Differential arrival times for source location with DAS arrays: tests on data selection and automatic weighting procedure

Emanuele Bozzi[1] (Speaker) N. Piana Agostinetti[1], G. Saccorotti[2], A. Fichtner[3], Lars Gebraad[3,4], Tjeerd Kiers[3], and Takeshi Nishimura[5]

[1] University of Milano Bicocca, Environmental and Earth Sciences Department (DISAT) , Milan, Italy (e.bozzi3@campus.unimib.it)

[2] Istituto Nazionale di Geofisica e Vulcanologia (INGV), Pisa, Italy

[3] ETH Zürich, Department of Earth Sciences, Institute of Geophysics, Zürich, Switzerland

[4] Mondaic AG, Zürich, Switzerland

[5] Tohoku University, Department of Geophysics, Tohoku, Japan

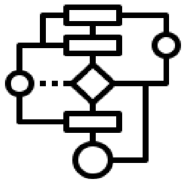
Vienna, EGU General Assembly, April 17, 2024.
SM 3.1 session

Outline

Information redundancy with DAS

Differential Arrival Times (DATs) for event location

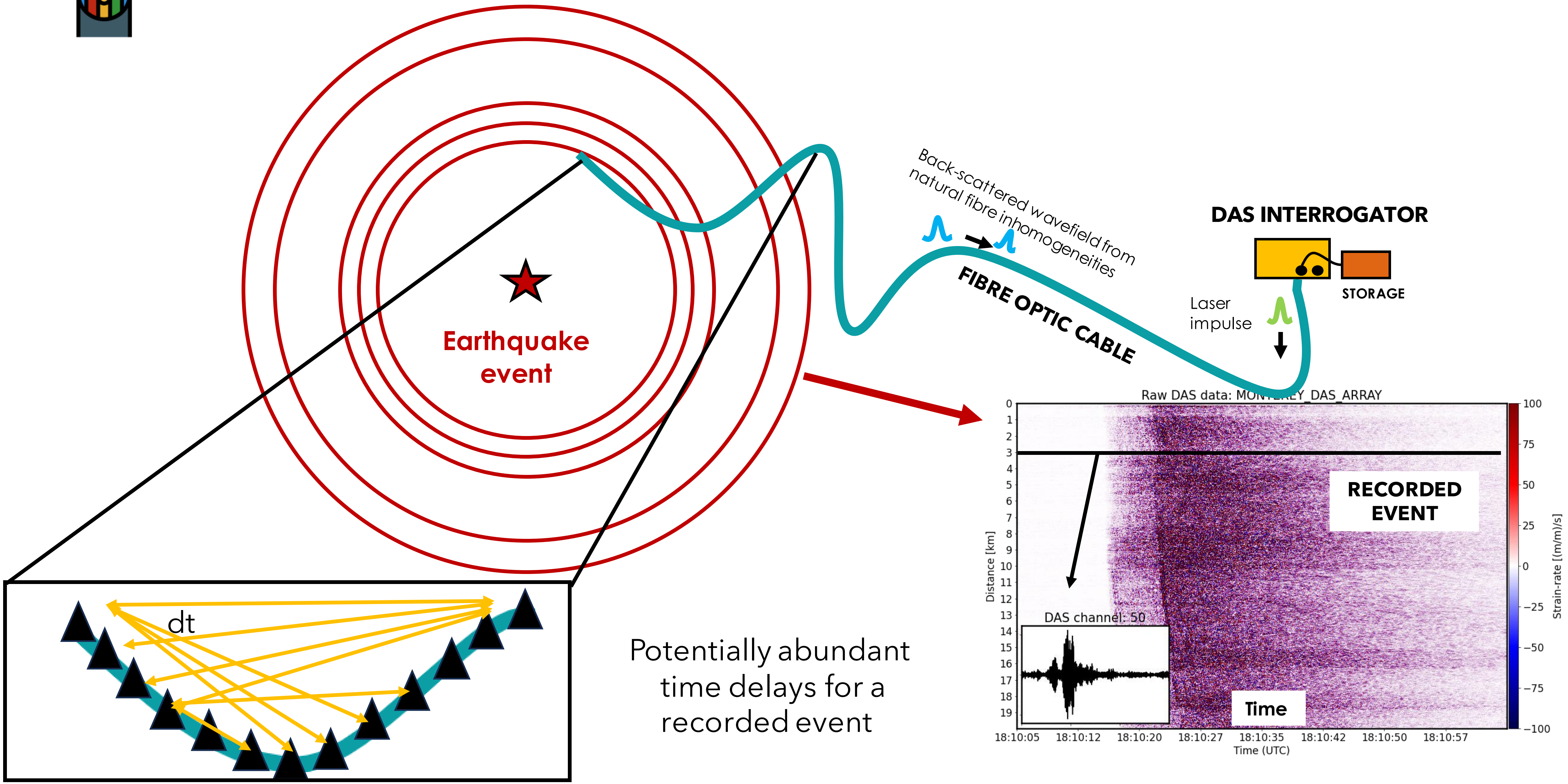
Weighting DATs (Hierarchical McMC)



Synthetic tests

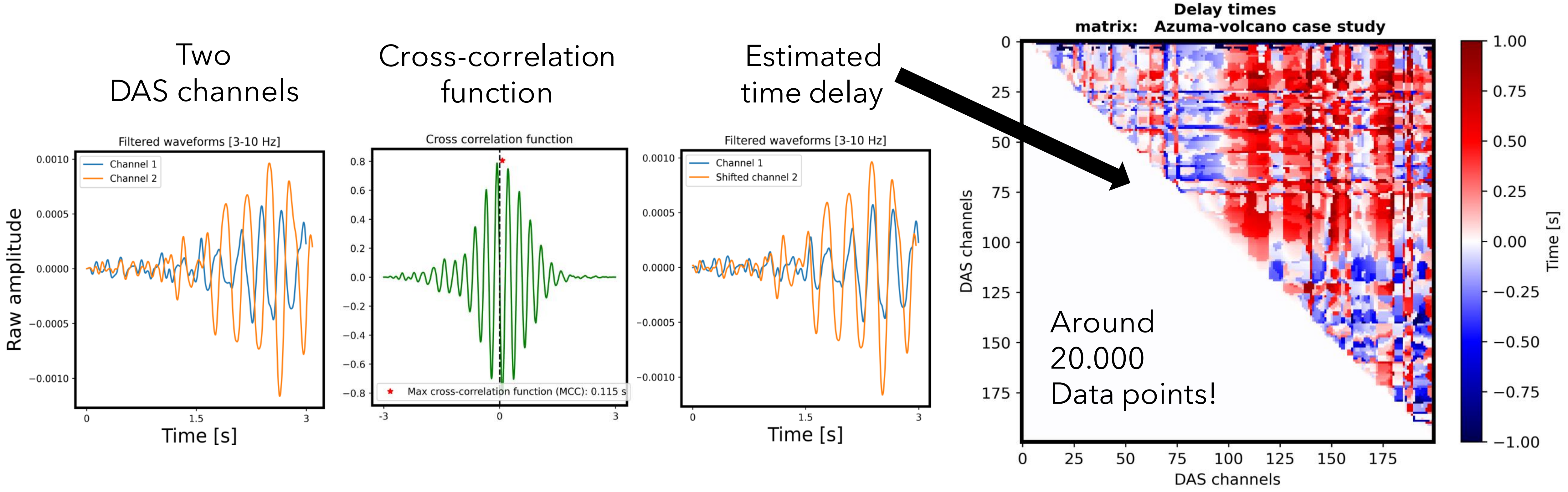
Applications to real data

Information redundancy with DAS



Potentially abundant time delays for a recorded event

Differential arrival times for event location

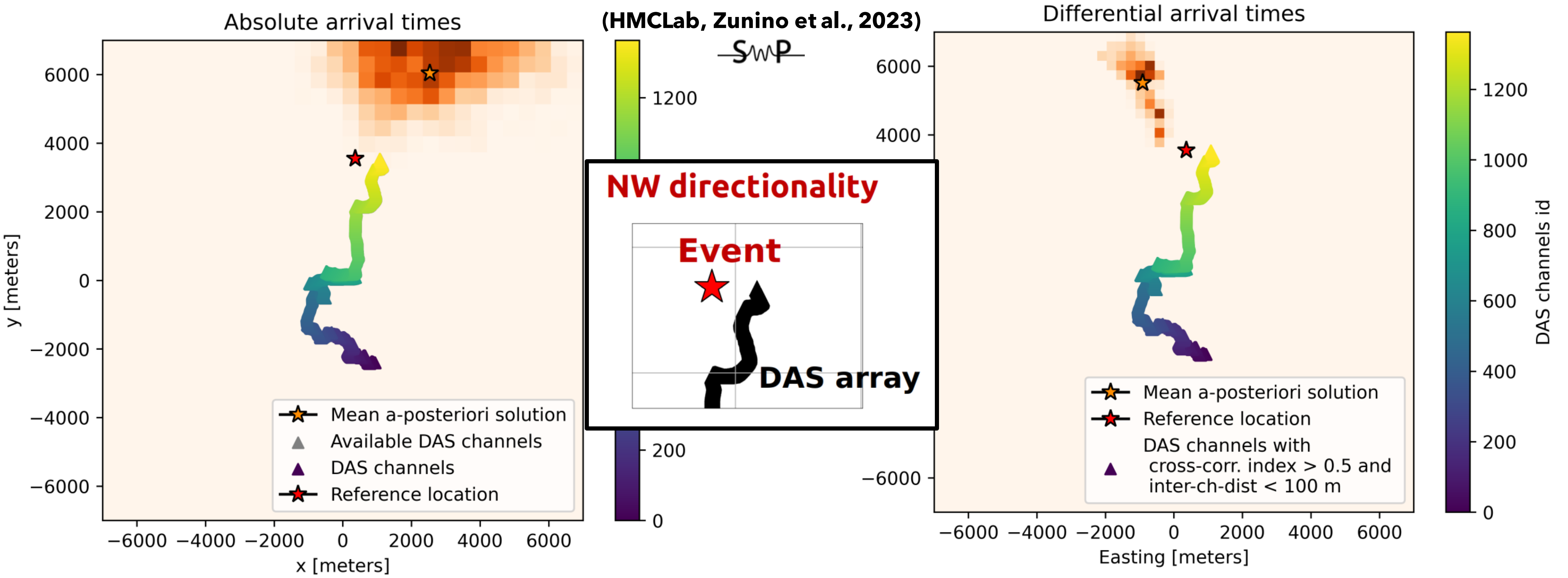


- Spatial variability of DAS waveforms + huge number of available DAS channel couples (no manual check)
- Inverting all DATs/assigning equal weight may lead to poor event location accuracy

The goal >> Test DATs selection or differential weighting in the inversion process

(First test) Selecting P-DATs

P-wave Absolute Arrival Times (P-AATs) | Selected P-DATs (same channels P-AATs)



Selected P-DATs (**Maximum value of the Cross-Correlation (MCC) function and interchannel Distance (INTER-DIST)**) >> better constrain on event location directionality (NW), compared to absolute arrival times.

This gave us an idea for P-DATs weighting ...

(Second test) Weighting P-DATs

- **MCC** and **INTER-DIST** seem **good candidates for weighting** differently the **time delays.**
- We adopted a similar procedure described in Piana Agostinetti et al., 2023 (**hierarchical McMC**)
- **Weighting entries in the covariance matrix** in the inversion scheme

$$C_e(m) = W^{-1}(m)\, C_e^*\, W^{-1}(m)$$

Channel pairs	MCC	INTER-DIST
1 <> 2	0.9	10
1 <> 3	0.6	20
1 <> 4	0.3	30
2 <> 3	0.4	10
2 <> 4	0.7	20
3 <> 4	0.9	10

[**hc(m)** + **hd(m)** + **(h0(m))**]***Ce**

Differential weights

Data features (MCC-INTER DIST) defining the error models (**hc(m)** and **hd(m)**)

Coherent weights

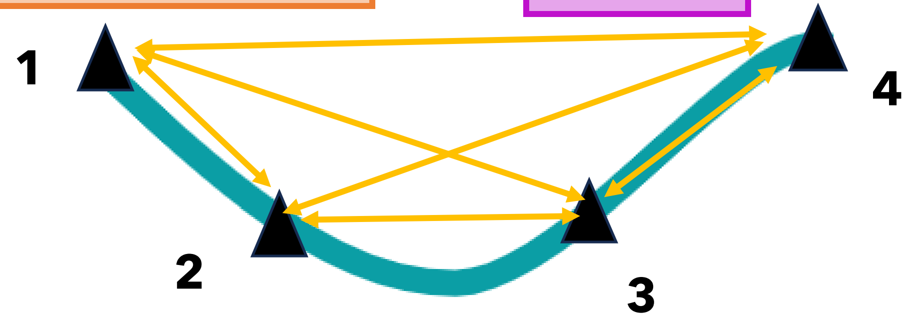
h0 (m), NOT defined by data features

Total weight for each phase (data point, time delay) **W^{-1}(m) = 10** (hc(m) + hd(m) + h0(m))**

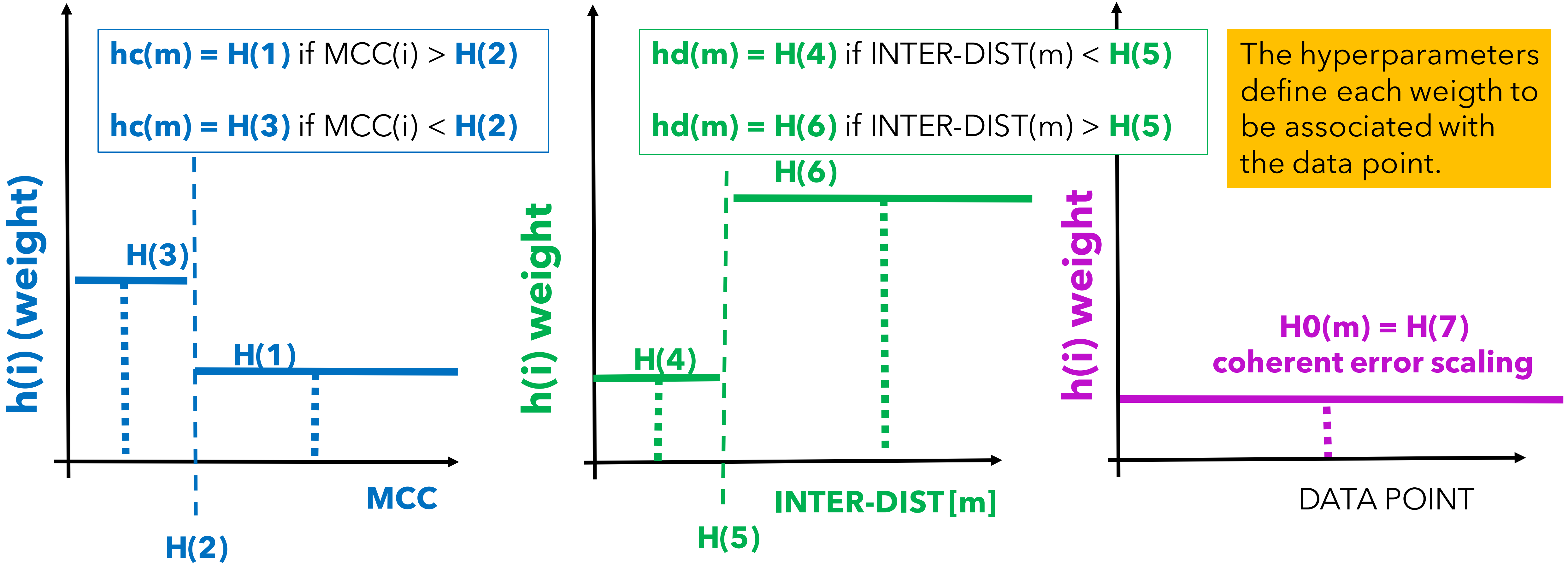
From MCC

From INTER-DIST

Coherent error scaling



Noise models: hc(m), hd(m) and h0(m)



$W^{-1}(m) = 10^{**}$ (hc(m) + hd(m) + h0(m))

H(1), H(2), H(3), H(4), H(5), H(6) and **H(7)** are hyperparameters sampled within a Markov Chain Monte Carlo approach, together with model parameters (event easting, northing, depth).

Weighting P-DATs

An example (e.g., a sampled model in the McMC):

- H(1) = 0.2 (Lower weight MCC)
- H(2) = 0.55 (Thr. MCC)
- H(3) = 0.5 (Upper weight MCC)
- H(4) = 0.2 (Lower weight INTER-DIST)
- H(5) = 15 m (Thr. INTER-DIST)
- H(6) = 0.5 (Upper weight INTER-DIST)
- H(7) = 0.2 (Coherent error scaling)

Channel pairs	MCC	INTER-DIST		hc(m) + hd(m)		h0(m)	
1 <> 2	0.9	10	=	[0.9	+	0.5]	*Ce
1 <> 3	0.6	20		1.5		0.5	
1 <> 4	0.3	30		1.5		0.5	
2 <> 3	0.4	10		1.2		0.5	
2 <> 4	0.7	20		1.2		0.5	
3 <> 4	0.9	10		0.9		0.5	

"More important" (green)

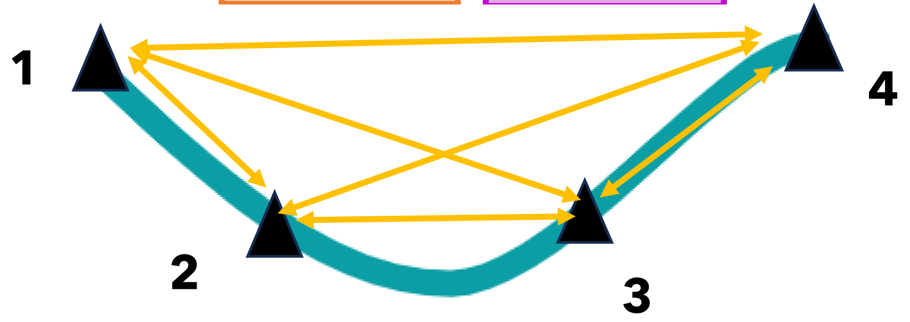
"Less Important" (red)

Total weight for each phase (data point, time delay) $\mathbf{W^{-1}(m) = 10^{**}(hc(m) + hd(m) + h0(m))}$

From MCC

From INTER-DIST

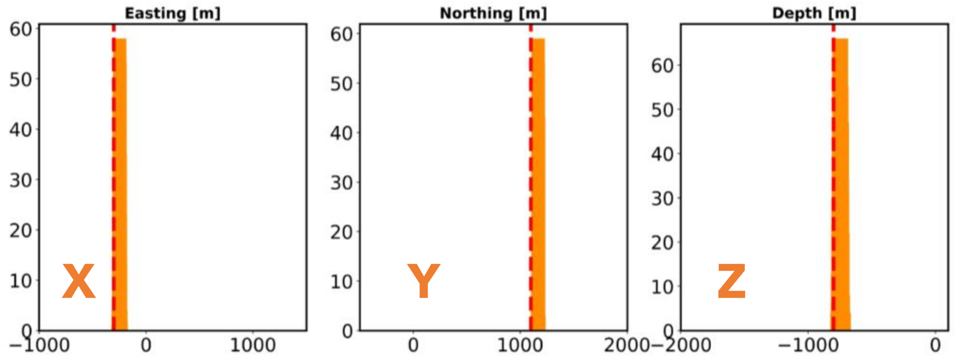
Coherent error scaling



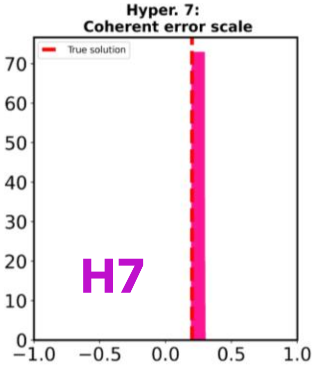
Synthetic tests

Model parameters + H7 (coherent error)

Model parameters

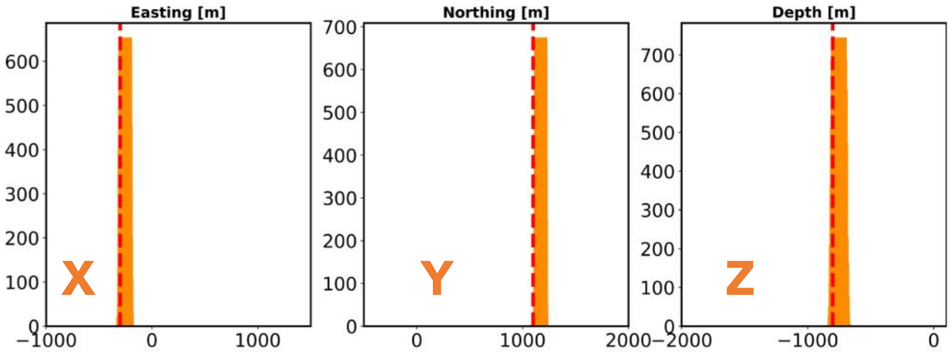


Hyperparameters

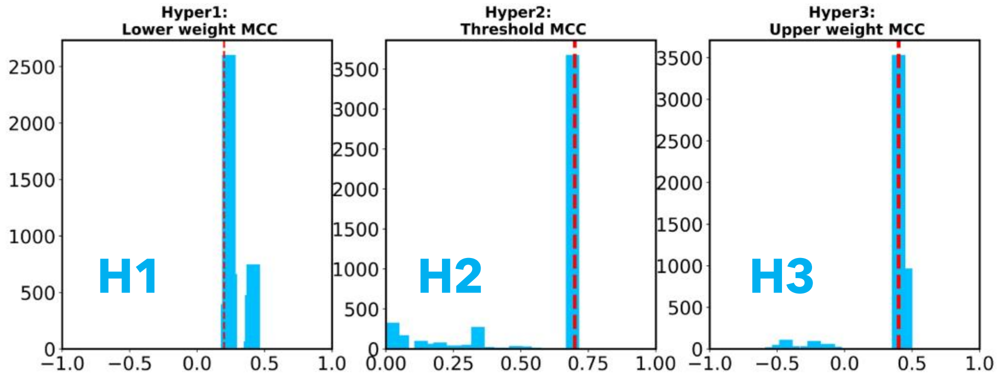


Model parameters + H1,H2,H3 (MCC hyperparameters)

Model parameters

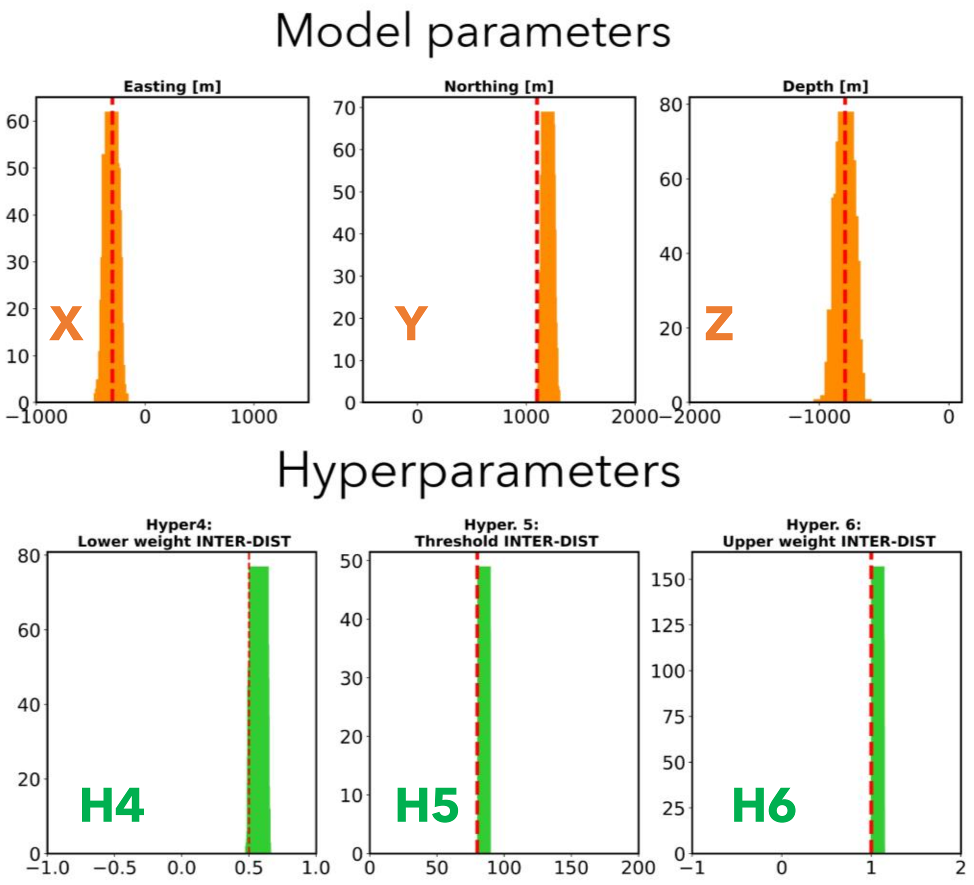


Hyperparameters

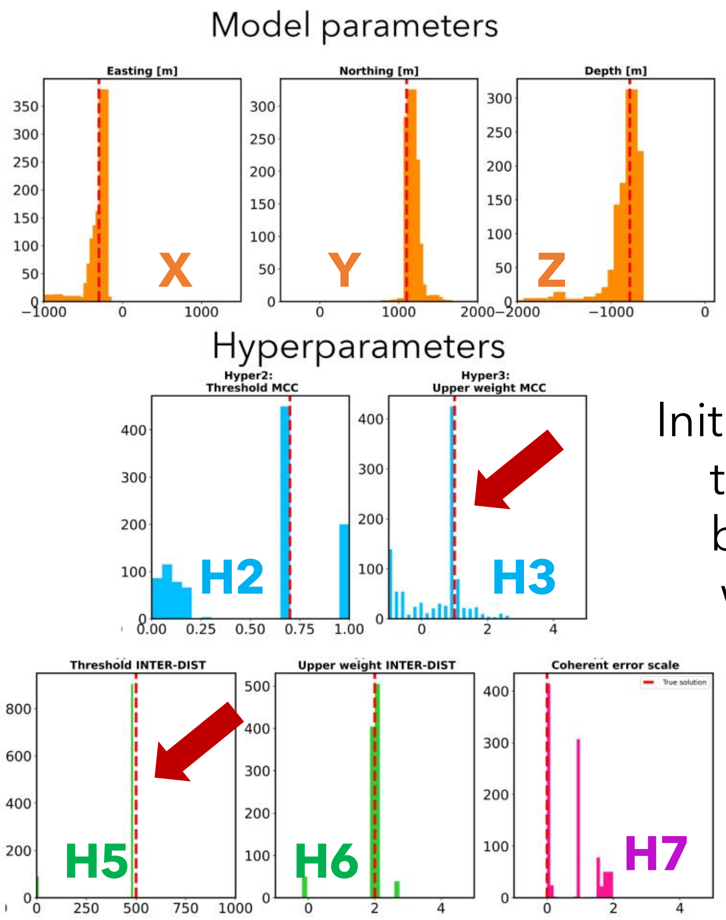


Synthetic tests

Model parameters + H4-H5-H6 (INTER-DIST hyperparameters)



Model parameters + H2,H3,H5,H6,H7 (thresholds MCC, INTER-DIST, upper weights and H7)



Initial signs of trade-off between weights

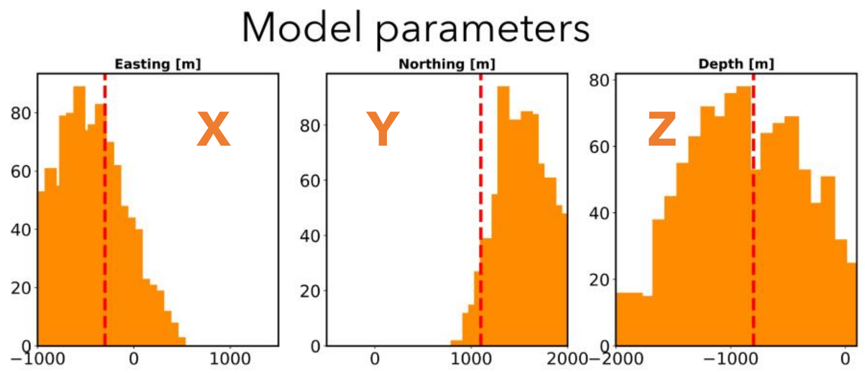
Synthetic tests

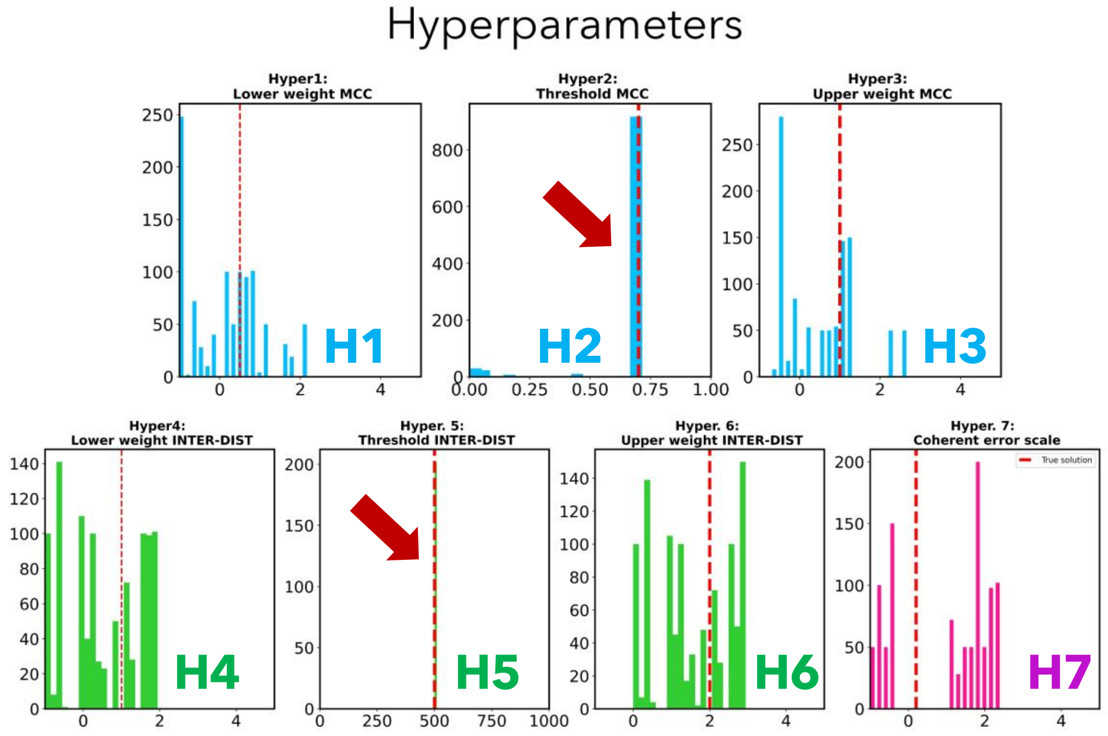
Model parameters
+ All hyperparameters

Total weight for each phase (data point, time delay) $W^{-1}(m) = 10^{**}(hc(m) + hd(m) + h0(m))$

From MCC **From INTER-DIST** **Coherent error scaling**

- Likely there is a **trade-off between the hyperparameter weights.**
- **MCC and INTER-DIST thresholds are correctly recovered.**

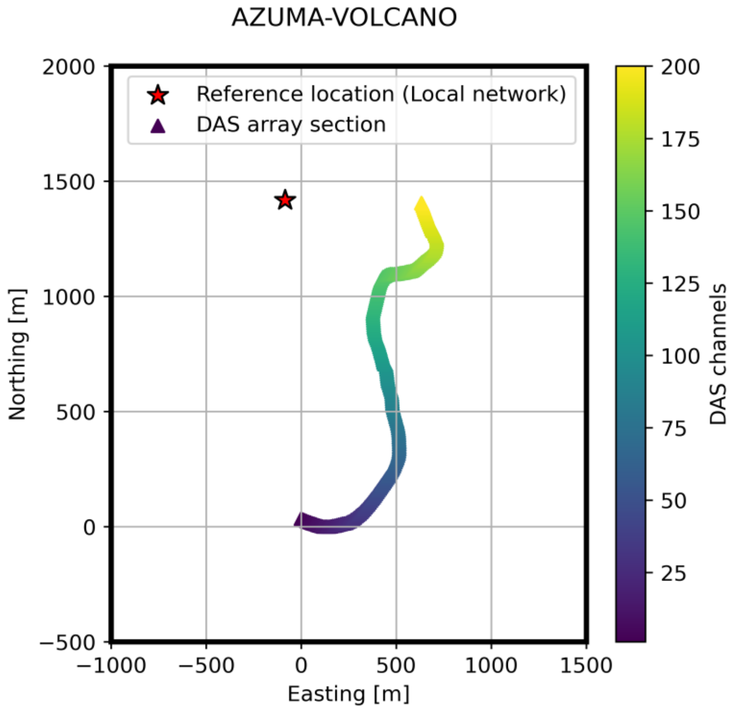




Application to real data: Azuma-Volcano and Cuolm da Vi

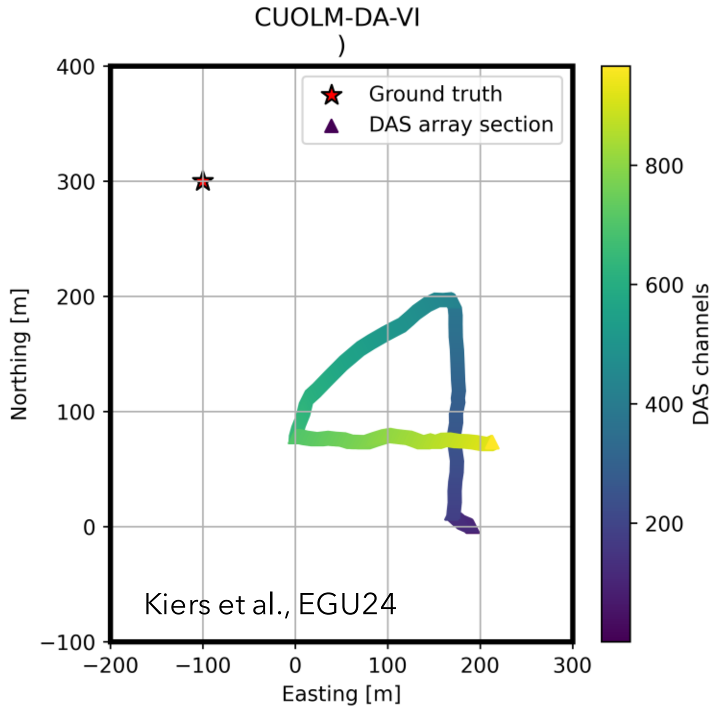
- **TEST-1 : Not-weighting**
- **TEST-2: Manual weighting** (hyperparameters are fixed)
- **TEST-3: Automatic weighting (only MCC or INTER-DIST)**
- **TEST-4: Automatic weighting (MCC + INTER-DIST + H7)**

Tectonic-volcanic event (Azuma-Volcano, Japan)



200 DAS channels (section)
Approx. 2 km length

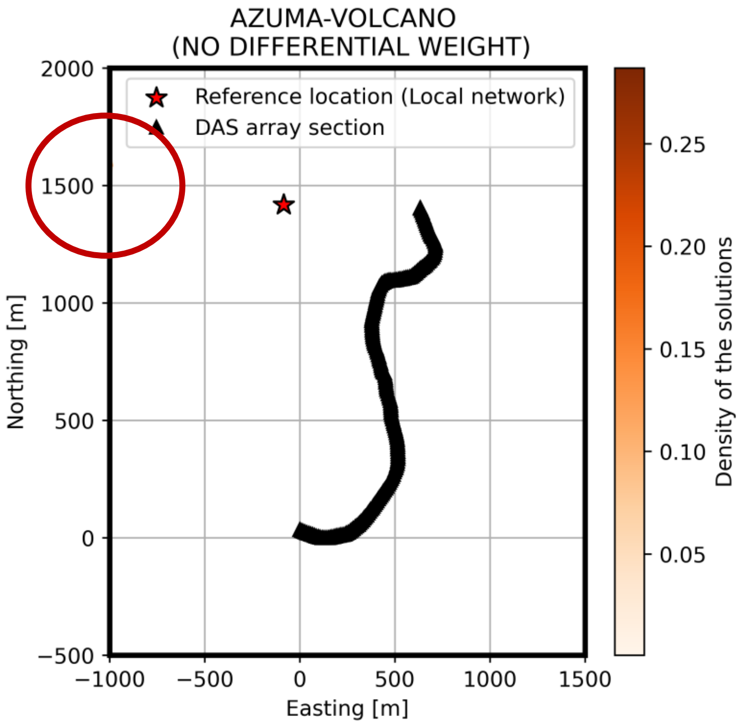
Active blast on a landslide (Cuolm da Vi, Switzerland)



966 DAS channels (section)
Approx. 1 km length

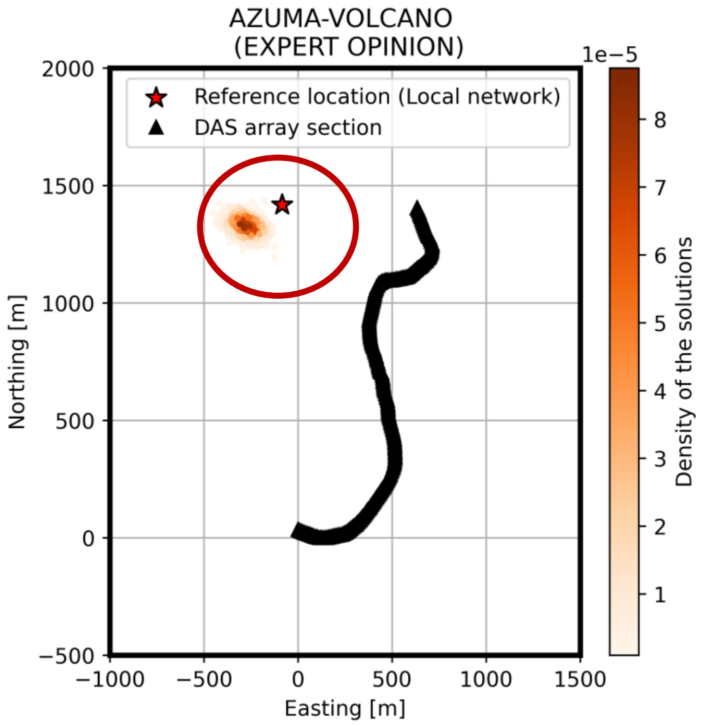
Azuma-Volcano

TEST-1 (NO WEIGHT)



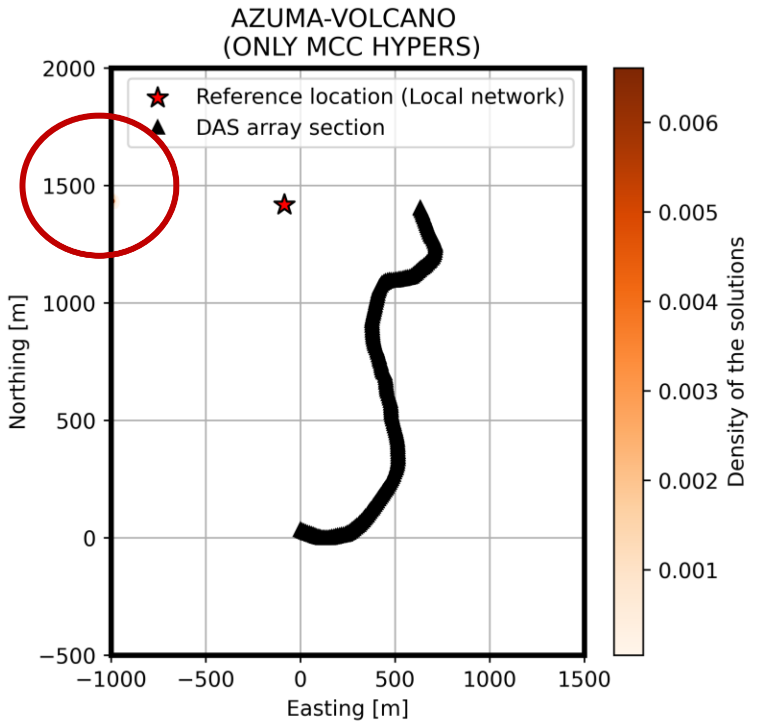
All the solutions are far from the reference location. Nevertheless the azimuth is correclty estimated.

TEST-2 (MANUAL WEIGHT)



H(1) = 0.2 (Lower weight)
H(2) = 0.5 (Thr. MCC)
H(3) = 2 (Upper weight)
H(4) = 0.2 (Lower weight)
H(5) = 200 (Thr. INTER-DIST)
H(6) = 2 (Upper weight)

TEST-3 (ONLY MCC)

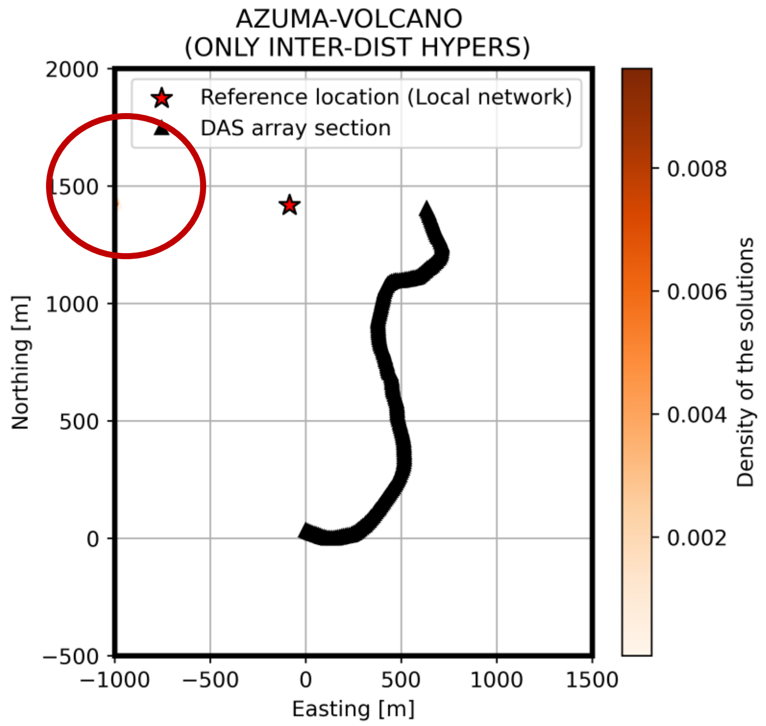


H(1) mean PPD = 0.6
H(2) mean PPD = 0.86
H(3) mean PPD = 1.56

(plausible values)

Azuma-Volcano

TEST-3 (ONLY INTER-DIST)



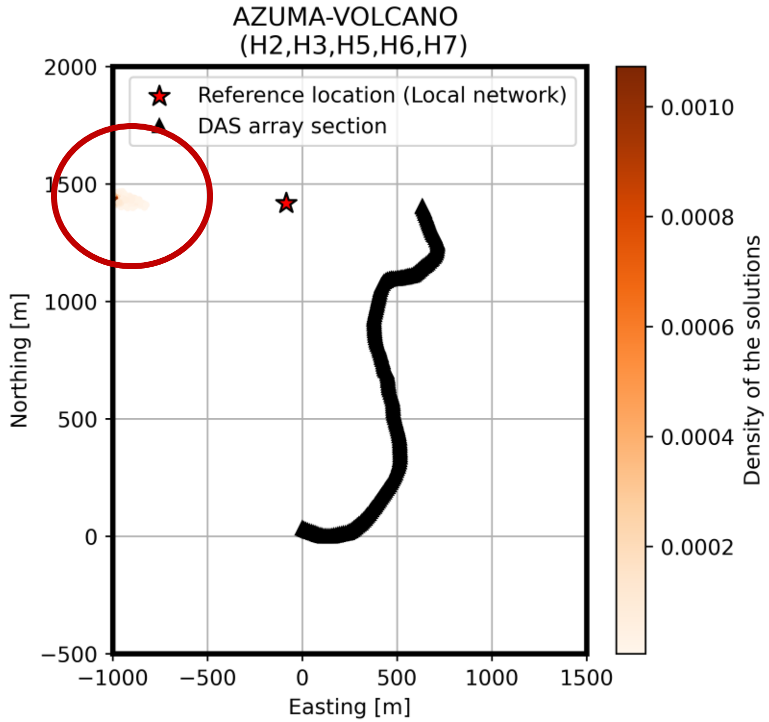
H(4) mean PPD = 1

H(5) mean PPD = 100.1 m

H(6) mean PPD = 1.6

(plausible values)

TEST-4 (H2,H3,H5,H6,H7)



H(2) mean PPD = 0.78

H(3) mean PPD = 0.90

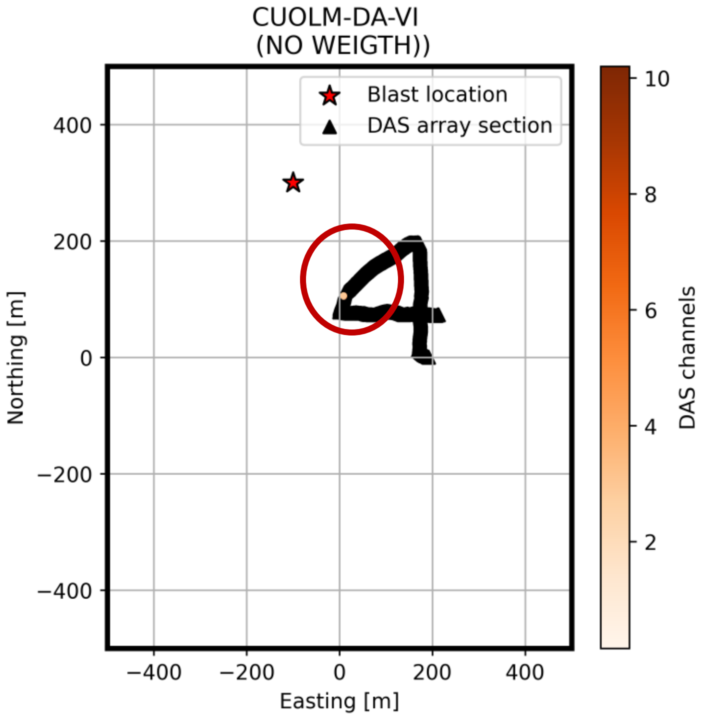
H(5) mean PPD = 168 m

H(6) mean PPD = 0.26

(plausible values)

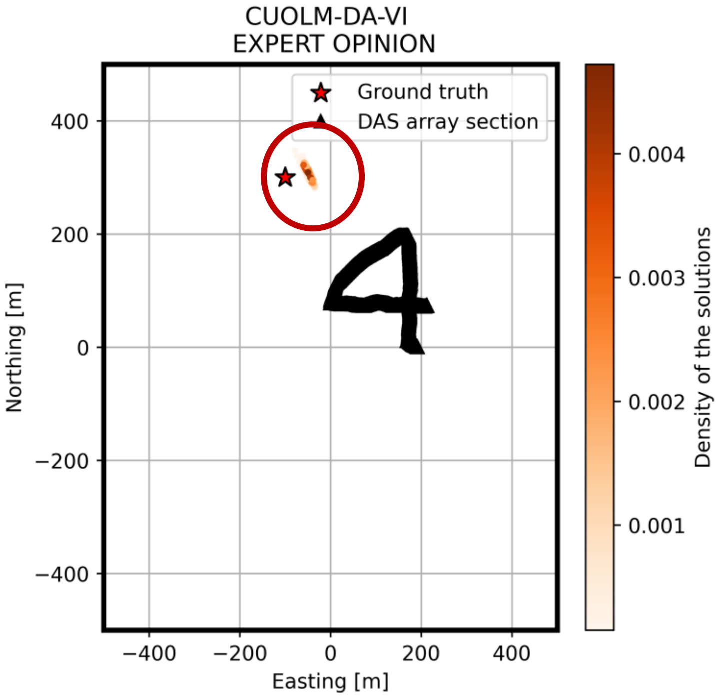
Cuolm-Da-Vi

TEST-1 (NO WEIGHT)



All the solutions are far from the reference location

TEST-2 (MANUAL WEIGHT)



H(1) = 0.2 (Lower weight)

H(2) = 0.7 (Thr. MCC)

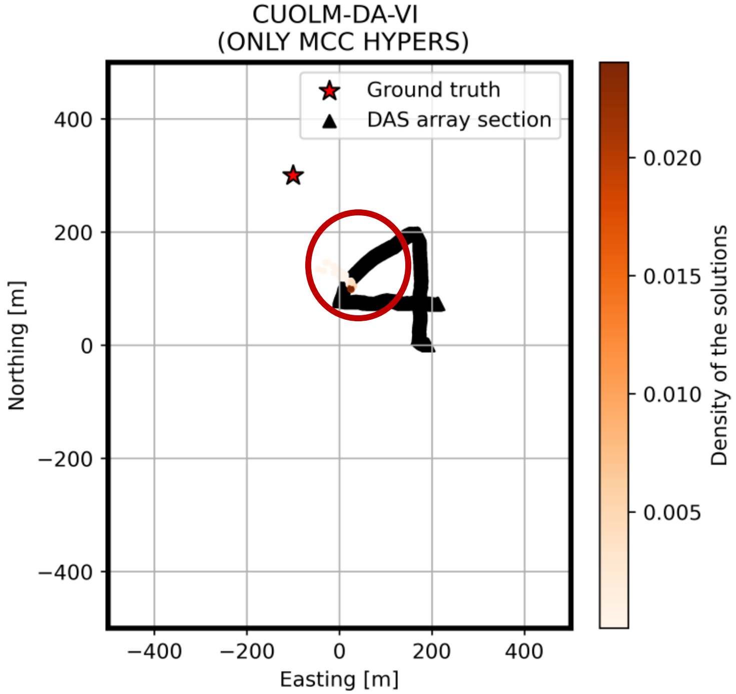
H(3) = 1 (Upper weight)

H(4) = 0.2 (Lower weight)

H(5) = 50 m (Thr. INTER-DIST)

H(6) = 1 (Upper weight)

TEST-3 (ONLY MCC)



H(1) mean PPD = 0.74

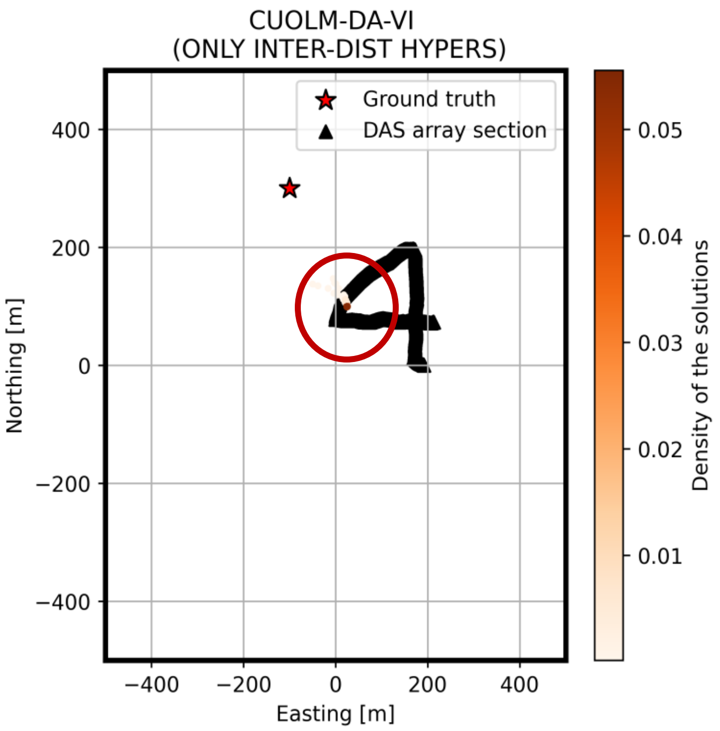
H(2) mean PPD = 0.78

H(3) mean PPD = 1.53

(plausible values)

Cuolm-Da-Vi

TEST-3 (ONLY INTER-DIST)



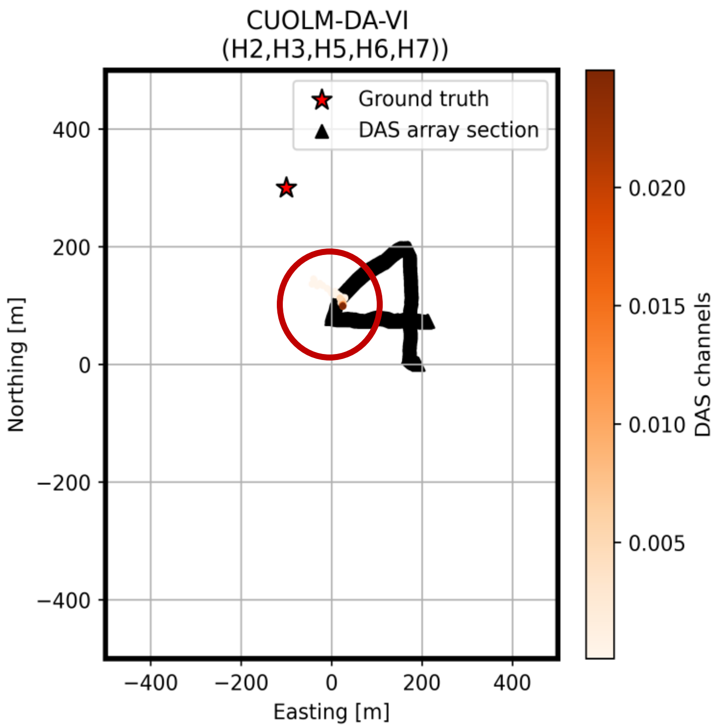
H(4) mean PPD = 0.94

H(5) mean PPD = 223 m

H(6) mean PPD = 1.18

(plausible values)

TEST-4 (H2,H3,H5,H6,H7)



H(2) mean PPD = 0.77

H(3) mean PPD = 1.03

H(5) mean PPD = 330 m

H(6) mean PPD = 0.91

Automatic weighting procedure

The results suggest that the **real data space might not strictly adhere to the basic assumptions of the algorithm** (higher cross-correlation index + lower interchannel distance indicate better data points).

Conclusions

- **What we did?** We tested DATs selection and developed a hierarchical McMC to weight the covariance matrix for event location with differential arrival times.

- **Does it work on synthetic tests?** The algorithm recovers the true values of the thresholds hyperparameters (MCC and INTER-DIST), but not more than two weights together (likely trade-off).

- **Does it work on "real-world" data?** The algorithm weights real data recovering the reference locations (manual weight) . However, automatically weighed solutions are not comparable to the reference solutions.

- **Possible explanations?** Not efficient noise models (thresholds + weights) + real data space not respecting our prior assumptions (highr MCC and lower INTER-DIST >> better data point)

- **Possible solutions?** A different formulation of the noise models is likely needed to avoid a trade-off between the hyperparameters + other real data test cases.

Differential arrival times for source location with DAS arrays: tests on data selection and automatic weighting procedure

Emanuele Bozzi[1] (Speaker) N. Piana Agostinetti[1], G. Saccorotti[2], A. Fichtner[3], Lars Gebraad[3,4], Tjeerd Kiers[3], and Takeshi Nishimura[5]

Thanks for your attention!

Vienna, EGU General Assembly, April 17, 2024.
SM 3.1 session