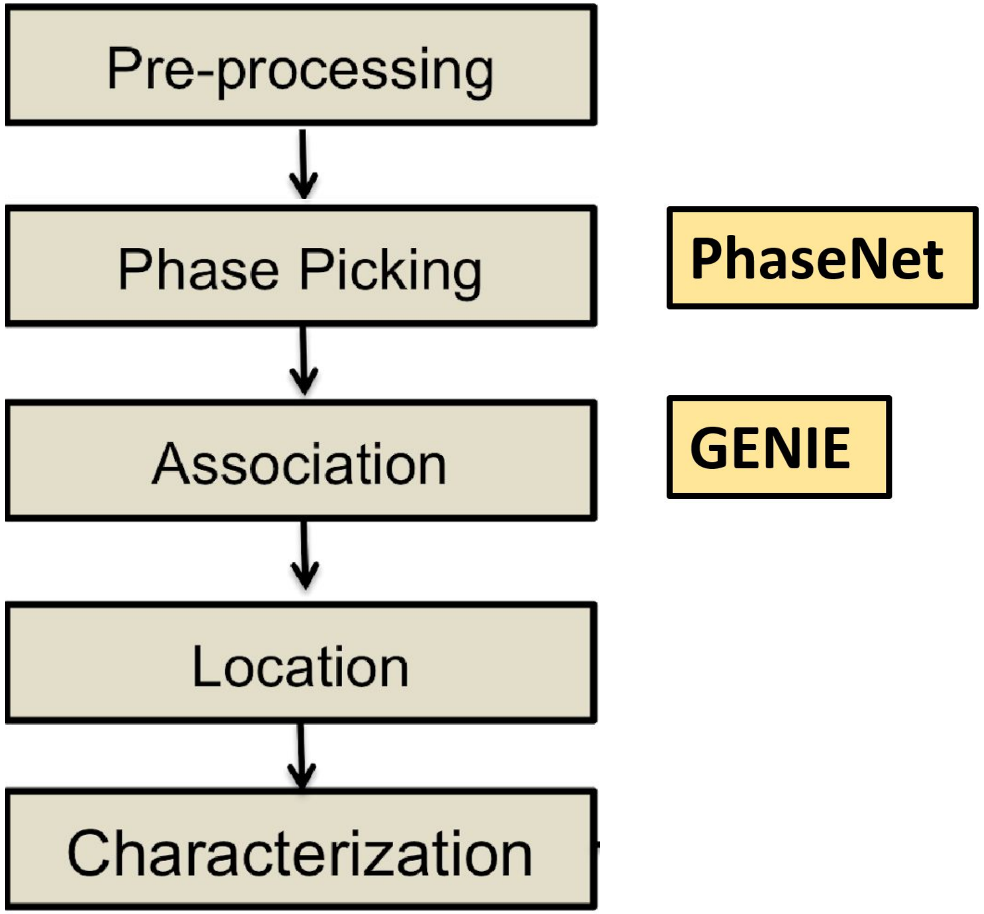
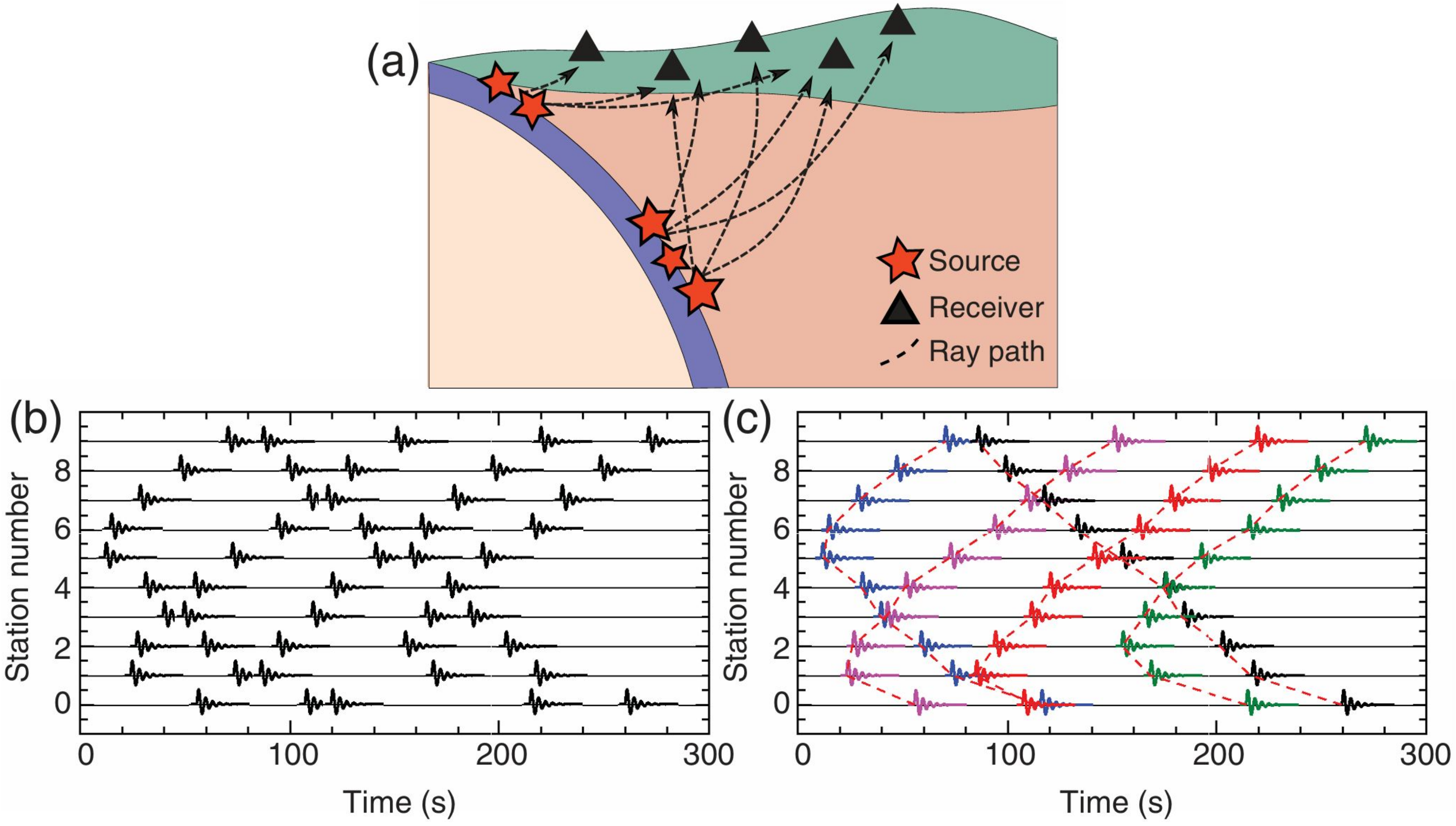
Earthquake Phase Association

Seismic Event Monitoring Workflow: From Continuous Waveforms to Catalogs



Association



“Associate” picks – (i.e., determine number of events and distinct assignments)

Location

Then use **associated picks** in a least squares optimization routine to find best fit location

Posterio r

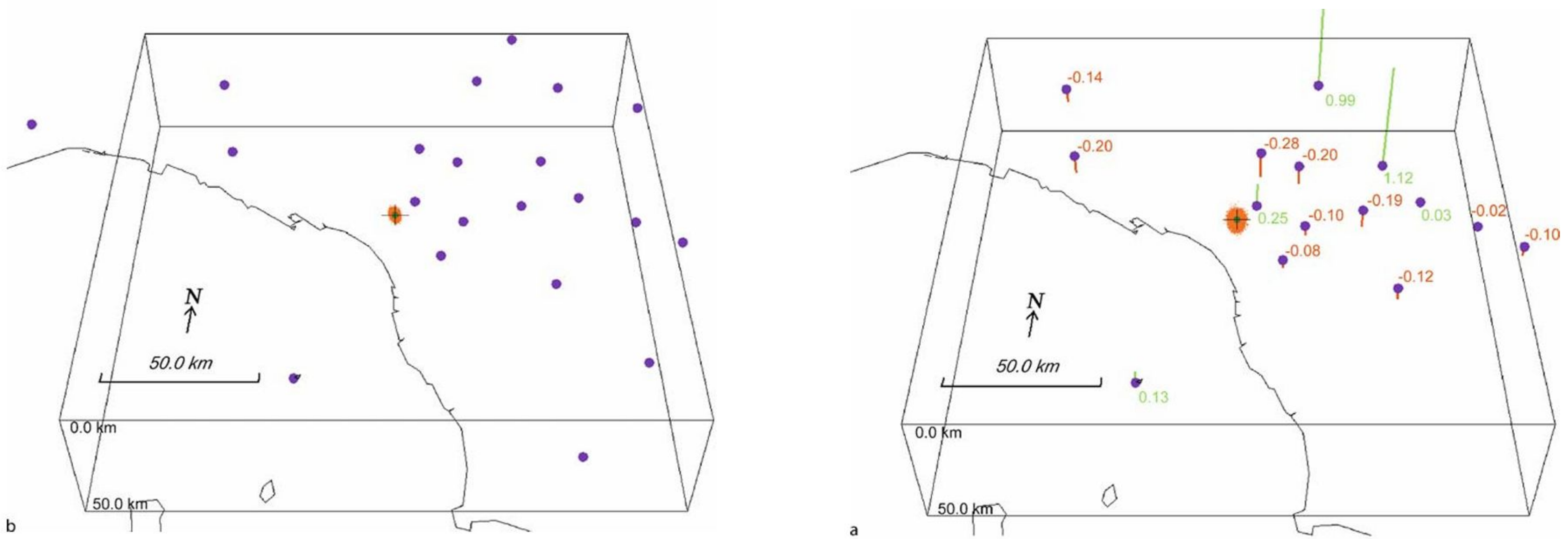
$$\theta(\boldsymbol{X}) \times \exp\Big(-\frac{(\bar{T}_i^r(\boldsymbol{X}) - \bar{\tau}_i^r)^T C_{cov}^{-1} (\bar{T}_i^r(\boldsymbol{X}) - \bar{\tau}_i^r)}{2}\Big)$$

$\bar{T}_i^r(\boldsymbol{X})$: travel time calculator

$\bar{\tau}_i^r$: arrival time

Location

Then use **associated picks** in a least squares optimization routine to find best fit location



Lomax et al., 2008

Posterior

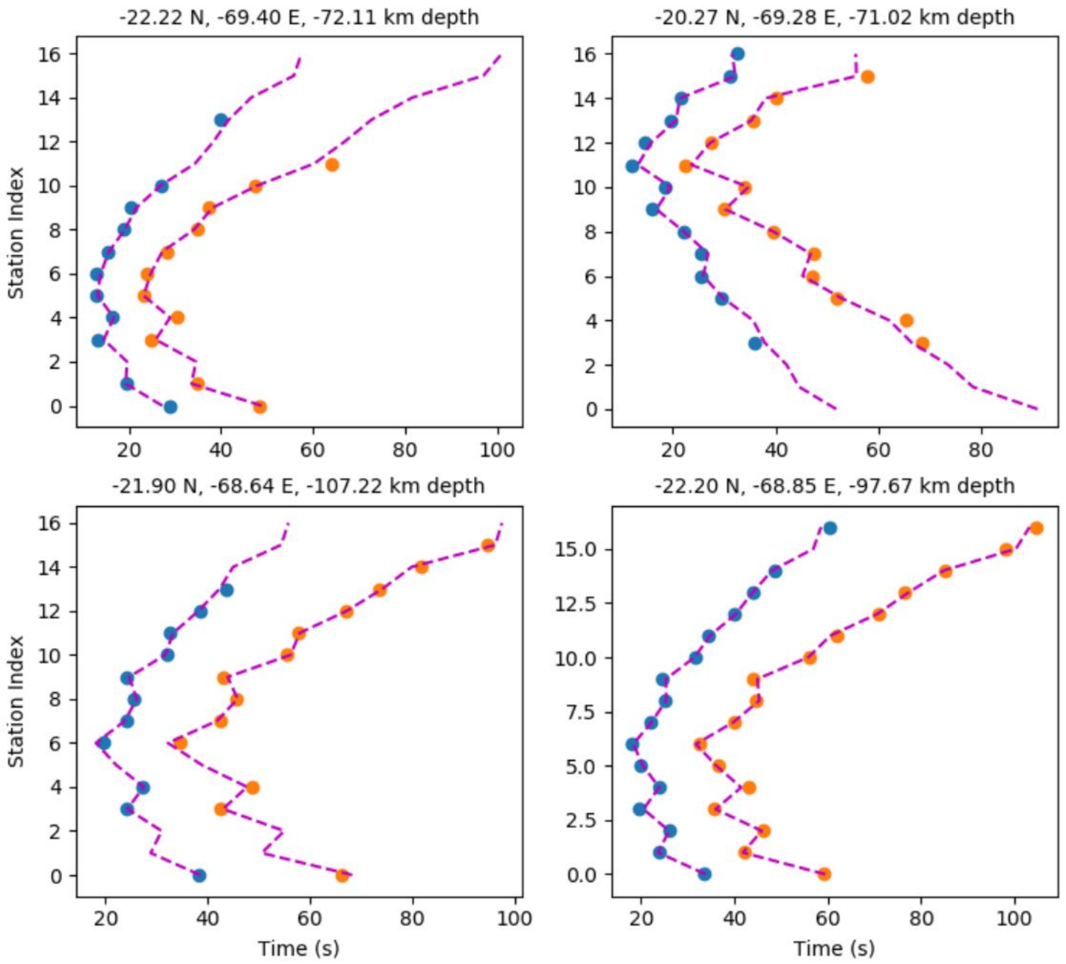
$$\theta(\boldsymbol{X}) \times \exp\left(-\frac{(\bar{T}_i^r(\boldsymbol{X}) - \bar{\tau}_i^r)^T C_{cov}^{-1}(\bar{T}_i^r(\boldsymbol{X}) - \bar{\tau}_i^r)}{2}\right)$$

$\bar{T}_i^r(\boldsymbol{X})$: travel time calculator

$\bar{\tau}_i^r$: arrival time

Location

Then use **associated picks** in a least squares optimization routine to find best fit location

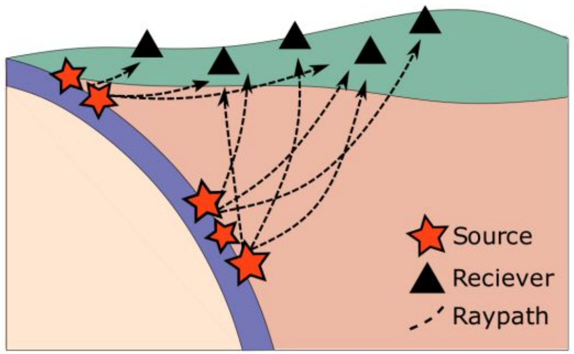


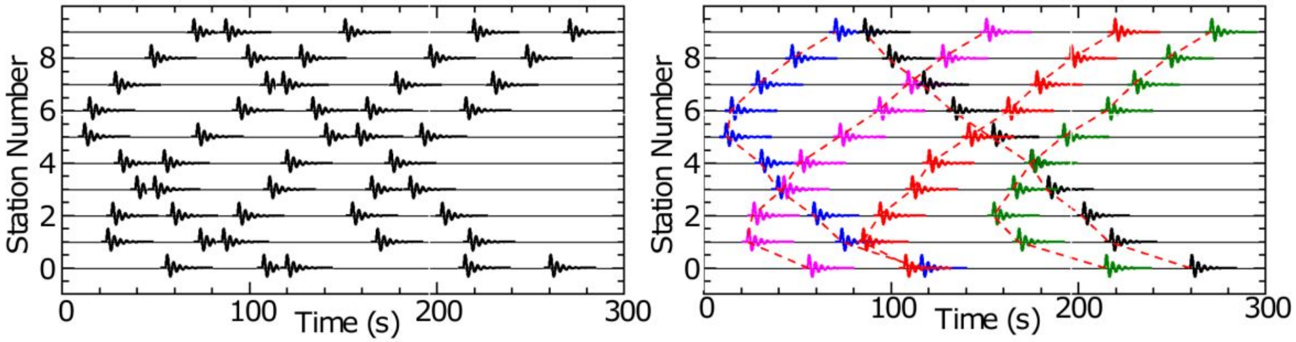
Posterior

$$\theta(\boldsymbol{X}) \times \exp\left(-\frac{(\bar{T}_i^r(\boldsymbol{X}) - \bar{\tau}_i^r)^T C_{cov}^{-1} (\bar{T}_i^r(\boldsymbol{X}) - \bar{\tau}_i^r)}{2}\right)$$

Phase Association connects waves with earthquakes.

Why is it challenging?

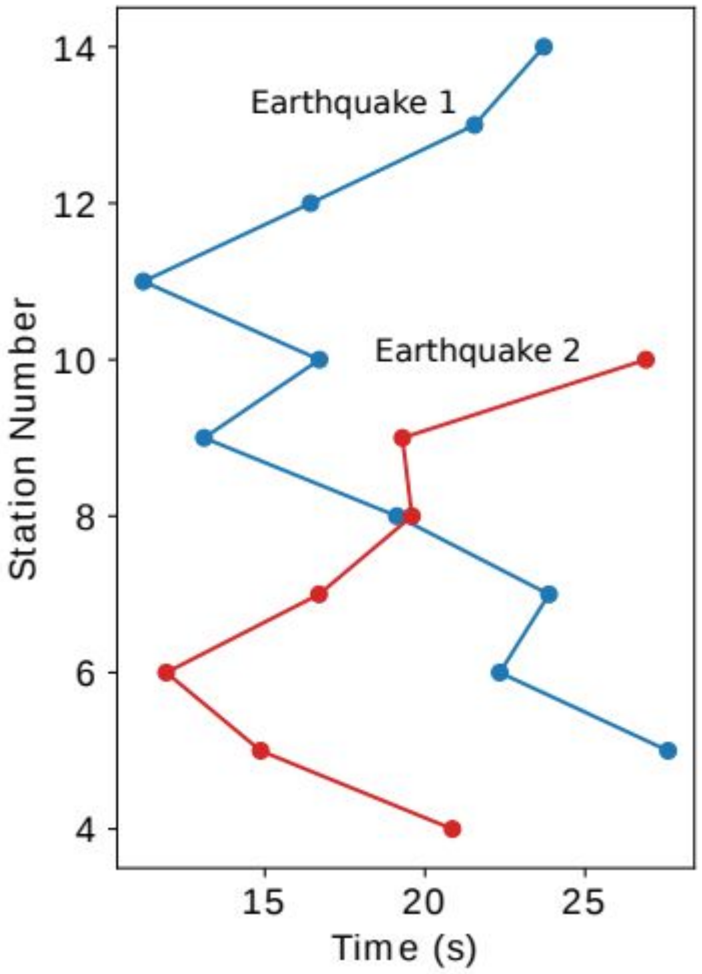


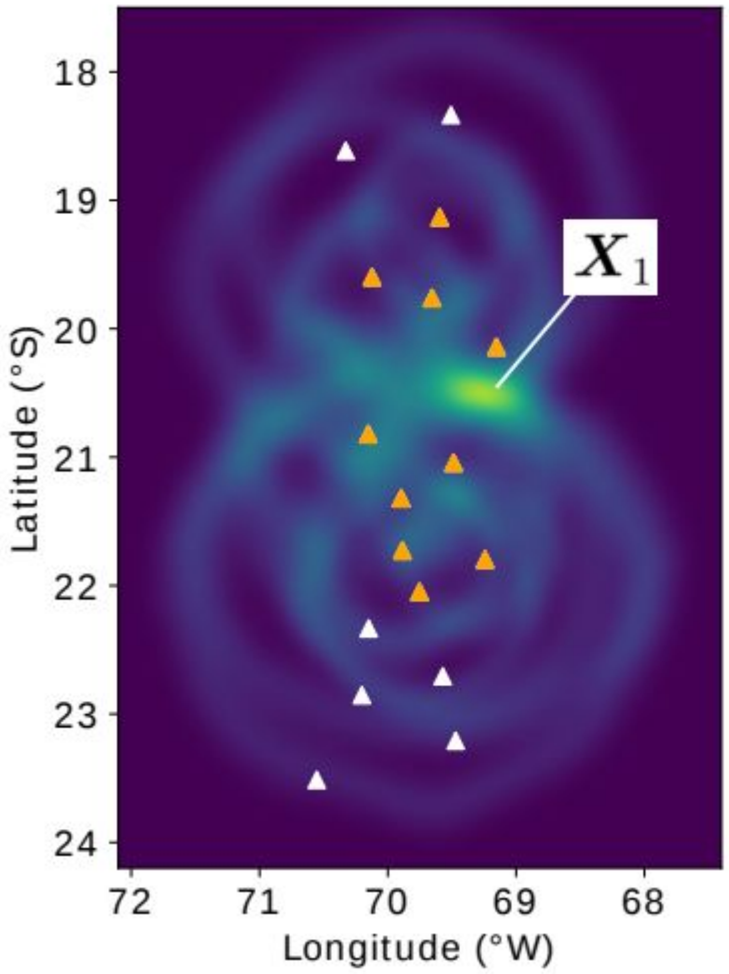


- Number of earthquakes is unknown
- Events close in time have overlapping waveforms
- Recording network is irregular and varies with time
- Small earthquakes are only recorded on a few stations

ML-based picks differ from traditional picks, which motivates another look at.

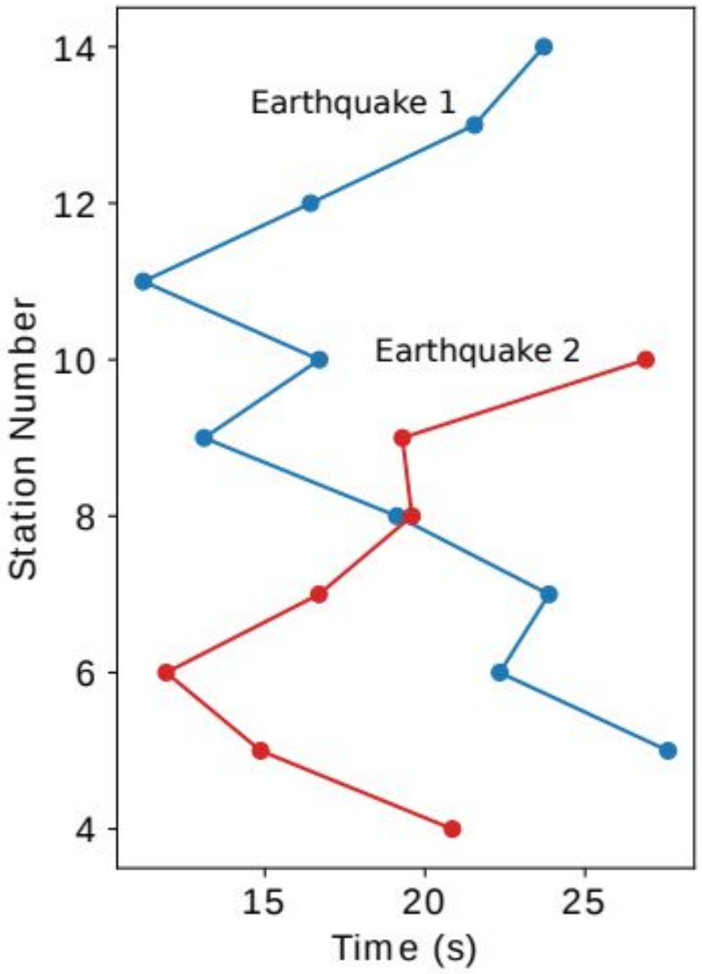
Ambiguity of Phase Association

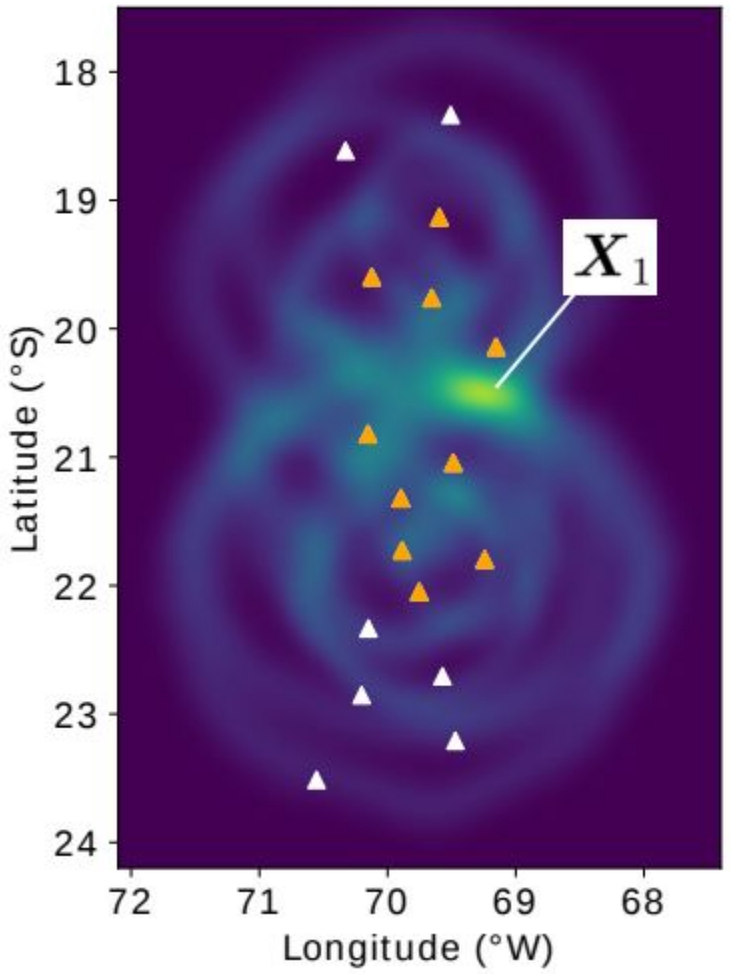


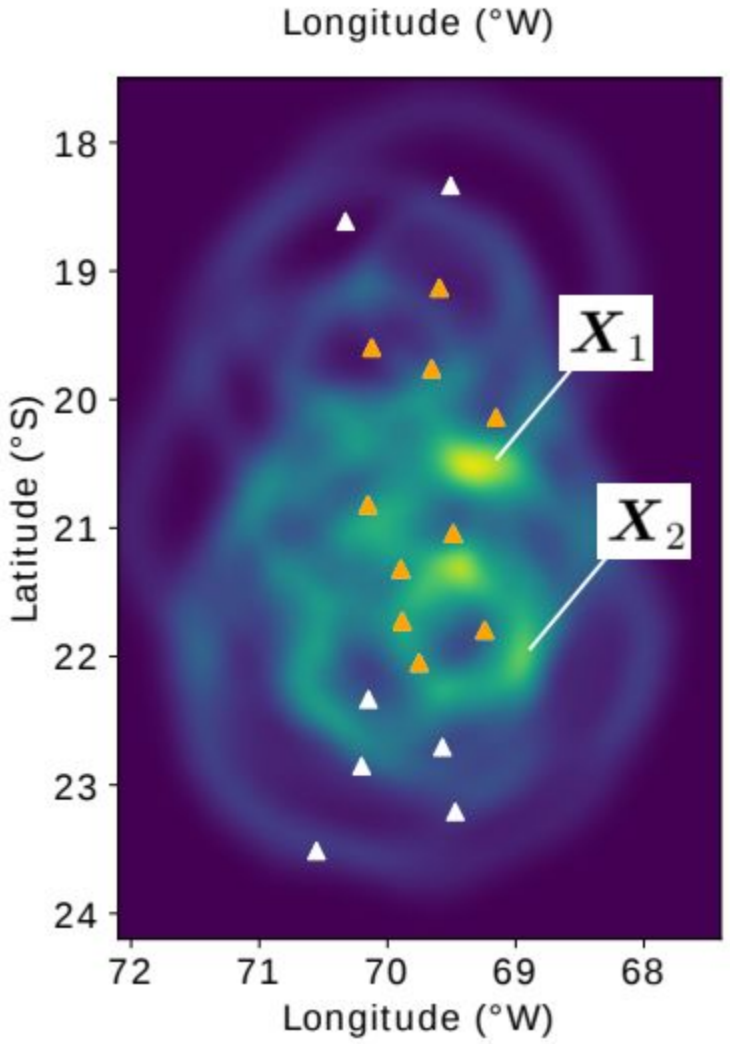


Backprojection: Time reverse picks and stack over stations (e.g., find moveout that fits observed picks)

Ambiguity of Phase Association







Backprojection: Time reverse picks and stack over stations (e.g., find moveout that fits observed picks)

Brief History

Brief History

Ringdall and
Kvearna
(1989)

1930 FRODE RINGDAL AND TORMOD KVÆRNA

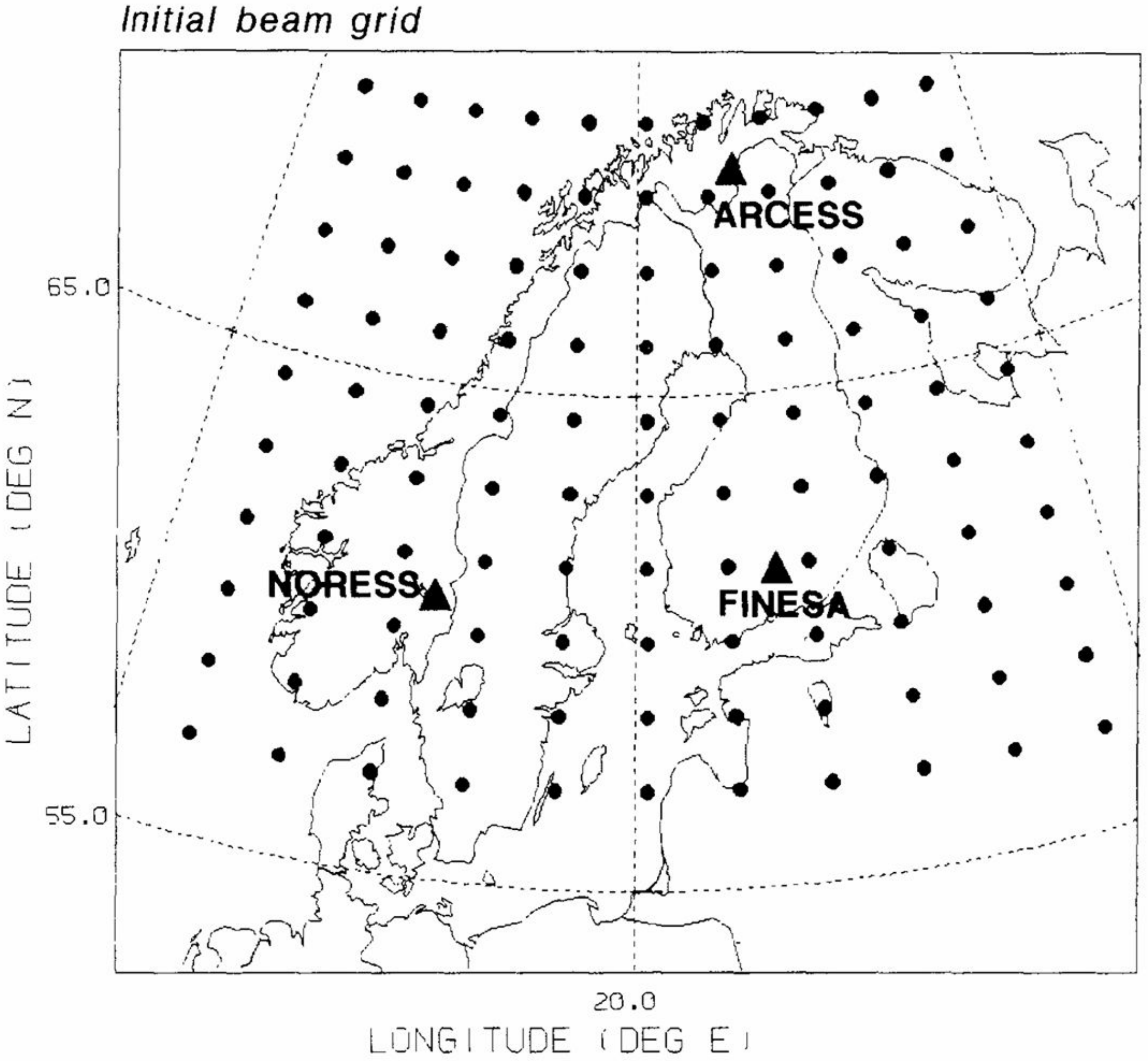
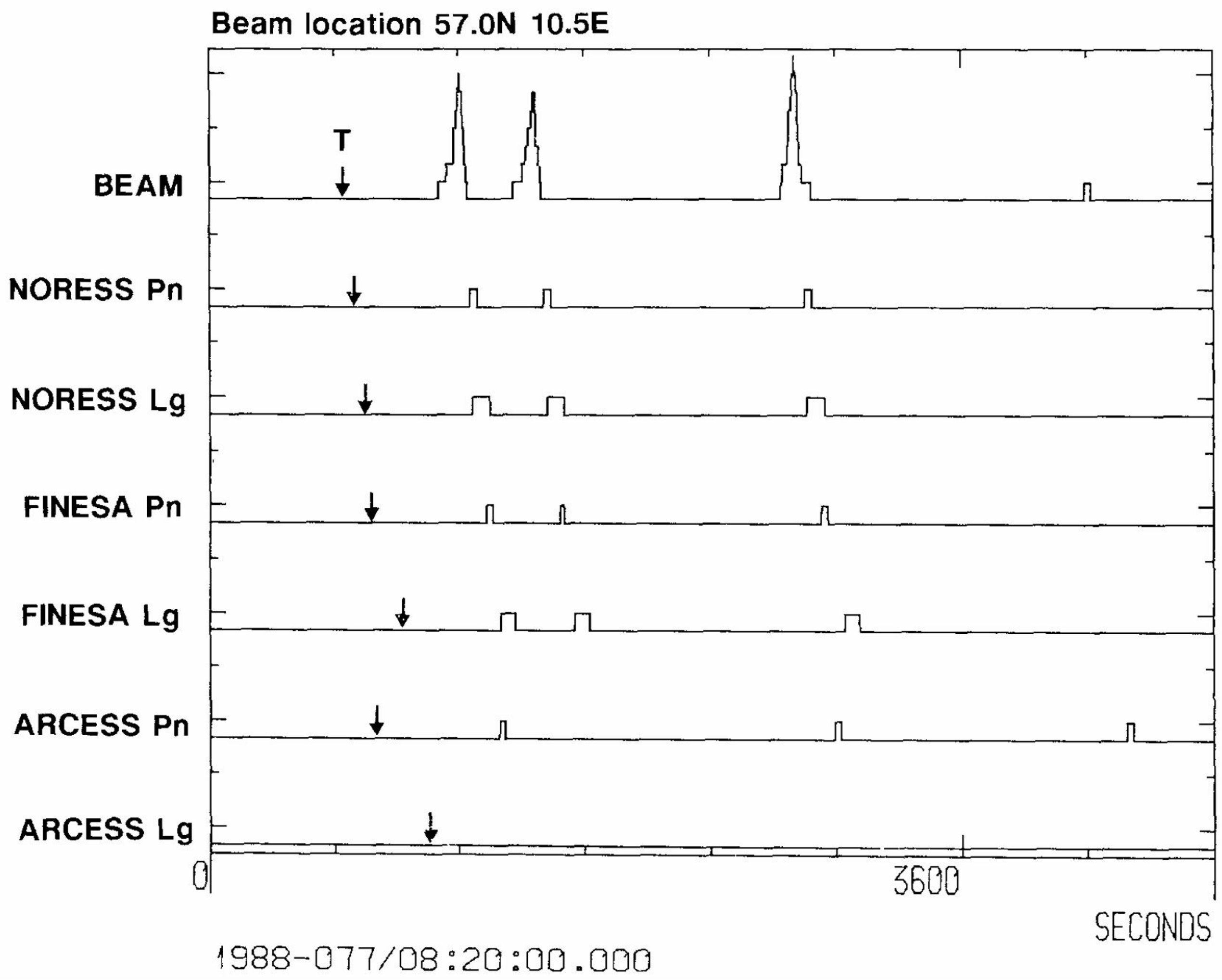


FIG. 1. Beam grid used in the generalized beamforming procedure for the purpose of associating regional phases from NORESS, ARCESS, and FINESA. The location of the three arrays is shown on the map.

Brief History

Ringdall and Kvearna (1989)



Standard travel-time tables are used in these computations. Thus, for the jth beam, we obtain a set of time-aligned channels:

$$\bar{s}_j(T) = \{s_{ijk}(T + \tau_{ijk})\} \quad k = 1, \ldots, K_{ij}; \quad i = 1, \ldots, N \tag{1}$$

Brief History

Ringdall and Kvearna (1989)

Events located by beampacking

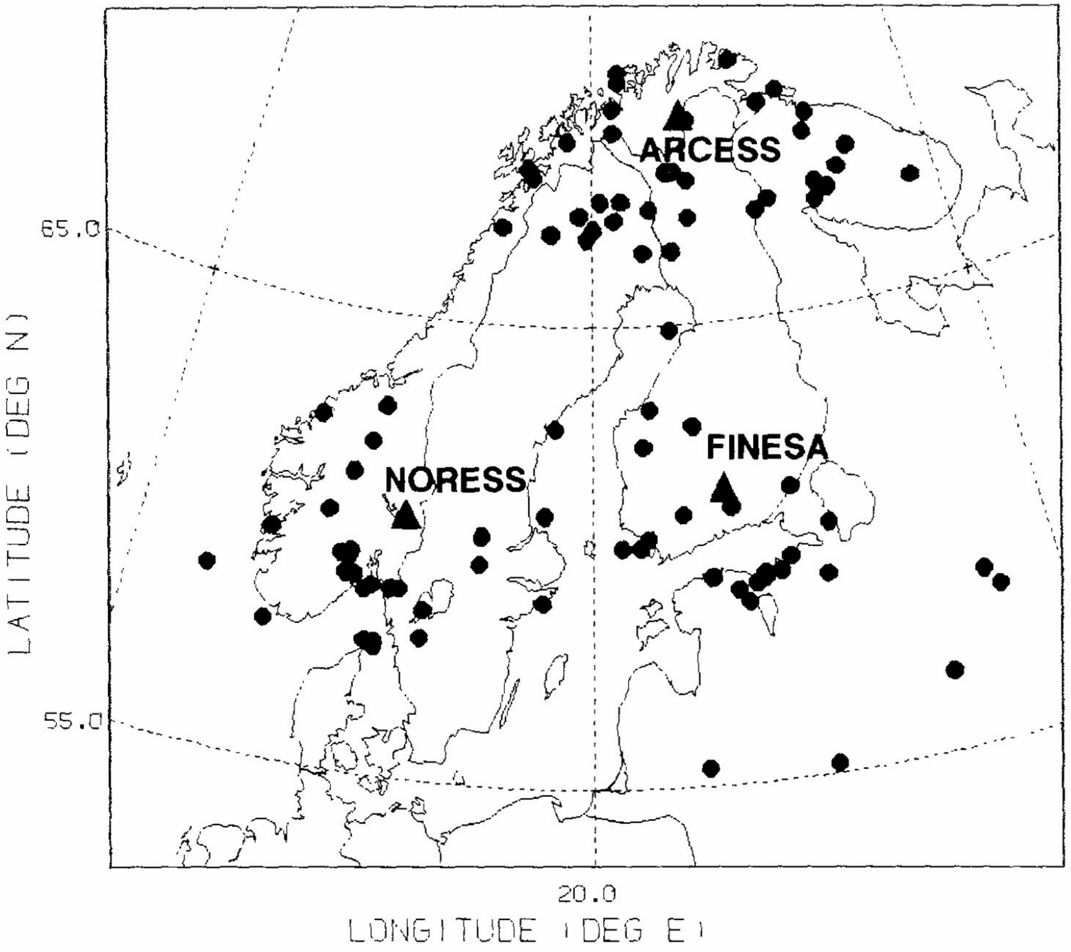


TABLE 3

LOCATION ESTIMATES

Event No.	Date (yy/mm/dd)	Time	Network Lat.	Network Lon.	Mag. M_L	No. of phases
1	88/03/17	08.40.25.0	57.73	11.03	2.5	7
2	88/03/17	08.46.18.7	58.07	11.36	2.6	6
3	88/03/17	09.07.10.3	58.08	11.43	2.7	8
4	88/03/17	10.21.23.0	69.6	29.9	2.9	8
5	88/03/17	10.27.20.0	59.2	27.6	2.3	4
6	88/03/17	10.46.21.0	59.2	27.6	<2	2
7	88/03/17	11.18.48.0	59.3	27.2	2.3	5
8	88/03/17	11.54.41.0	65.8	24.7	<2	5
9	88/03/17	11.57.57.9	60.57	8.36	1.8	2
10	88/03/17	12.02.36.0	59.4	28.5	2.1	3
11	88/03/17	12.42.22.9	59.78	10.76	2.3	3
12	88/03/17	14.13.14.0	58.33	6.28	2.4	4
13	88/03/17	14.21.08.0	60.9	29.4	2.3	3
14	88/03/17	14.33.58.3	59.06	5.88	2.2	2
15	88/03/17	18.58.08.1	59.68	5.57	3.2	7

Location estimates obtained automatically from the beampacking pro ent network locations from the Helsinki and Bergen bulletins. Note the events with more than one detecting array.

Brief History

Johnson et al., (1997)

$$\mathrm{Phs}_i := \{\ldots, \mathrm{phs}_{i-2}, \mathrm{phs}_{i-1}, \mathrm{phs}_i\}.$$

$$\mathrm{Hyp}_i := \{\ldots, \mathrm{hyp}_{i,j-2}, \mathrm{hyp}_{i,j-1}, \mathrm{hyp}_{i,j}\},$$

Formulate problem as a discrete assignment problem..

Brief History

Johnson et al., (1997)

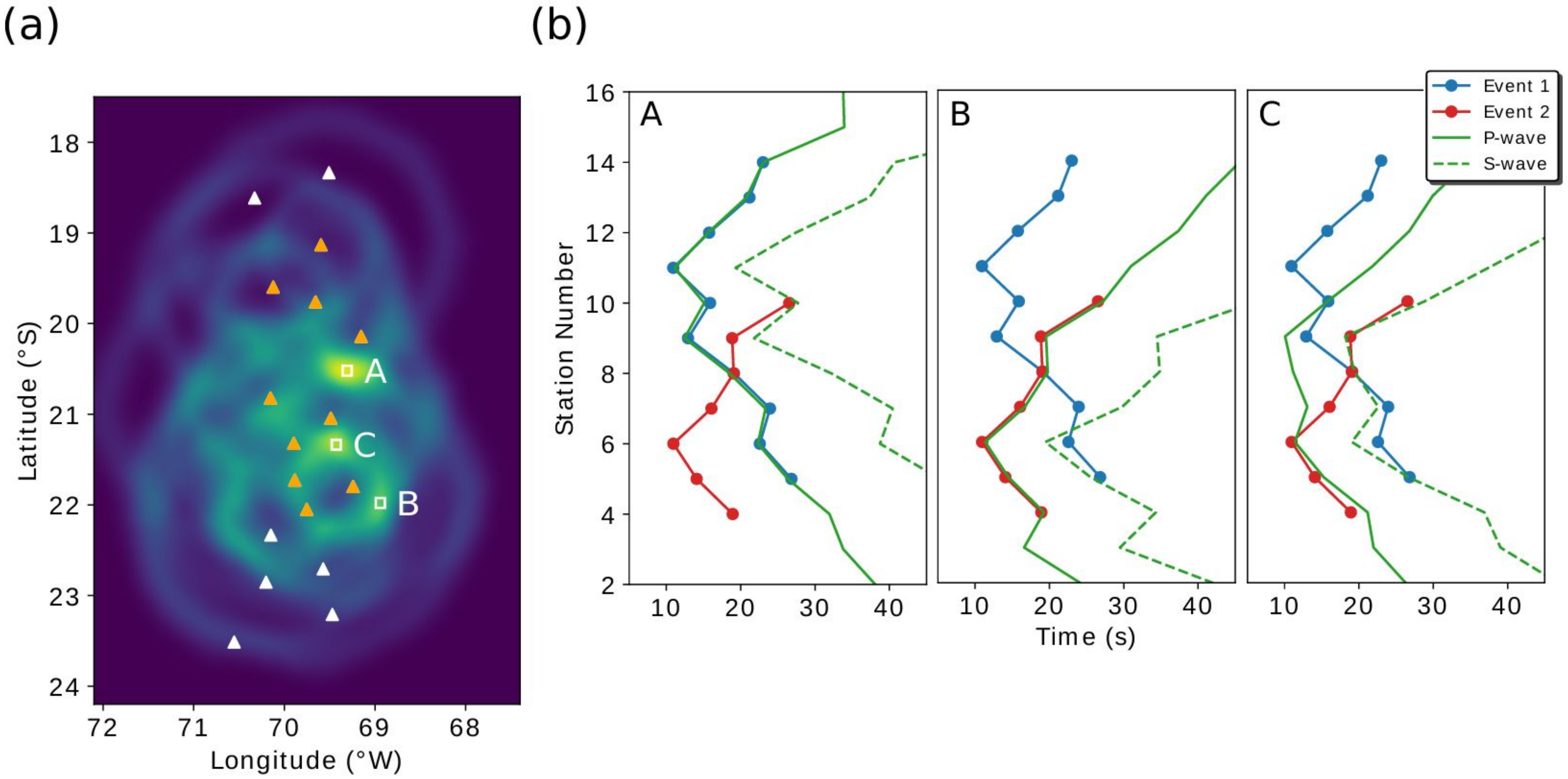
$$\text{Phs}_i := \{\ldots, \text{phs}_{i-2}, \text{phs}_{i-1}, \text{phs}_i\}.$$

$$\text{Hyp}_i := \{\ldots, \text{hyp}_{i,j-2}, \text{hyp}_{i,j-1}, \text{hyp}_{i,j}\},$$

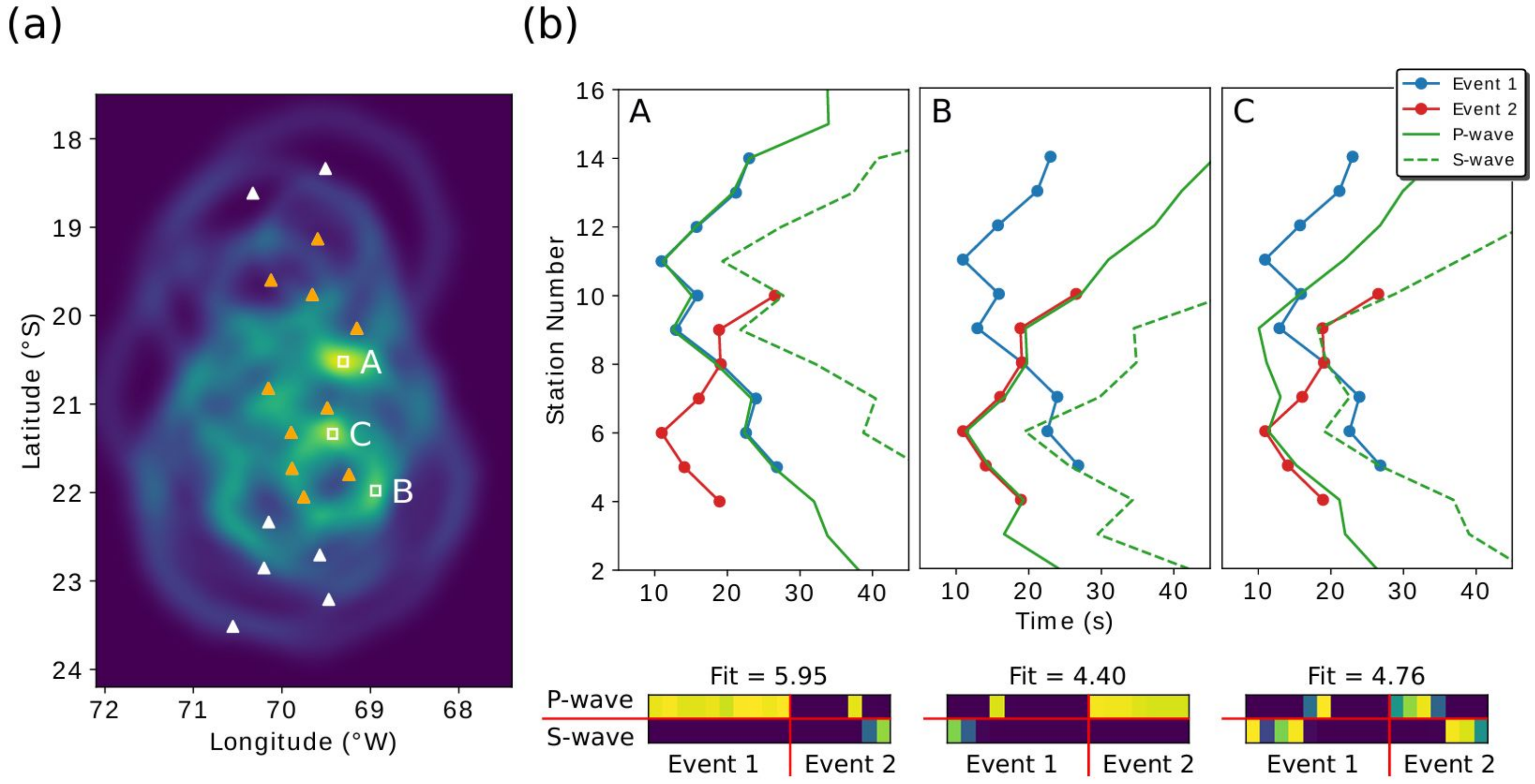
$$A_{i,j} = \begin{cases} \dfrac{W}{\left(W+N_j\right)} \left| \dfrac{Tobs_i - Tcal_{i,j}}{\Delta(r_{ij})} \right| & \textit{If; } phs_i \textit{ is associated with } hyp_j \\ 0 & \textit{Otherwise} \end{cases}$$

Formulate problem as a discrete assignment problem..

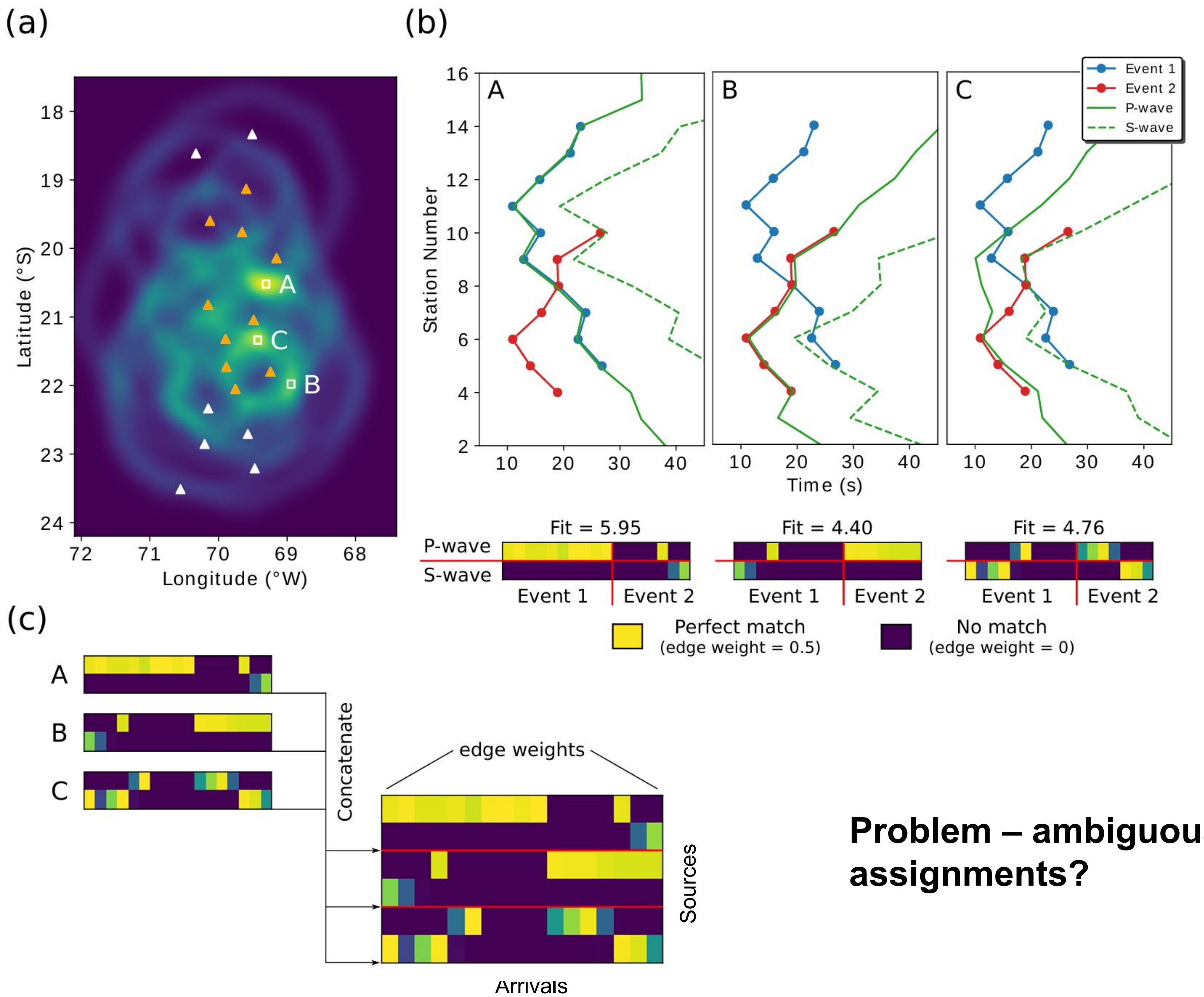
Brief History



Brief History

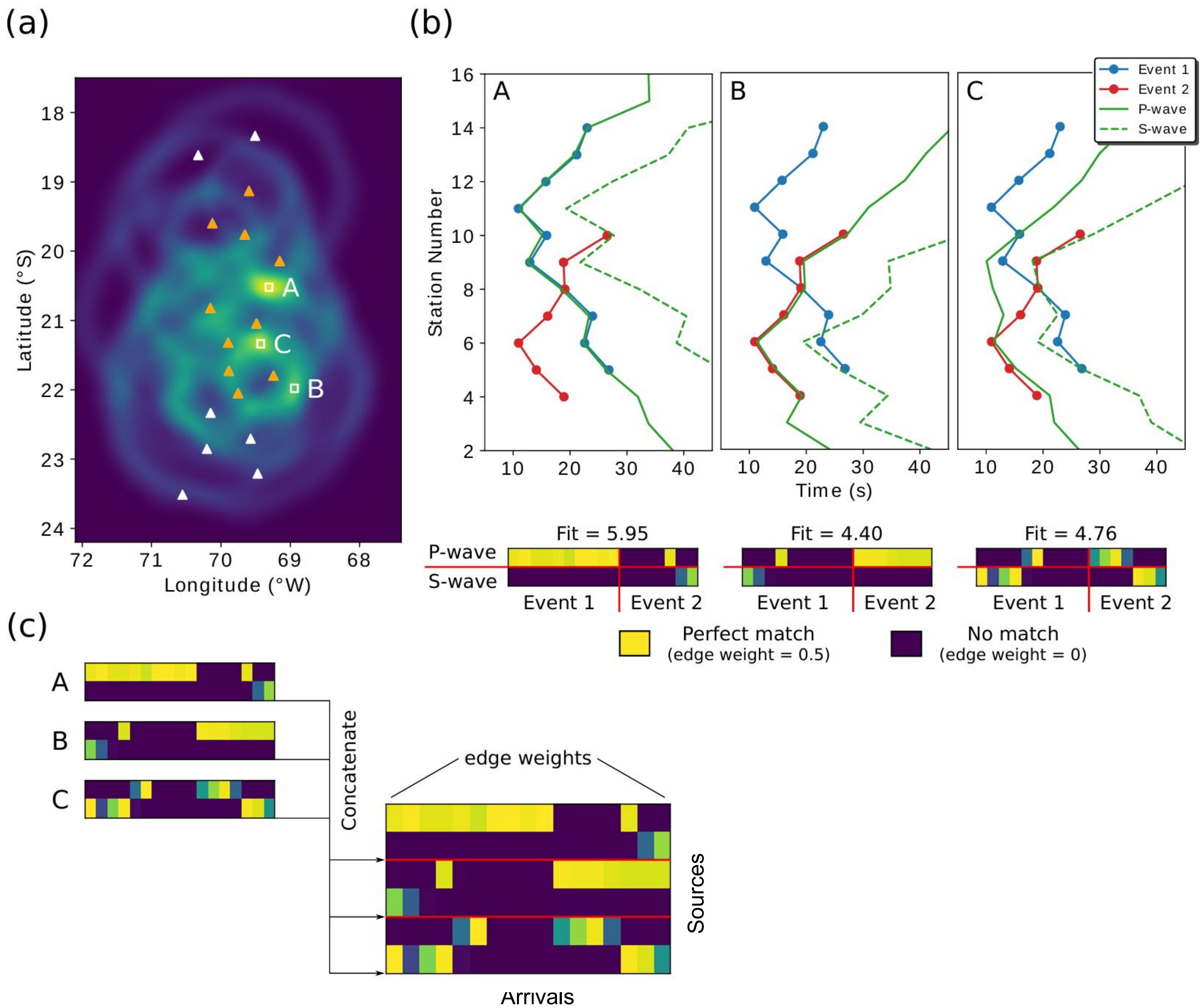


Brief History



Problem – ambiguous assignments?

Brief History



Brief History

Johnson et al., (1997)

$$\text{Phs}_i := \{\ldots, \text{phs}_{i-2}, \text{phs}_{i-1}, \text{phs}_i\}.$$

$$\text{Hyp}_i := \{\ldots, \text{hyp}_{i,j-2}, \text{hyp}_{i,j-1}, \text{hyp}_{i,j}\},$$

$$A_{i,j} = \begin{cases} \frac{W}{(W+N_j)} \left| \frac{Tobs_i - Tcal_{i,j}}{\Delta(r_{ij})} \right| & \textit{If phs}_i \textit{ is associated with hyp}_j \\ 0 & \textit{Otherwise} \end{cases}$$

Formulate problem as a discrete assignment problem..

Algorithms:

$$Norm = \Sigma_{i,j} A_{i,j}.$$

$$Norm_n^r = \sqrt{\frac{n}{(n-r)}} \sum_{ij} A_{ij}$$

Brief History

Johnson et al., (1997)

$$\text{Phs}_i := \{\ldots, \text{phs}_{i-2}, \text{phs}_{i-1}, \text{phs}_i\}.$$

$$\text{Hyp}_i := \{\ldots, \text{hyp}_{i,j-2}, \text{hyp}_{i,j-1}, \text{hyp}_{i,j}\},$$

$$A_{i,j} = \begin{cases} \frac{W}{(W+N_j)}\left|\frac{Tobs_i - Tcal_{i,j}}{\Delta(r_{ij})}\right| & If; phs_i \text{ is associated with } hyp_j \\ 0 & Otherwise \end{cases}$$

Formulate problem as a discrete assignment problem..

Perspective:

The association problem divides into three interconnected subsets problems. They are:

1. Identification
2. Location, and
3. Selection, or association,

Steps 2 and 3 will be looped through many times, as more arrivals associated with an event, the event relocated, etc..

A fourth step should probably be included, which can be roughly as:

4. Clean-up.

Algorithms:

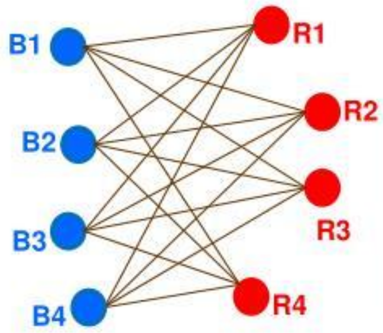
$$Norm = \Sigma_{i,j} A_{i,j}.$$

$$Norm_n^r = \sqrt{\frac{n}{(n-r)}} \sum_{ij} A_{ij}$$

Brief History

Johnson et al., (1997)

Hungarian method - Example



	R1	R2	R3	R4
B1	8	17	3	23
B2	39	4	11	20
B3	13	2	41	6
B4	22	8	9	2

Step 1

First for each row we subtract the row minimum from the rest of the row

	R1	R2	R3	R4	
B1	8	17	3	23	-3
B2	39	4	11	20	-4
B3	13	2	41	6	-2
B4	22	8	9	2	-2

Step 2

Then for each column we subtract the column minimum from the rest of the column

	R1	R2	R3	R4
B1	5	14	0	20
B2	35	0	7	16
B3	11	0	39	4
B4	20	6	7	0
	-5	-0	-0	-0

Brun et al., 2008

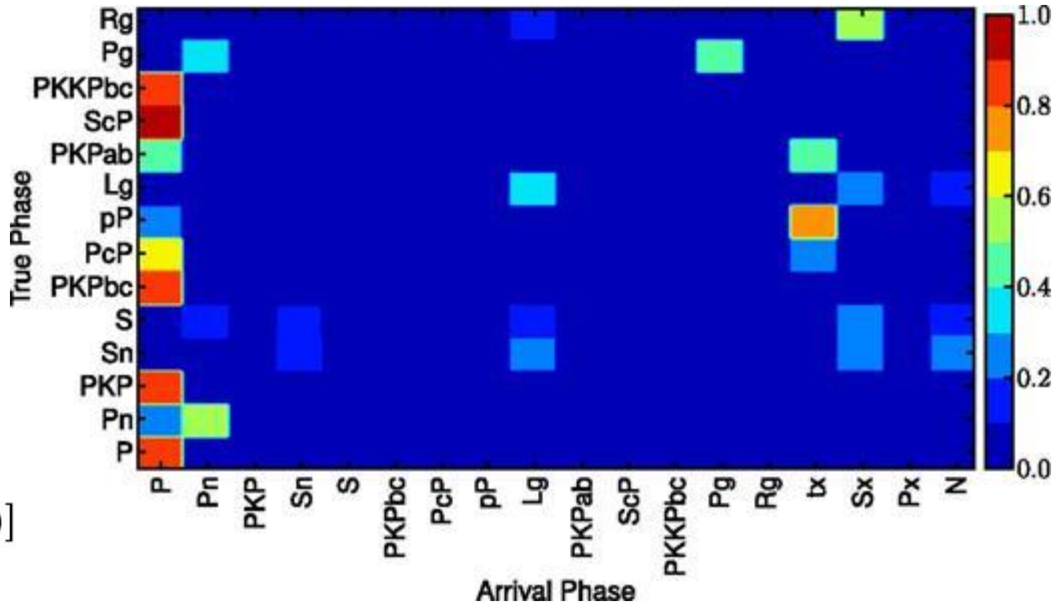
The **Hungarian method** is a combinatorial optimization algorithm that solves the assignment problem in polynomial time and which anticipated later primal–dual methods. It was developed and published in 1955 by Harold Kuhn, who gave it the name "Hungarian

Brief History

Arora et al.,
(2013)

Net-Visa: Probabilistic method – possibly more "accurate", but difficult to implement, and still rule-based, iterative processing of data

$$P_\theta(e) = \exp(-\lambda_e T) \prod_{i=1}^{|e|} P_{\theta,l}(e_l^i) \frac{1}{D} \lambda_e \lambda_m \exp[-\lambda_m(e_m^i - 2)]$$



Brief History

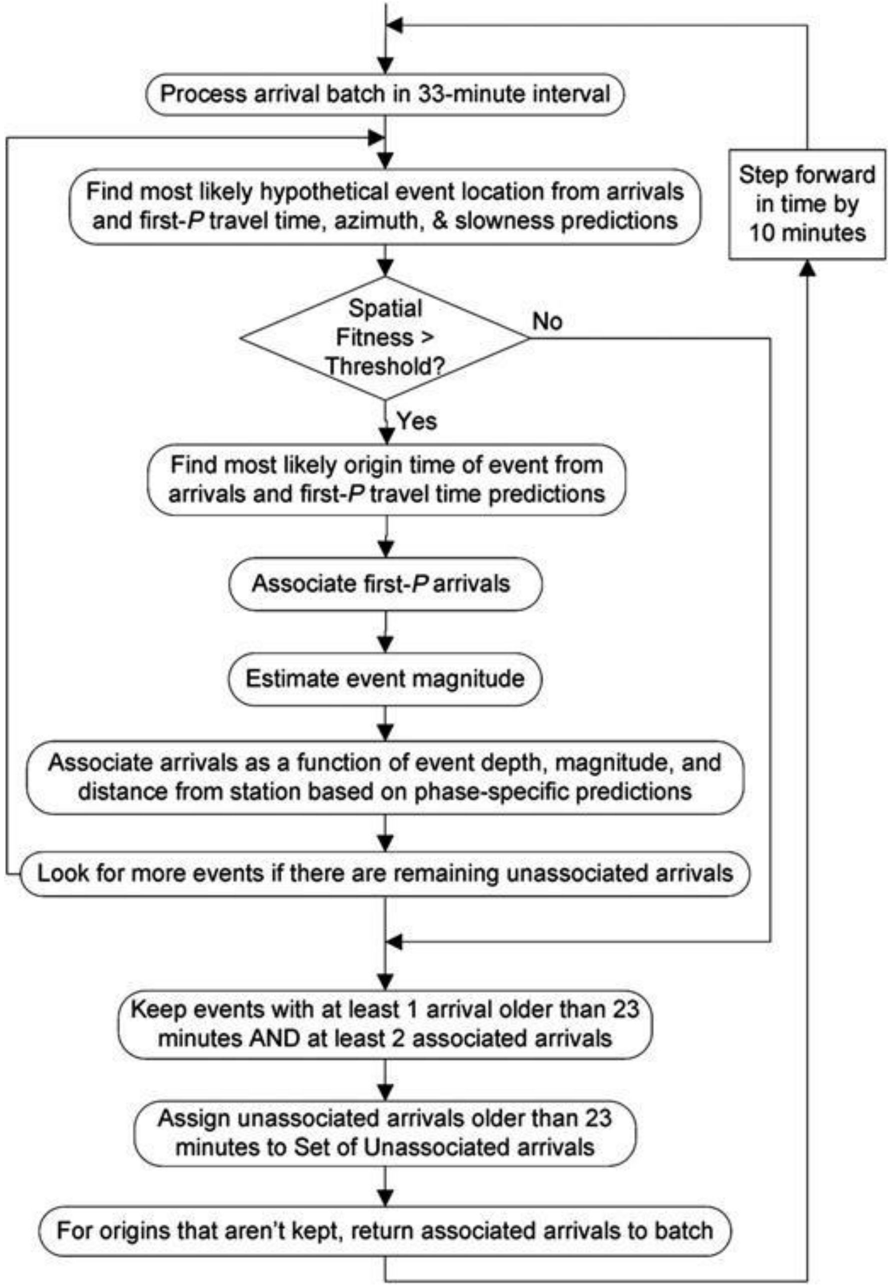
Draelos et al., 2015)

Pedal (similar to GA; 1994**):** temporal energy stack, misfit tables, iterative processing logic/thresholding

$$g_{i,j} = P(d_{s_i}|E_\omega) \times P(d_{s_j}|E_\omega) \times Q_{s_i} \times Q_{s_j},$$

$$w_{i,j} = \frac{\frac{\sqrt{2}}{3} N_{tt}}{(\sigma^2_{tt,i,\omega} + \sigma^2_{tt,j,\omega})^{1/2}} + \frac{\frac{1}{6} N_{az}}{\sigma_{az,i,\omega}} + \frac{\frac{1}{6} N_{az}}{\sigma_{az,j,\omega}} + \frac{\frac{1}{6} N_{sh}}{\sigma_{sh,i,\omega}} + \frac{\frac{1}{6} N_{sh}}{\sigma_{sh,j,\omega}},$$

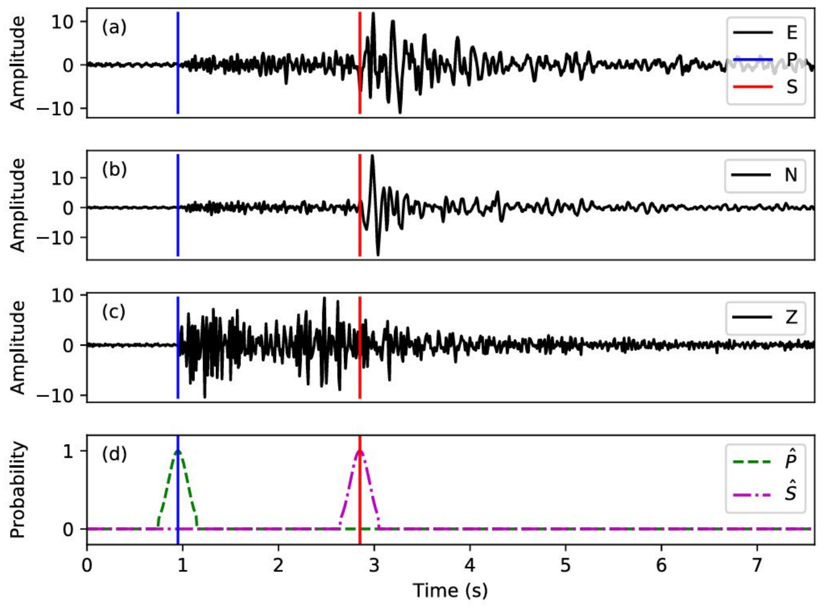
and $$r^2_{i,j} = \frac{[(T_i - T_j) - (p_{tt,i,\omega} - p_{tt,j,\omega})^2]}{\sigma^2_{tt,i,\omega} + \sigma^2_{tt,j,\omega}} + \frac{(az_i - p_{az,i,\omega})^2}{\sigma^2_{az,i,\omega}} + \frac{(az_j - p_{az,j,\omega})^2}{\sigma^2_{az,j,\omega}} + \frac{(sh_i - p_{sh,i,\omega})^2}{\sigma^2_{sh,i,\omega}} + \frac{(sh_j - p_{sh,j,\omega})^2}{\sigma^2_{sh,j,\omega}},$$

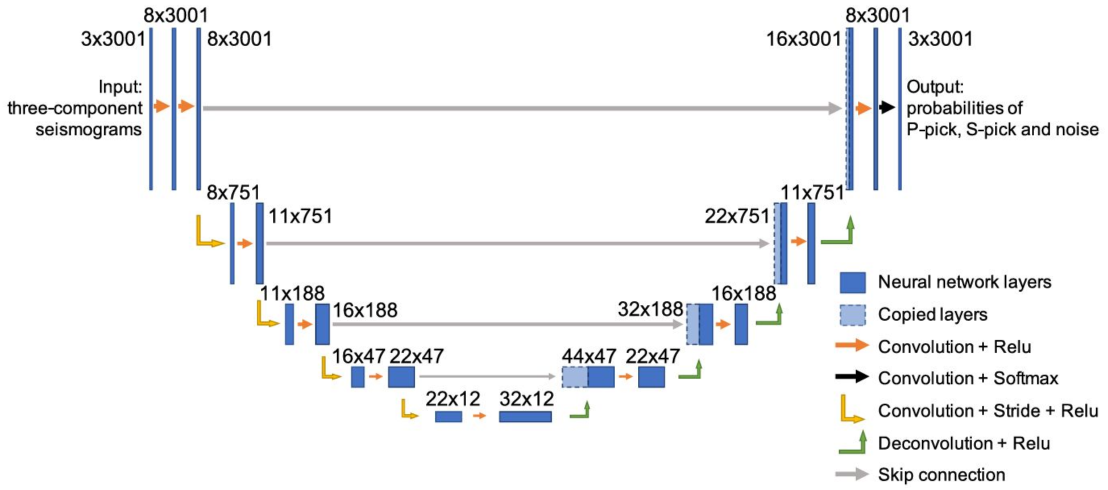


Brief History

Zhu et al., (2019)

PhaseNet: More picks, more data □ harder association challenge

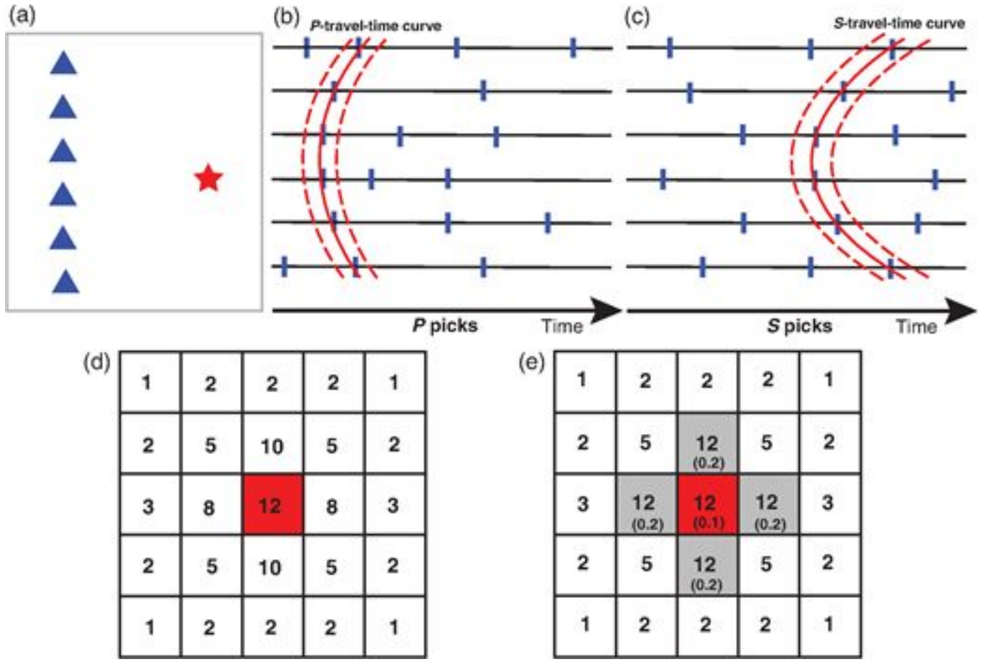


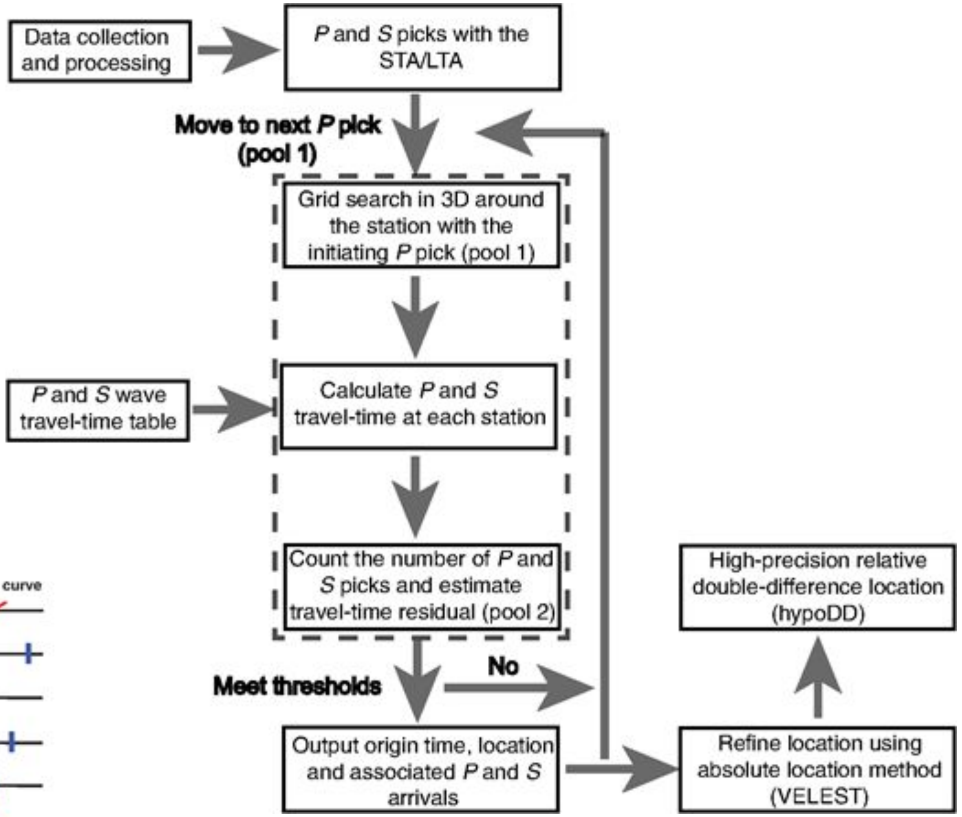


Brief History

Zhang et al.,
(2019)

REAL: back-projection based association and greedy association assignment





Brief History

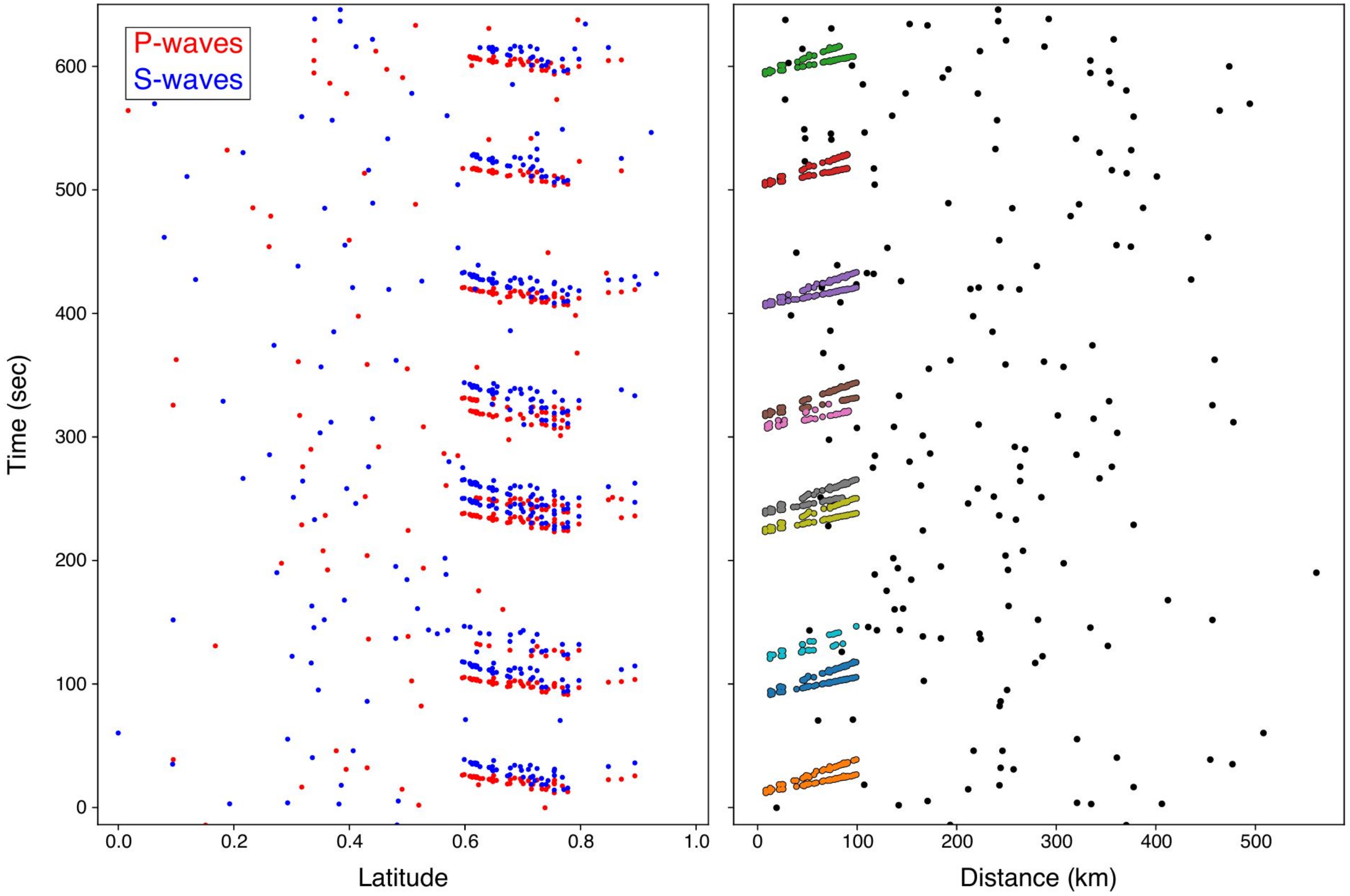
Ross et al.,
(2019)

PhaseLink: RNN based association

- Train on synthetic examples and learn solution

Input

Targe
t

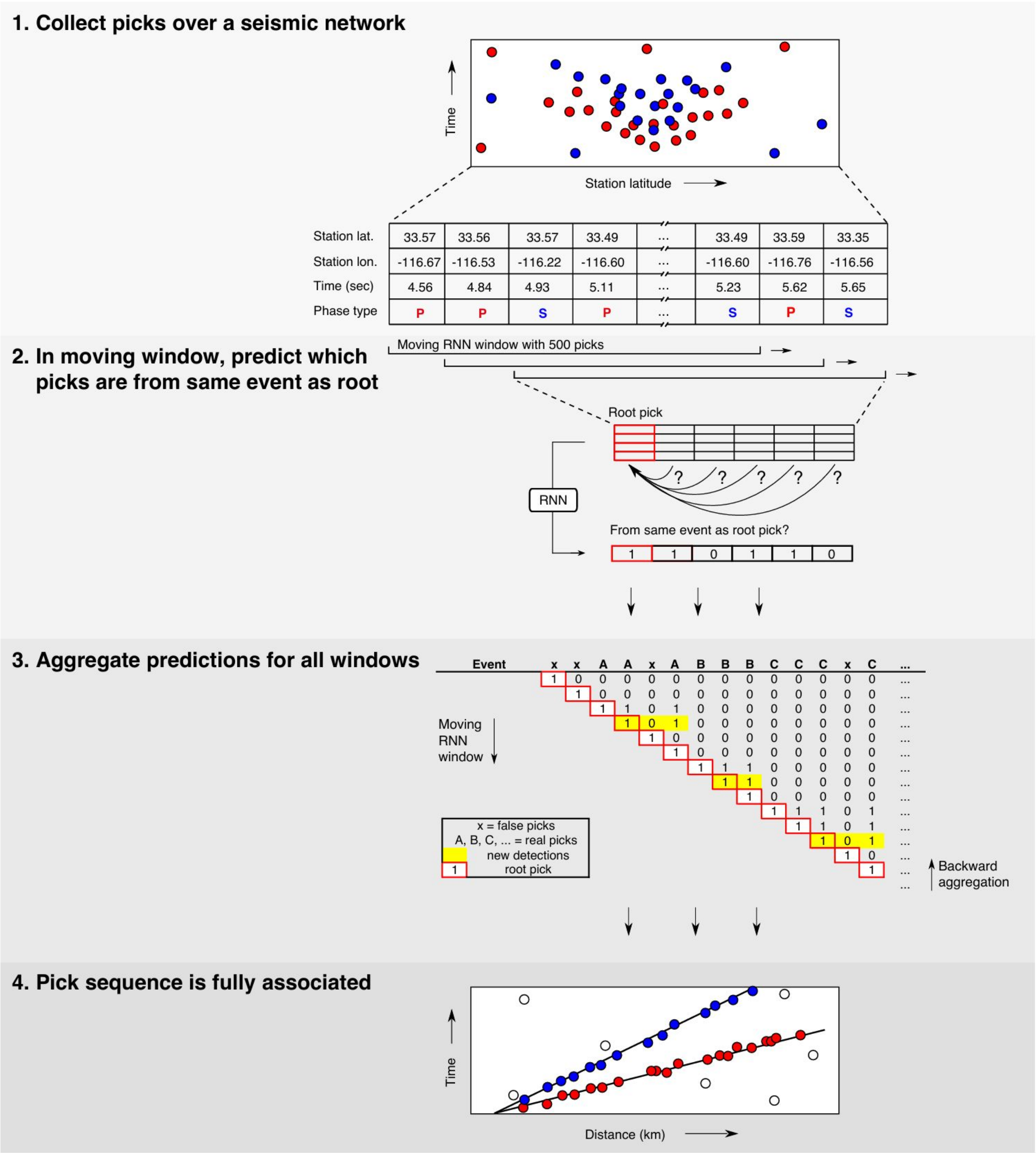


Brief History

Ross et al.,
(2019)

PhaseLink: RNN based association

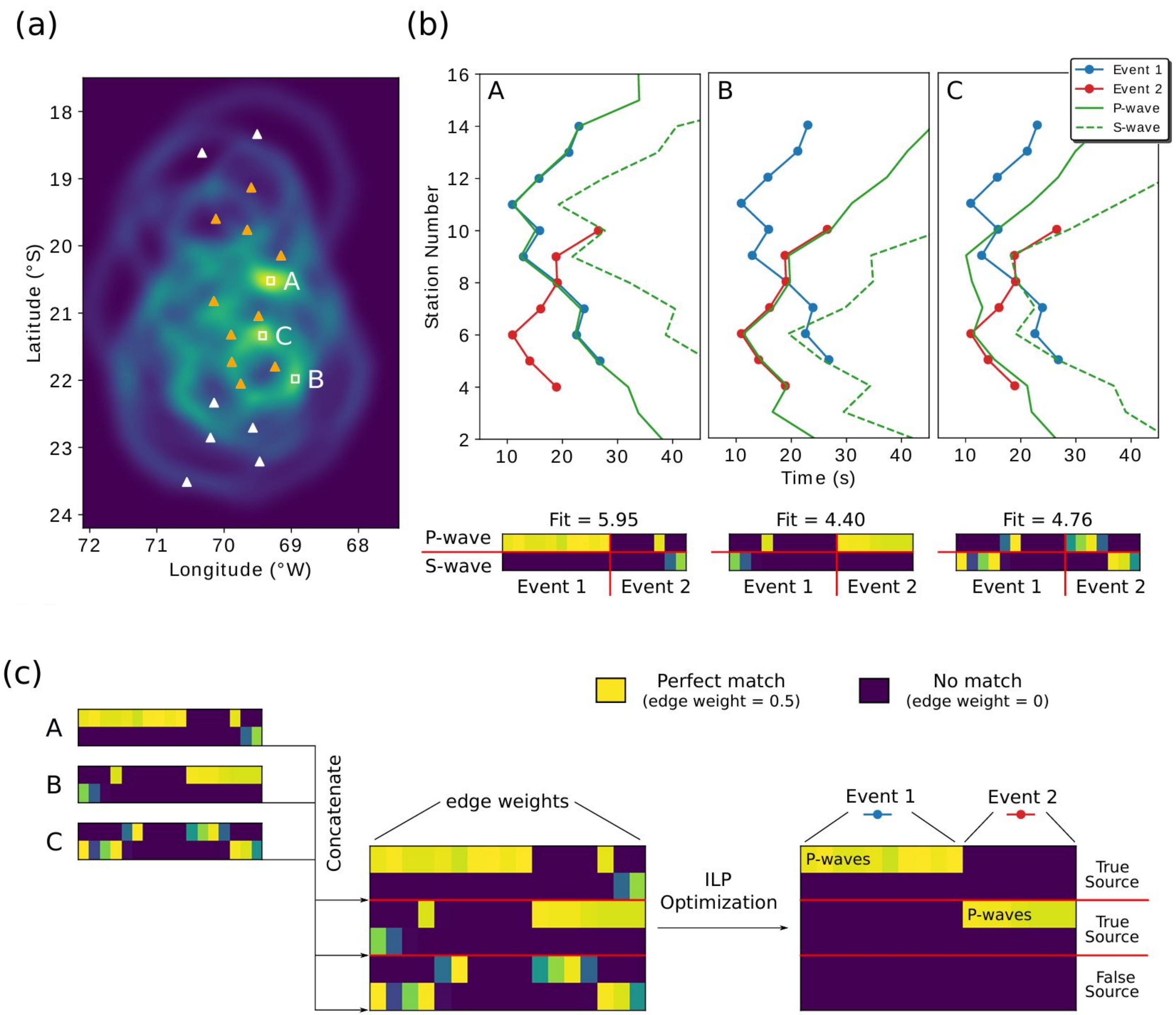
- Train on synthetic examples and learn solution



Brief History

McBrearty et al., (2019)

Use **Backprojection** and **Integer Linear Optimization**

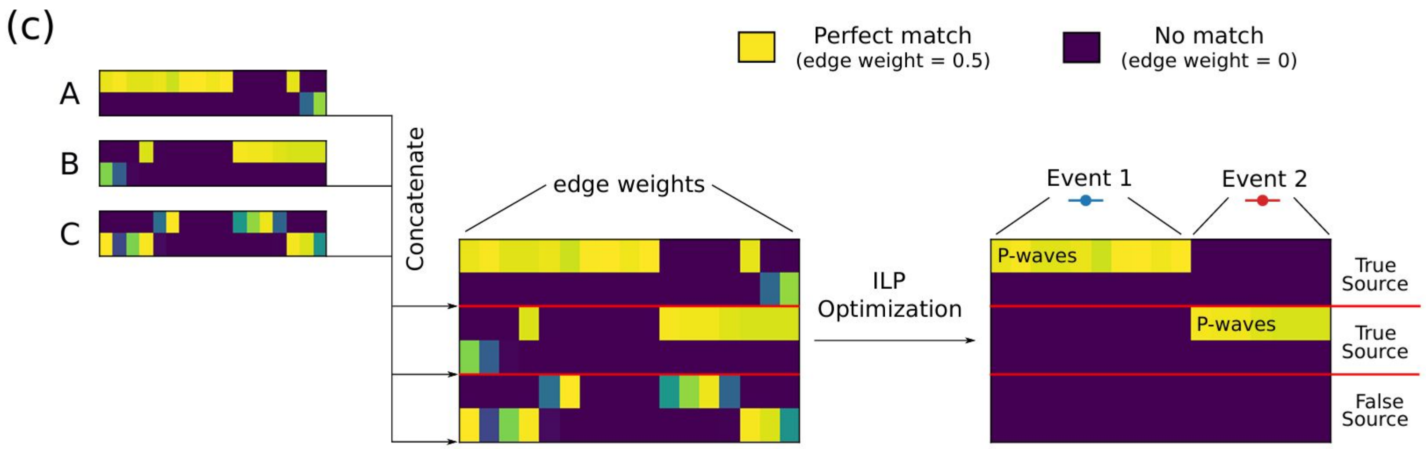


Brief History

McBrearty et al., (2019)

Use **Backprojection** and **Integer Linear Optimization**

$$\max_{x} c^T x$$
$$\text{s.t. } Ax \leq b$$
$$x \in \{0,1\}$$
$$\{A, b, c\} \xleftarrow{Algs.\ S1-S3} \{\mathcal{D}, \mathcal{S}\}\ ,$$



- Explicit optimization (more robust than Hungarian algorithm)
- Still must determine sources/scaling issues

Brief History

McBrearty et al., (2019)

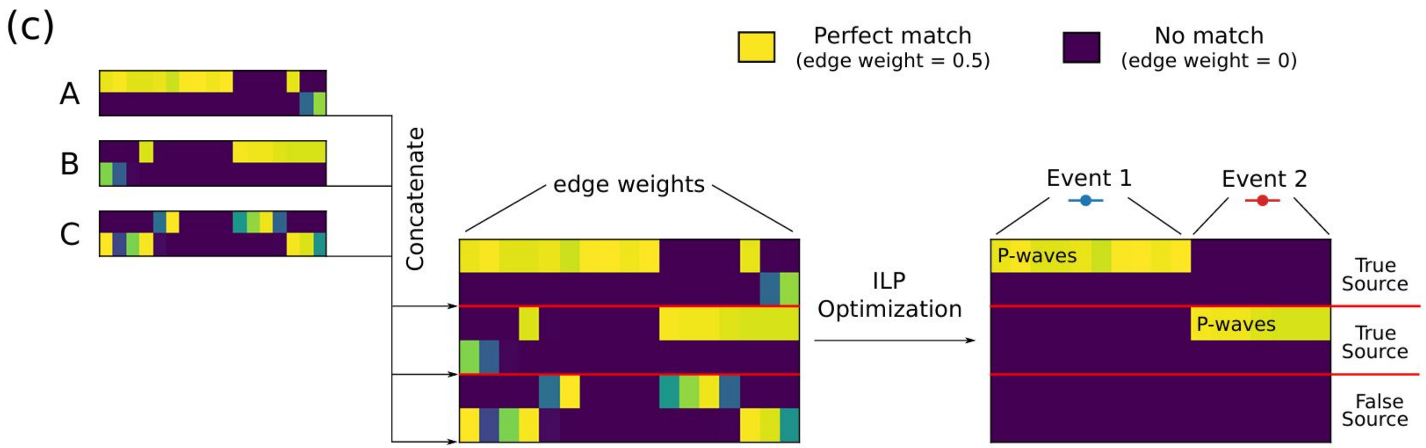
Use **Backprojection** and **Integer Linear Optimization**

$$\max_{x} c^T x$$

```
# Solve ILP

x = cp.Variable(n_phases*n_srcs*n_arvs + n_srcs, integer = True)

prob = cp.Problem(cp.Minimize(c.T@x), constraints = [A@x <= b.reshape(-1), 0 <= x, x <= 1])

prob.solve()

assert prob.status == 'optimal', 'competitive assignment solution is not optimal'
```

Python: Cvxpy package

- Explicit optimization (more robust than Hungarian algorithm)
- Still must determine sources/scaling issues



Brief History

Zhu et al., 2022

GaMMA: Bayesian Gaussian Mixture model association (**unsupervised clustering)**

- ”Iteratively” solve association and event location (i.e., “soft assignment”)

E-step:

$$\gamma_{ik} = \frac{\phi_k \mathcal{N}(\mathbf{x}_i | \mu_k, \mathbf{\Sigma}_k)}{\sum_{k=1}^{K} \phi_k \mathcal{N}(\mathbf{x}_i | \mu_k, \mathbf{\Sigma}_k)}$$

M-step:

1. Effective number of picks assigned to the k-th earthquake:

$$N_k = \sum_{i=1}^{N} \gamma_{ik}$$

$$\phi_k = \frac{N_k}{N}$$

2. Earthquake location, origin time, and magnitude of the k-th earthquake:

$$\underset{(x_k, y_k, z_k, t_k)}{\text{minimize}}\ l(x_k, y_k, z_k, t_k) = \sum_{i=1}^{N} \gamma_{ik} \mathcal{L}\left(t_i, \hat{t}_{ik}(x_k, y_k, z_k, t_k)\right)$$

$$m_k = \frac{1}{N_k} \sum_{i=1}^{N} \gamma_{ik} \mathcal{F}'_a(a_i, d_{ik})$$

3. Theoretical travel time, amplitude, and statistics of residuals:

$$\mu_k = \begin{bmatrix} \hat{t}_{ik} \\ \hat{a}_{ik} \end{bmatrix} = \begin{bmatrix} \mathcal{F}_t(x_k, y_k, z_k, t_k) \\ \mathcal{F}_a(m_k, d_{ik}) \end{bmatrix}$$

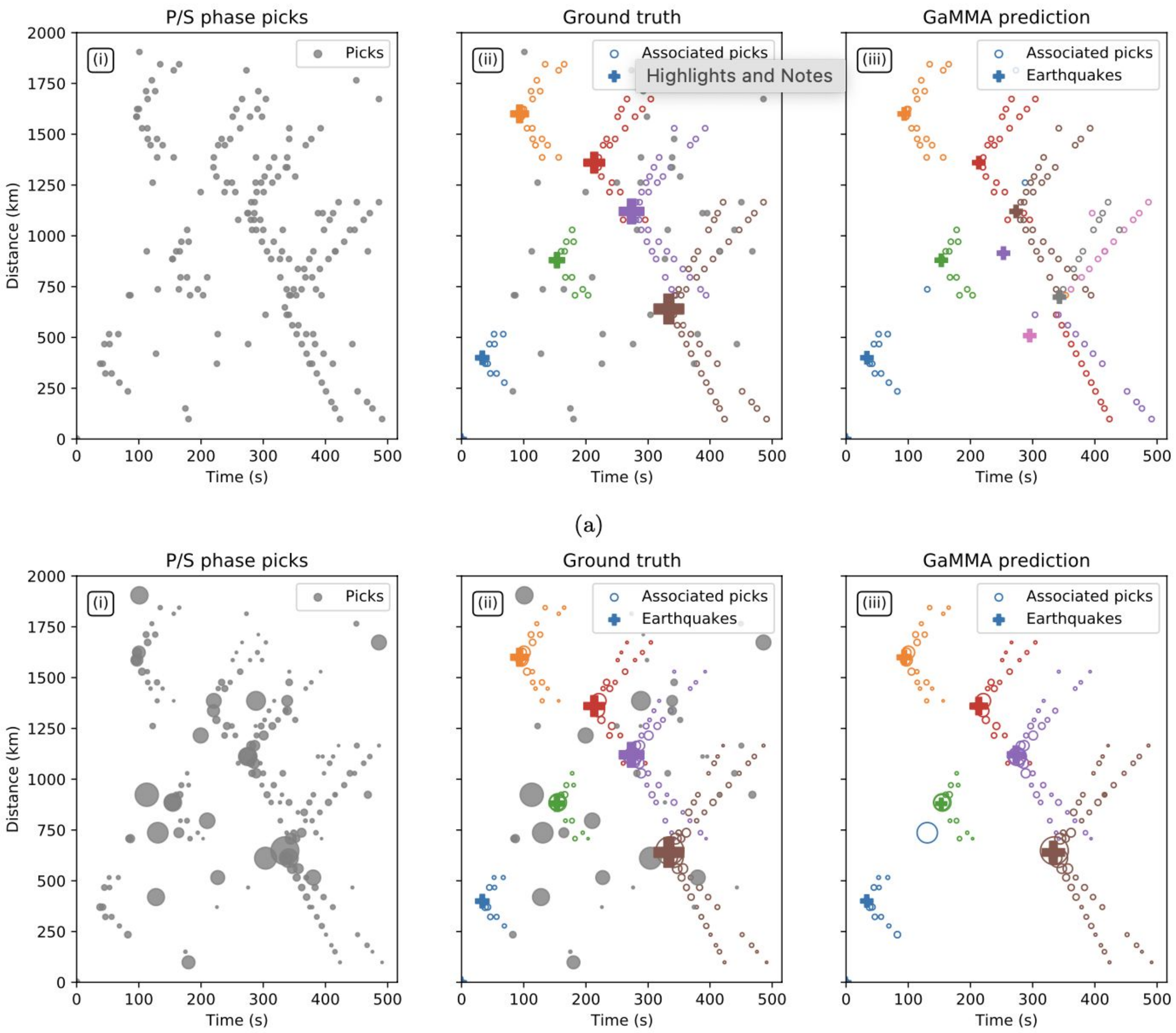
$$\mathbf{\Lambda}_k^{-1} = \frac{1}{N_k} \sum_{i=1}^{N} \gamma_{ik} (\mathbf{x}_i - \mu_k)(\mathbf{x}_i - \mu_k)^T$$

Brief History

Zhu et al.,
2022

GaMMA: Bayesian Gaussian Mixture model association (**unsupervised clustering)**

- ”Iteratively” solve association and event location (i.e., “soft assignment”)

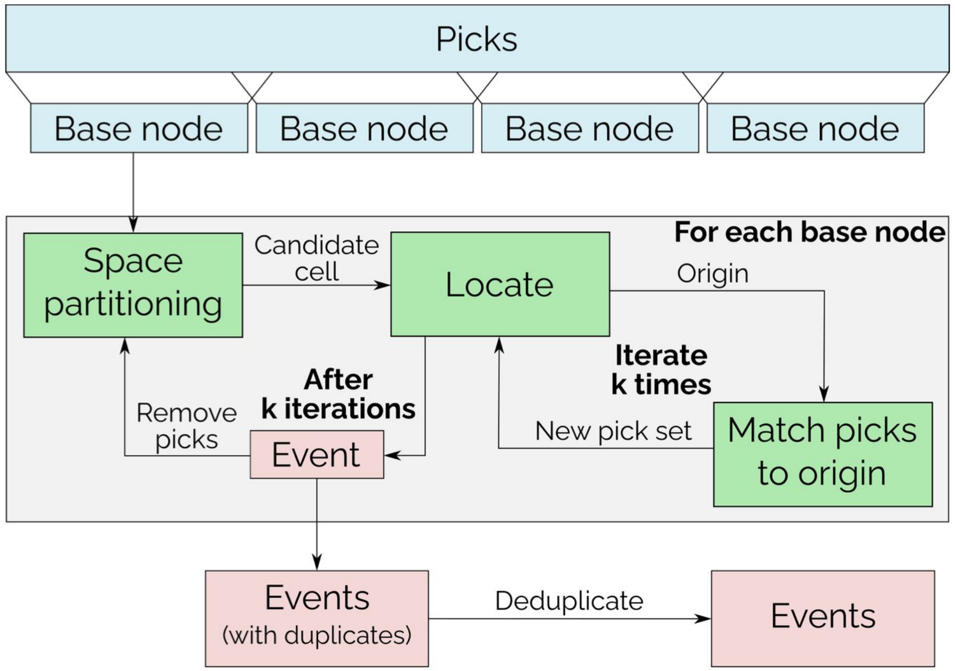


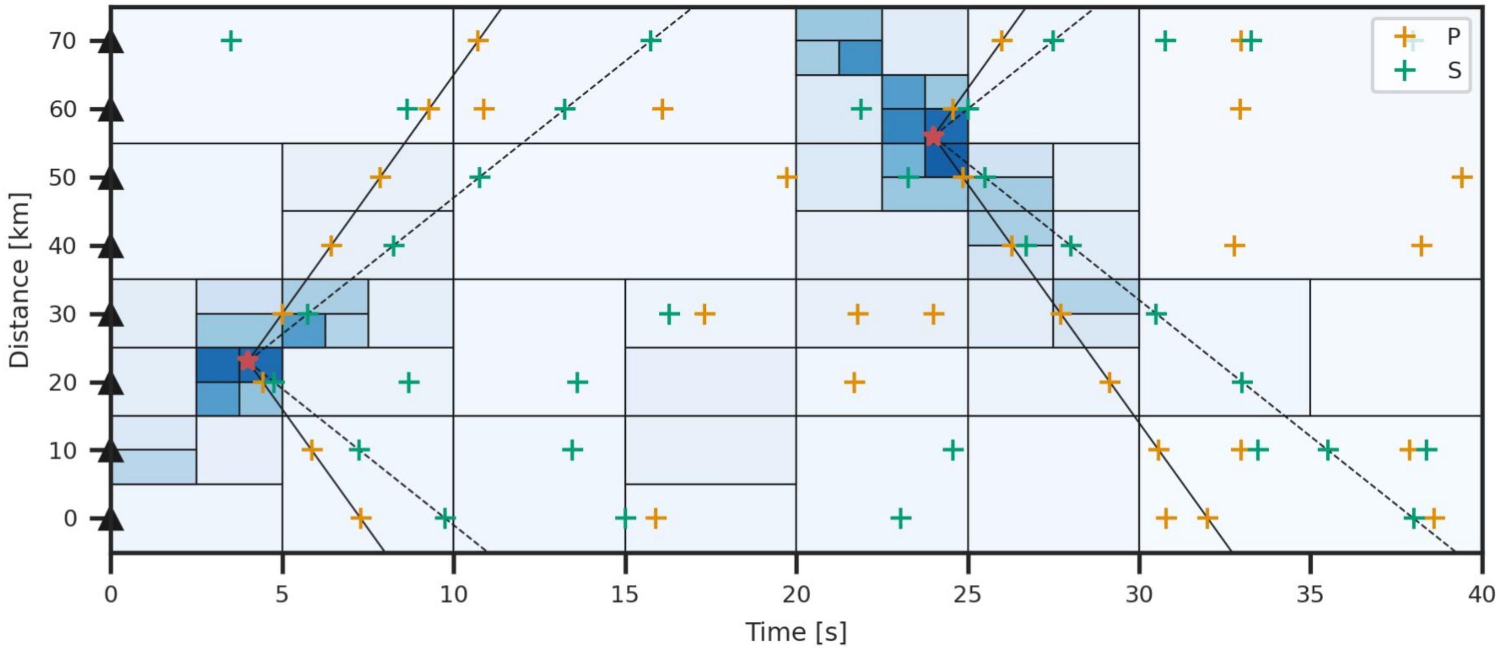
Brief History

Munchmeyer, 2024

PyOcto: Efficient back-projection search with Oct-Tree

- Check optimal sources first; assign picks; iterate

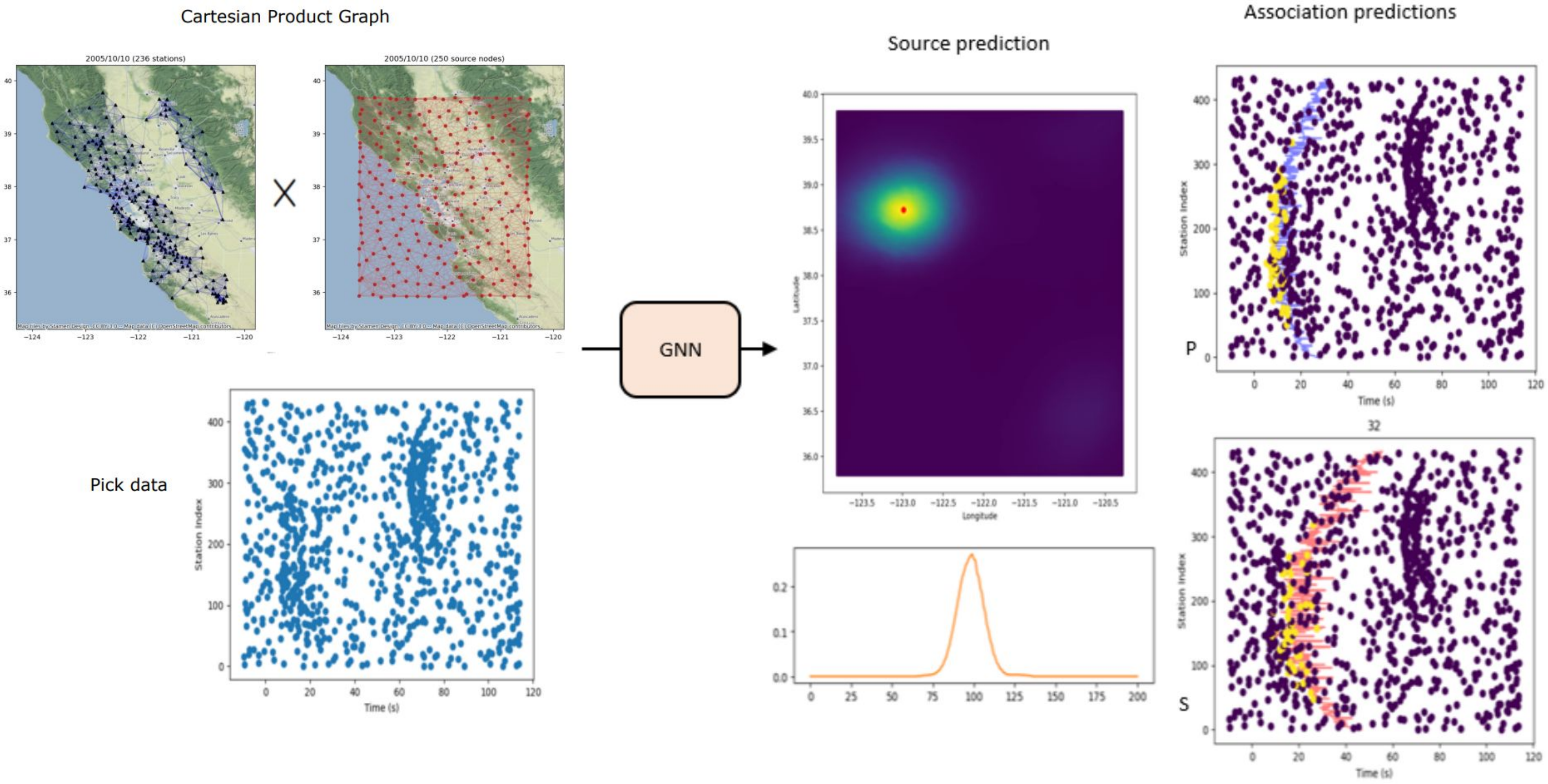




Brief History

McBrearty and Beroza, 2023

GENIE: GNN based source location and phase association

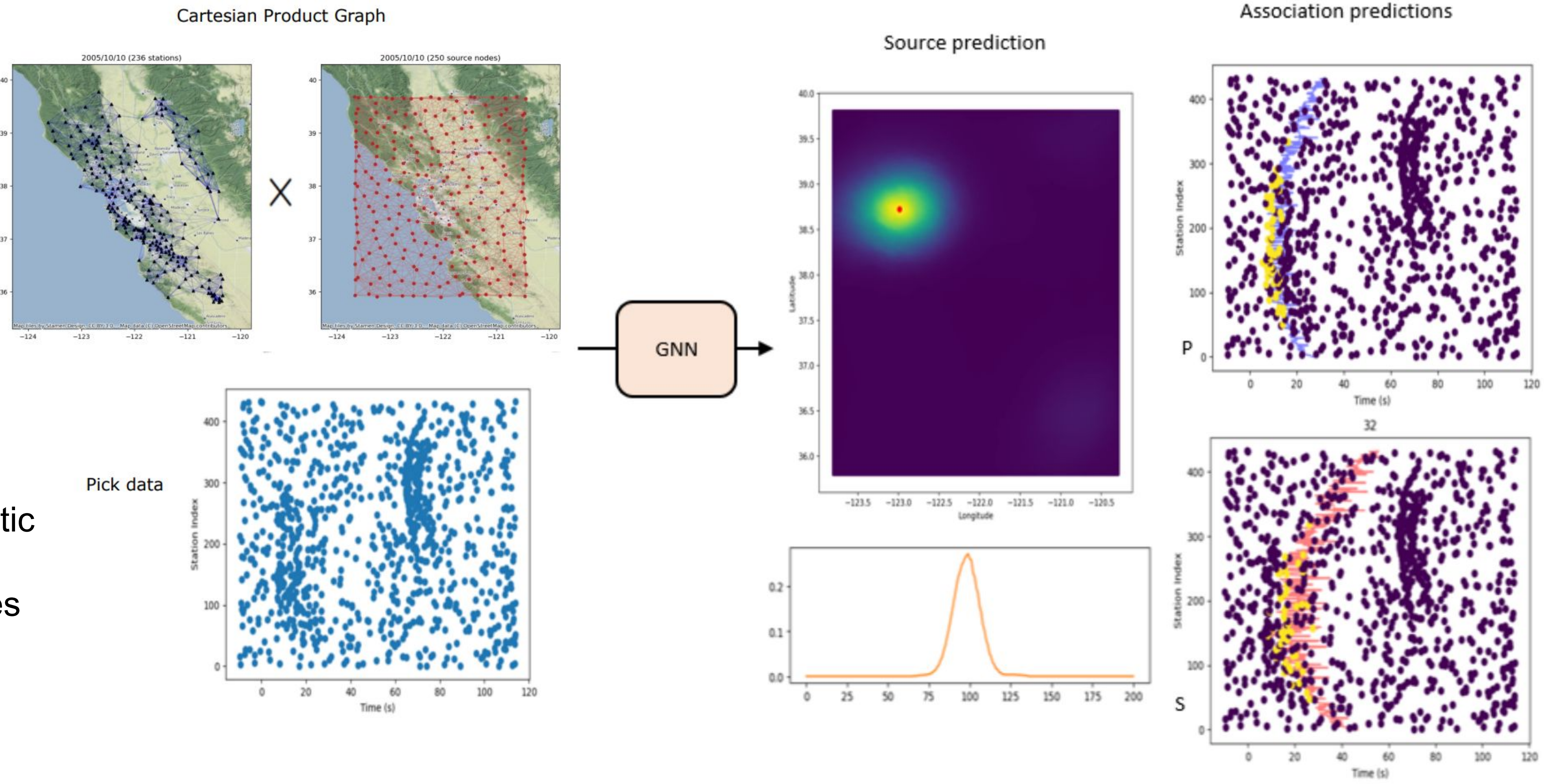


Brief History

McBrearty and
Beroza, 2023

GENIE: GNN based source location and phase association

- Trained on synthetic data; simultaneous prediction of sources and associations

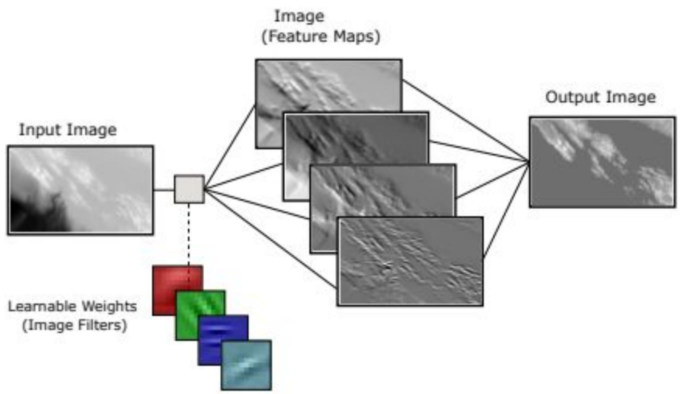


Association predictions are conditioned on source predictions

Why GNNs?

Convolutional Neural Networks

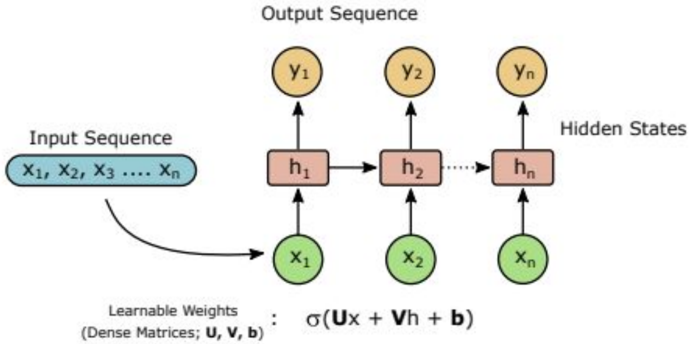
Effective for Euclidean data
(e.g., time series, images)



Relies on the distribution and type of spatial features (e.g., edges, shapes, gradients).

Recurrent Neural Networks

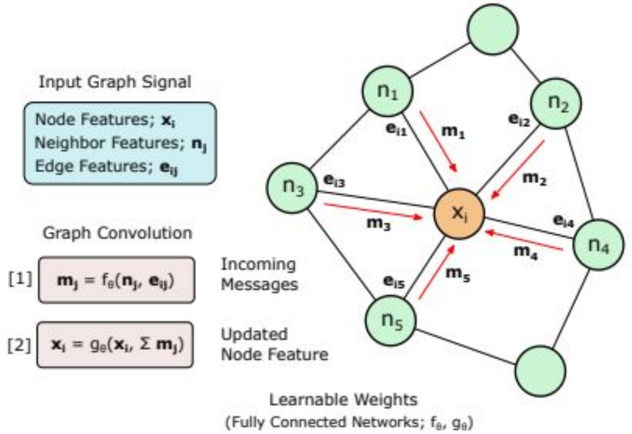
Effective for Euclidean data
(e.g., time series, text)



Relies on the timing/sequencing and strength of temporal signals.

Graph Neural Networks

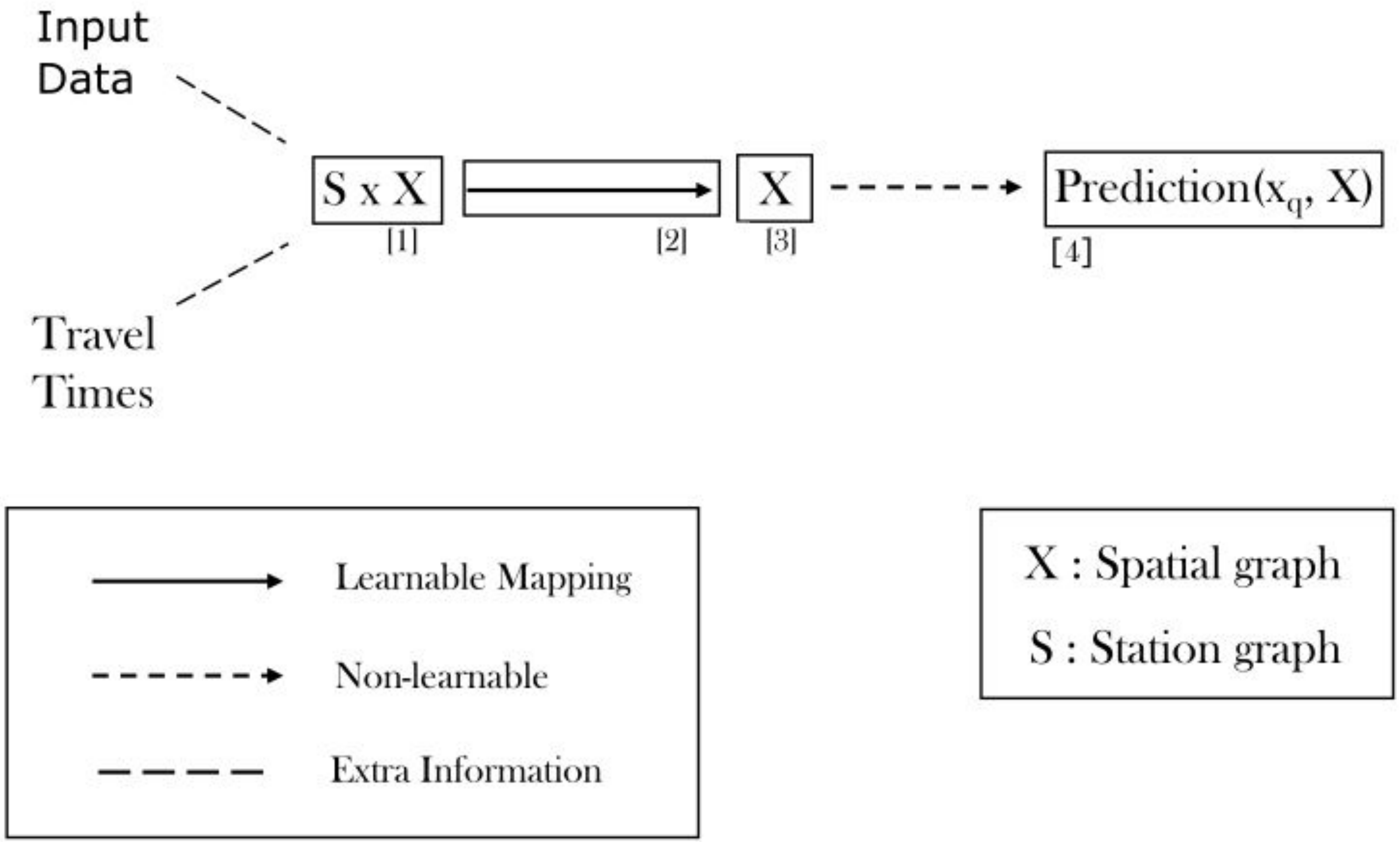
Effective for non-Euclidean data
(e.g., graphs, sensor arrays)



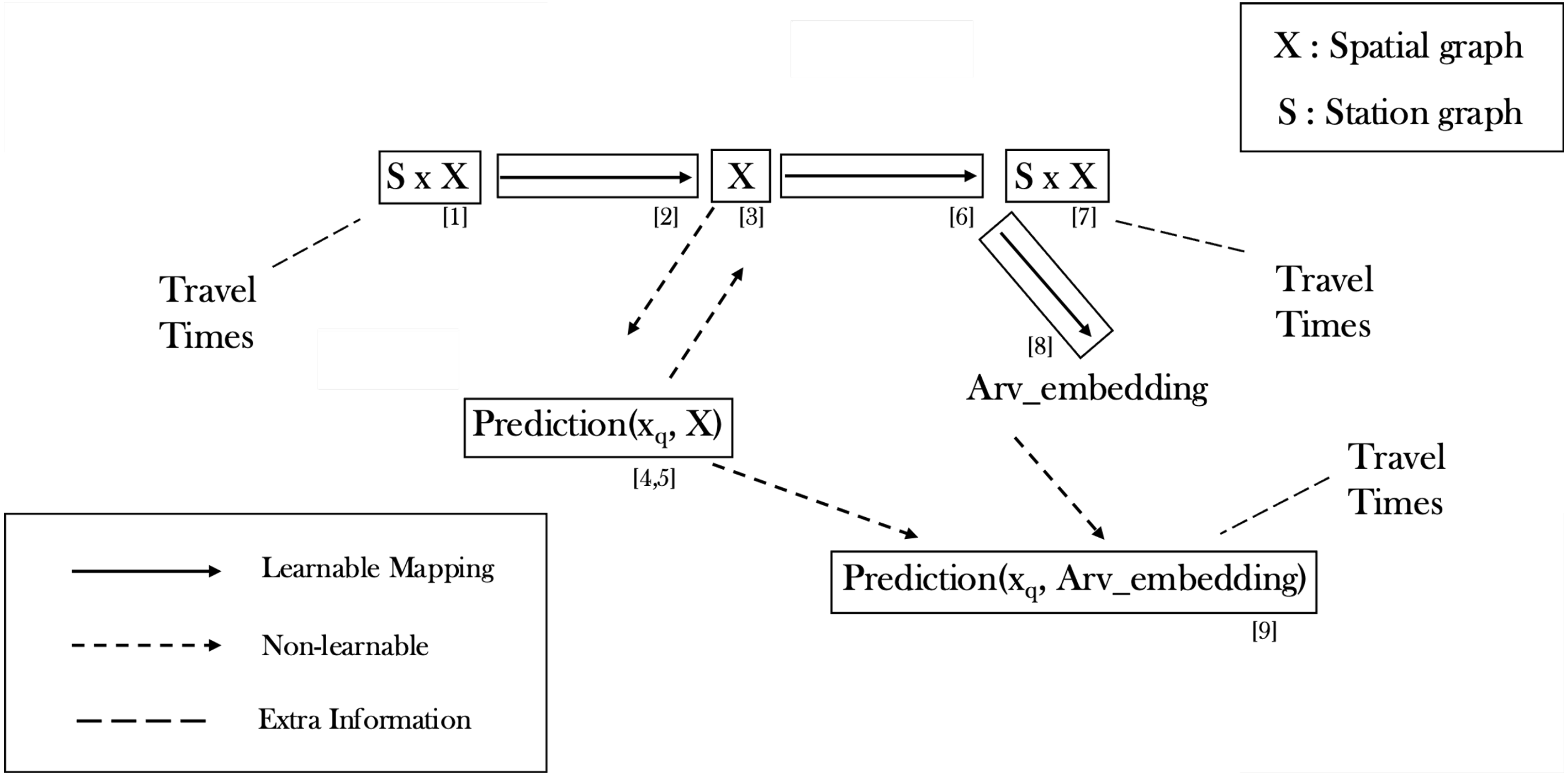
Relies on local information passing between nodes.

Relaxed conditions on the spatial regularity of data.

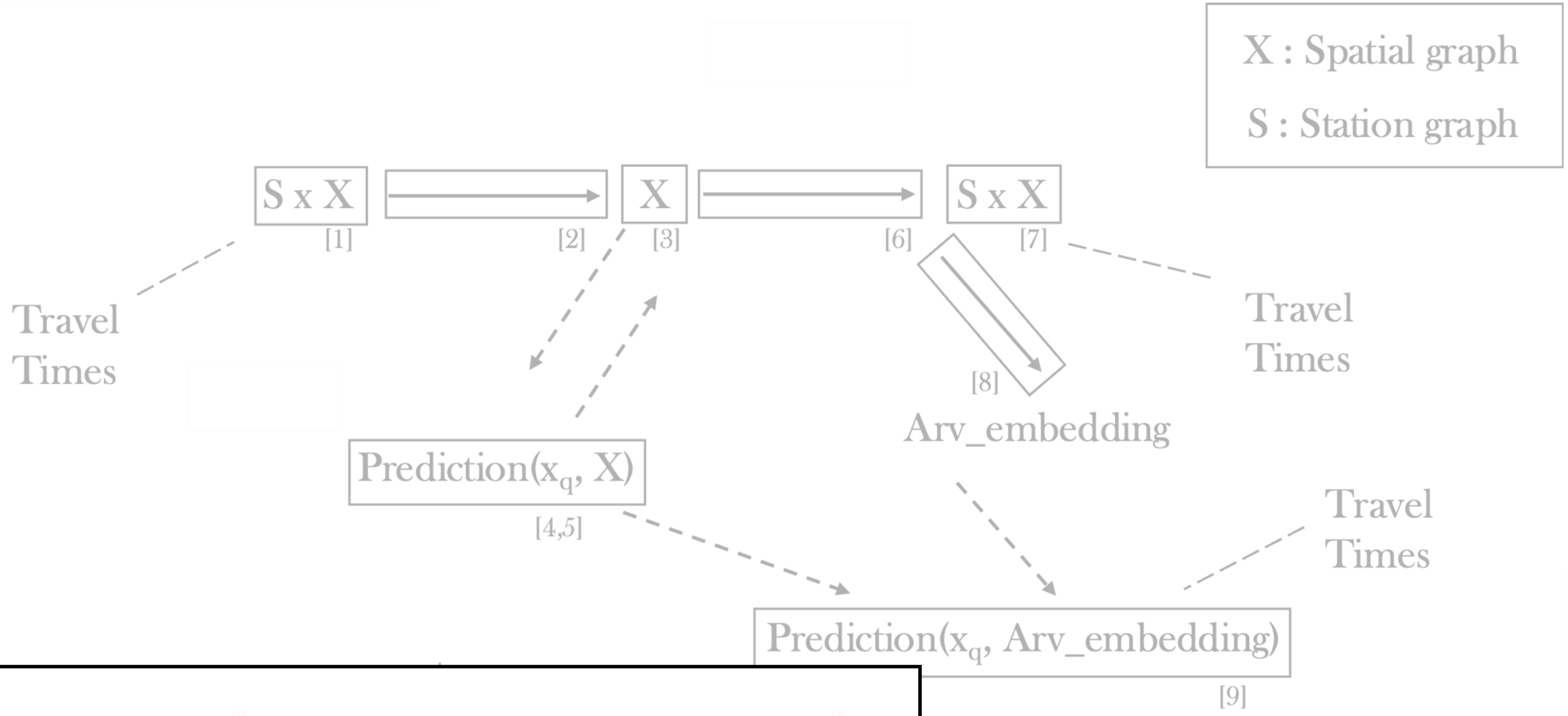
GNN: Architecture



GENIE: Architecture



GENIE: Architecture



Input feature:

$$\boldsymbol{h}_k^{\mathcal{X}}(\boldsymbol{s}_i, \boldsymbol{x}) = \exp\left(-\frac{(t_0 + T_k(\boldsymbol{s}_i, \boldsymbol{x}) - \tau_i^k)^2}{2\sigma_t^2}\right)$$

pre-stack BP metric

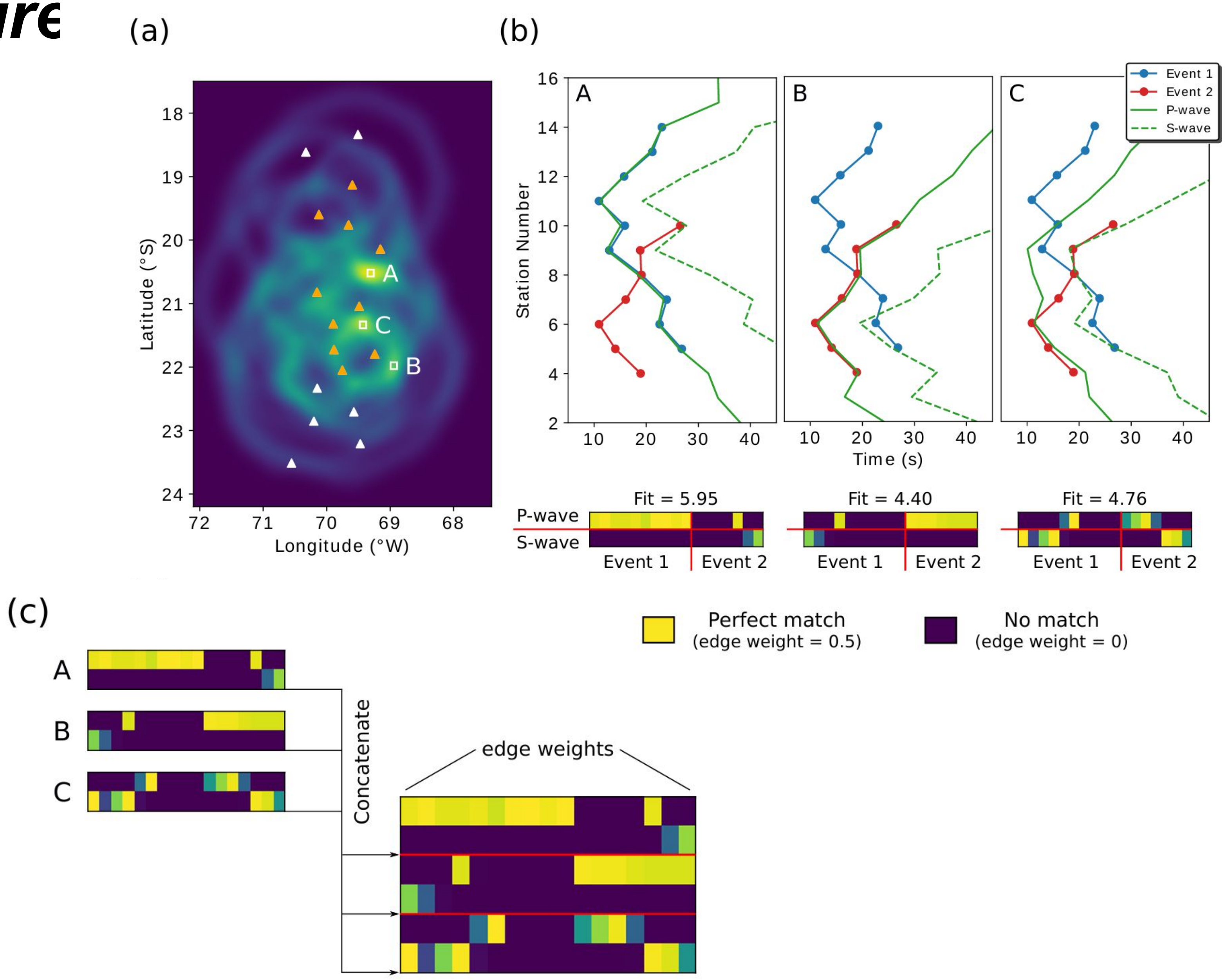
GENIE: Architecture

Strengths:

- Input feature is the **misfit** between observed and theoretical arrivals

- Doesn't have to "learn" velocity model (unlike PhaseLink)

- "Knows" the relative position of stations, and weights them differently (unlike back-projection)



GENIE: Architecture

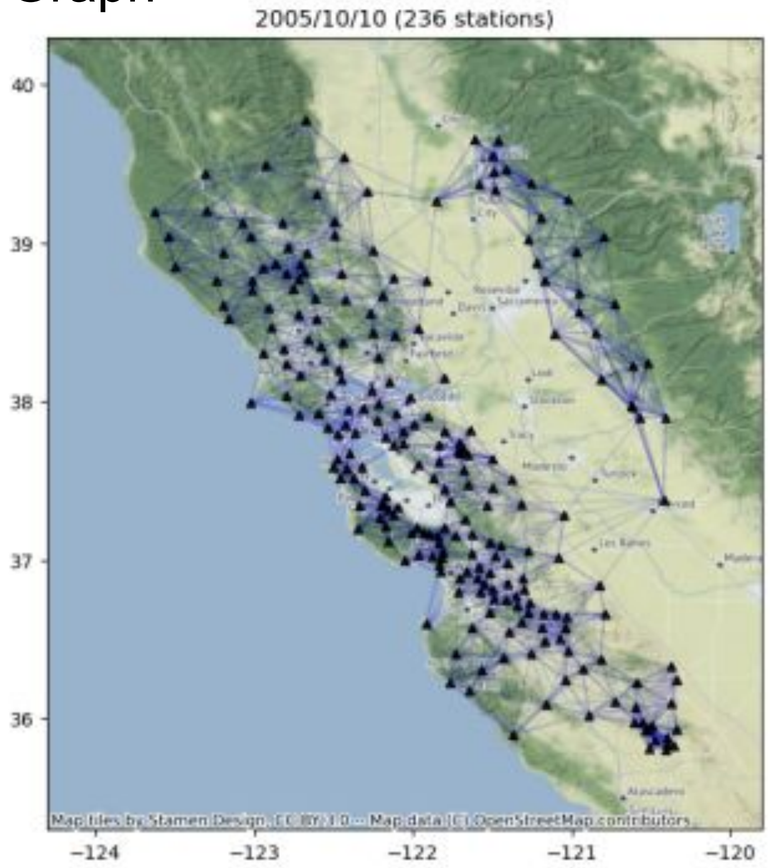
Strengths:

- Input feature is the **misfit** between observed and theoretical arrivals

- Doesn't have to ”learn” velocity model (unlike PhaseLink)

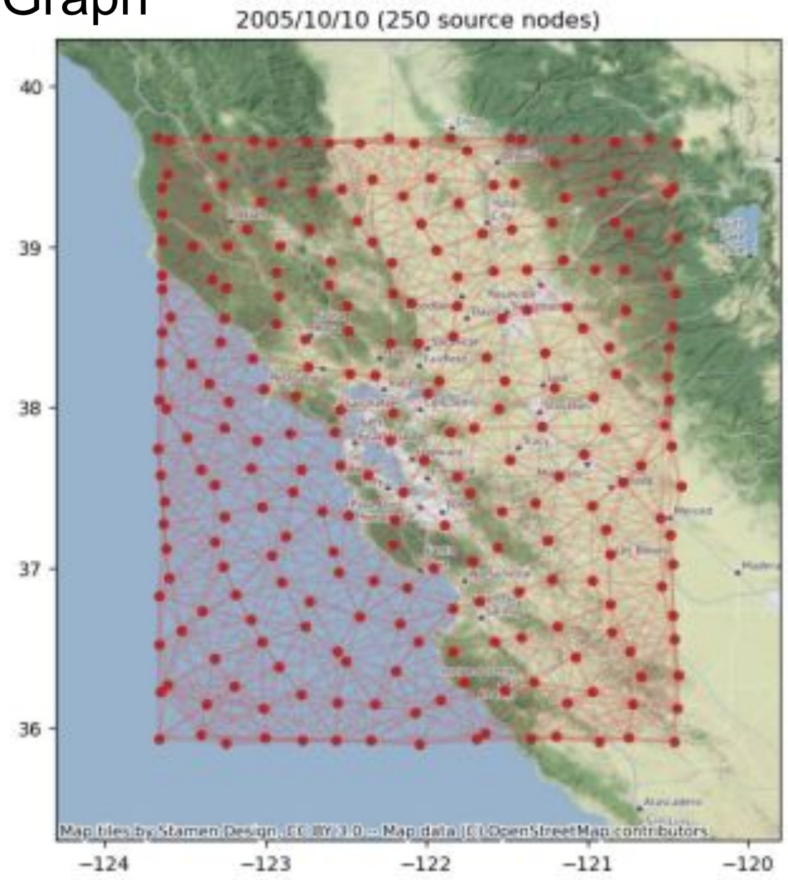
- “Knows” the relative position of stations, and weights them differently (unlike back-projection)

Station Graph



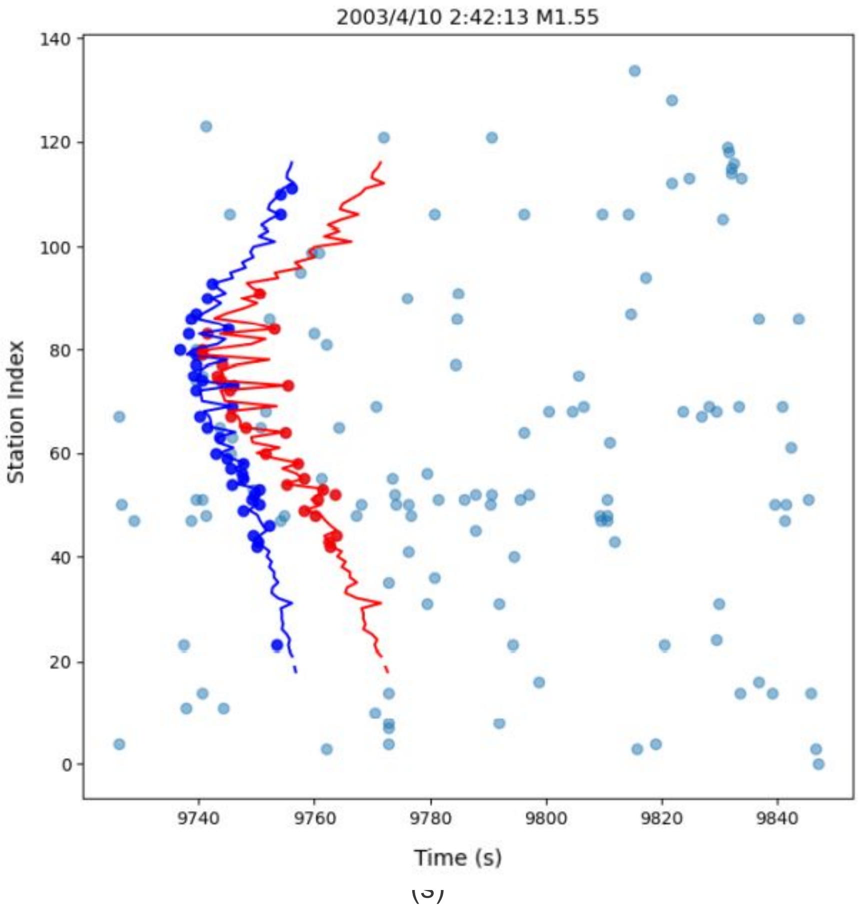
8-nearest-neighbors

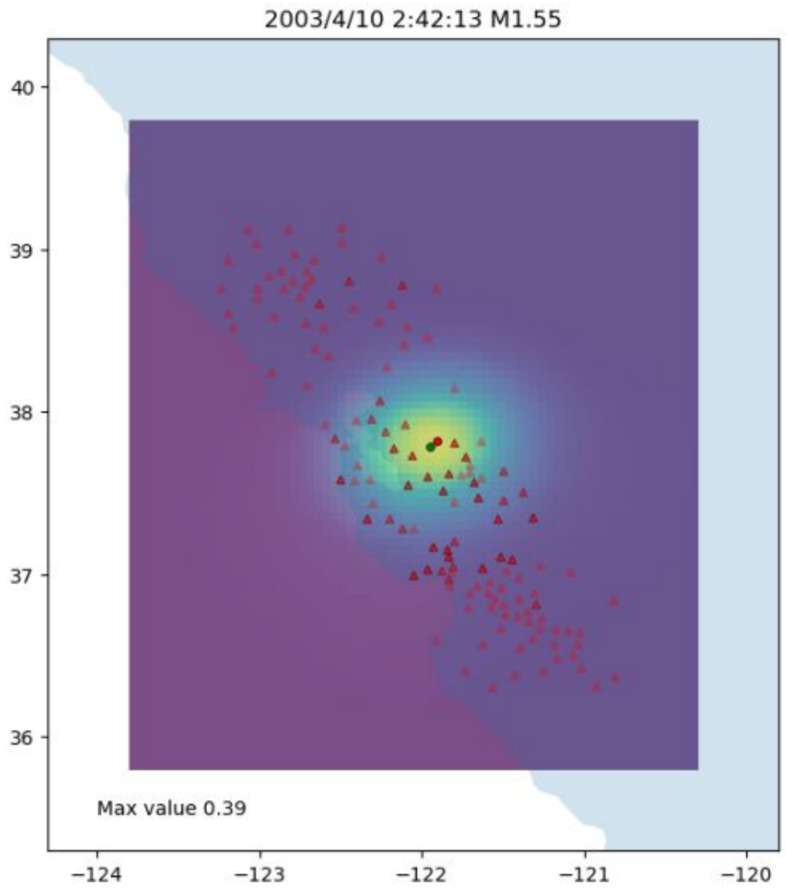
Source Graph

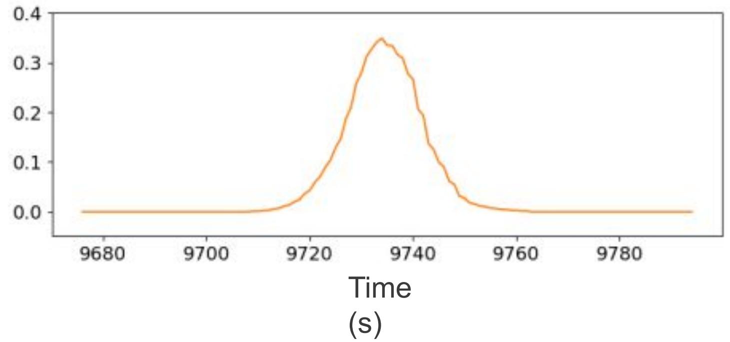


15-nearest-neighbors

Example Detections

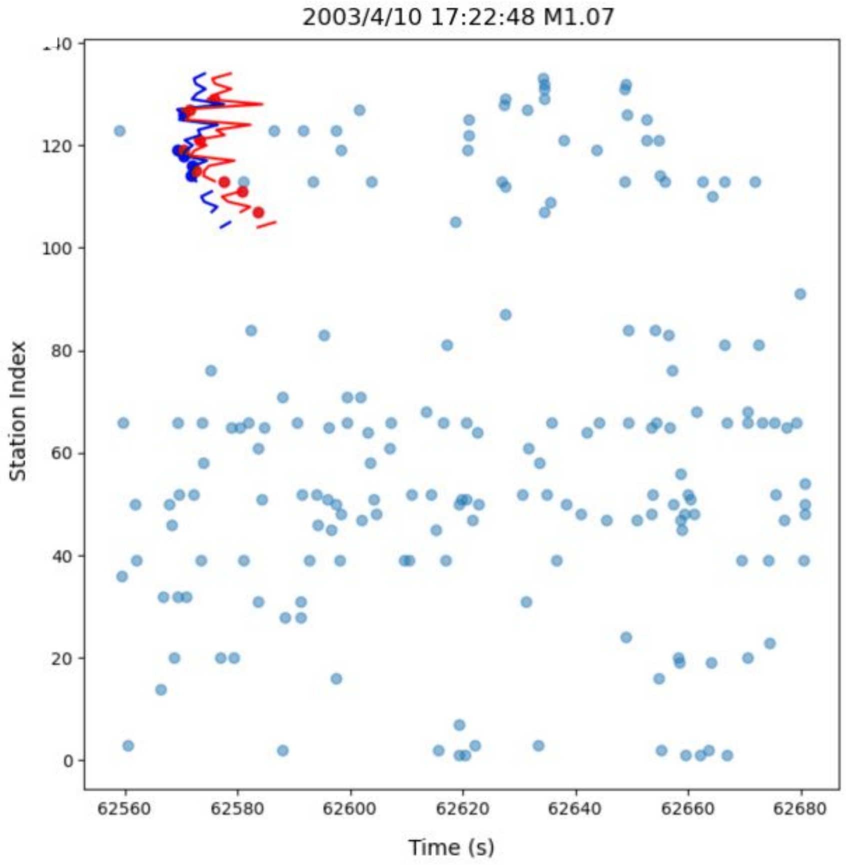




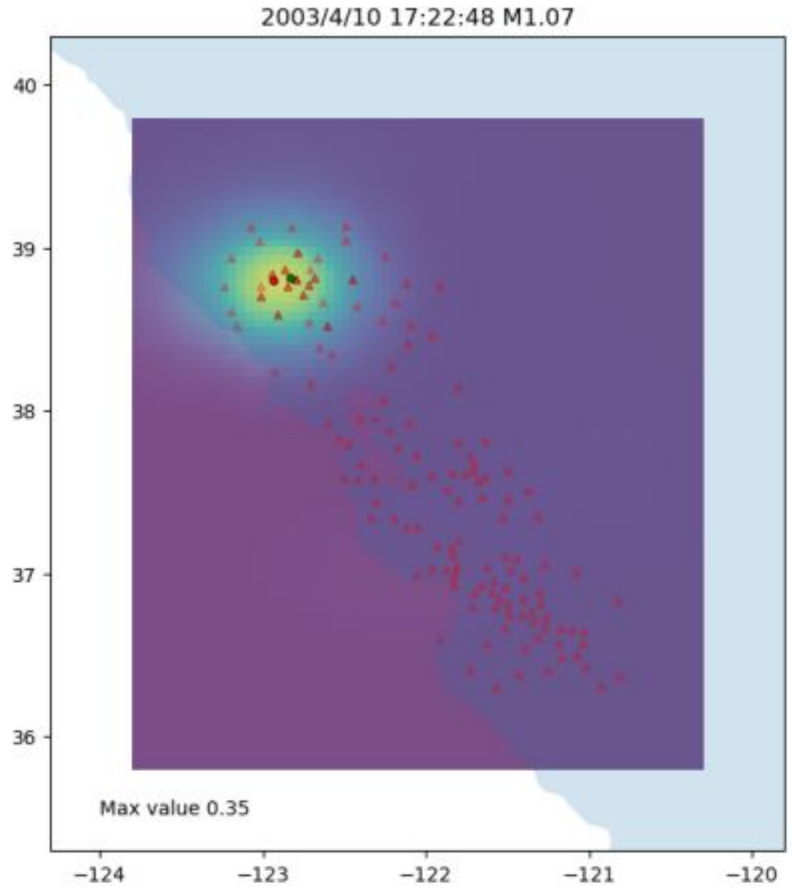


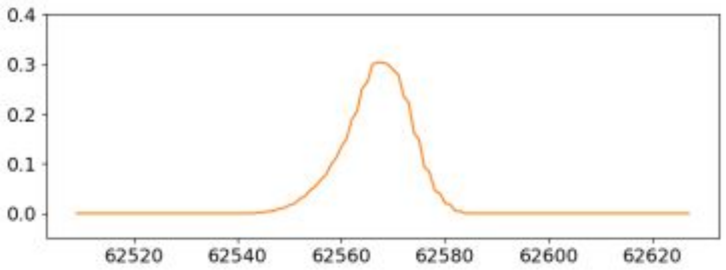
Spatio-temporally localized known **M1.5** earthquake on **Calaveras Fault**, and obtained P and S wave associations

Example Detections



(s)





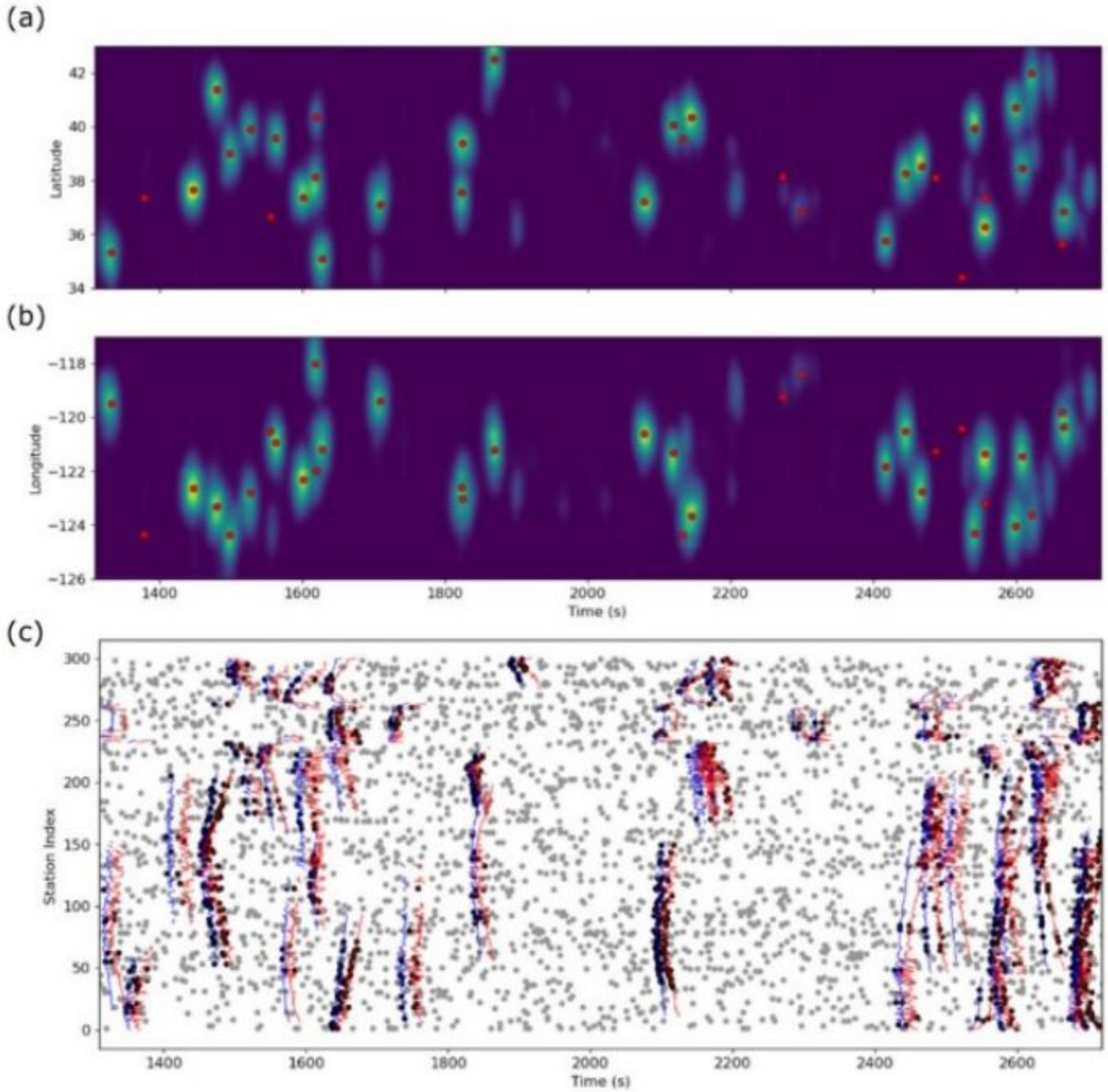
Time
(s)

Spatio-temporally localized known **M1.1** earthquake at **Geysers**, and obtained P and S wave associations

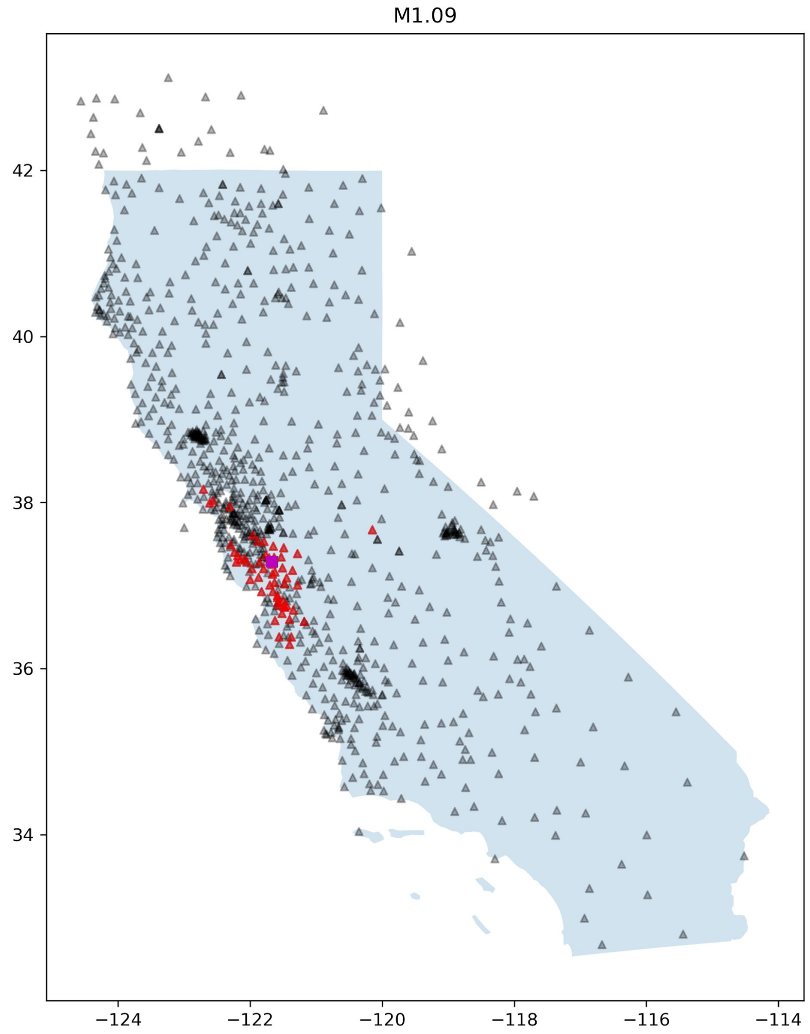
Example Detections

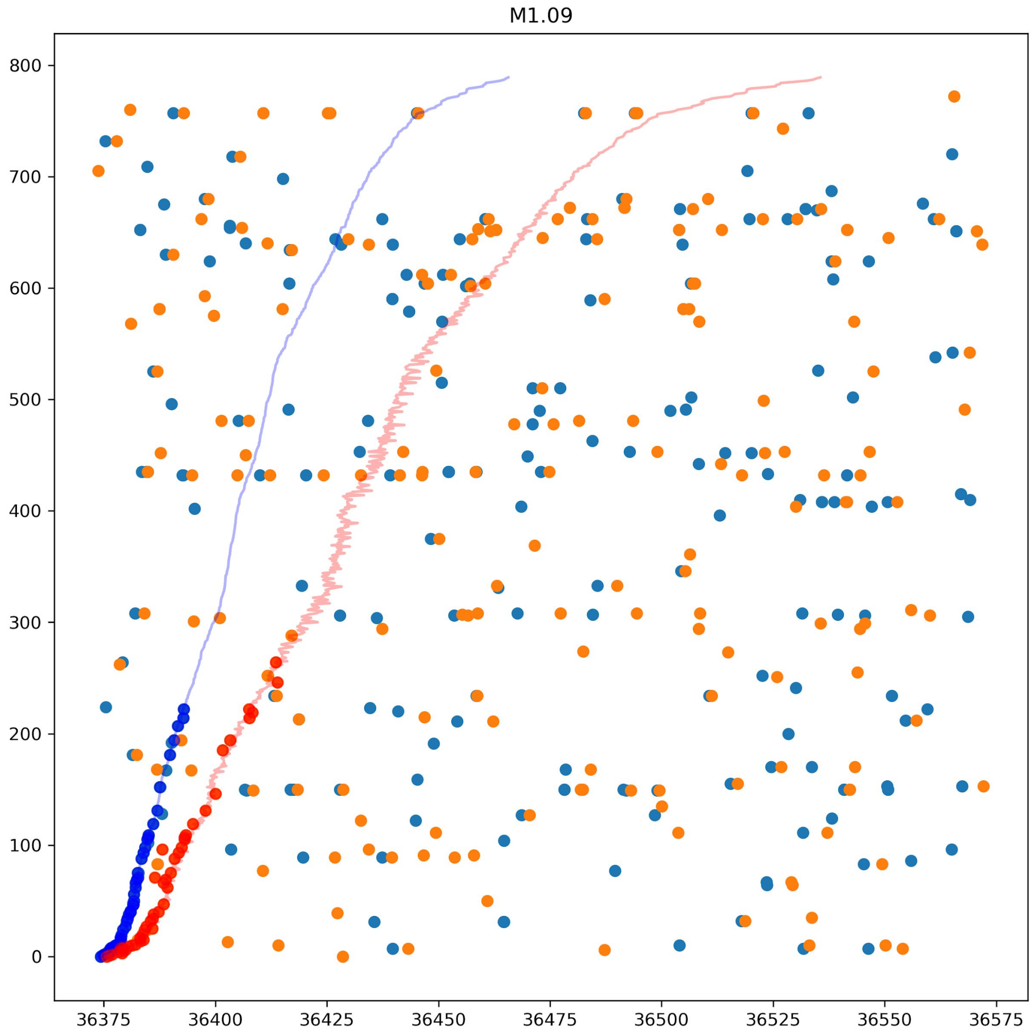
• Continuous space-time output

• Can handle even closely overlapping events and many false/noisy picks



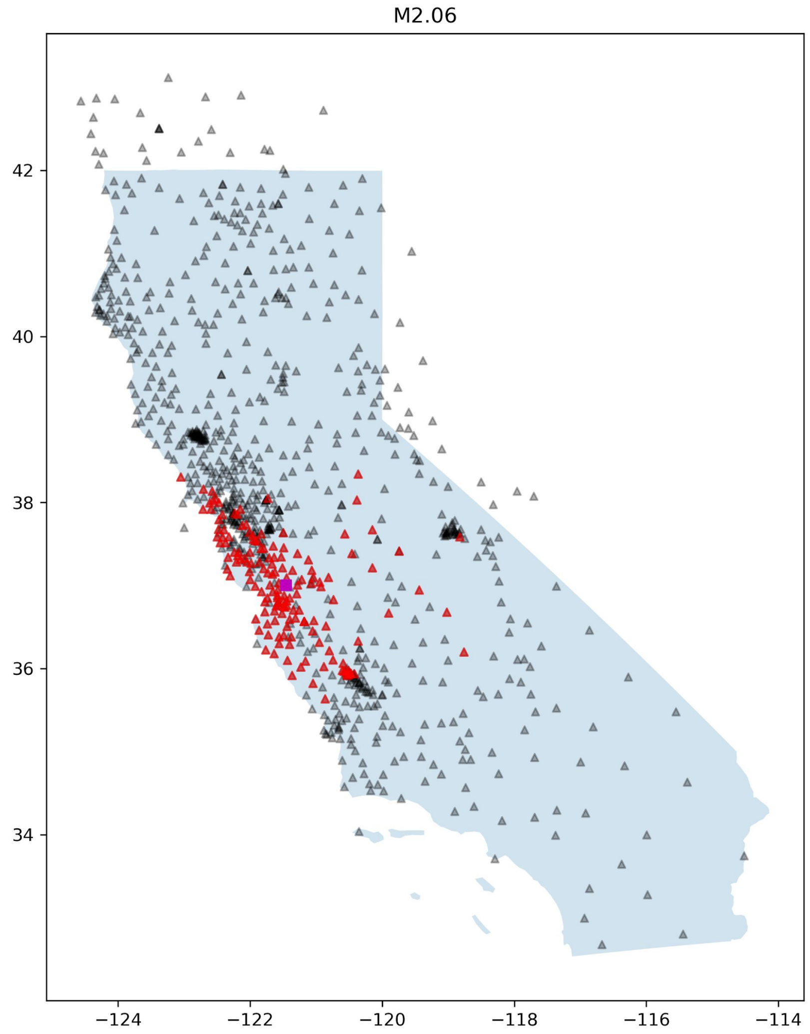
Example Detections

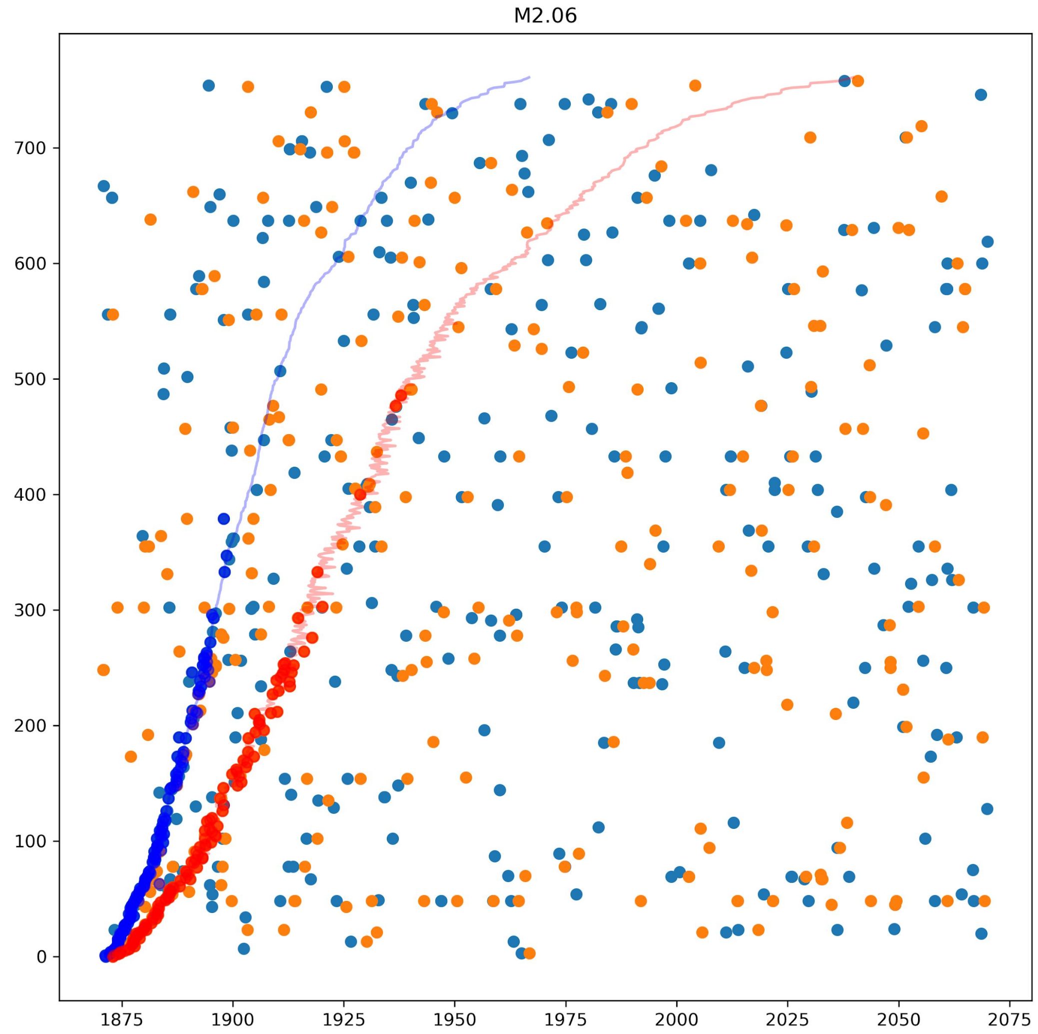




M1, Bay
Area

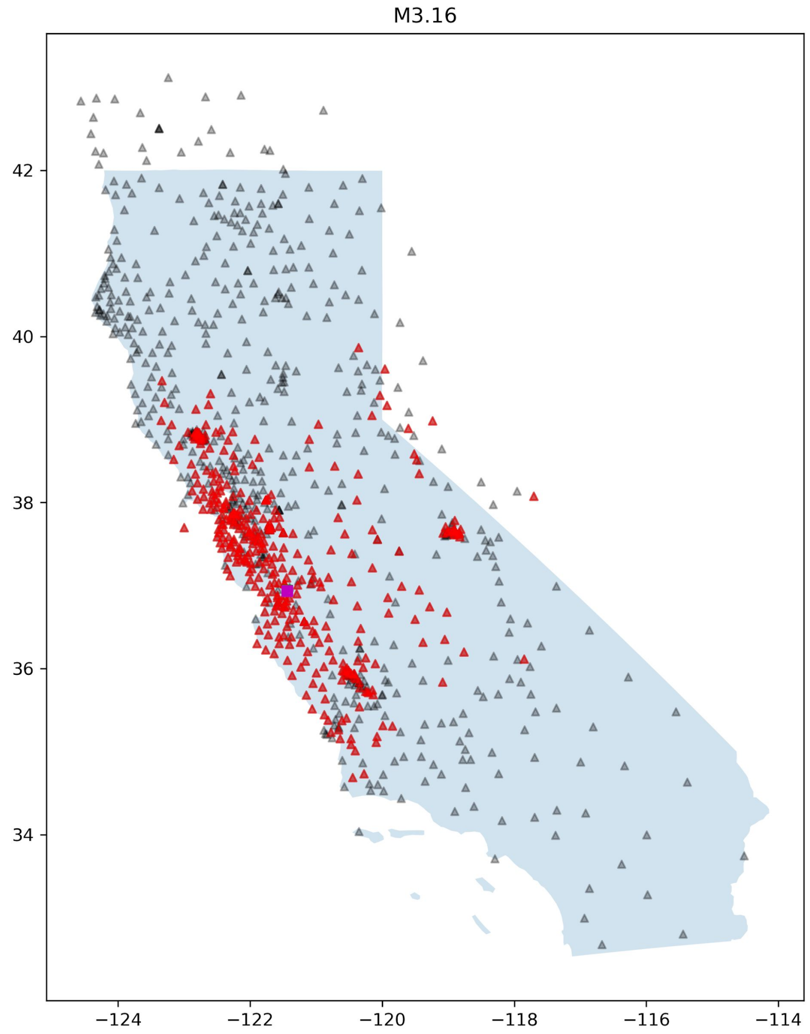
Example Detections

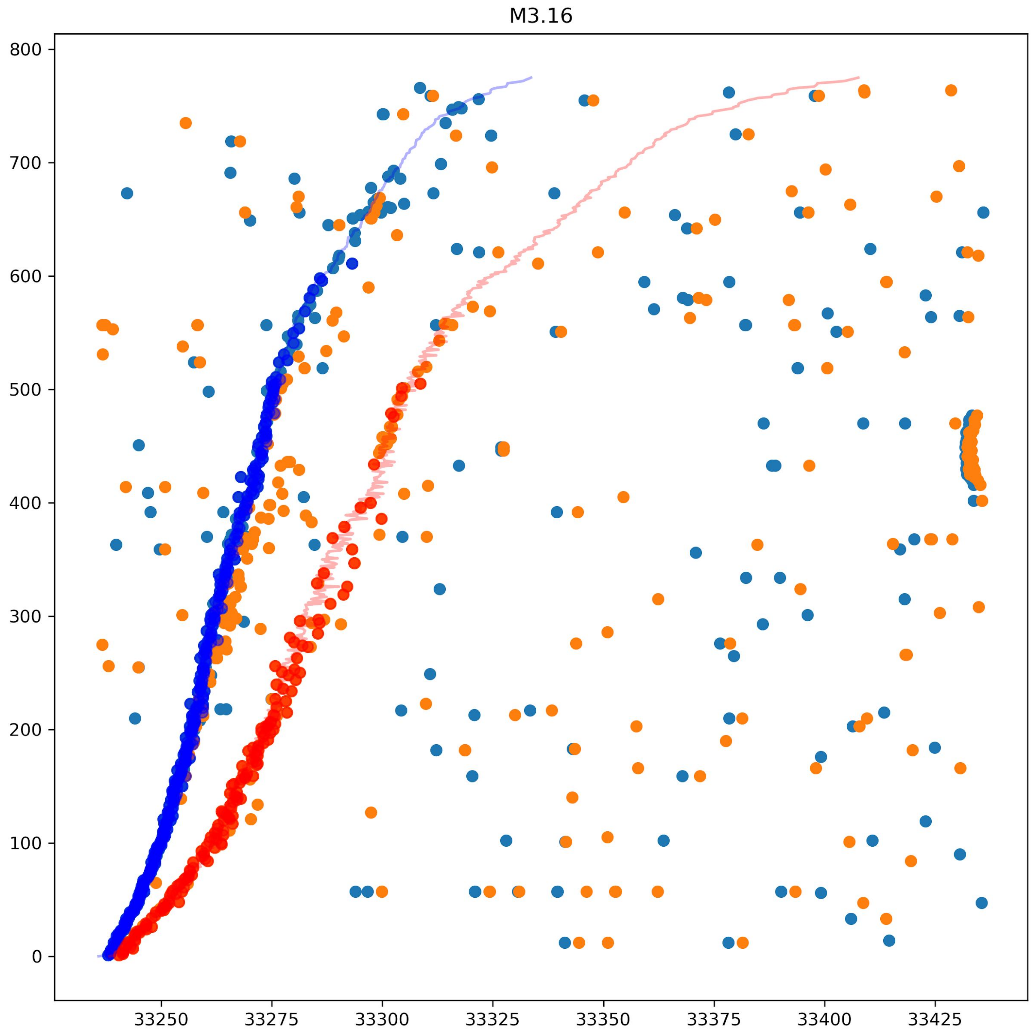




M2, Bay
Area

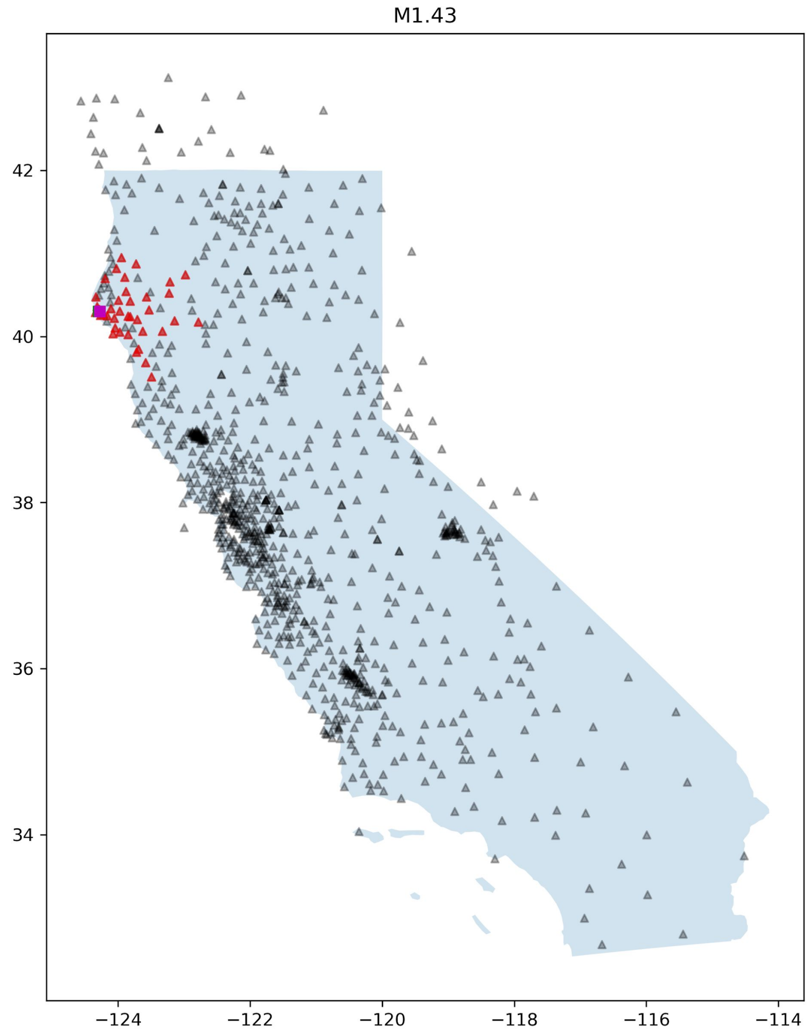
Example Detections

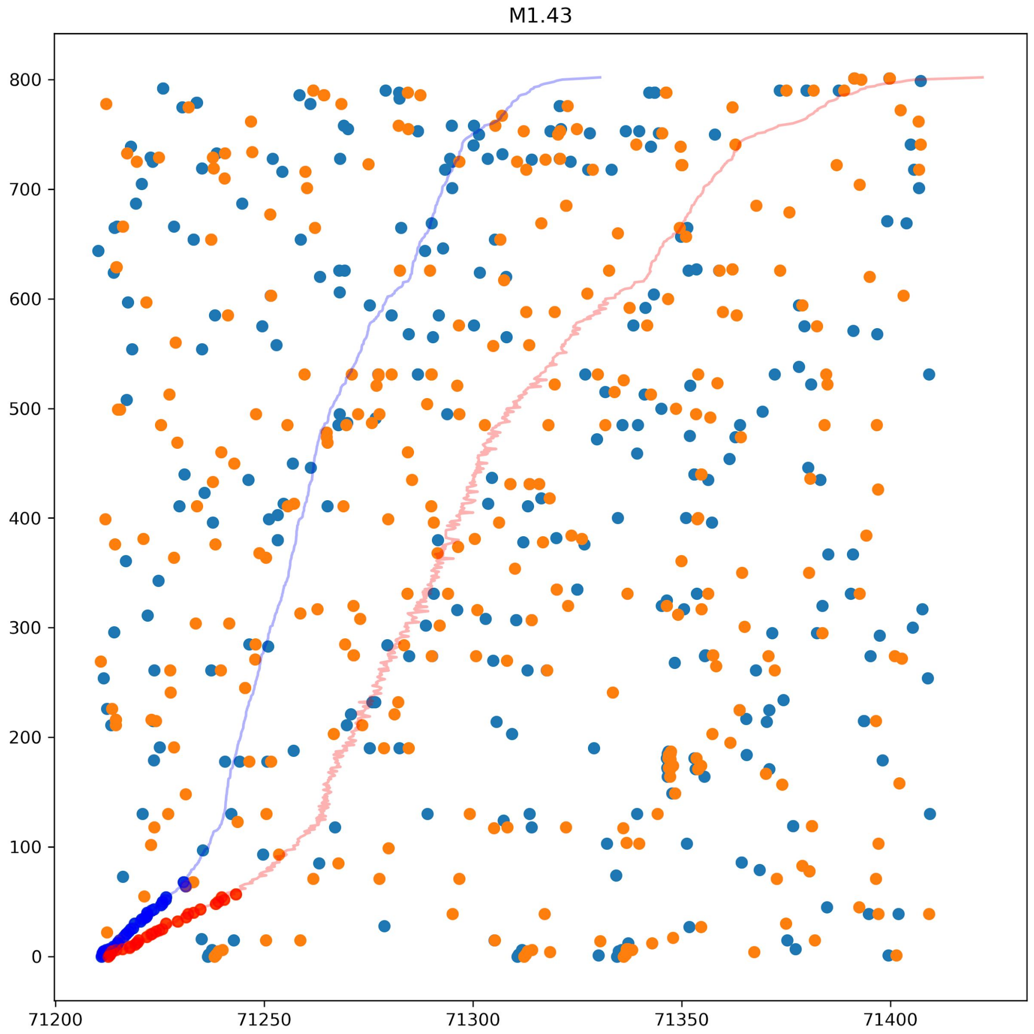




M3, Bay
Area

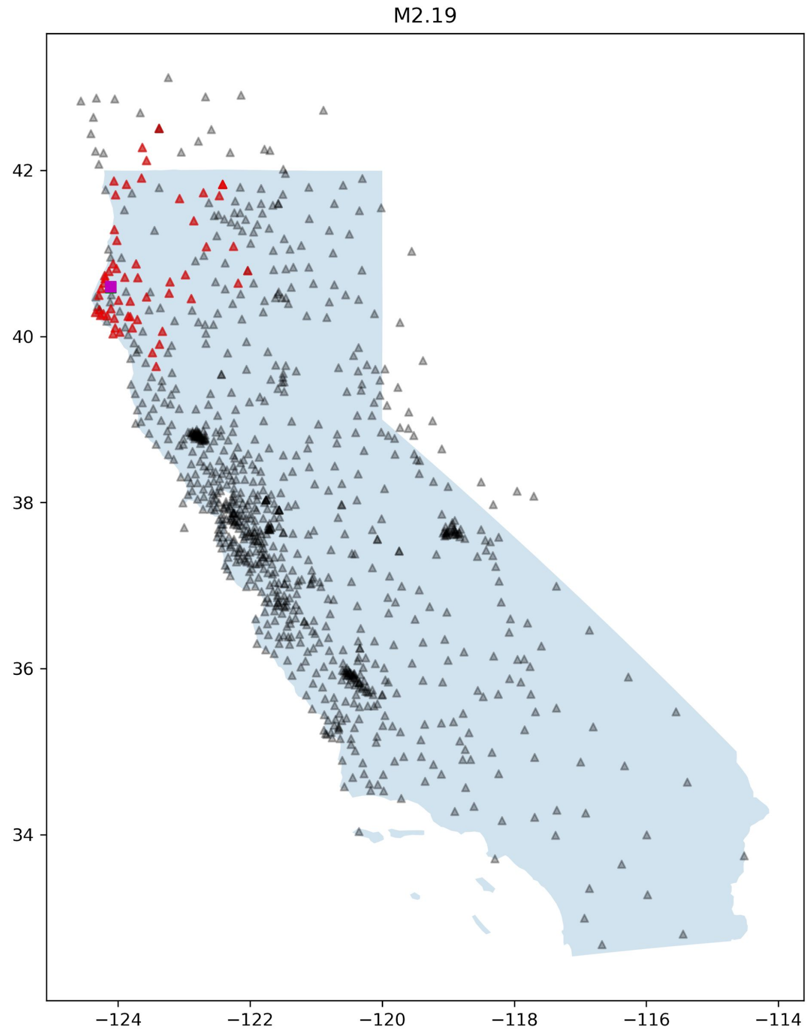
Example Detections

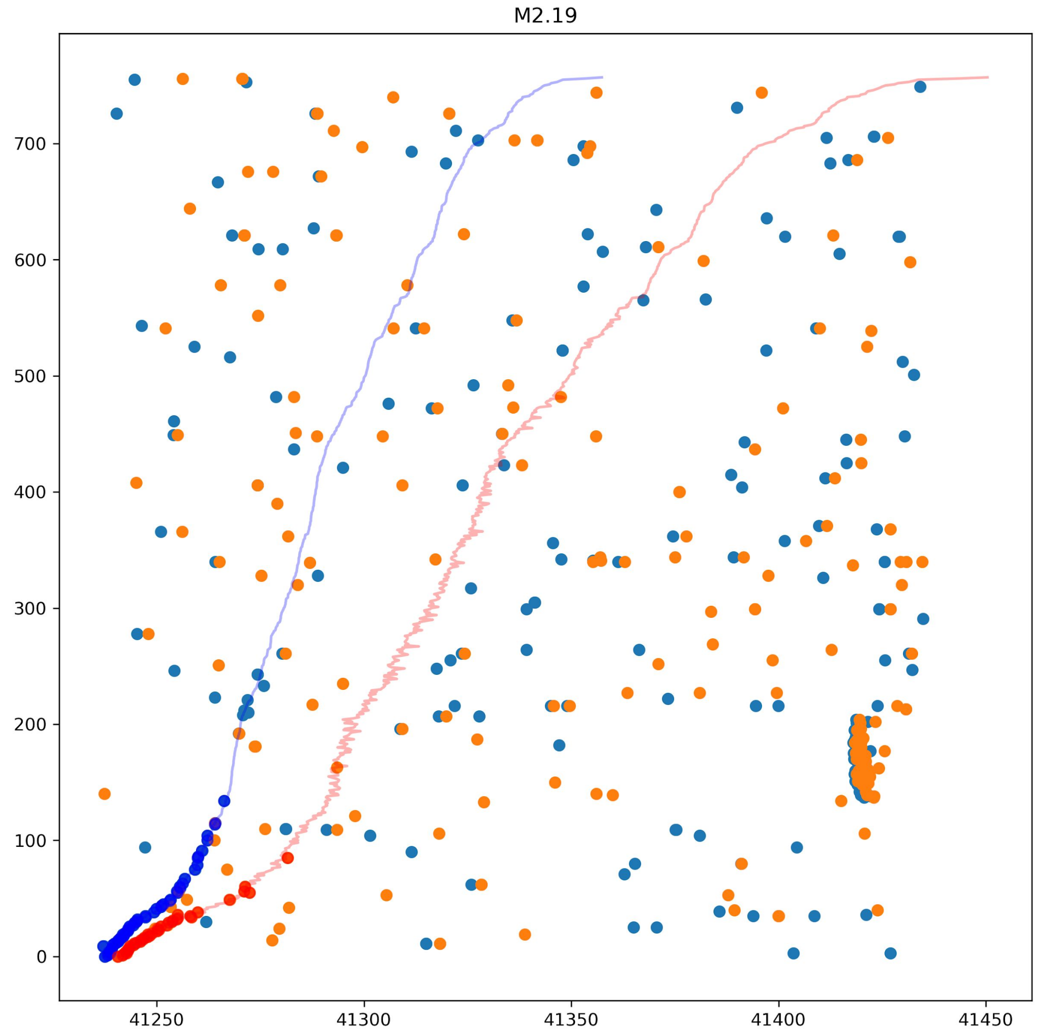




M1, Mendocino Triple Junction

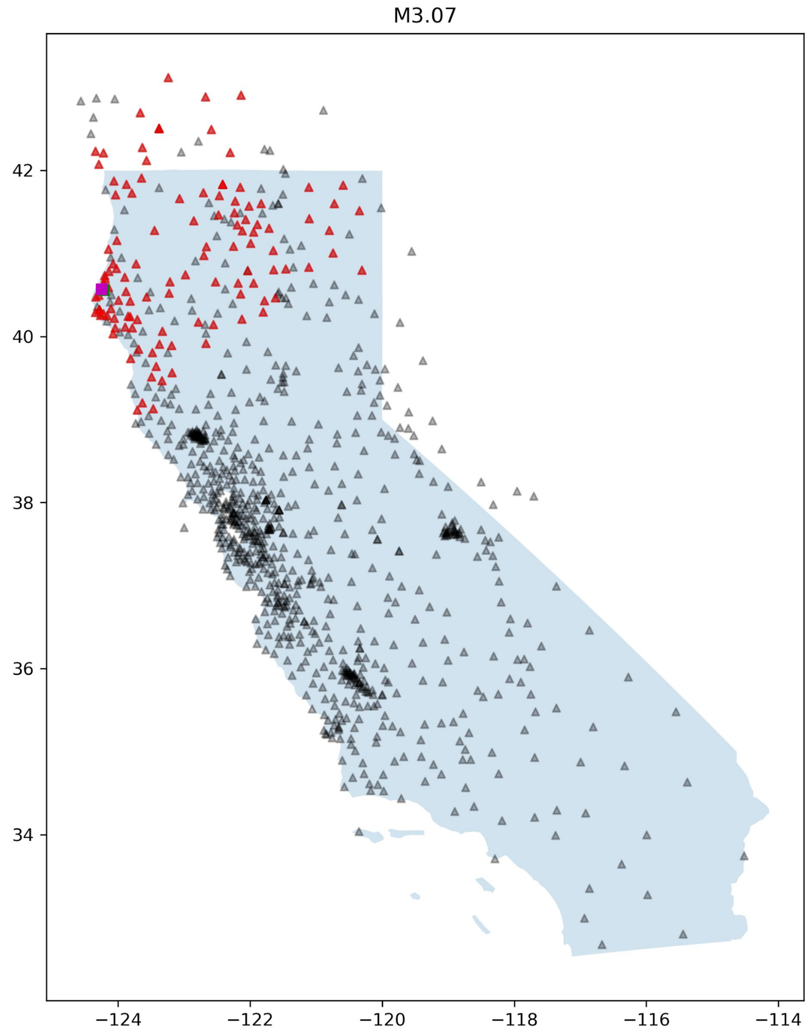
Example Detections

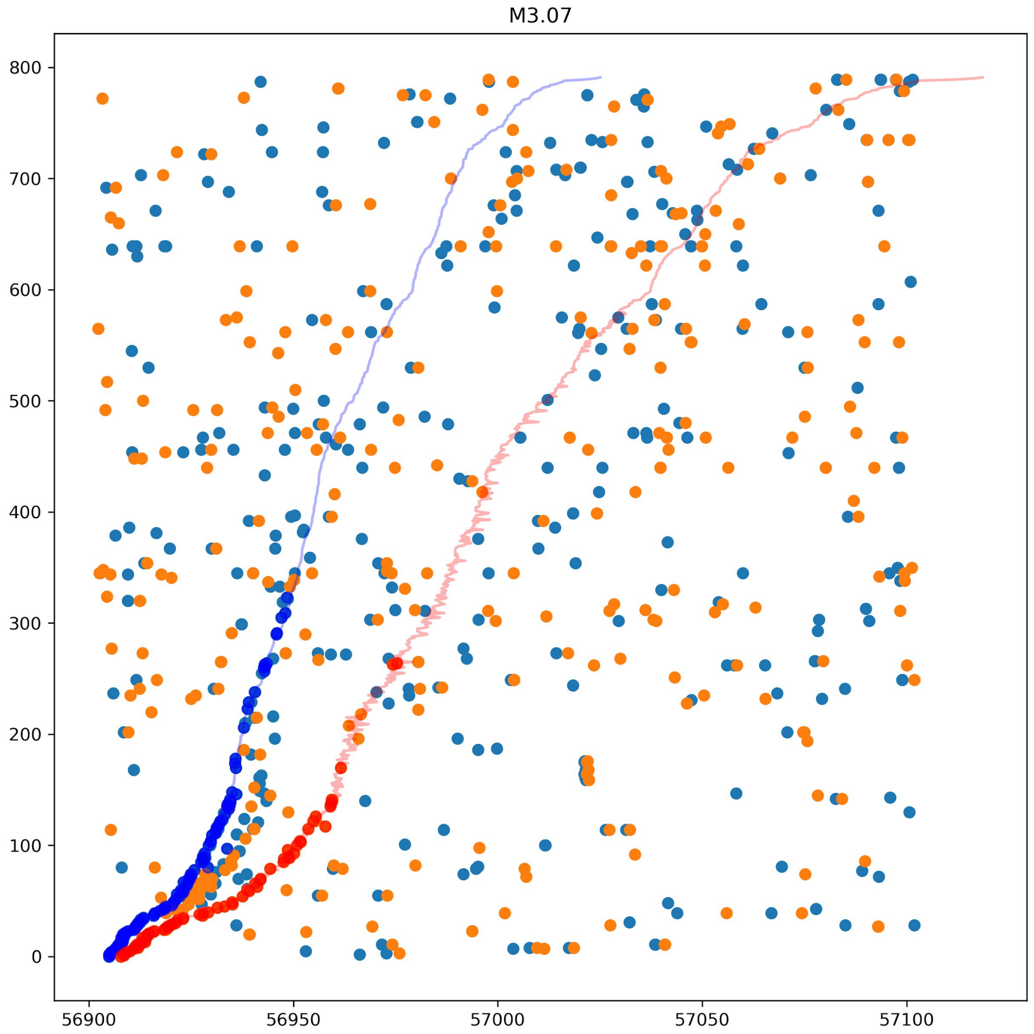




M2, Mendocino Triple
Junction

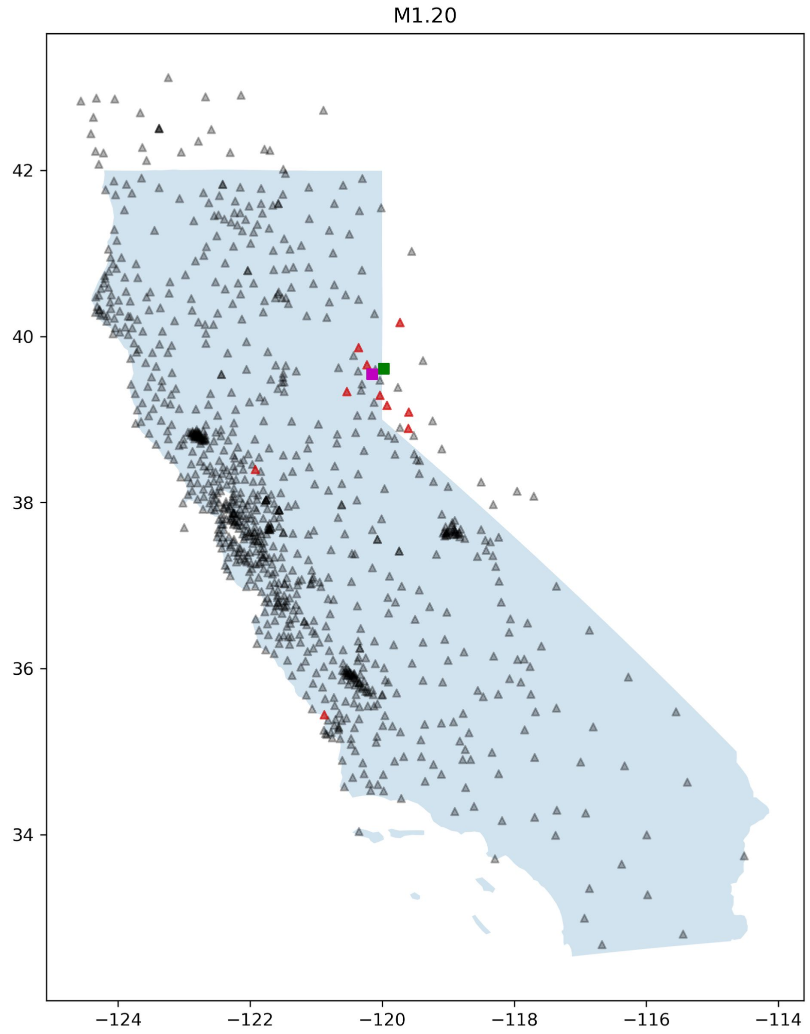
Example Detections

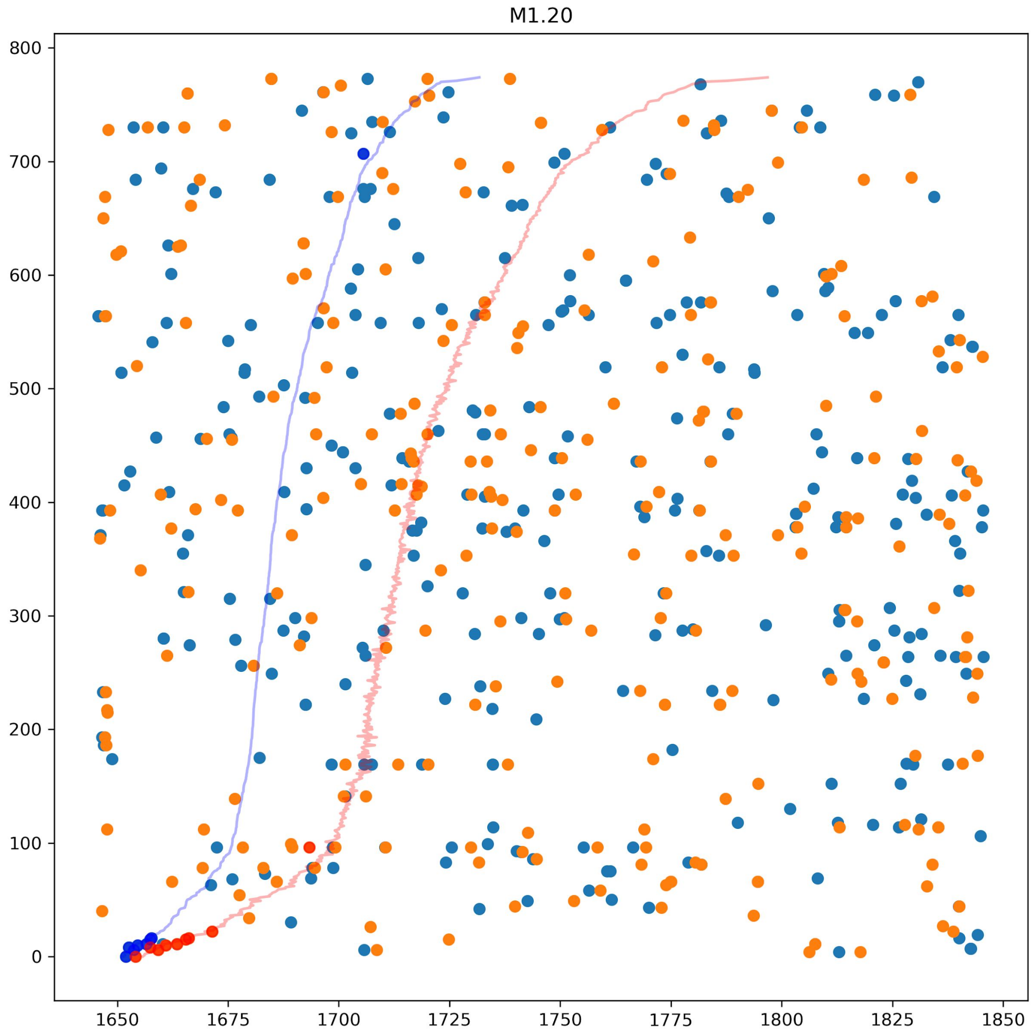




M3, Mendocino Triple Junction

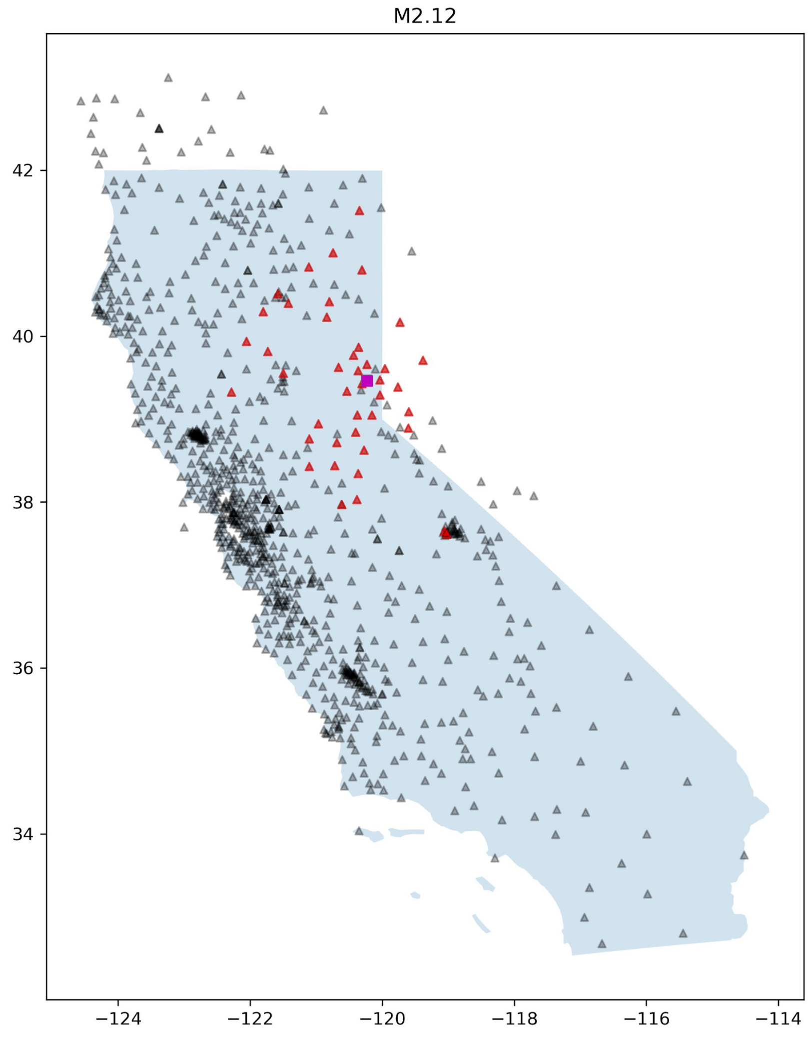
Example Detections

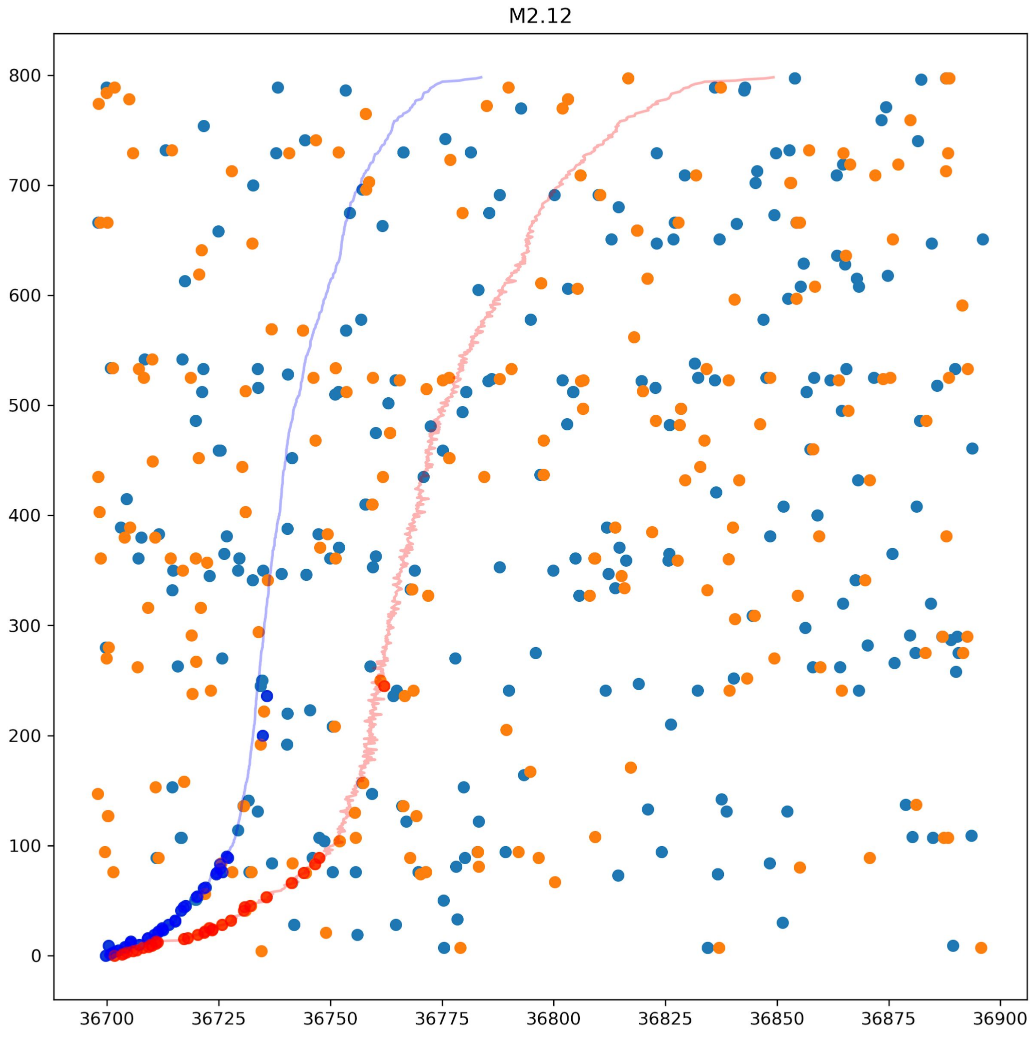




M1, California-Nevada Border

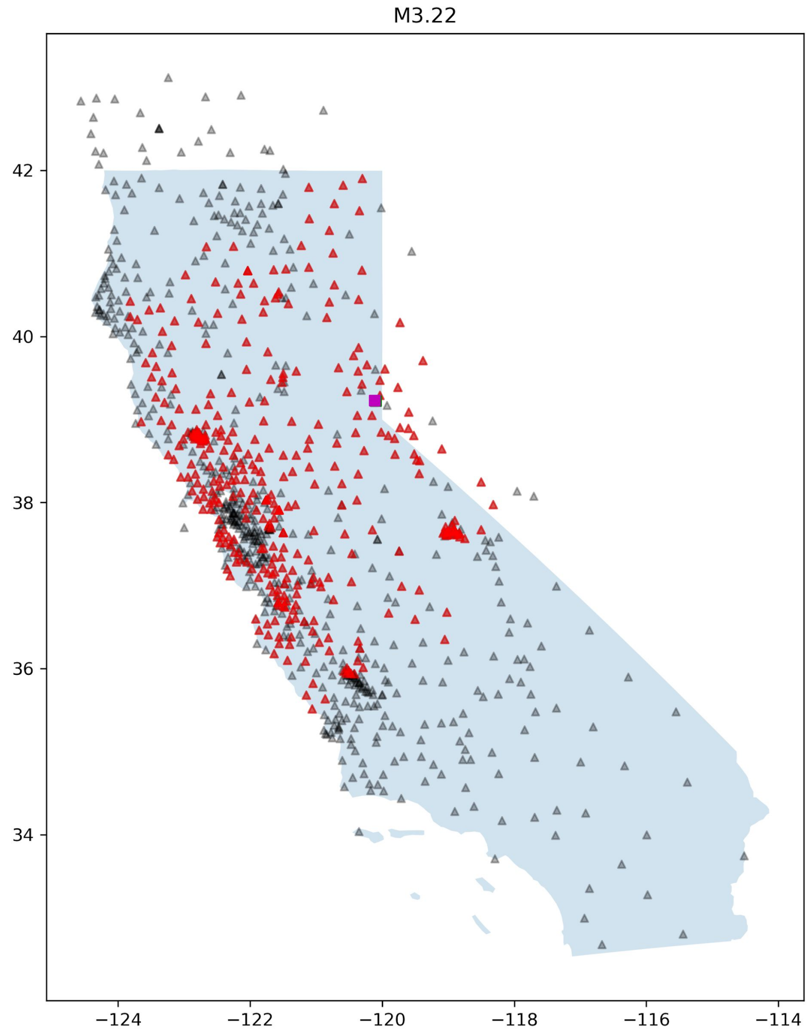
Example Detections

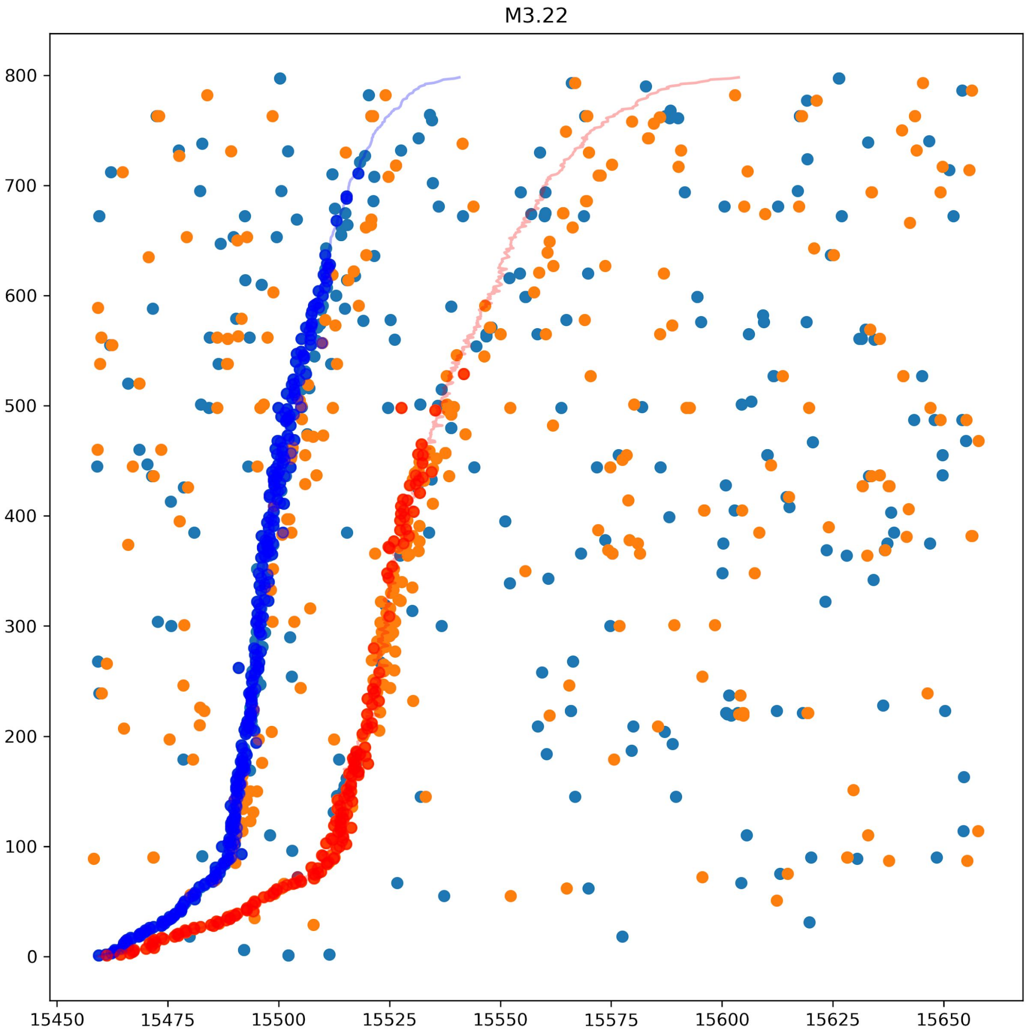




M2, California-Nevada Border

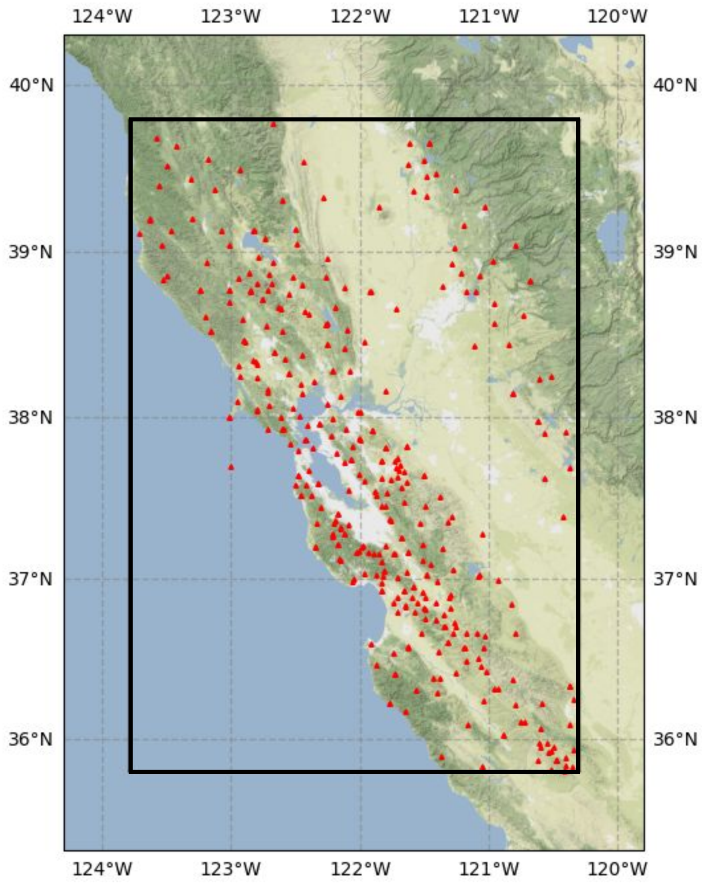
Example Detections





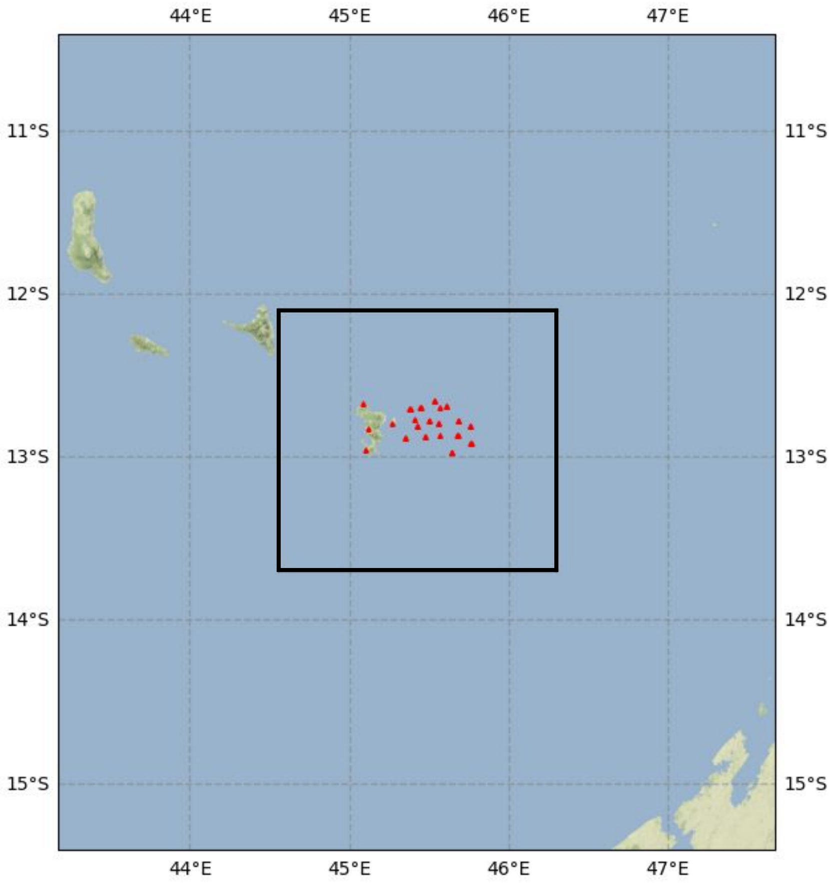
M3, California-Nevada Border

Comparisons of Associators



~250 stations

~100's events per day

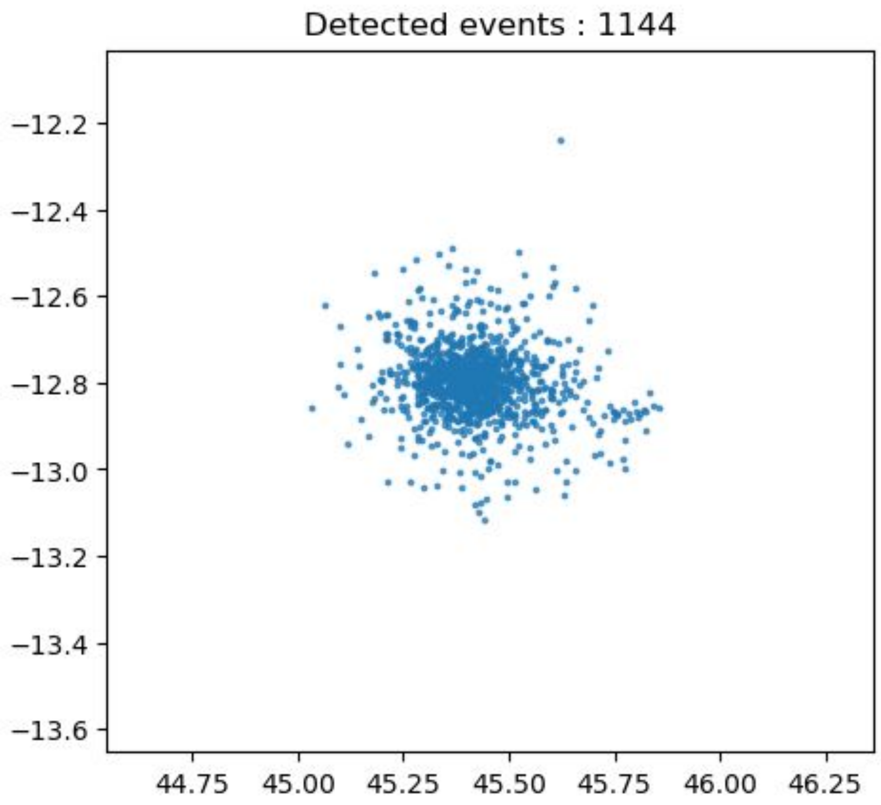


~1/4 scale, 25 stations

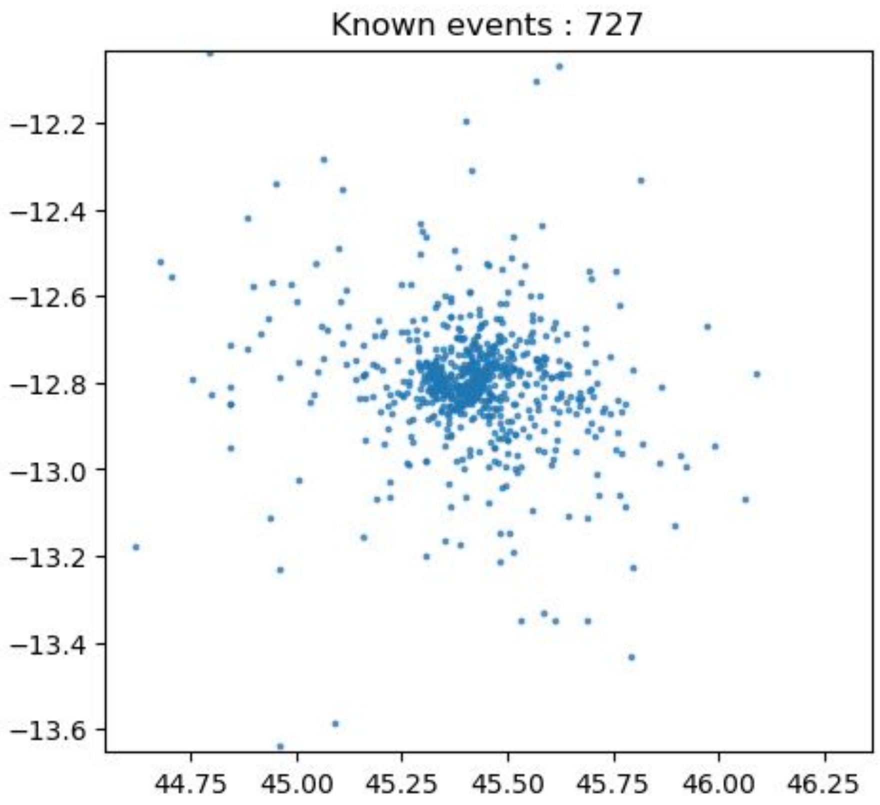
~1000's events per day

Spatial Localization

GENIE

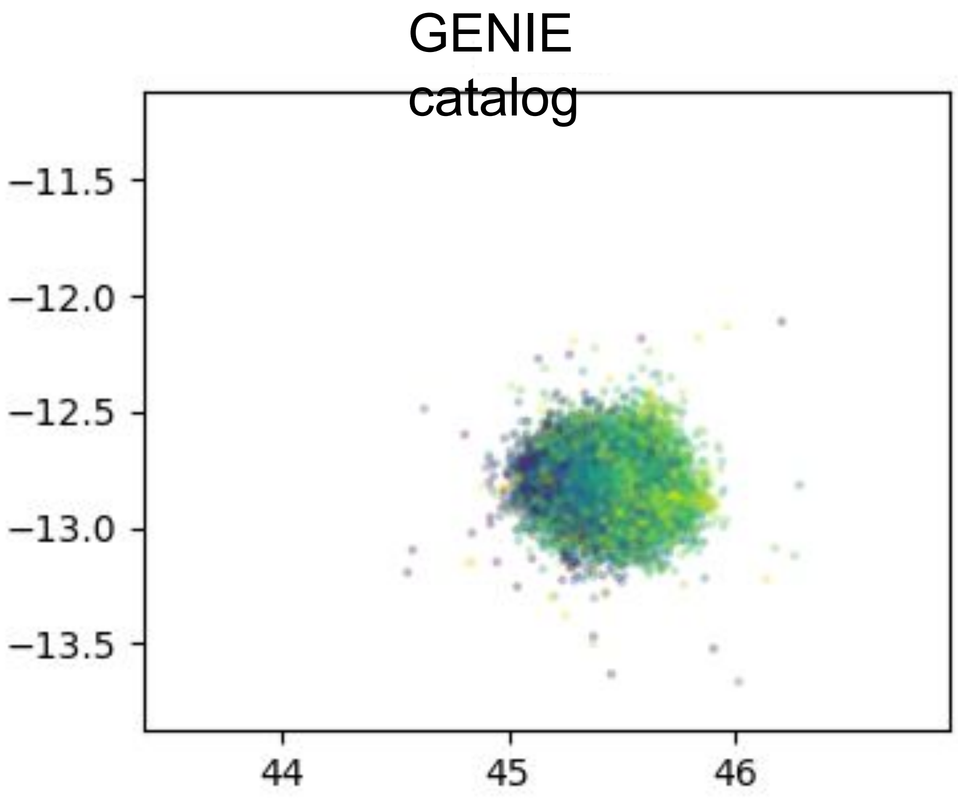


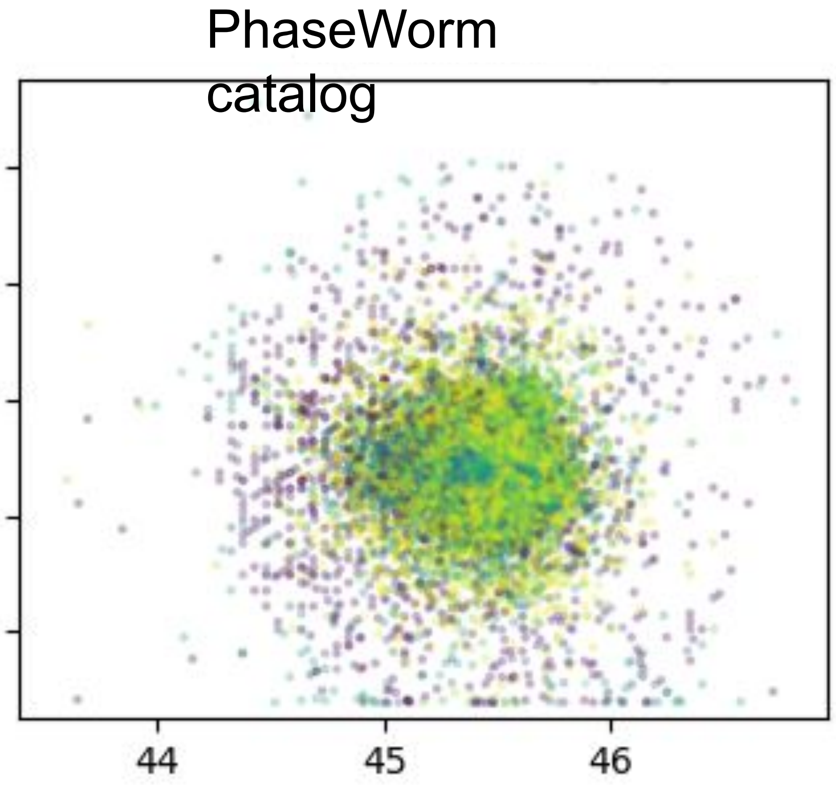
PhaseWorm



Less scatter in GENIE
catalog

Spatial Localization (Full catalog)

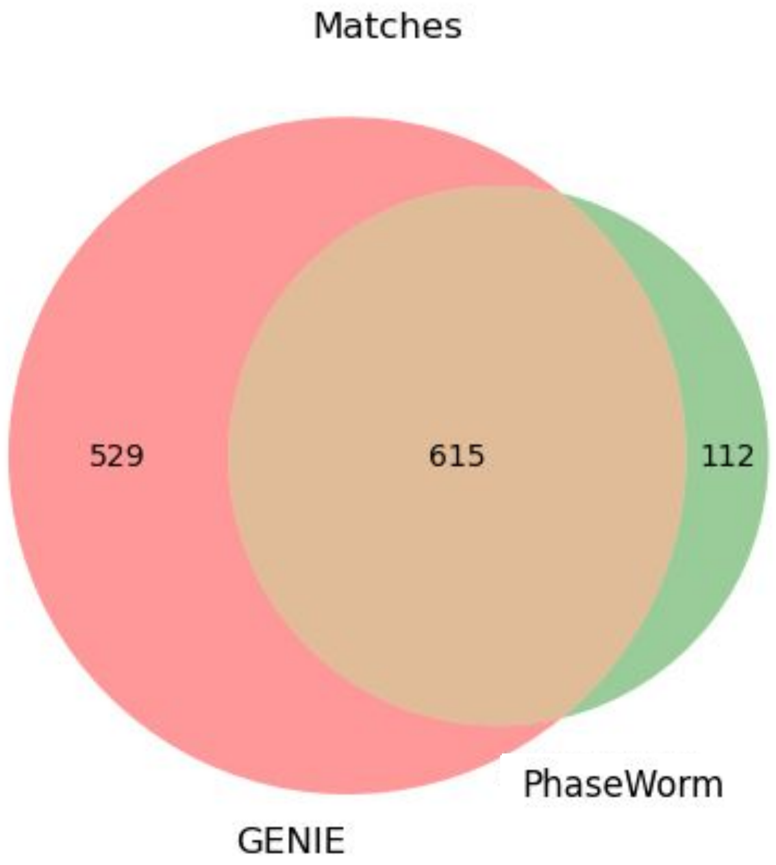




Event Comparison: Number of Events

- Increased detection rate to 1.5x PhaseWorm catalog

- Re-detected ~85% of PhaseWorm catalog



(Using events with spatial window: 150 km)
(Temporal window: 8 s)

LP events in Earthworm catalog

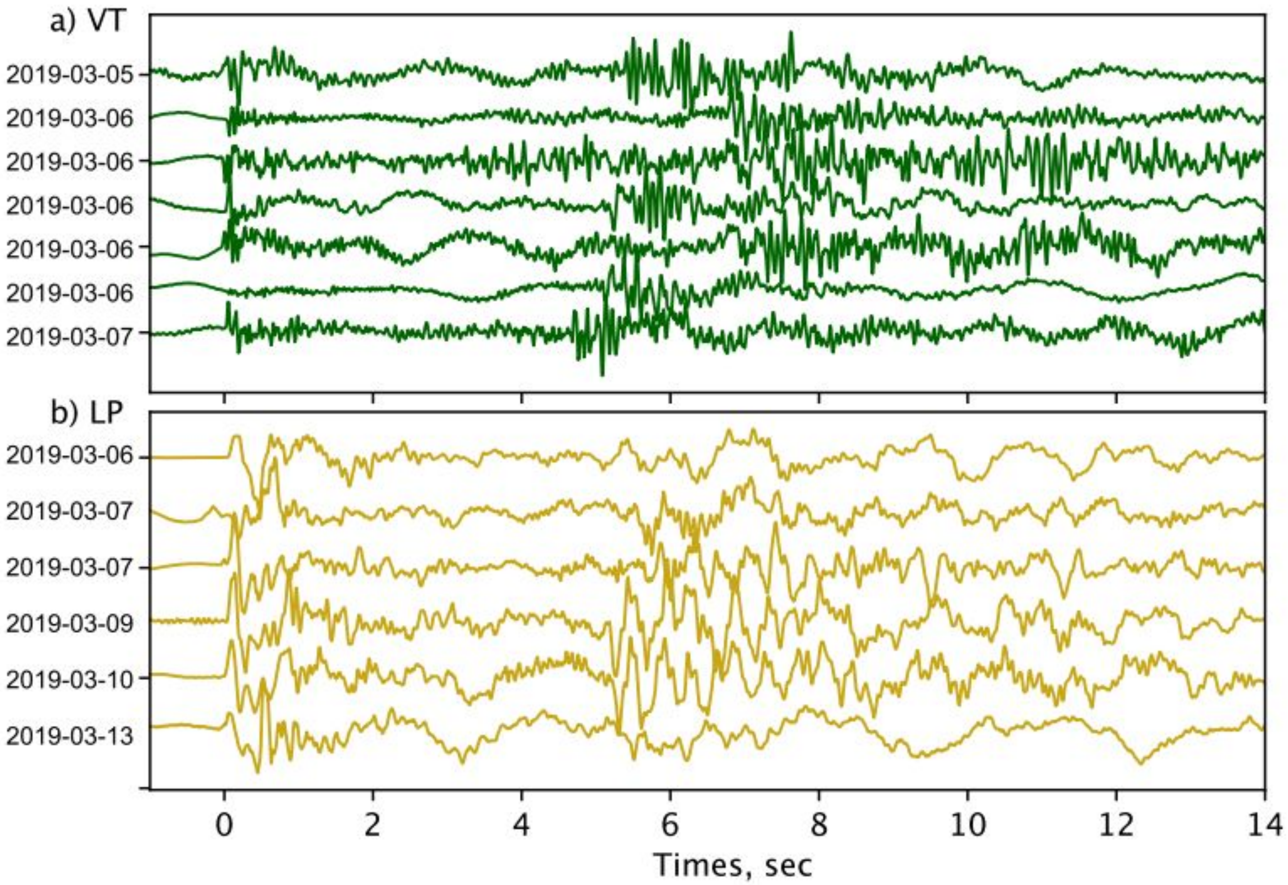
- VT events were well detected, but LP events harder to detect by PhaseWorm

- Increased rate of S vs. P phase picks

- Earthworm nucleates events based on P waves only

VT

LP



(Retailleau et al., 2022)

LP events in Earthworm catalog

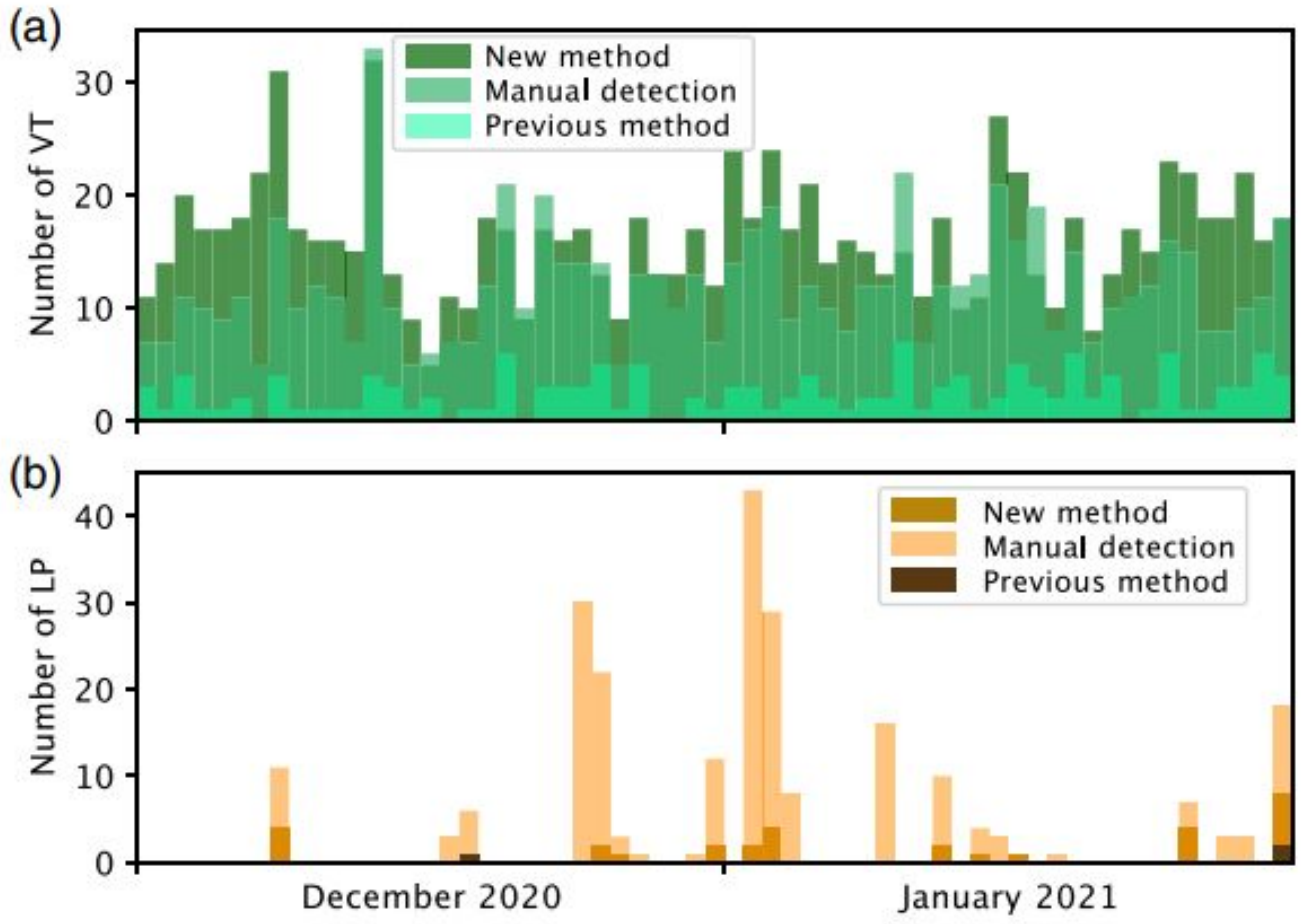
- VT events were well detected, but LP events harder to detect by PhaseWorm

- Increased rate of S vs. P phase picks

- Earthworm nucleates events based on P waves only

V
T

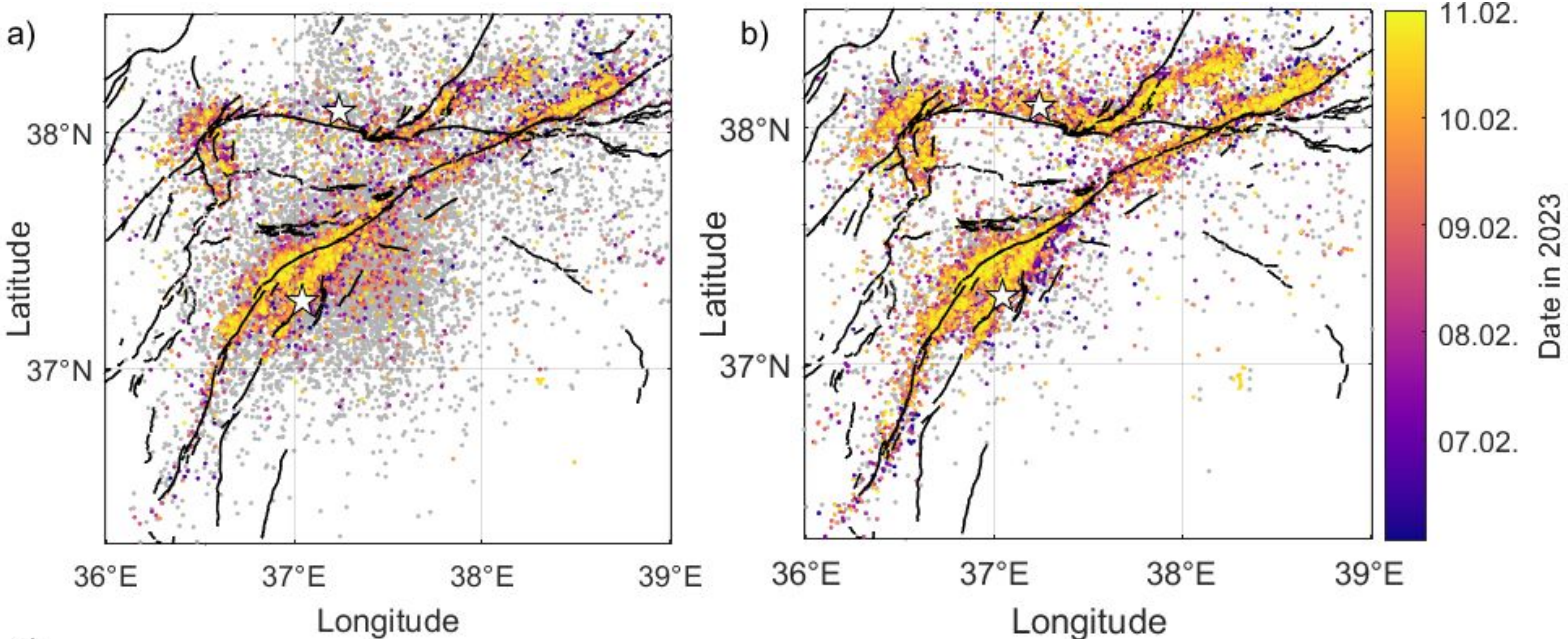
L
P



(Retailleau et al., 2022)

Kahramanmaraş Aftershock Sequence

GaMMA GENIE



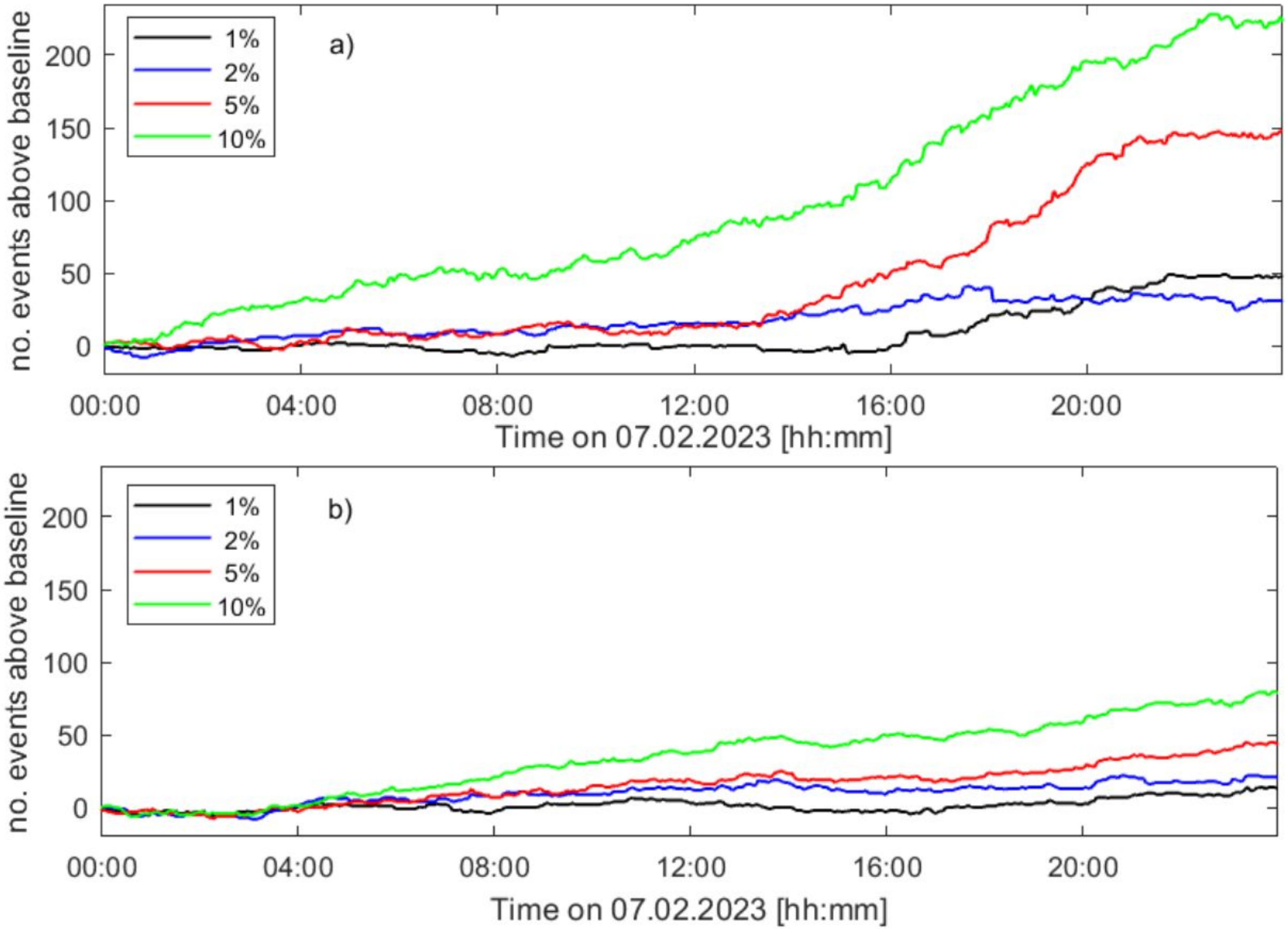
For this data set, GaMMA finds more events, while GENIE associates more phase *per* event

Becker et al. (2024)

Influence of adding random picks

GaMM
A

GENIE



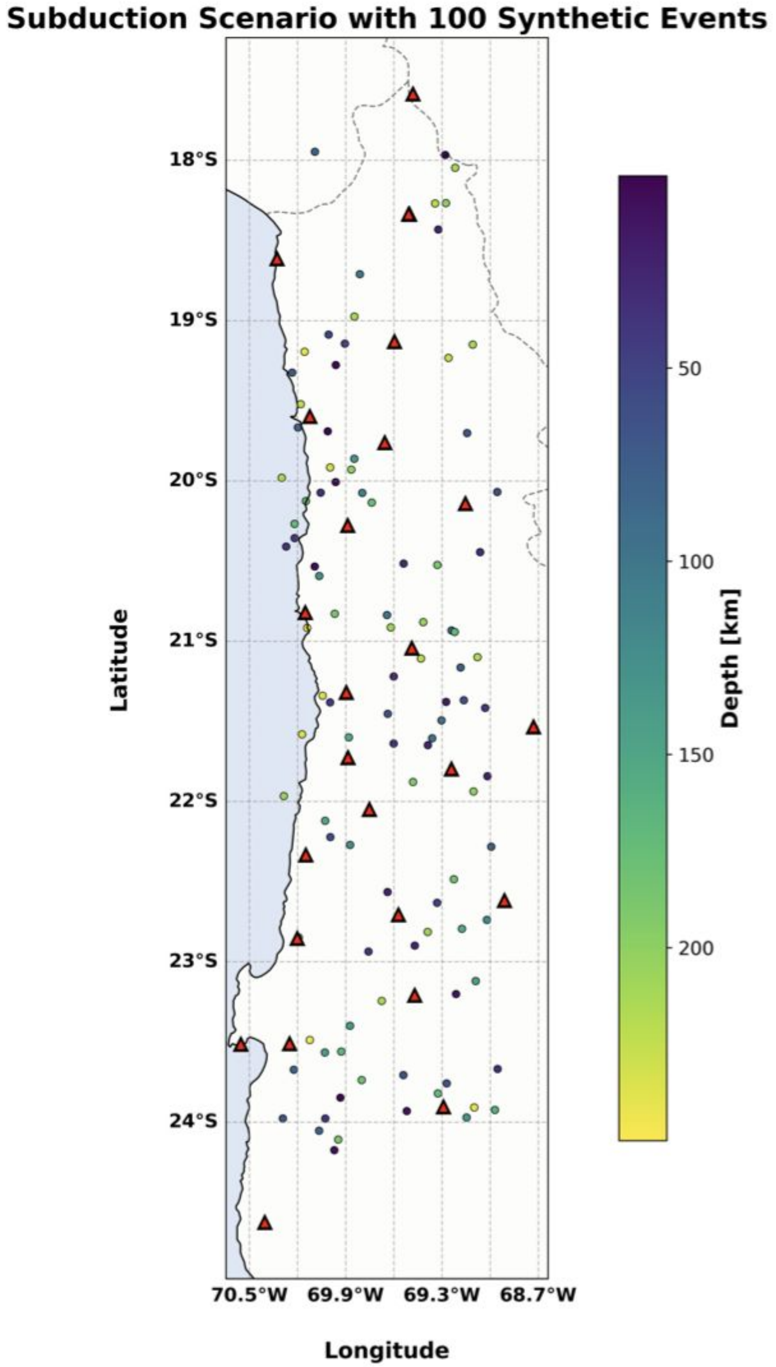
For this data set, GaMMA seems more prone to mis-association.

Becker et al. (2024)

Associator Comparisons

• Tested performance of five different associators (GaMMA, PhaseLink, REAL, GENIE, PyOcto) on synthetic scenarios

• Found similar performance for low complexity cases, but large differences for high complexity data

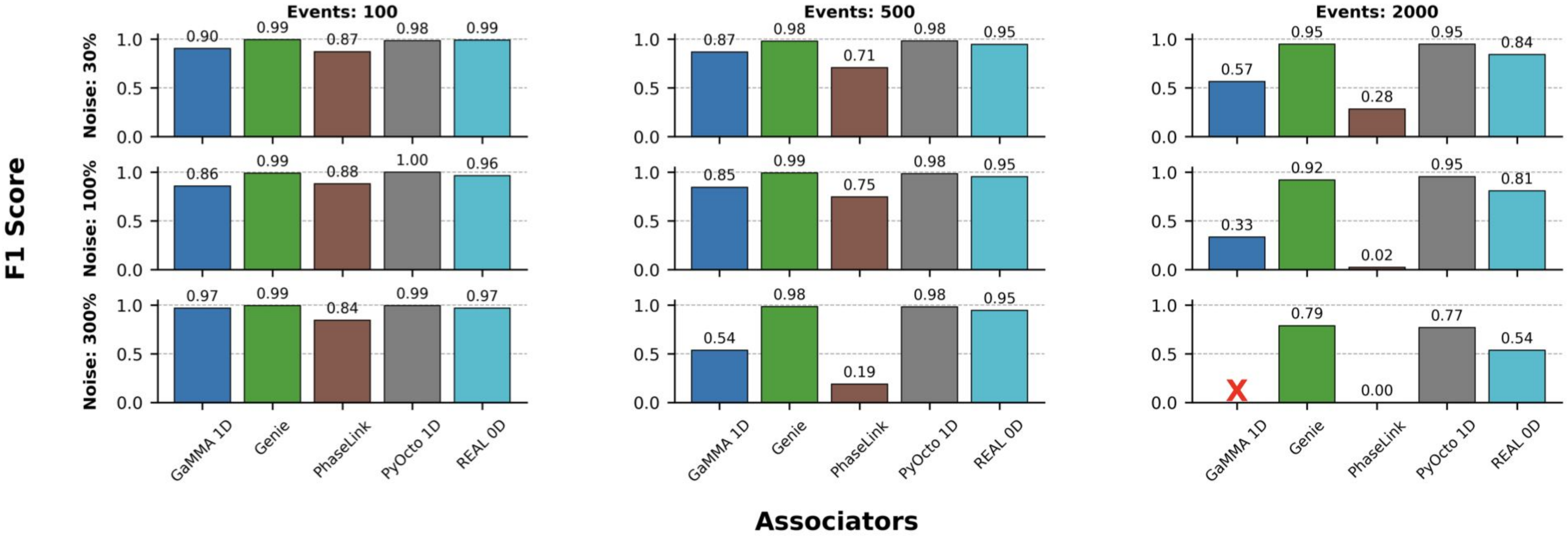


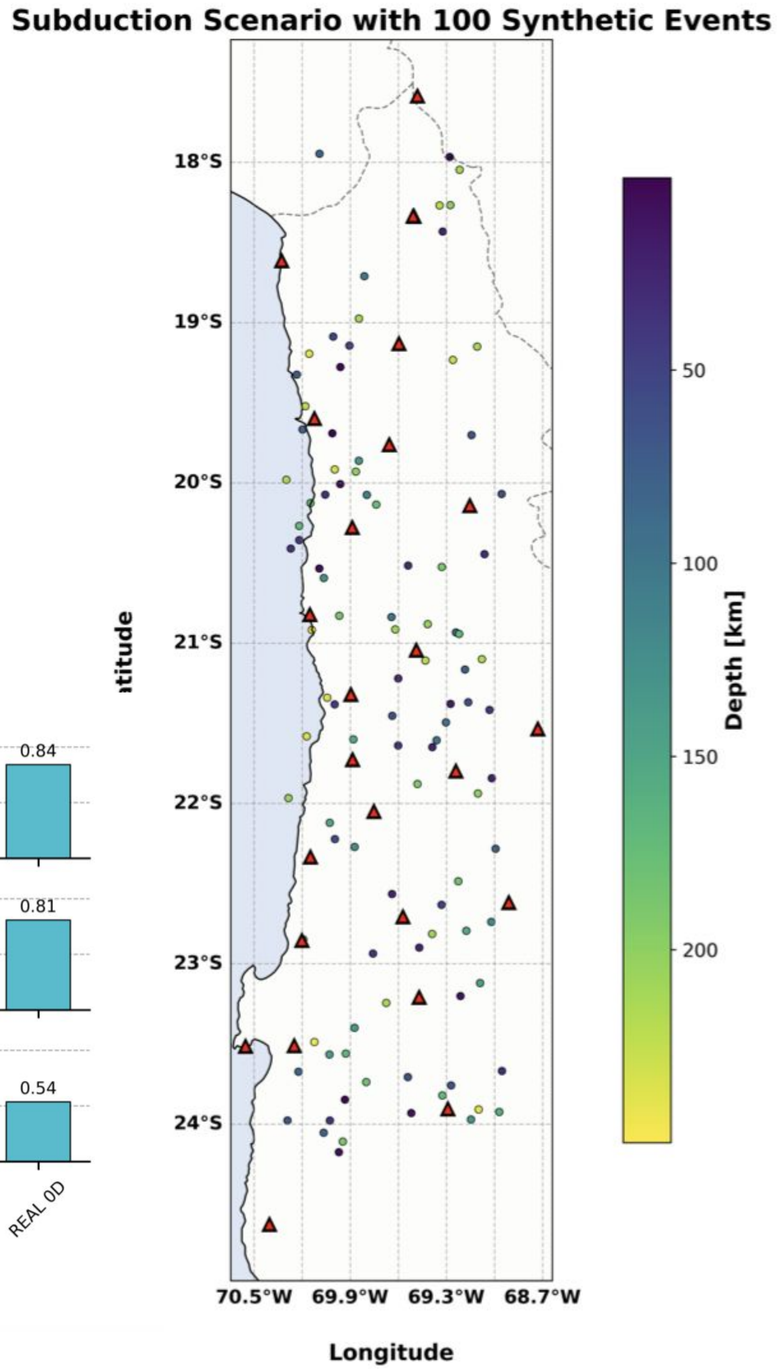
Puenta et al., (submitted)

Associator Comparisons

- Tested performance of five different associators (GaMMA, PhaseLink, REAL, GENIE, PyOcto) on synthetic scenarios

- Found similar performance for low complexity cases, but large differences for high complexity data

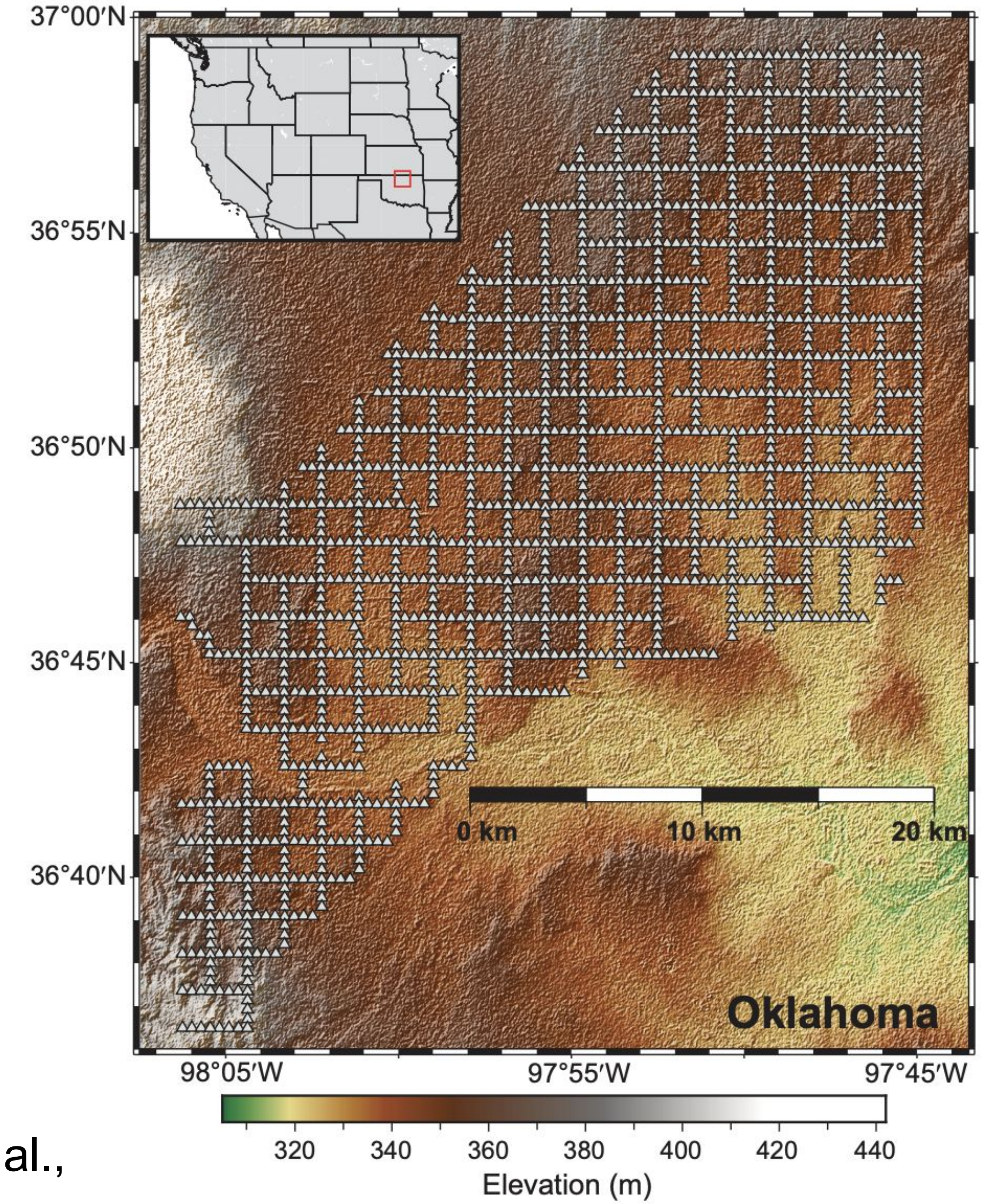




Puenta et al.,

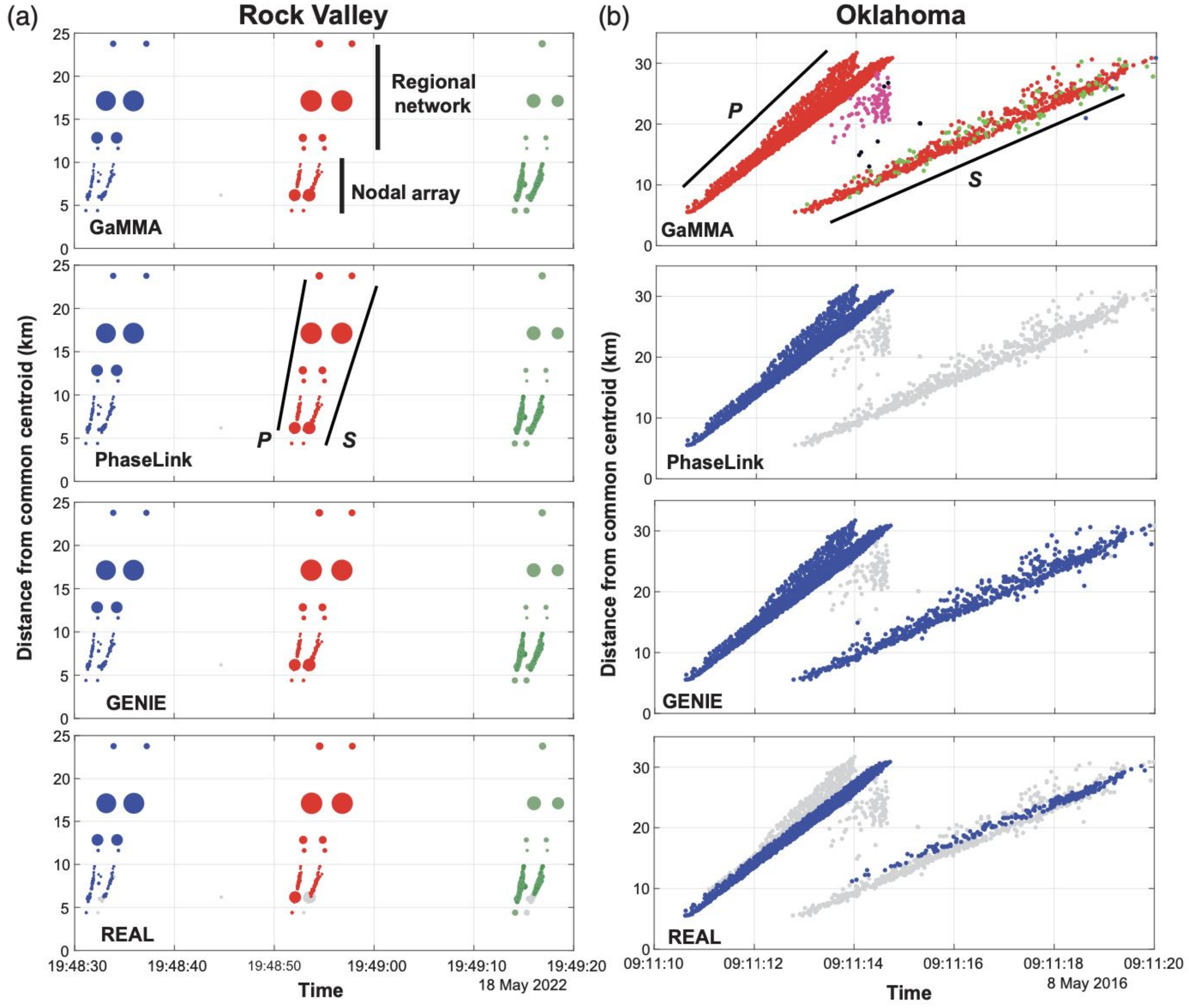
Associators applied to dense nodal arrays

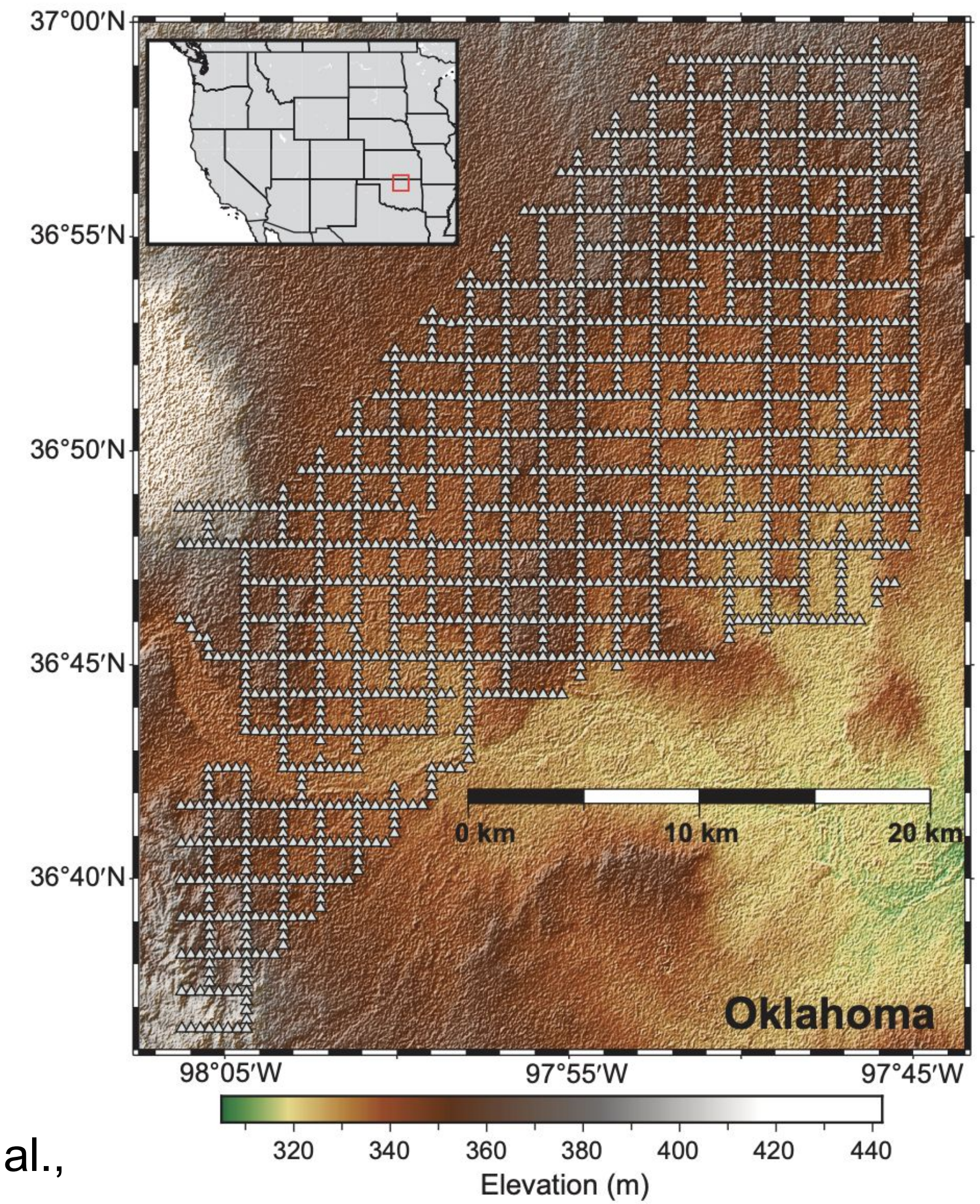
- Tested performance of four different associators (GaMMA, PhaseLink, REAL, GENIE) on data from Rock Valley (52 nodes + 9 regional broadband sensors) and ~1800 geophones at LASSO



Pennington et al., (2025)

Associators applied to dense nodal arrays

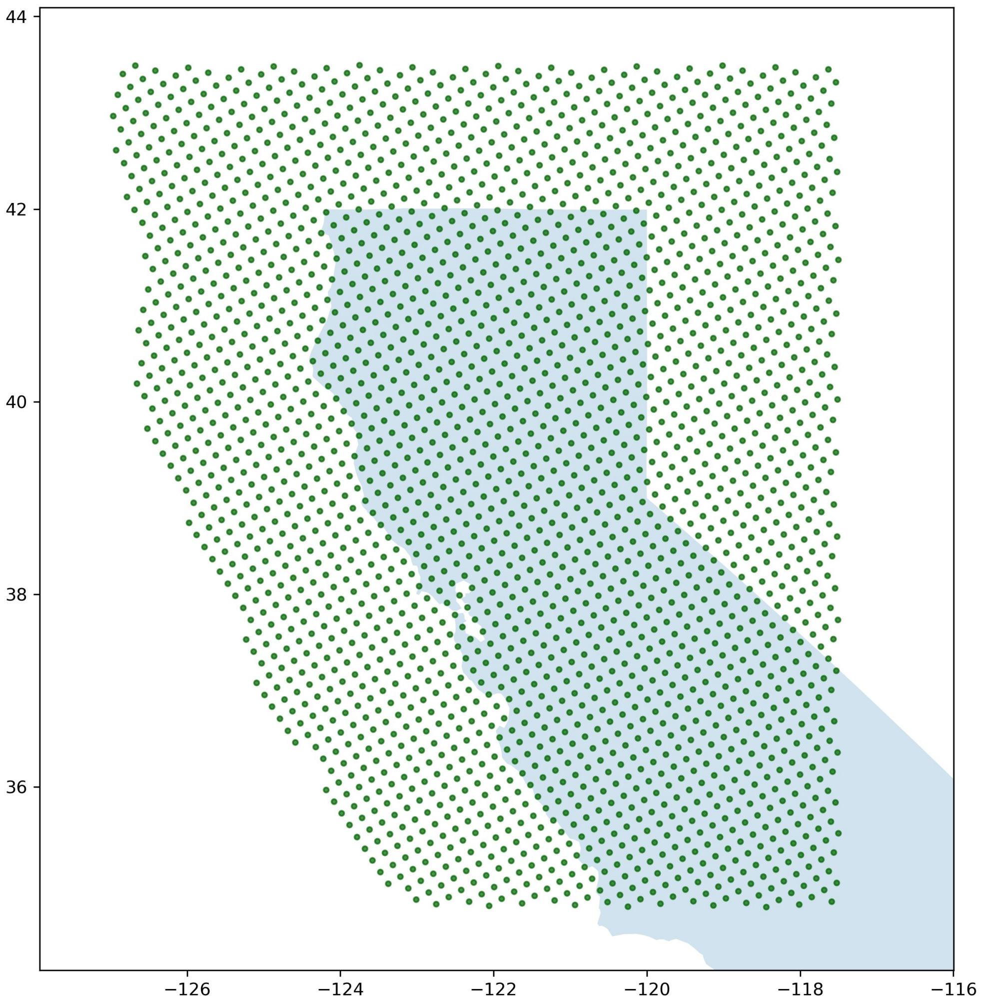


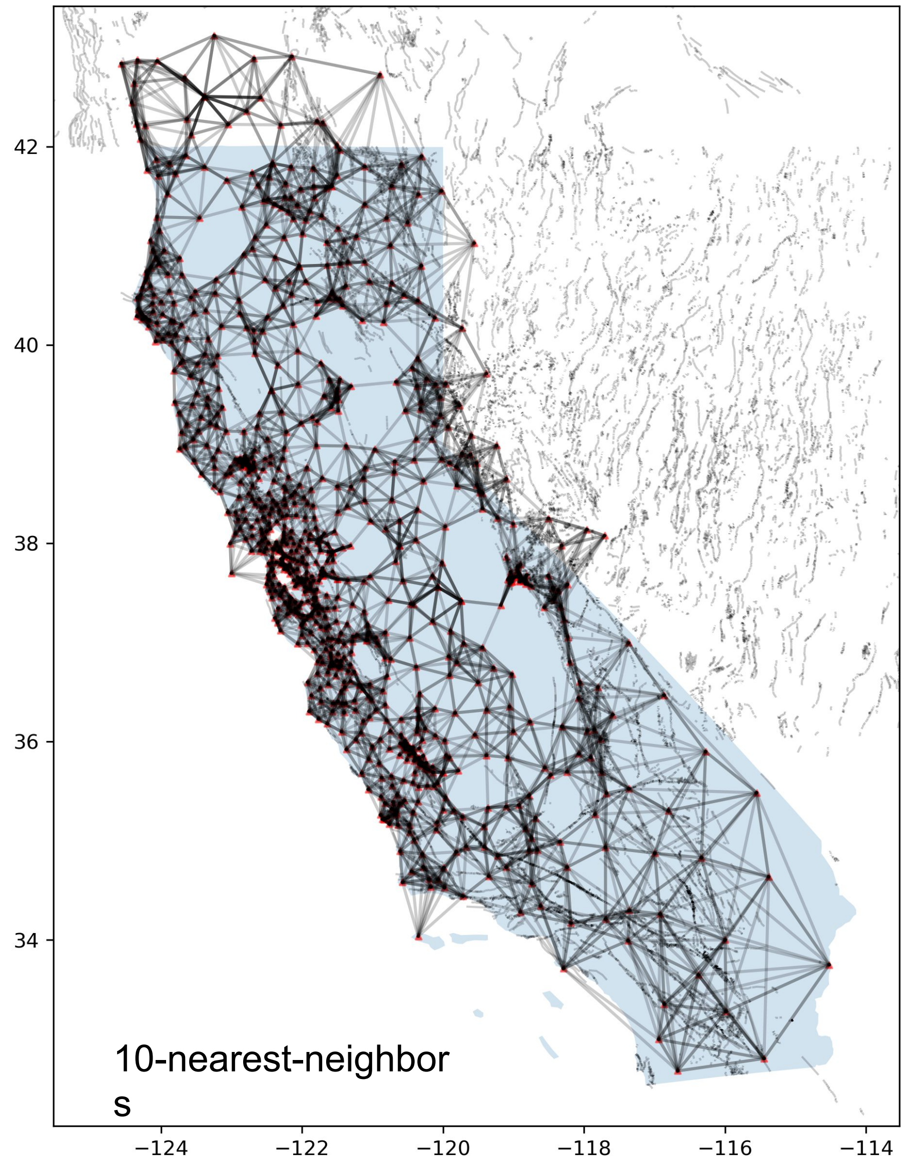


Pennington et al., (2025)

Northern California

Northern California





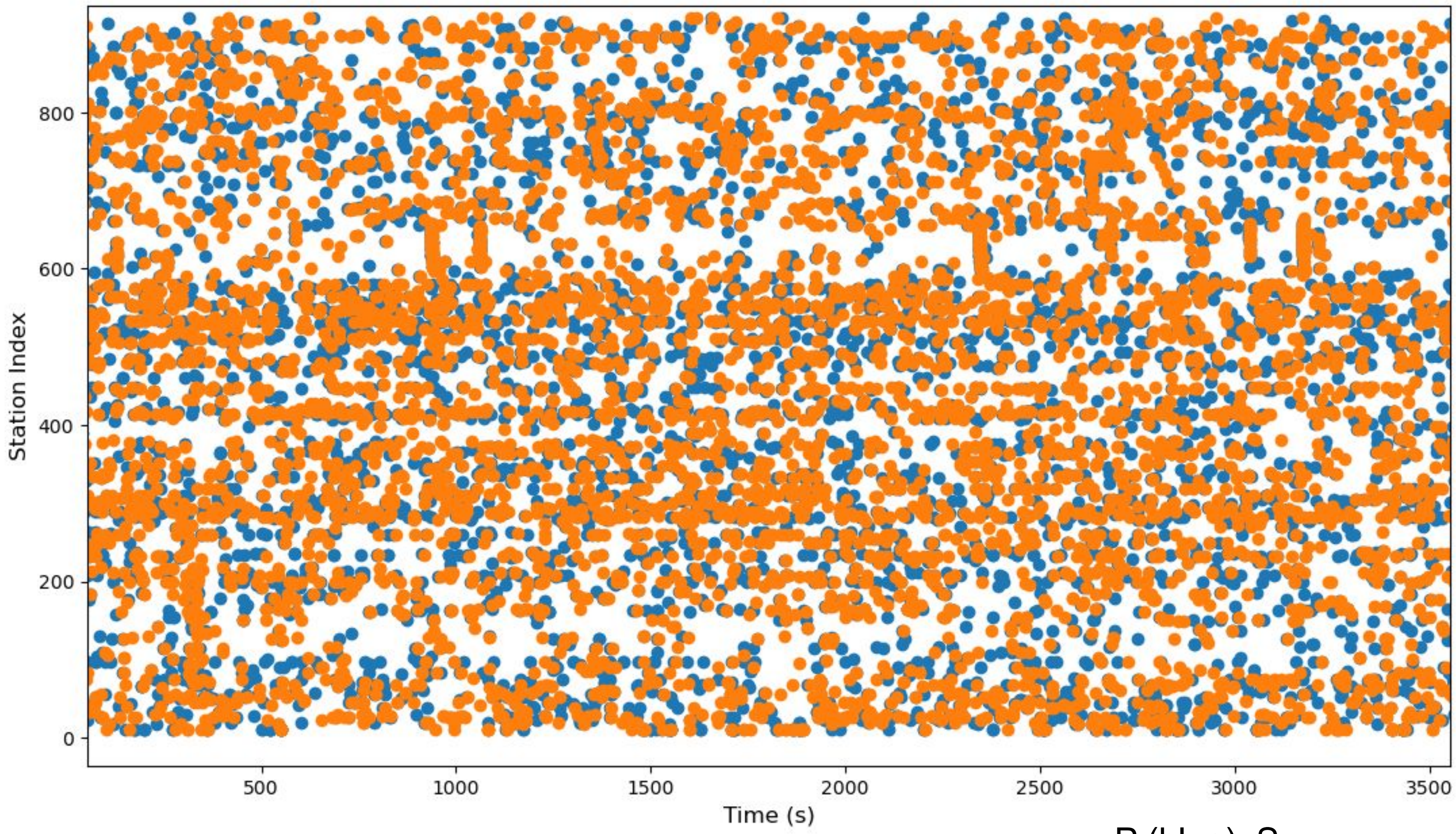
~1000 stations from many networks: NC, BK, PN, BG, UW, NN

Picks

Average:
240,000
picks per day

P-waves
121,000
picks per day

S-waves
117,000
picks per day



P (blue), S (orange)

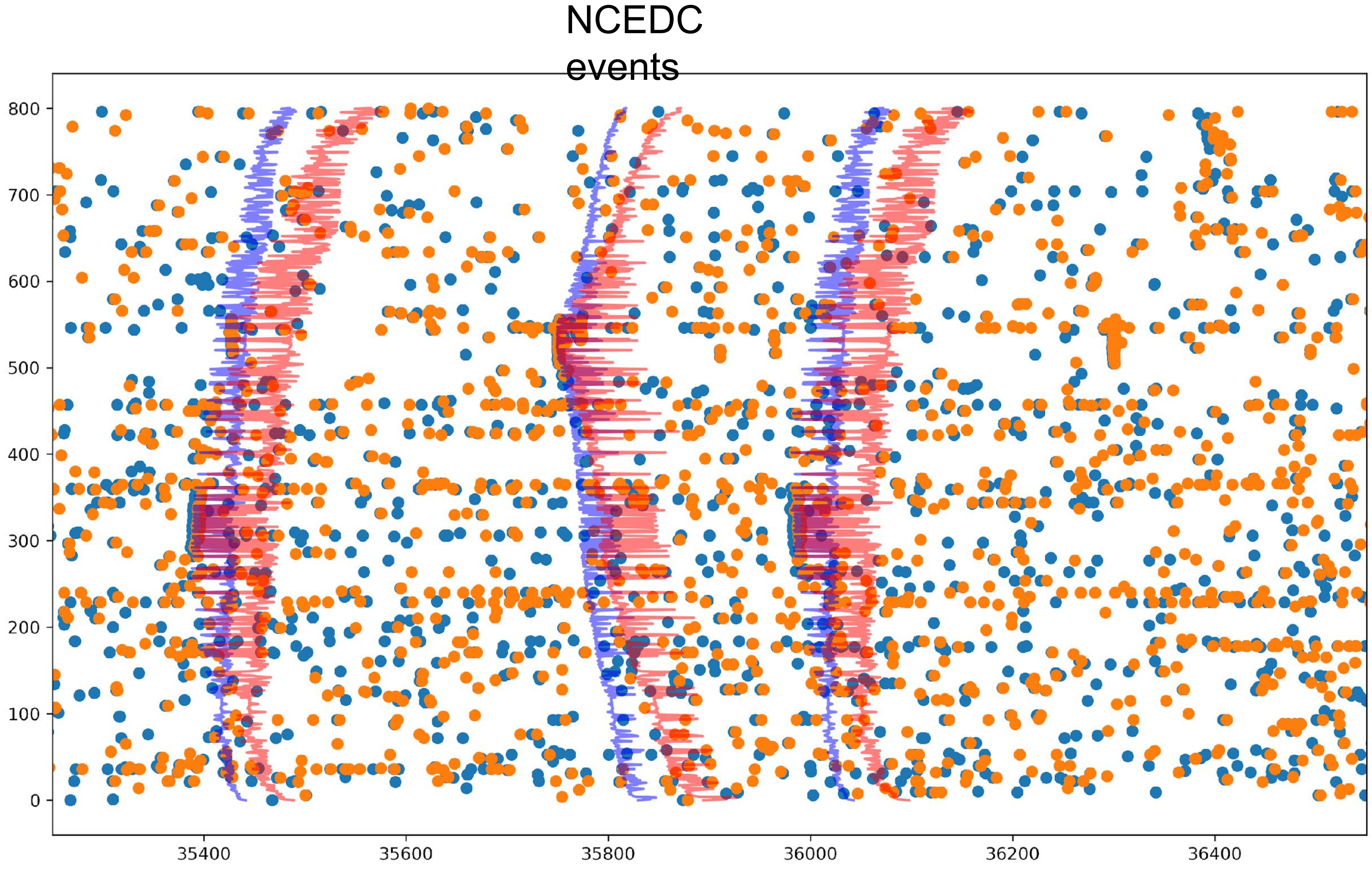
- Collaboration with Weiqiang Zhu to obtain PhaseNet picks

Picks

Average:
240,000
picks per day

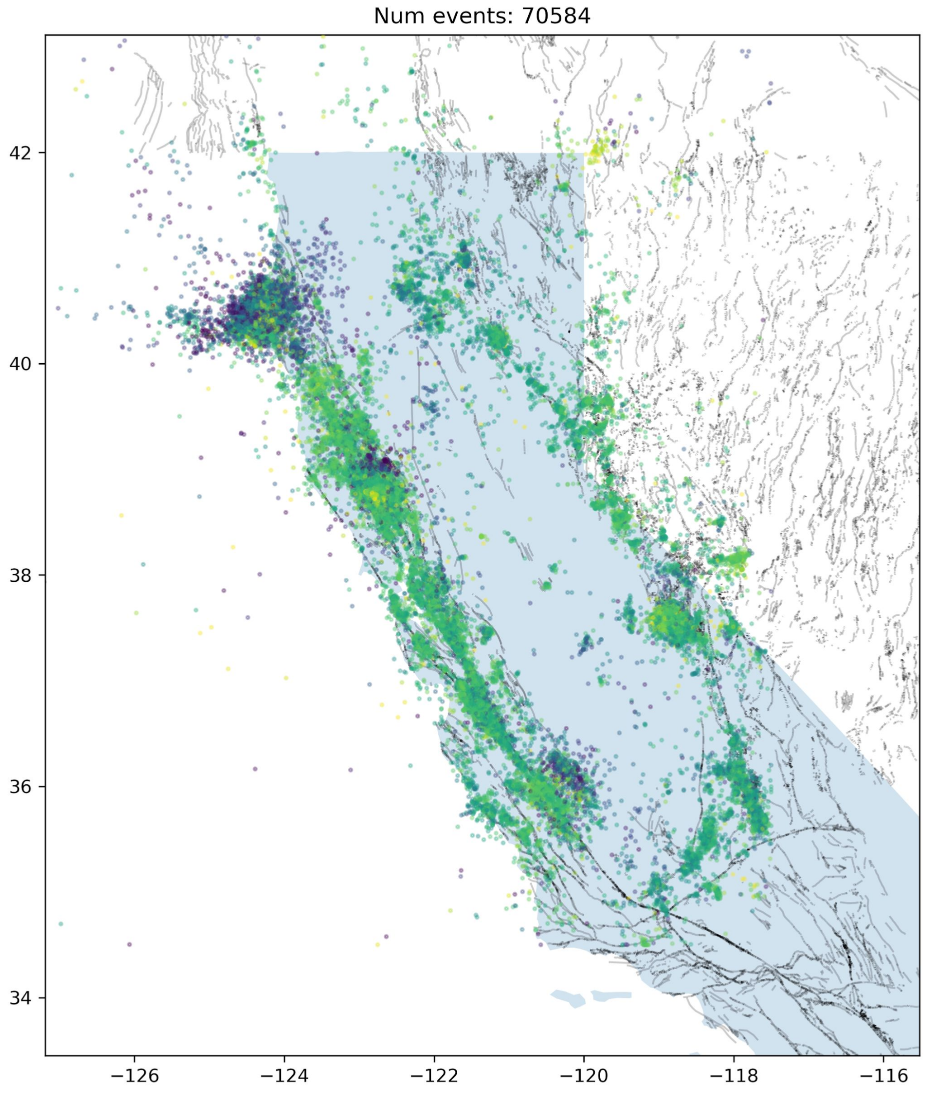
P-waves
121,000
picks per day

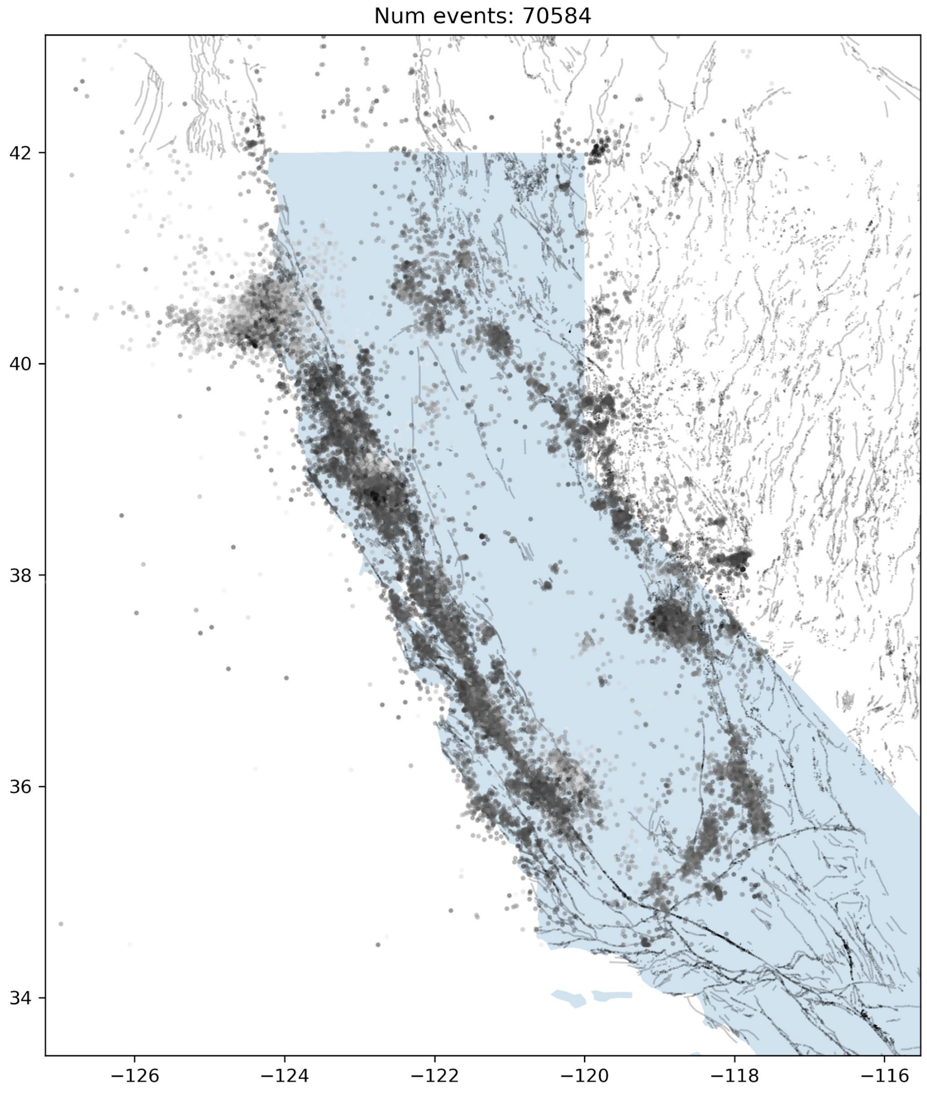
S-waves
117,000
picks per day



• Collaboration with Weiqiang Zhu to obtain PhaseNet picks

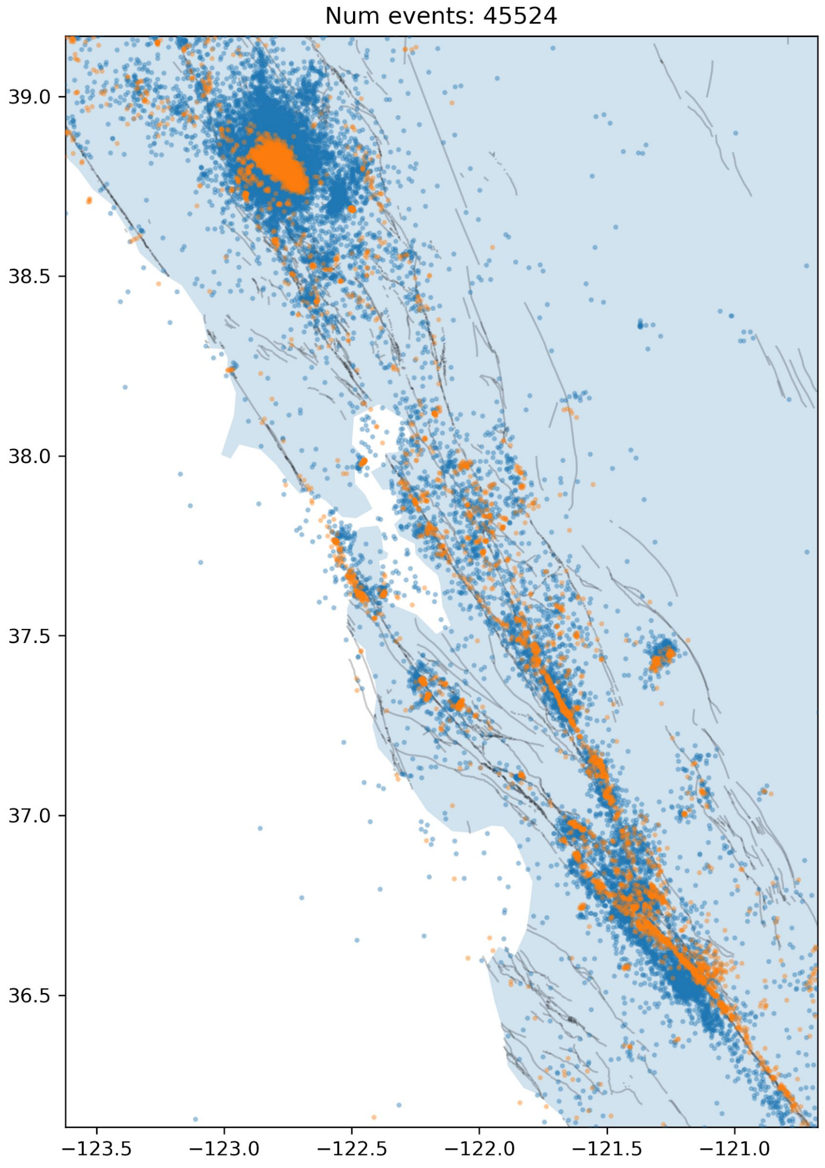
Initial Catalog (2023)





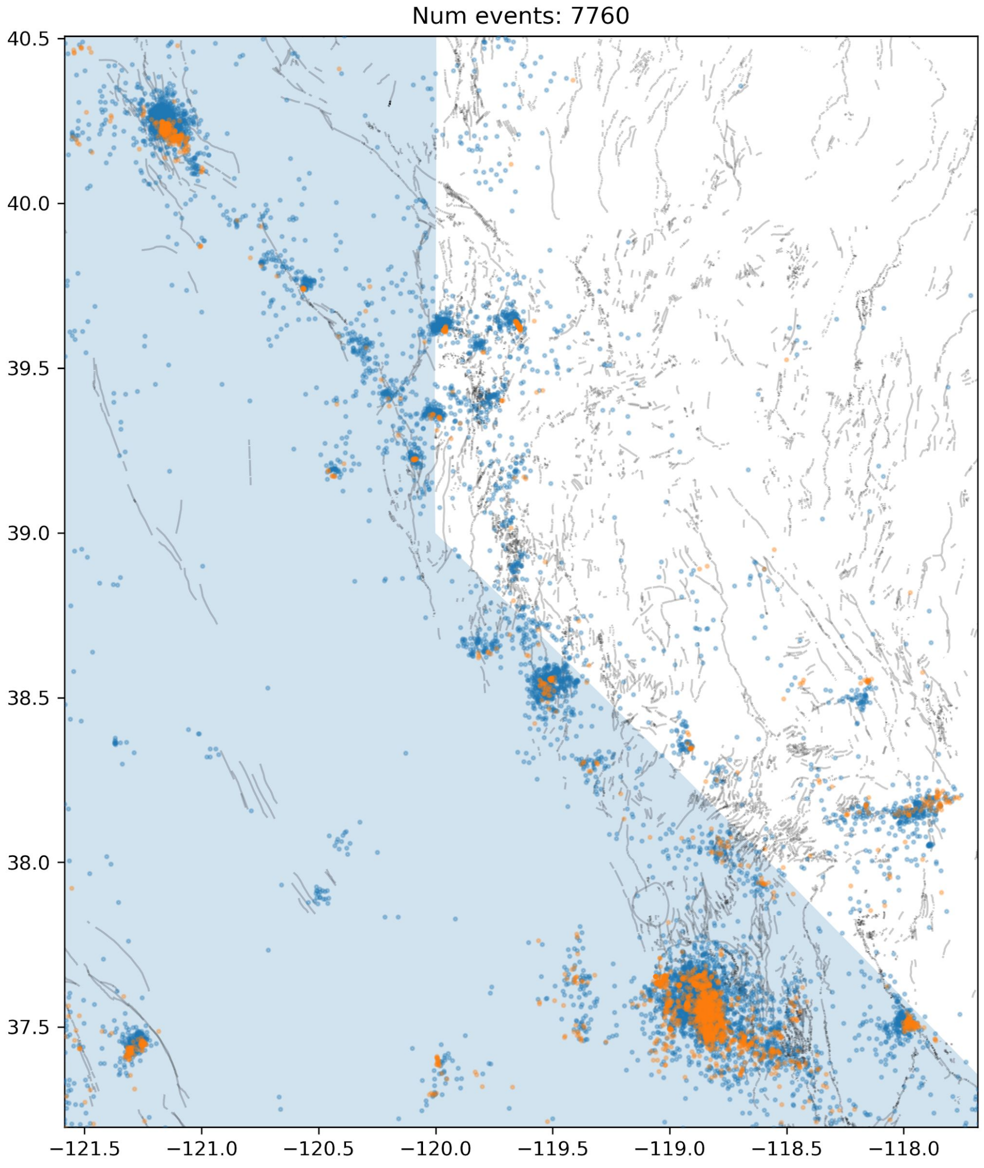
Example Catalog

Comparison of NCEDC (orange) and our initial catalog (blue)



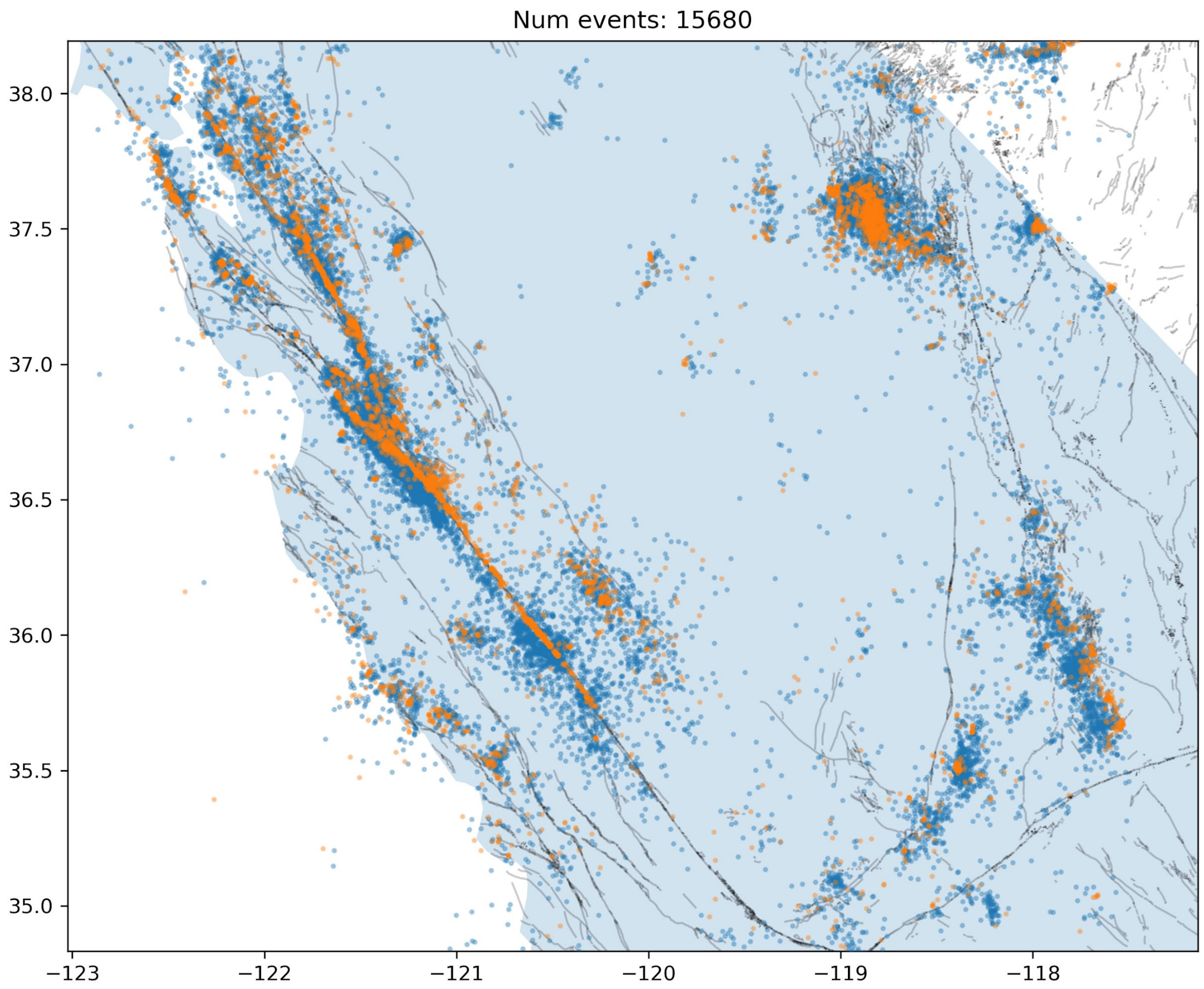
Example Catalog

Comparison of NCEDC (orange) and our initial catalog (blue)

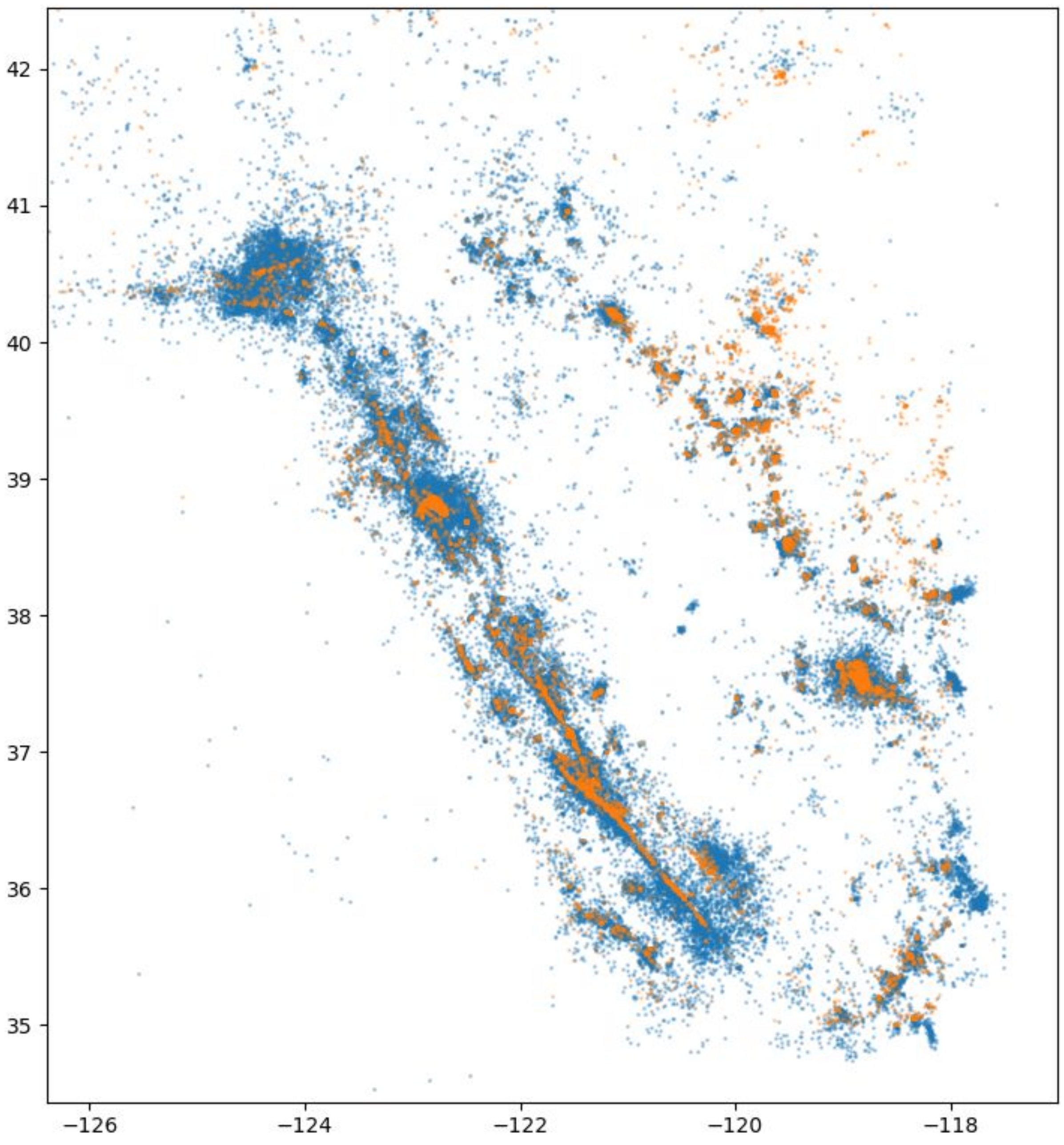


Example Catalog

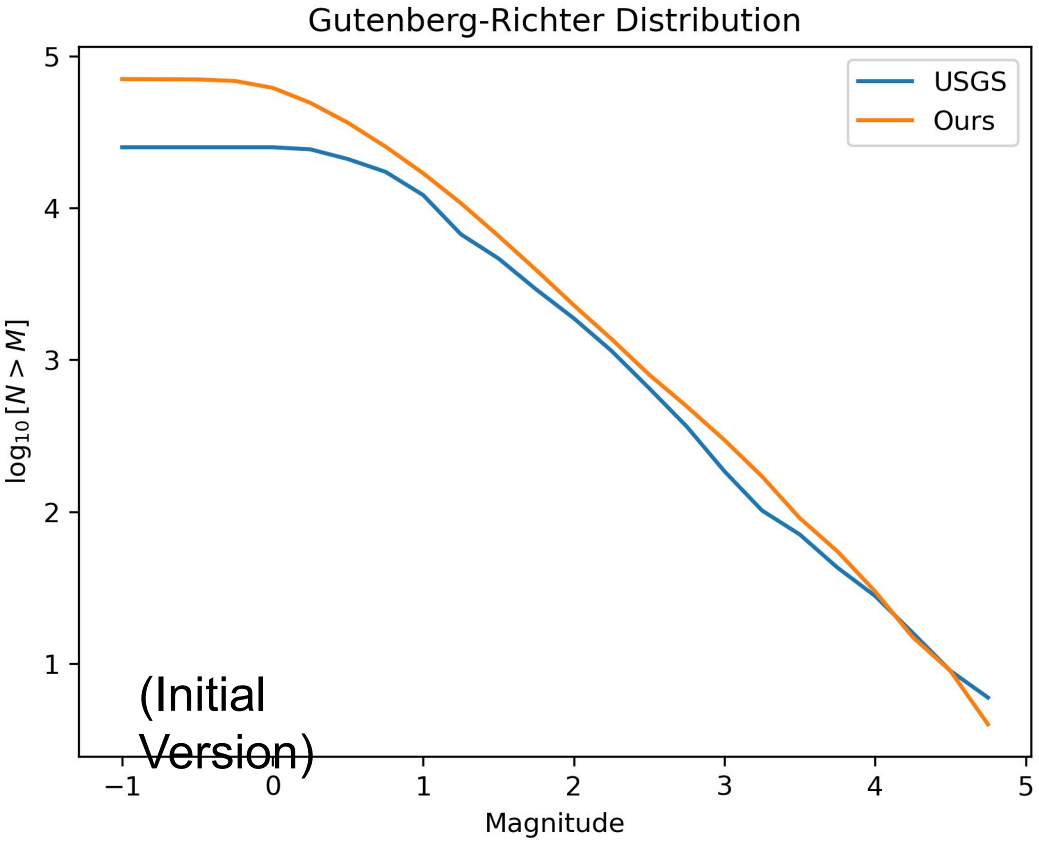
Comparison of NCEDC (orange) and our initial catalog (blue)



Increased Catalog (2023)

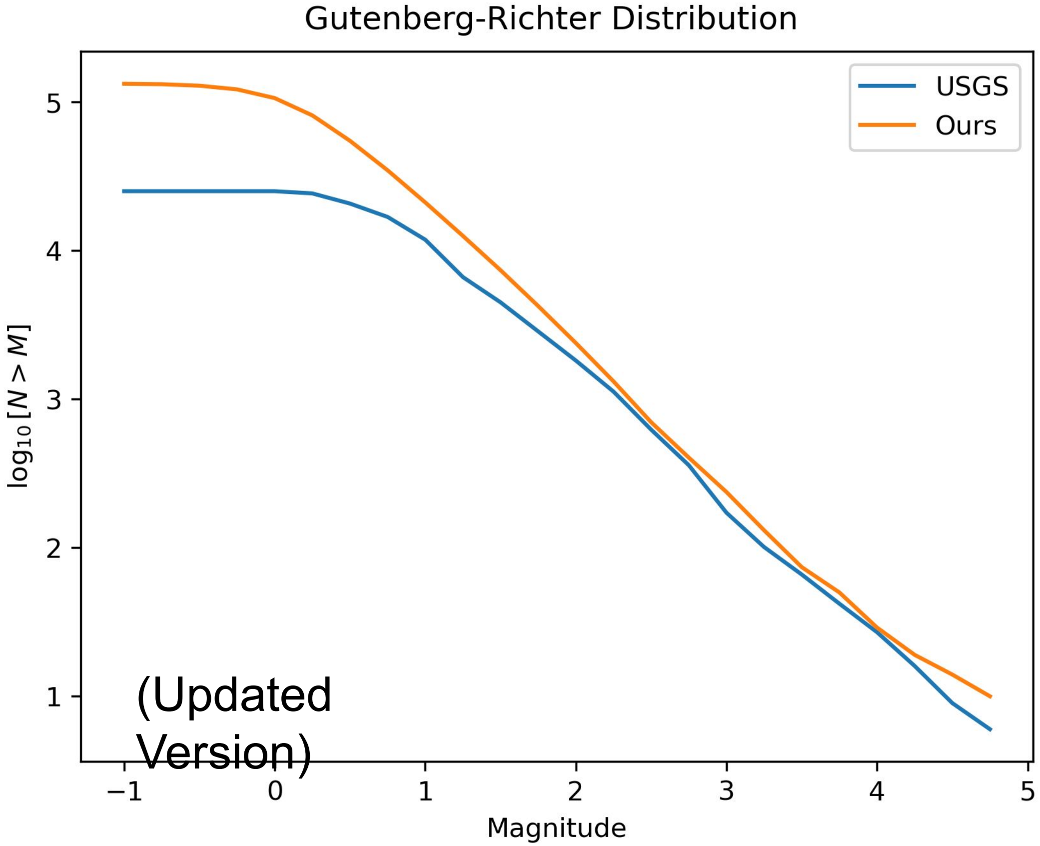


Event Counts



(Initial Version)

- 2.8x NCEDC catalog



(Updated Version)

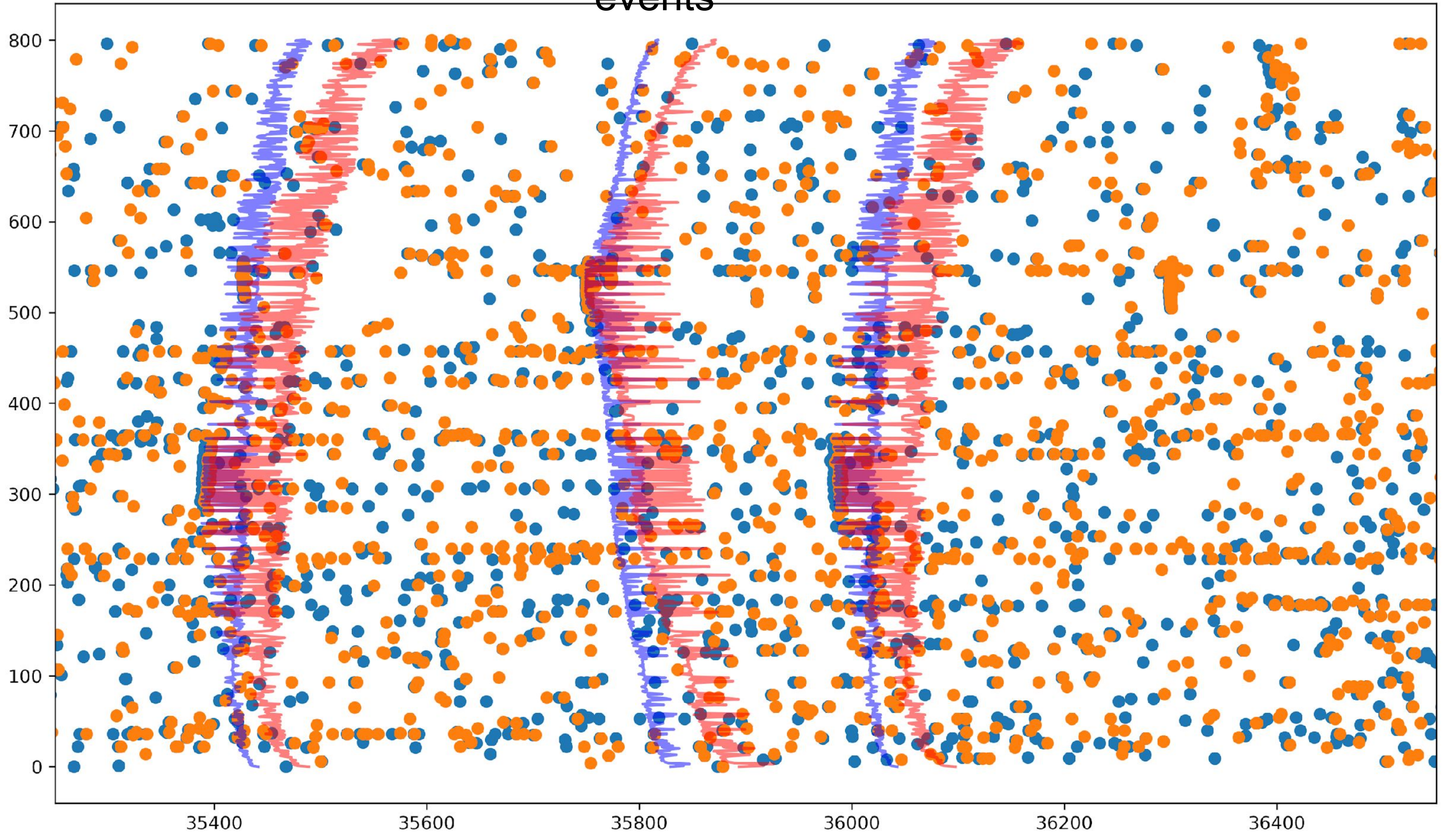
- 5.3x NCEDC catalog

Picks

Average:
240,000
picks per day

P-waves
121,000
picks per day

S-waves
117,000
picks per day



• Collaboration with Weiqiang Zhu to obtain PhaseNet picks

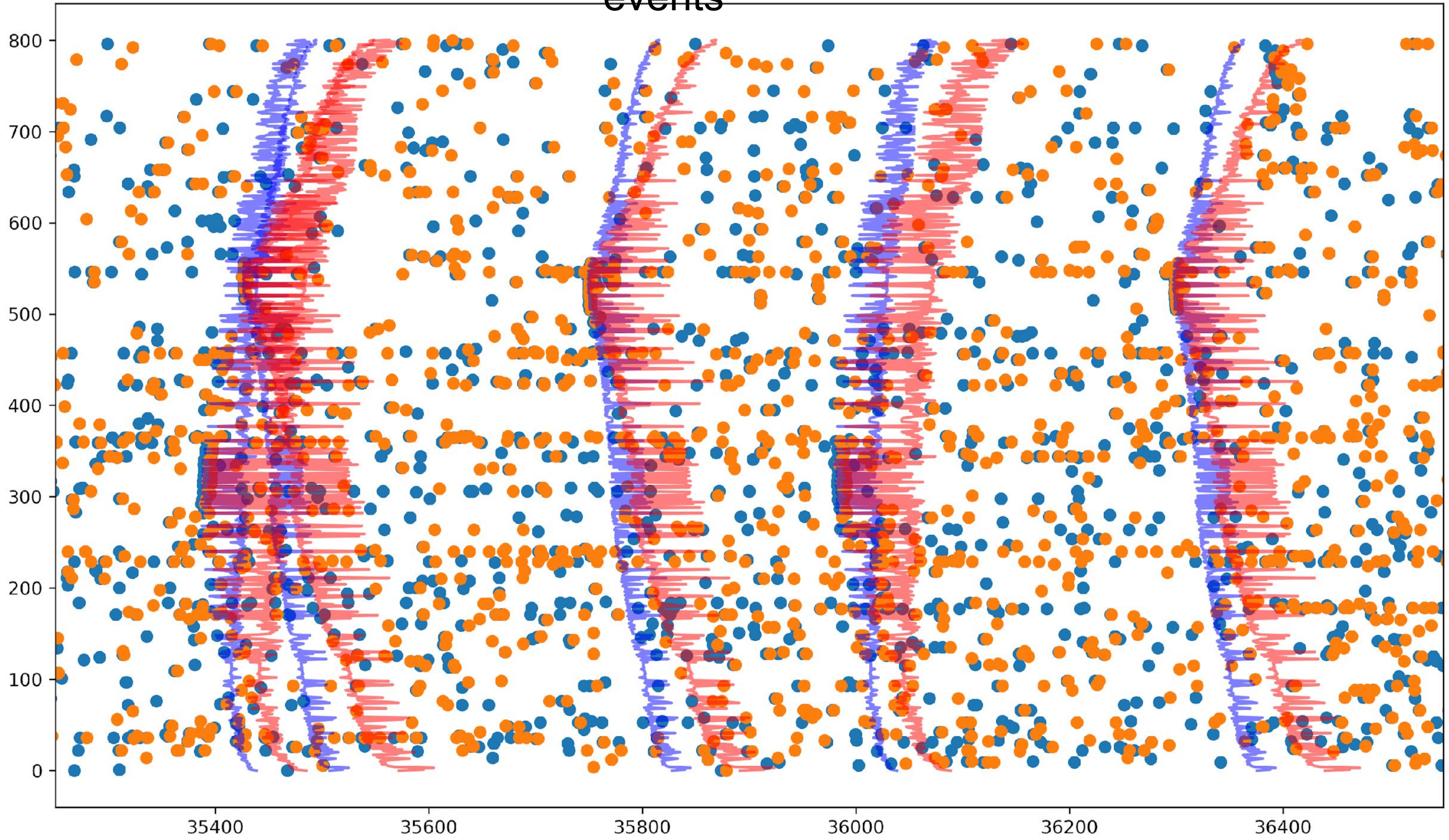
Picks

Average:
240,000
picks per day

P-waves
121,000
picks per day

S-waves
117,000
picks per day

GENIE
events



5
events

- Collaboration with Weiqiang Zhu to obtain PhaseNet picks

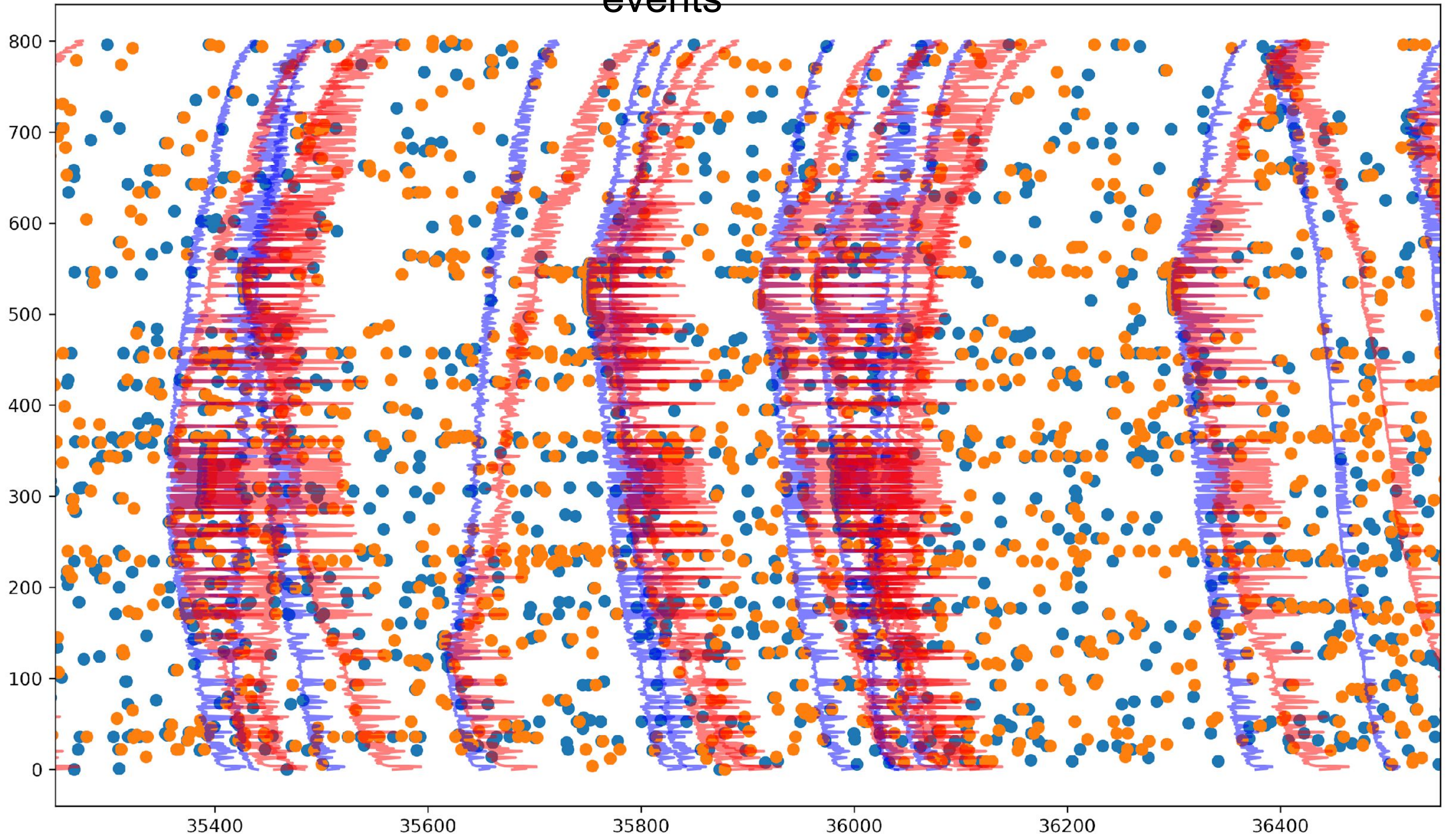
Picks

Average:
240,000
picks per day

P-waves
121,000
picks per day

S-waves
117,000
picks per day

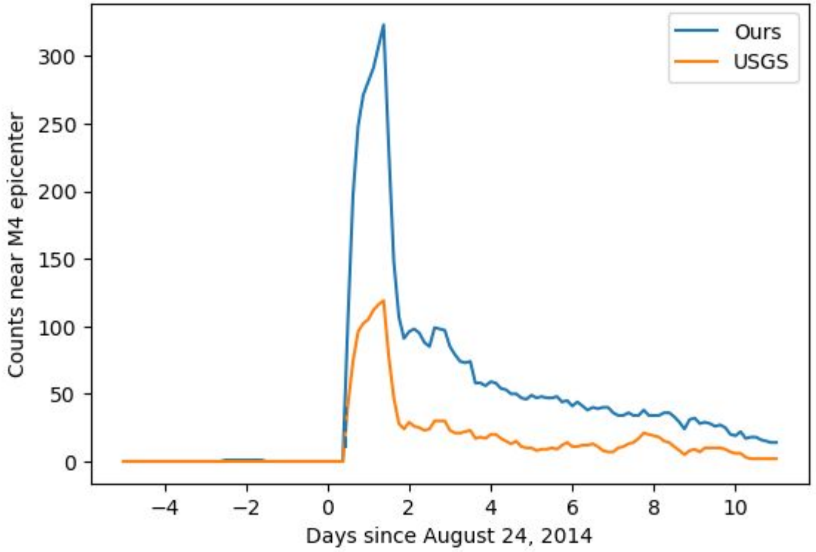
GENIE
events

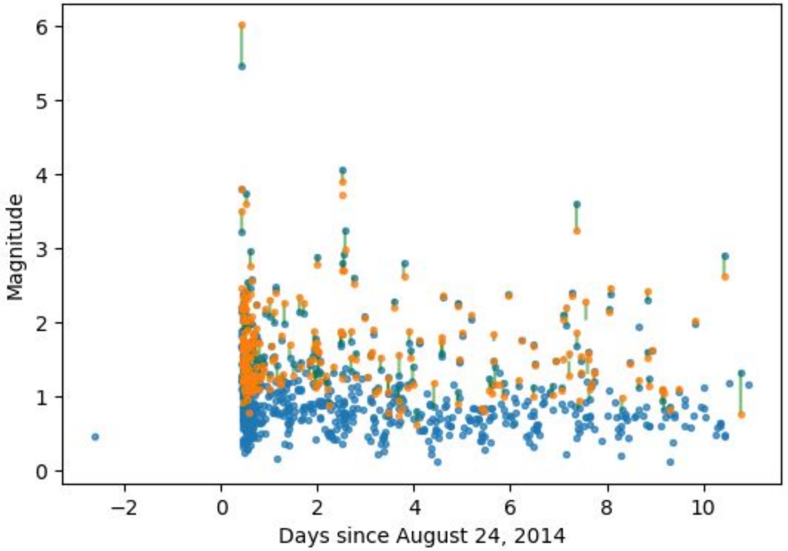


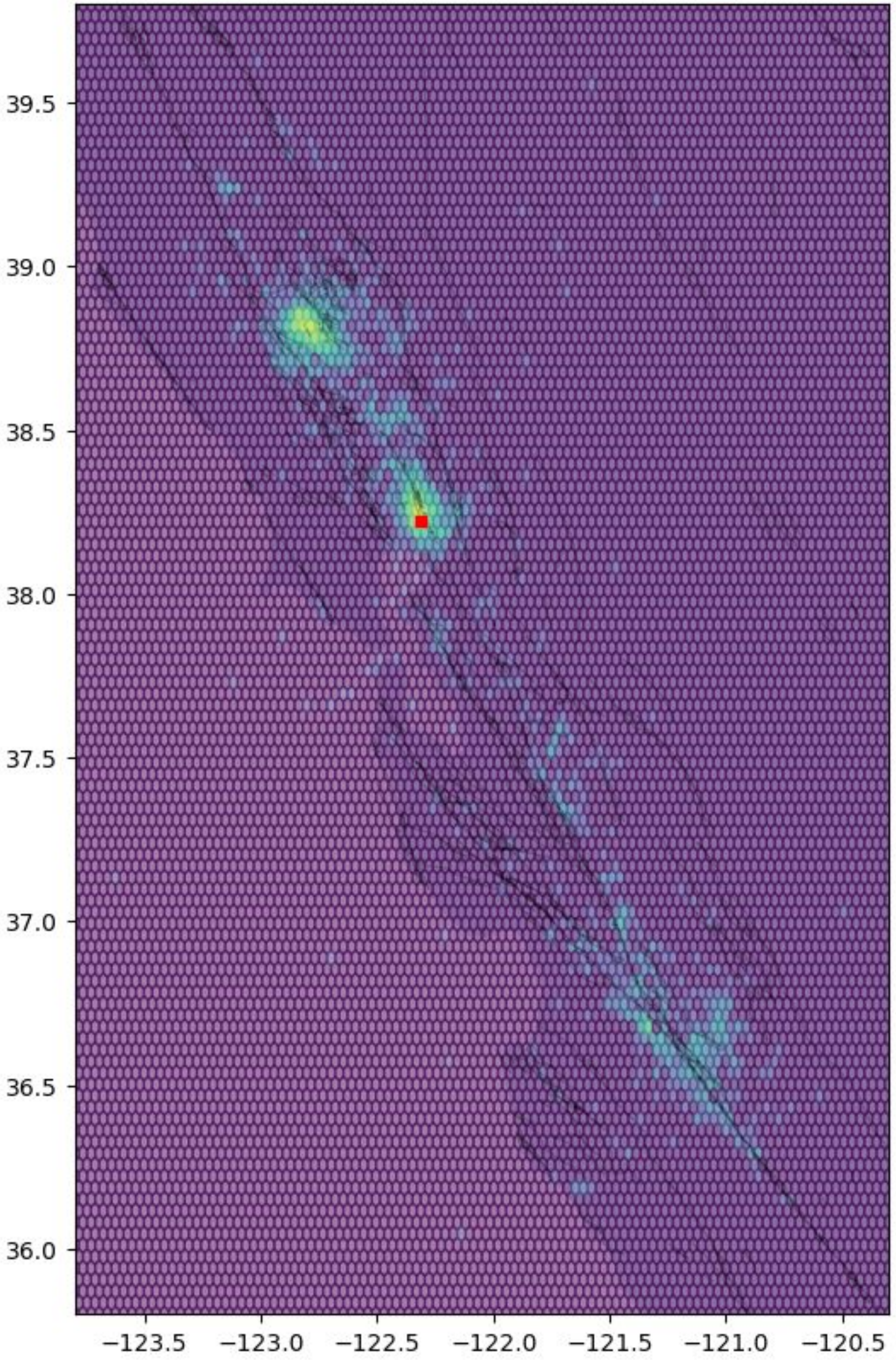
13
events

- Collaboration with Weiqiang Zhu to obtain PhaseNet picks

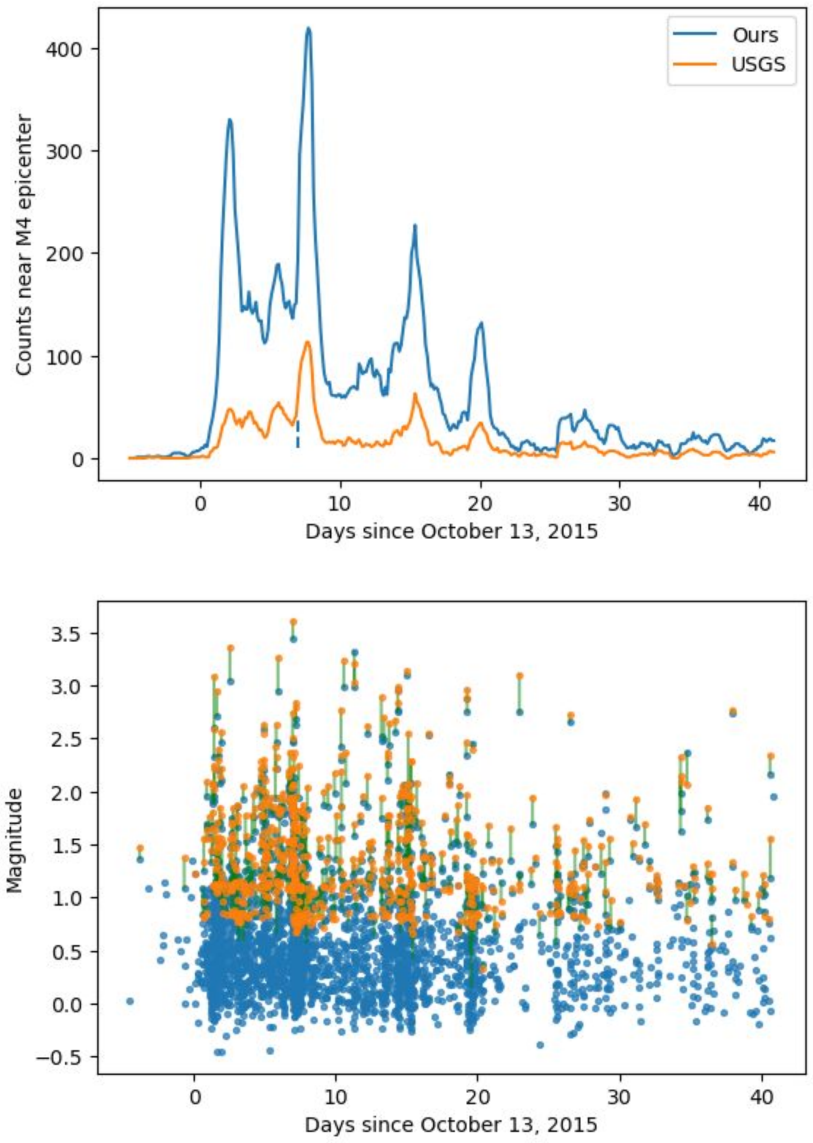
Mw 6.0 Napa Earthquake, 2014

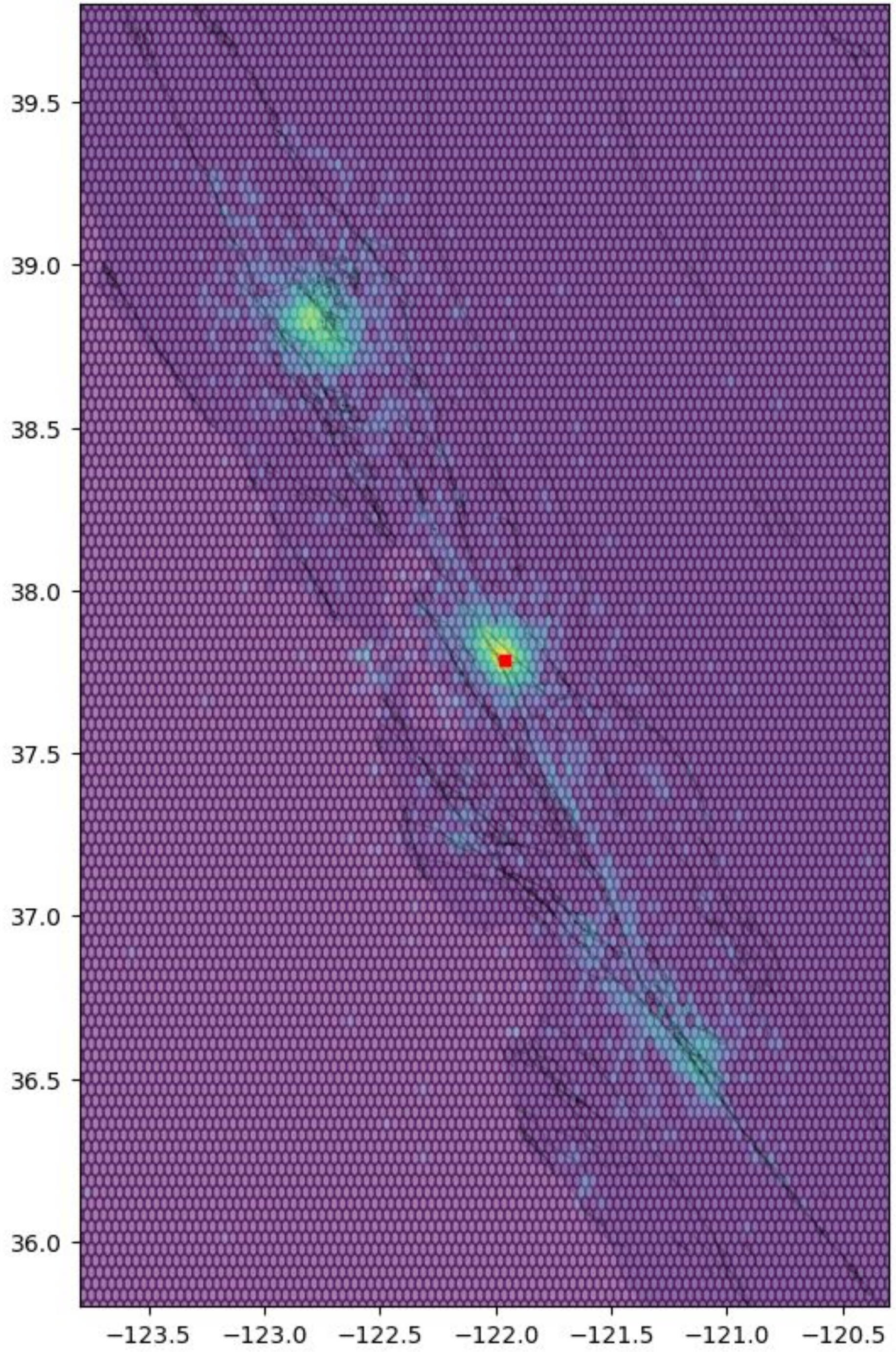




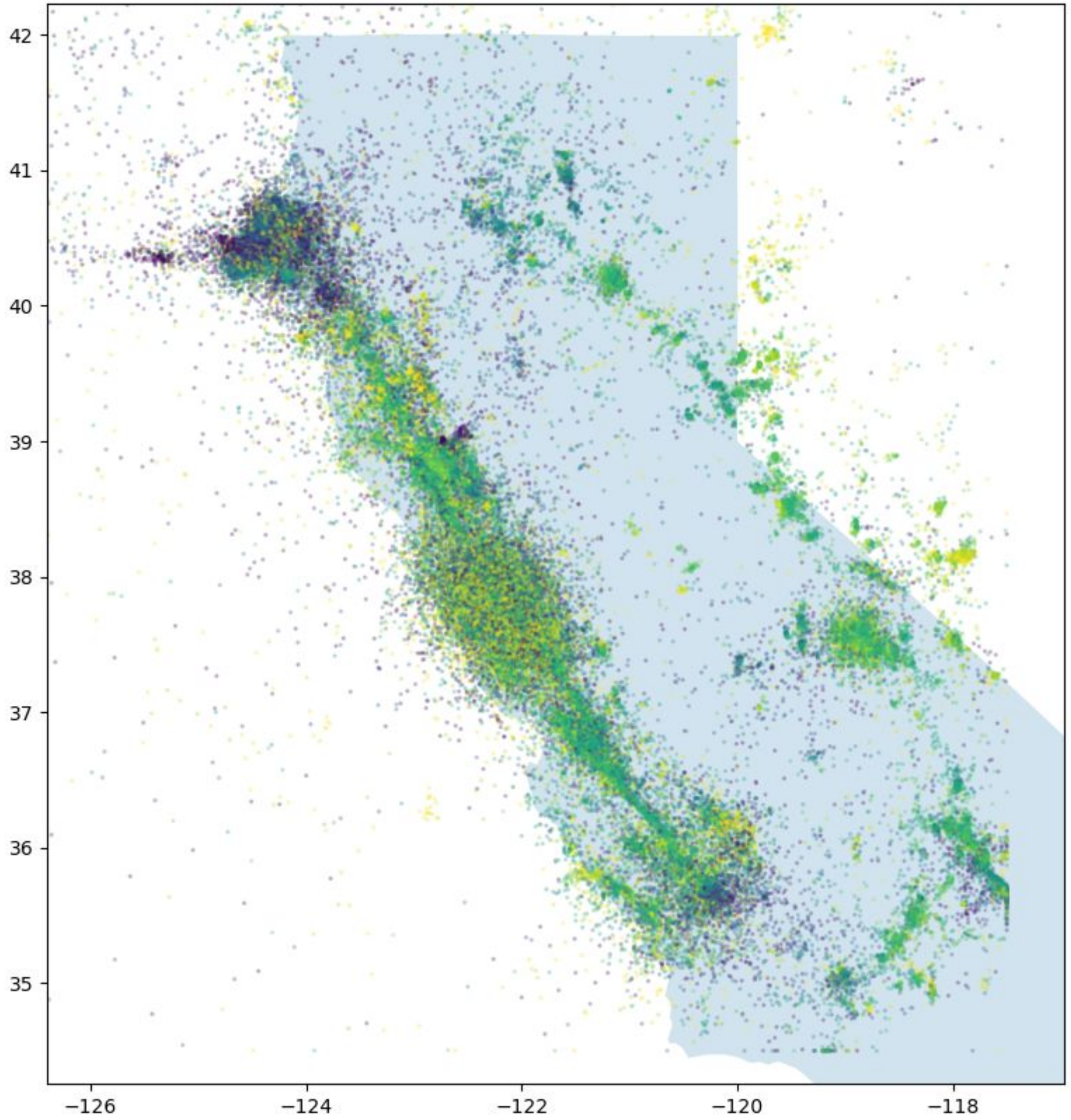


San Ramon Swarm, 2015





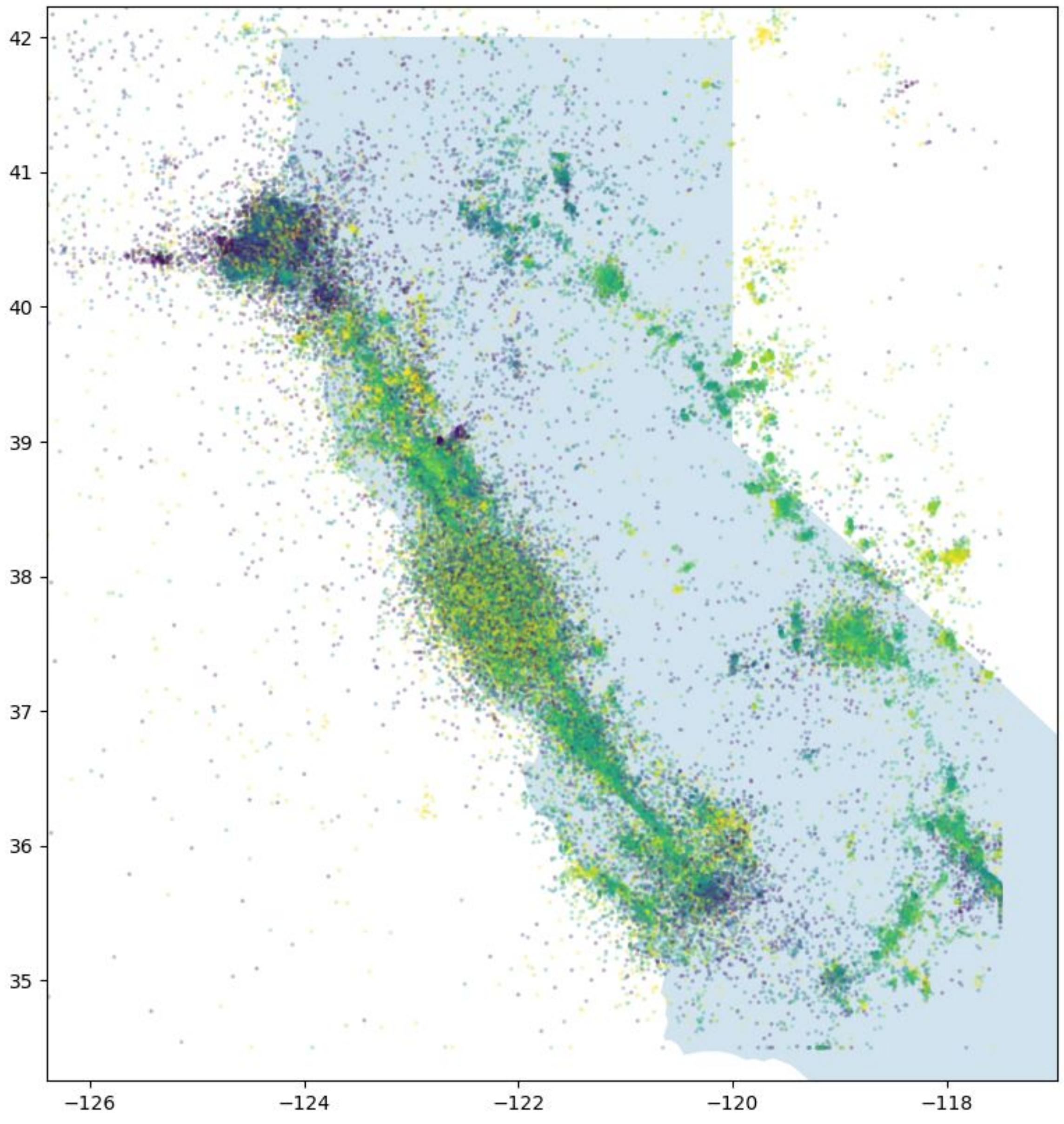
Bad Catalog



Bad Catalog

Training data too many events

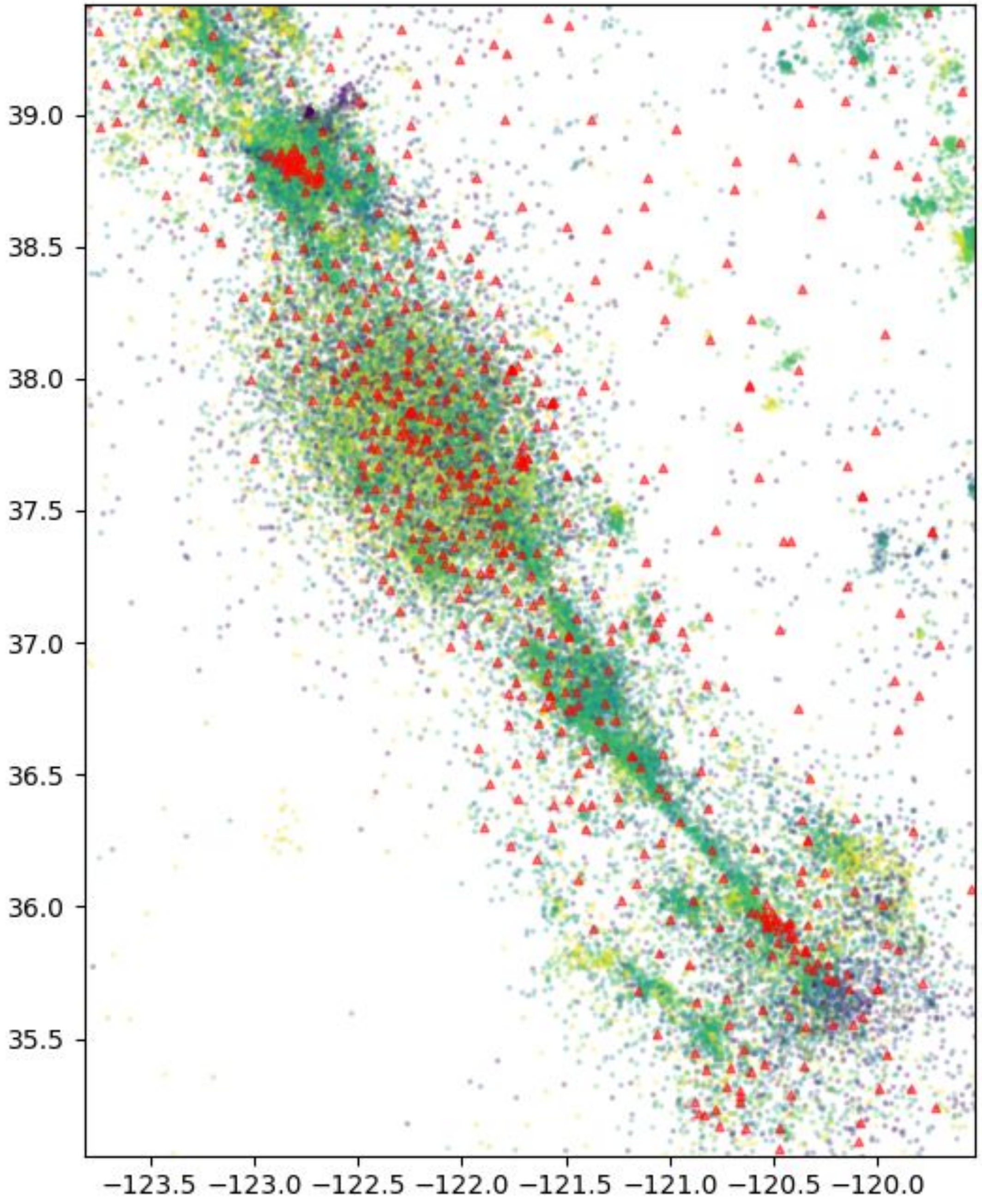
Also – set of associated picks in training, too random.



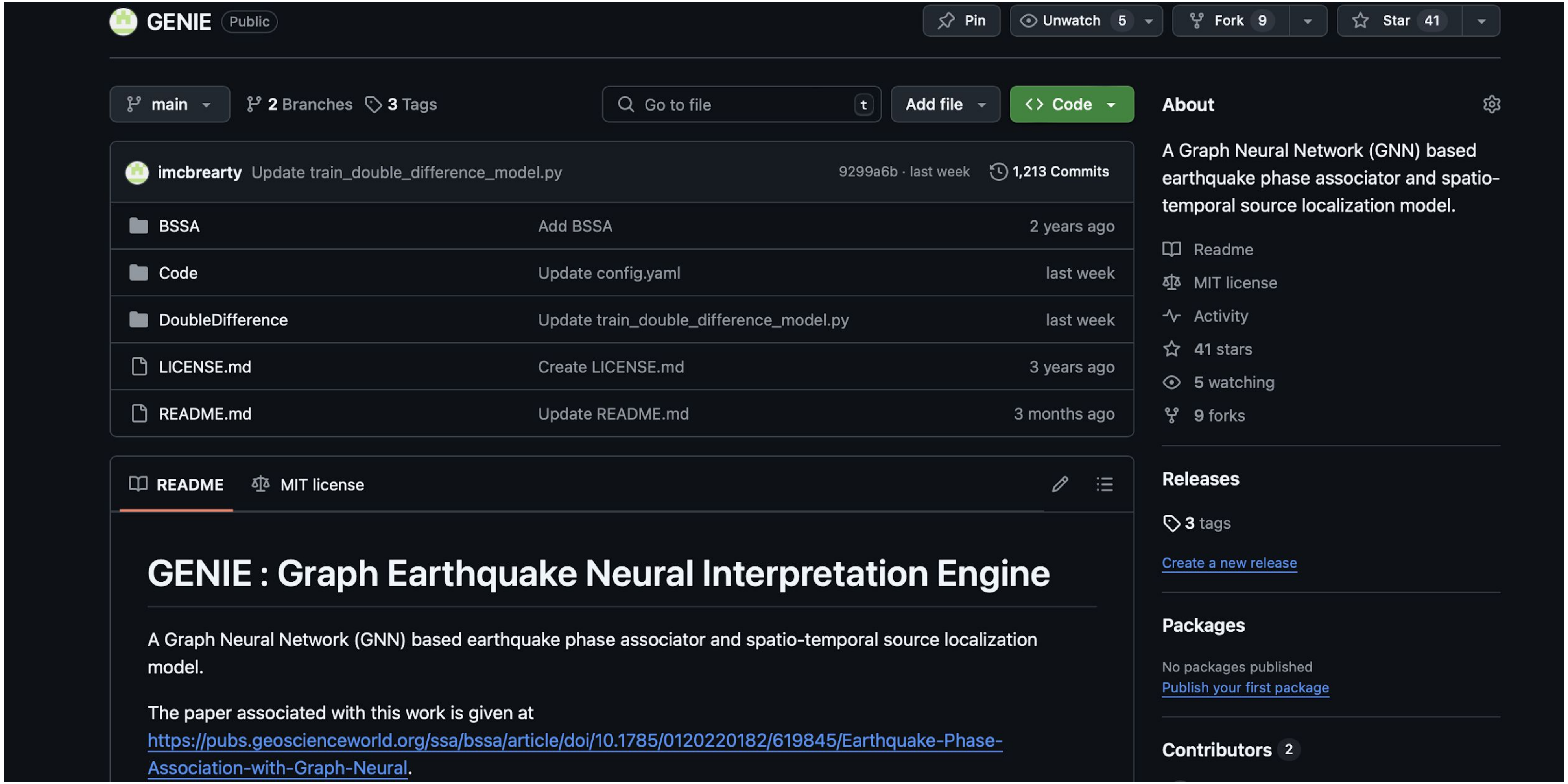
Bad Catalog

Training data too many events

Also – set of associated picks in training, too random.



How to use GENIE



github.com/imcbrearty/GENI
E

How to use GENIE

(1). Set region and station file

(2). Set velocity model

(3). Compute travel times

(4). Choose synthetic data parameters and train

(5). Apply

How to use GENIE

GENIE / Code / config.yaml

imcbrearty Update config.yaml

Code Blame 90 lines (73 loc) · 6.42 KB Code 55% faster with GitHub Copilot

```
name_of_project: 'Mayotte'
num_cores: 1 # How many cores would you like to use for travel time calculations? (it w
vel_model_ver: 1 ## Which travel time version to save when running calculate travel tim

## Note, when running continuous days processing a number of parameters are also set in

latitude_range: [18.8, 20.3] # Latitude range of the region that will be processed
longitude_range: [-156.1, -154.7] # Longitude range of the region that will be processe
depth_range: [-40000, 5000] # Note: depths are in meters, positive above sea level, neg
time_range: # This sets up the Catalog and Pick files to have these years initialized
  start: '2018-01-01'
  end: '2023-01-01'
```

Set
region

How to use GENIE

```
[In [14]: z = np.load('stations.npz')

[In [15]: list(z.keys())
 Out[15]: ['locs', 'stas']

[In [16]: print(z['locs'][0:10])
 [[  39.04446   -123.541092   687.        ]
  [  39.11709   -123.70883    144.        ]
  [  39.12745   -122.82347    858.        ]
  [  39.12997   -123.07651   1077.        ]
  [  39.133171 -123.46788    370.        ]
  [  39.17902   -122.63618    975.        ]
  [  39.1853    -123.2109     193.        ]
  [  39.20074   -123.63514    327.        ]
  [  39.205704 -123.301003   654.        ]
  [  39.30477   -123.19748    264.7       ]]

[In [17]: print(z['stas'][0:10])
 ['GHO.NC' 'GBL.NC' 'GHGB.NC' 'GCWB.NC' 'GMR.NC' 'GSR.NC' '79666.CE'
  'GNR.NC' 'GWR.NC' 'BARR.BK']
```

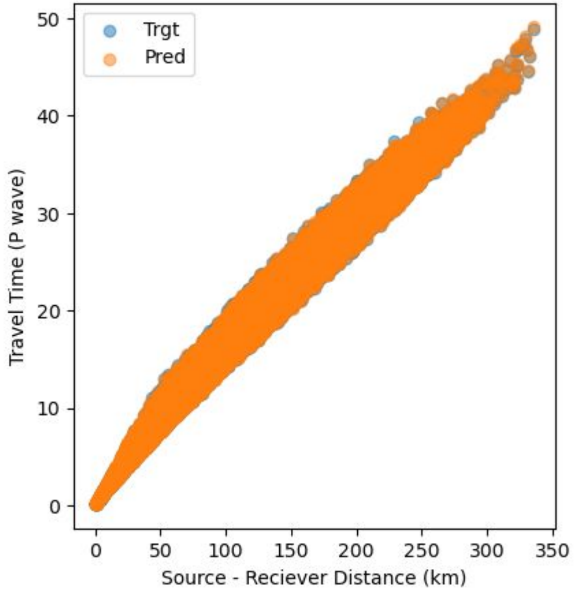
Set
stations

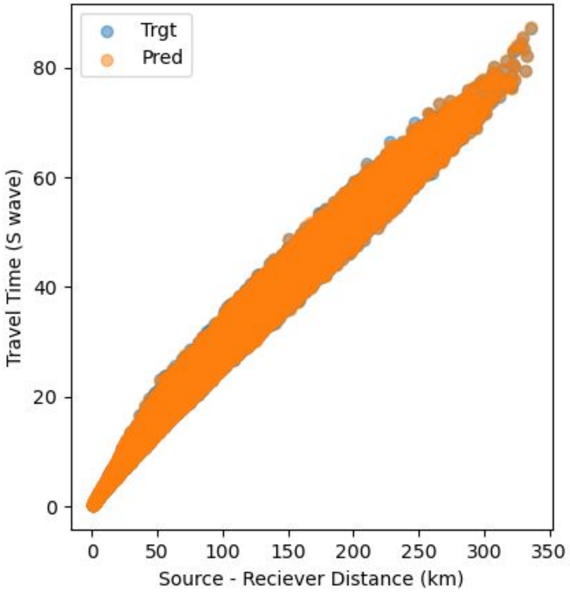
How to use GENIE

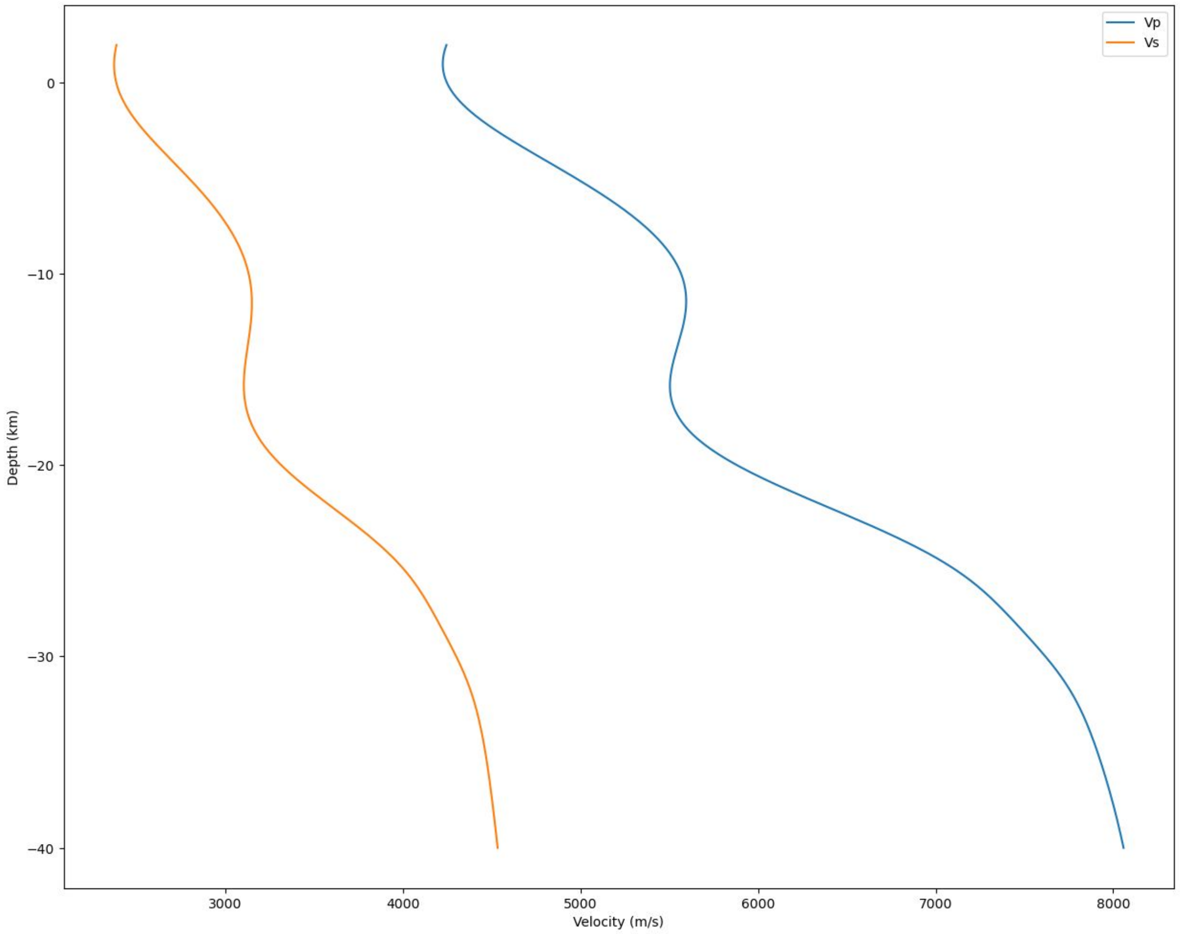
```
[In [30]: z = np.load('3d_velocity_model.npz')

[In [31]: list(z.keys())
 Out[31]: ['X', 'Vp', 'Vs']

[In [32]: print(z['X'][0:10,:])
 [[ 3.94651464e+01 -1.26163064e+02 -4.14937689e+04]
  [ 3.94652337e+01 -1.26162896e+02 -4.09940675e+04]
  [ 3.94653211e+01 -1.26162729e+02 -4.04943661e+04]
  [ 3.94654085e+01 -1.26162561e+02 -3.99946646e+04]
  [ 3.94654958e+01 -1.26162393e+02 -3.94949631e+04]
  [ 3.94655831e+01 -1.26162226e+02 -3.89952615e+04]
  [ 3.94656705e+01 -1.26162058e+02 -3.84955599e+04]
  [ 3.94657578e+01 -1.26161891e+02 -3.79958582e+04]
  [ 3.94658450e+01 -1.26161723e+02 -3.74961565e+04]
  [ 3.94659323e+01 -1.26161556e+02 -3.69964547e+04]]

[In [33]: print(z['Vp'][0:10])
 [8027.197  8028.1963 8029.197  8029.8975 8029.9995 8029.9995 8029.9995
  8029.9995 8029.9995 8030.2964]

[In [34]: print(z['Vs'][0:10])
 [4509.899  4510.299  4510.8984 4511.0005 4511.0005 4511.0005 4511.
  4511.     4511.     4511.2964]
```

Set velocity
model

How to use GENIE

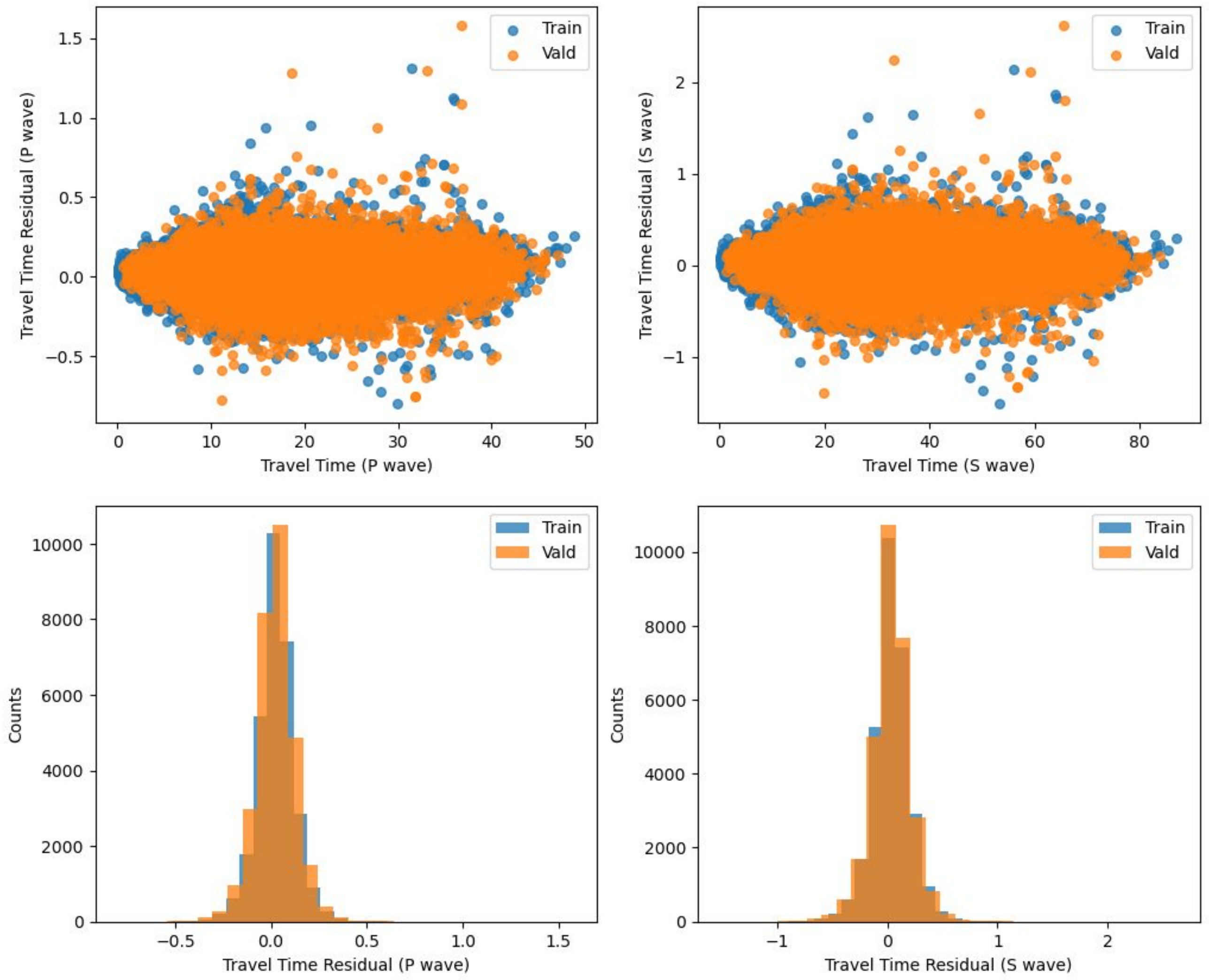






Train travel time PINN neural network

How to use GENIE



Accurate even for 3D velocity models

Train travel time PINN neural network

How to use GENIE

Set scale and event rate dependent training parameters

Set training data

```
## Prediction params
## These parameters should somewhat scale with the size of the application
kernel_sig_t: 3.5 # Kernel to embed arrival time - theoretical time misfit (
src_t_kernel: 3.5 # Kernel of origin time label (s)
src_t_arv_kernel: 3.5 # Kernel for arrival association time label (s)
src_x_kernel: 30000. # Kernel for source label, horizontal distance (m)
src_x_arv_kernel: 30000. # Kernel for arrival-source association label, hori
src_depth_kernel: 30000. # Kernel of source label in Cartesian projection, v

## Training params list 2
spc_random : 3500 # Spatial scale to randomly remove true picks from station
sig_t : 0.03 # Percent of travel time error on pick times (e.g., 3%)
spc_thresh_rand : 3500 # Spatial scale to randomly shift threshold distance
min_sta_arrival : 6 # Min number of unique stations required for a positive
coda_rate : 0.1 # Percent of picks with false coda picks (e.g., 3.5%)
coda_win : [0, 10.0] # Window that code picks can occur over (e.g., 25 s)
max_num_spikes : 2 # Number of possible network wide spikes per window T of
spike_time_spread : 0.15 # The temporal spread of the network wide spikes
s_extra : 0.0 # If this is non-zero, it can increase (or decrease) the total
use_stable_association_labels : True # This flag only allows positive associ
thresh_noise_max : 2.5 # ratio of sig_t*travel time considered excess noise
min_misfit_allowed: 1.0 # The minimum time (in seconds), beneath which, diff
total_bias: 0.03 ## Total possible bias on travel times (uniform across stat
# training_params_2 = [spc_random, sig_t, spc_thresh_rand, min_sta_arrival,

## Training params list 3
dist_range : [20000, 225000] # This is the distance range over which to simu
max_rate_events : 200 # 350 # 450 # Average rate of events per T window of t
max_miss_events : 204 # 225 # 325 # Average rate of missed picks per station
max_false_events : 2.25 # Now by default represents the ratio of false picks
miss_pick_fraction : [0.05, 0.35] # False # Average ratio of missed picks (i
T : 10800 # Time window to simualate synthetic data. More variability occurs
dt : 30 # Time resolution to allow synthetic data parameters to vary in time
tscale : 3600 # Time scale that synthetic data parameters vary in time, duri
n_sta_range : [0.75, 1.0] # The ratio of possible stations from full set con
use_sources : False
use_full_network : False
fixed_subnetworks : True ## If True, this uses realistic sets of stations av
use_preferential_sampling : True ## This concentrates more of the samples ar
use_extra_nearby_moveouts : True ## This up-samples the amount of sources wi
use_shallow_sources : False
```