks per event* as *ppe*.

Scenario	Events	Noise	Associator	Precision	Recall	F1	Missing ppe	Additional ppe	Run time [s]
subduction	500	1.00	REAL1D	0.98	0.91	0.94	1.57	1.13	1803.82
subduction	500	3.00	Gamma	0.61	0.47	0.53	4.51	1.62	11758.45
subduction	500	3.00	Gamma1D	0.71	0.56	0.62	0.86	1.12	1514.86
subduction	500	3.00	PyOcto	0.95	0.93	0.94	3.66	1.55	21.26
subduction	500	3.00	PyOcto1D	0.99	0.98	0.99	1.00	1.09	7.56
subduction	500	3.00	REAL	0.97	0.90	0.93	3.29	1.55	3467.77
subduction	500	3.00	REAL1D	0.98	0.94	0.96	2.08	2.26	4470.67
subduction	2000	0.30	Gamma	0.54	0.43	0.48	4.85	1.38	2842.45
subduction	2000	0.30	Gamma1D	0.82	0.64	0.72	1.07	0.96	382.85
subduction	2000	0.30	PyOcto	0.89	0.81	0.85	4.51	1.45	109.50
subduction	2000	0.30	PyOcto1D	0.98	0.89	0.93	1.48	0.98	93.24
subduction	2000	0.30	REAL	0.92	0.72	0.80	4.21	1.56	3210.97
subduction	2000	0.30	REAL1D	0.97	0.78	0.87	2.60	2.00	3574.75
subduction	2000	1.00	Gamma1D	0.49	0.45	0.47	1.06	2.03	140152.07
subduction	2000	1.00	PyOcto	0.83	0.77	0.80	4.51	2.46	176.07
subduction	2000	1.00	PyOcto1D	0.96	0.89	0.92	1.90	1.83	146.68
subduction	2000	1.00	REAL	0.89	0.70	0.78	4.54	2.73	5571.98
subduction	2000	1.00	REAL1D	0.93	0.76	0.84	3.34	3.42	6602.46
subduction	2000	3.00	PyOcto	0.27	0.51	0.36	5.74	5.47	636.98
subduction	2000	3.00	PyOcto1D	0.46	0.74	0.57	3.35	4.20	679.37
subduction	2000	3.00	REAL	0.43	0.43	0.43	5.87	5.70	9675.44
subduction	2000	3.00	REAL1D	0.51	0.54	0.53	5.02	6.57	11269.04

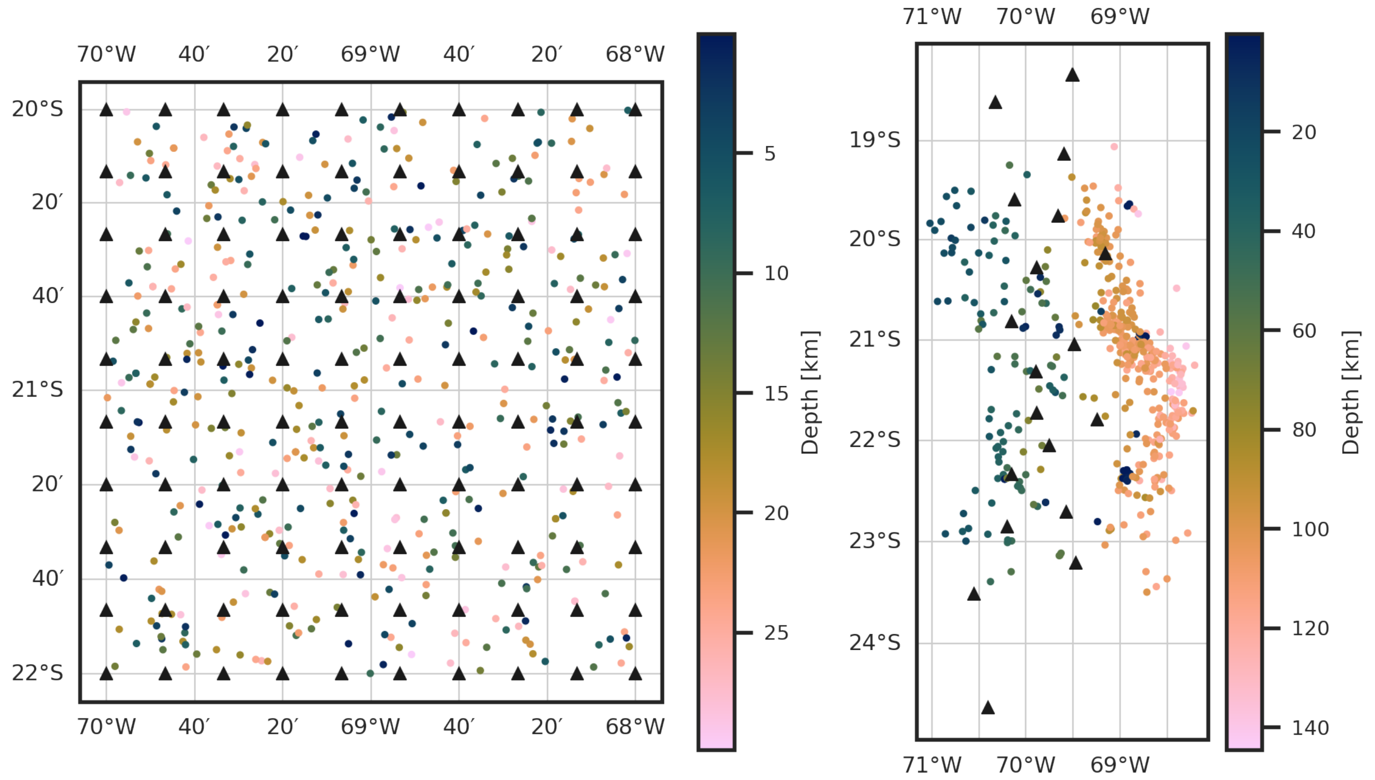

Figure S1: Example maps for the shallow (left) and subduction (right) scenario. Both examples show the station configuration with black triangles. In addition, they visualize an example distribution of 500 seismic events. The station configuration is kept stable across all experiments while the realisation of events changes.

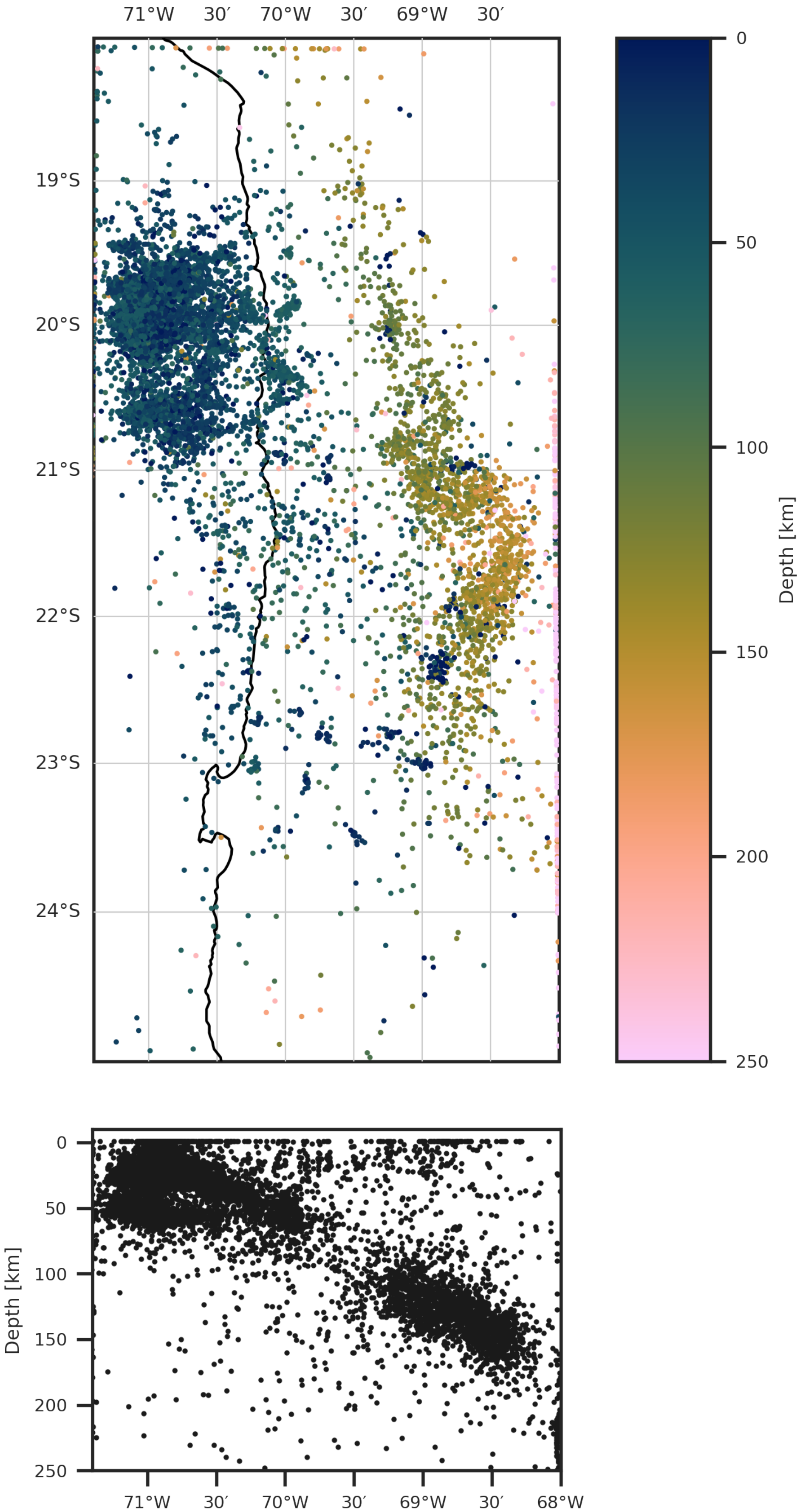

Figure S2: Visualisation of the catalog for the Iquique sequence generated with GaMMA using a 1D velocity model. Note that the catalog only captures 22 of 31 days as GaMMA did not converge for the other days.

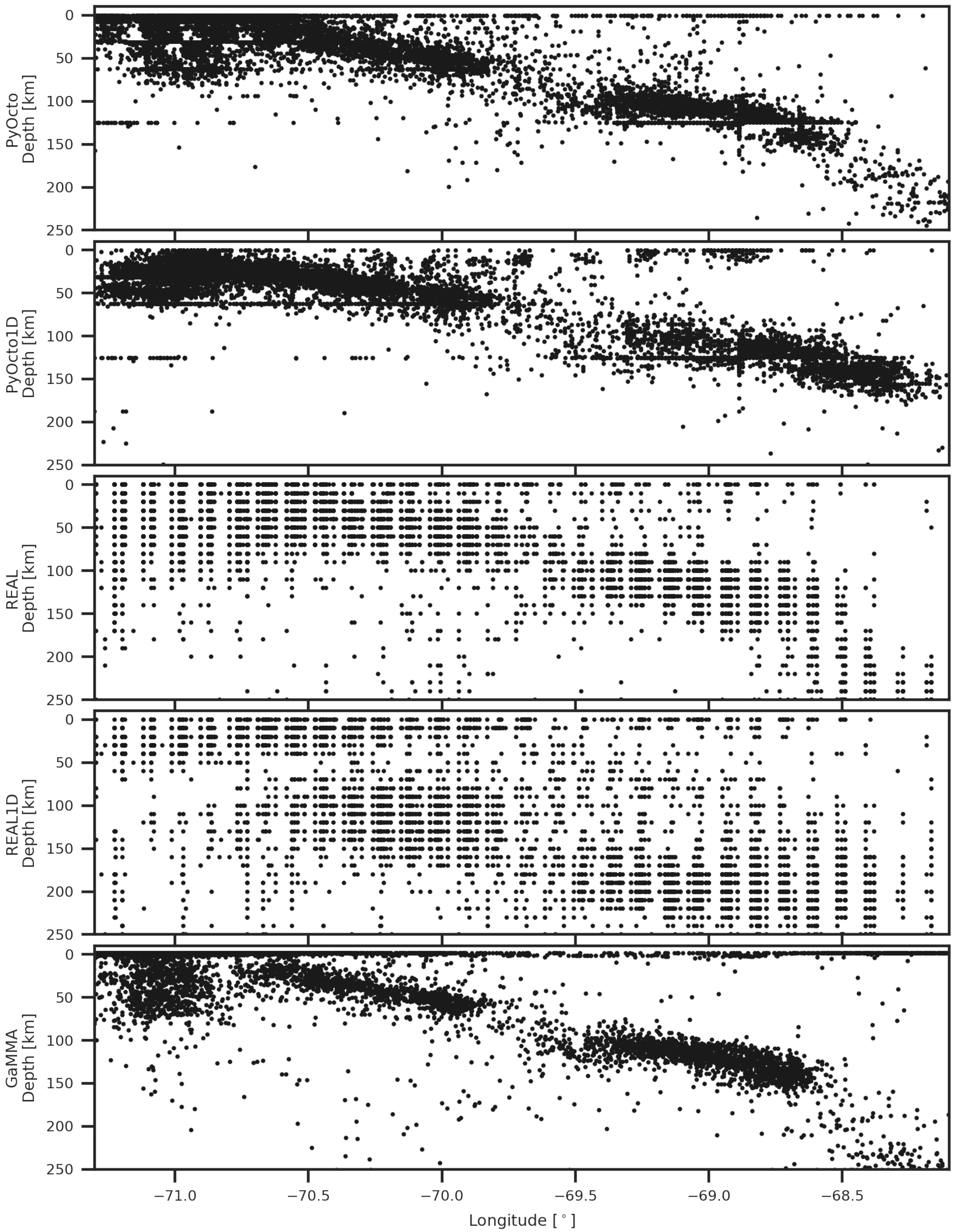

Figure S3: Cross section views of the catalogs generated for the Iquique sequence (15th March 2014 to 15th April 2014) using different phase associators. We visualize the output locations as provided by the associators. Please note that in a comprehensive workflow, absolute and relative relocation techniques should be used as a refinement step. Gridding for REAL is caused by the grid-search approach, stripes for PyOcto by local minima in the EDT loss for localization.

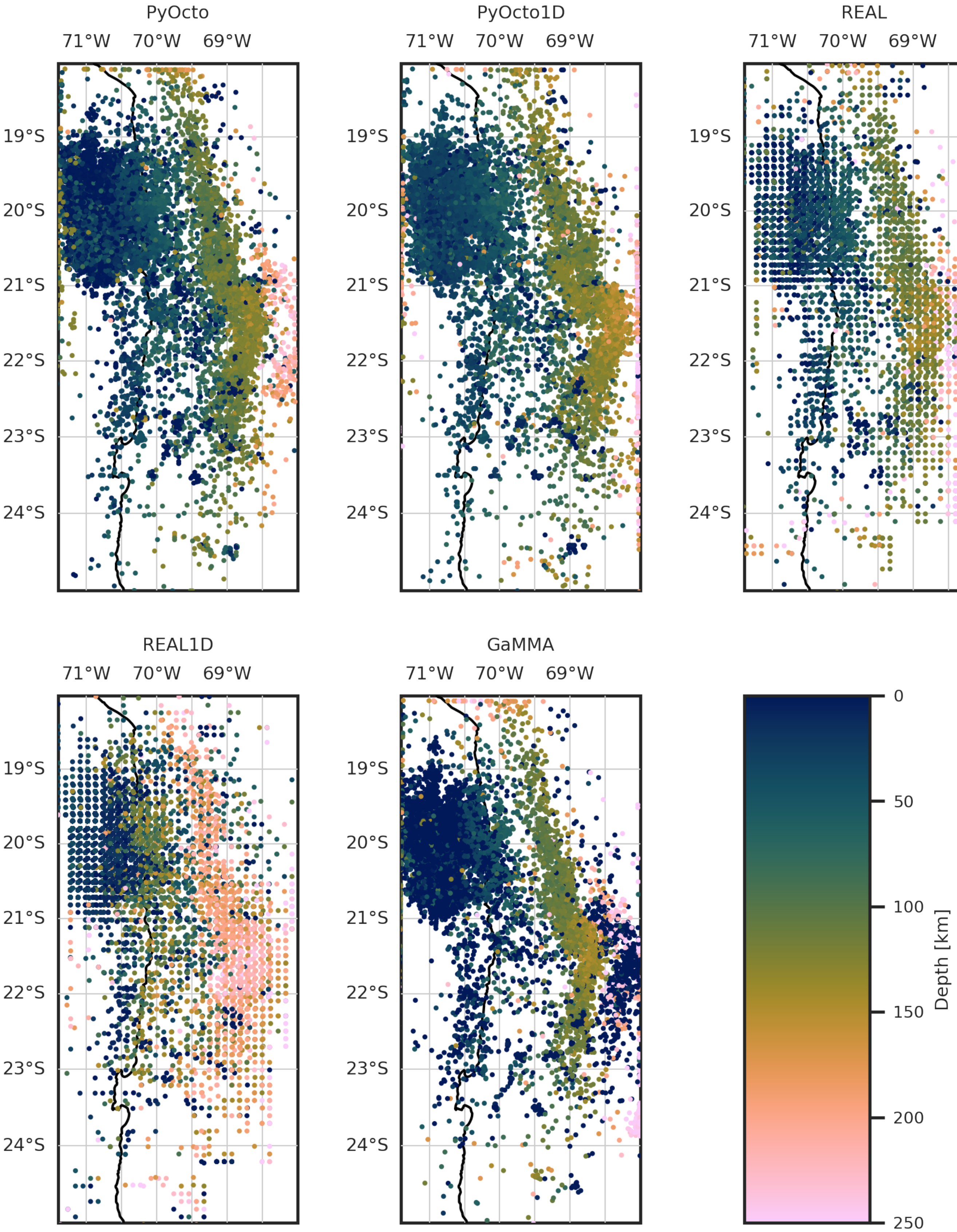

Figure S4: Same as Figure 5 but requiring only 7 total picks and 3 stations with P and S picks.

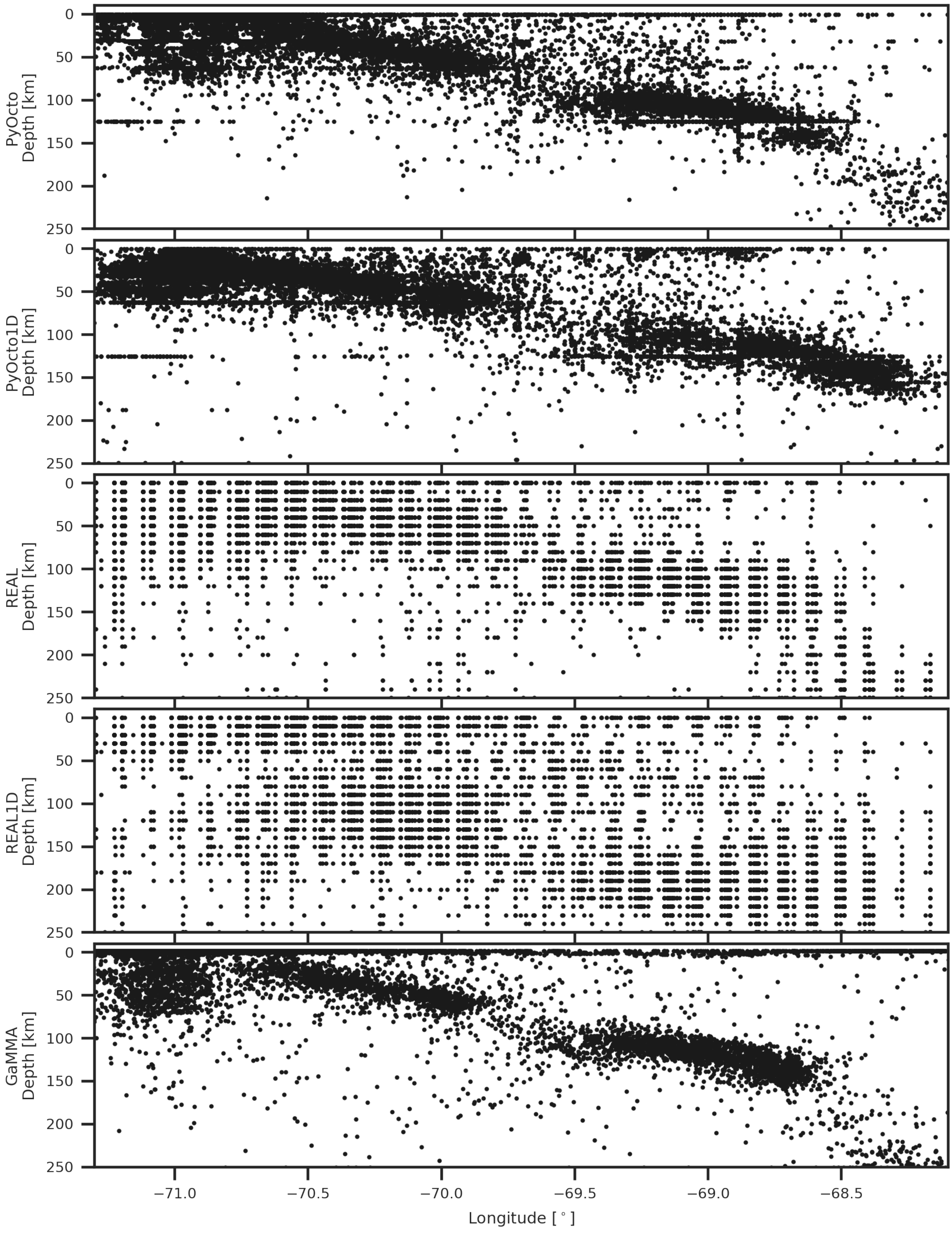

Figure S5: Same as Figure S3 but requiring only 7 total picks and 3 stations with P and S picks.

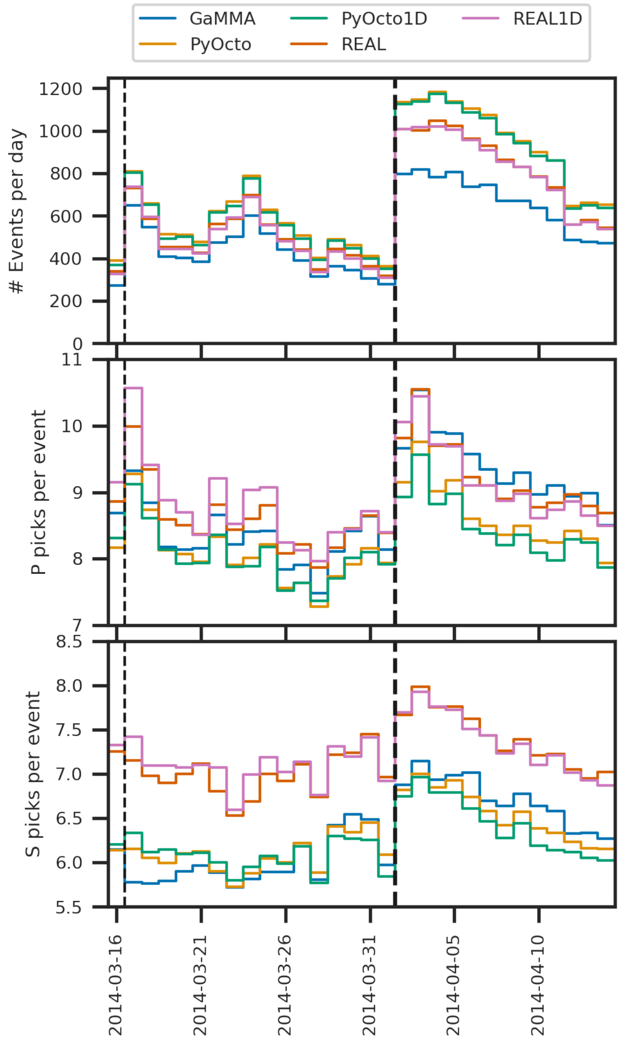

Figure S6: Same as Figure 6 but requiring only 7 total picks and 3 stations with P and S picks.