Exploration of Distributed Acoustic Sensing (DAS) data-space using a trans-dimensional algorithm, for locating geothermal induced microseismicity

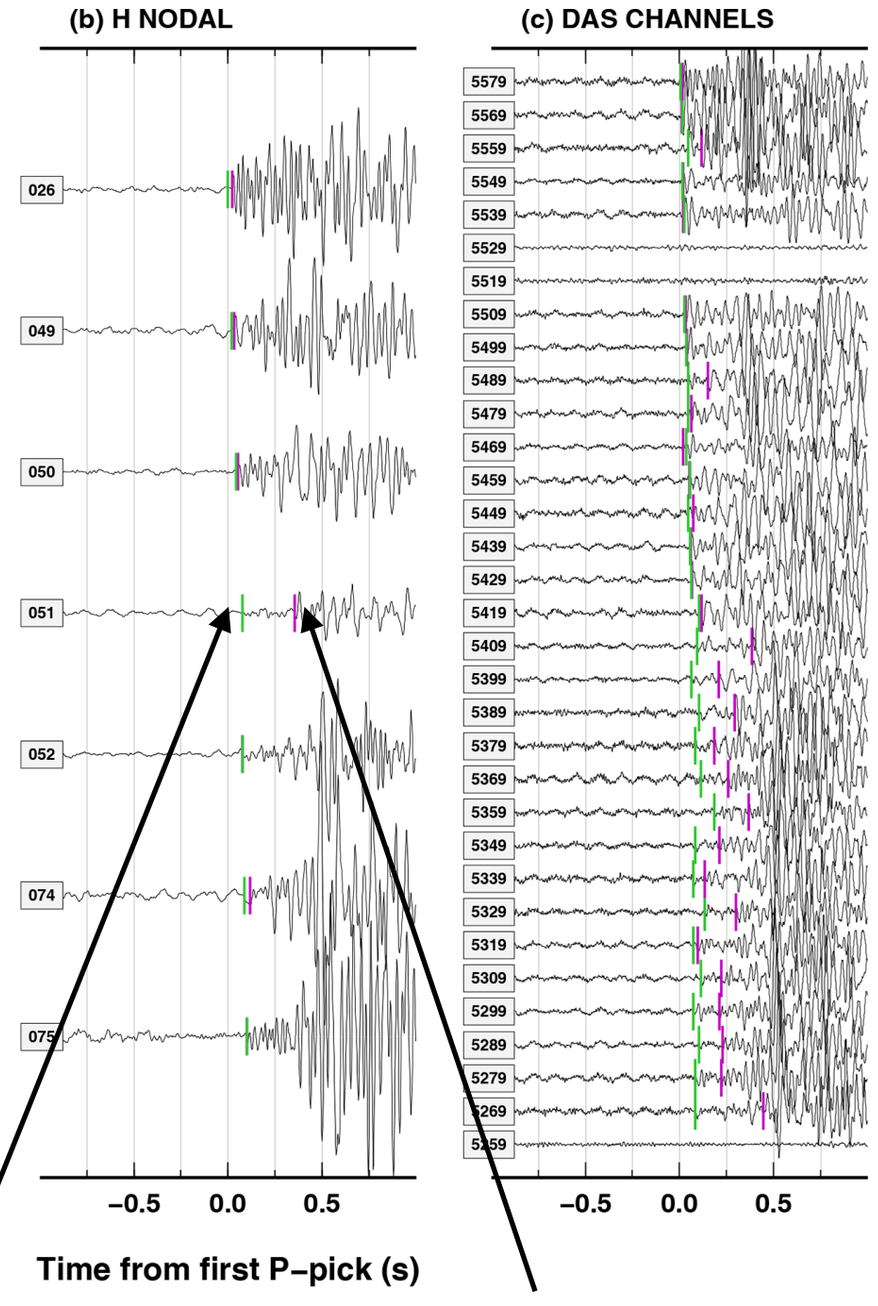
Nicola Piana Agostinetti [1] , Emanuele Bozzi [1], Alberto Villa [1]
Gilberto Saccorotti [2]

Vienna, May 2022

1 Dipartimento di Scienze dell'Ambiente e della Terra, Università di Milano Bicocca
2 Istituto Nazionale di Geofisica E vulcanologia, Sezione di Pisa

Key-observation, scientific question

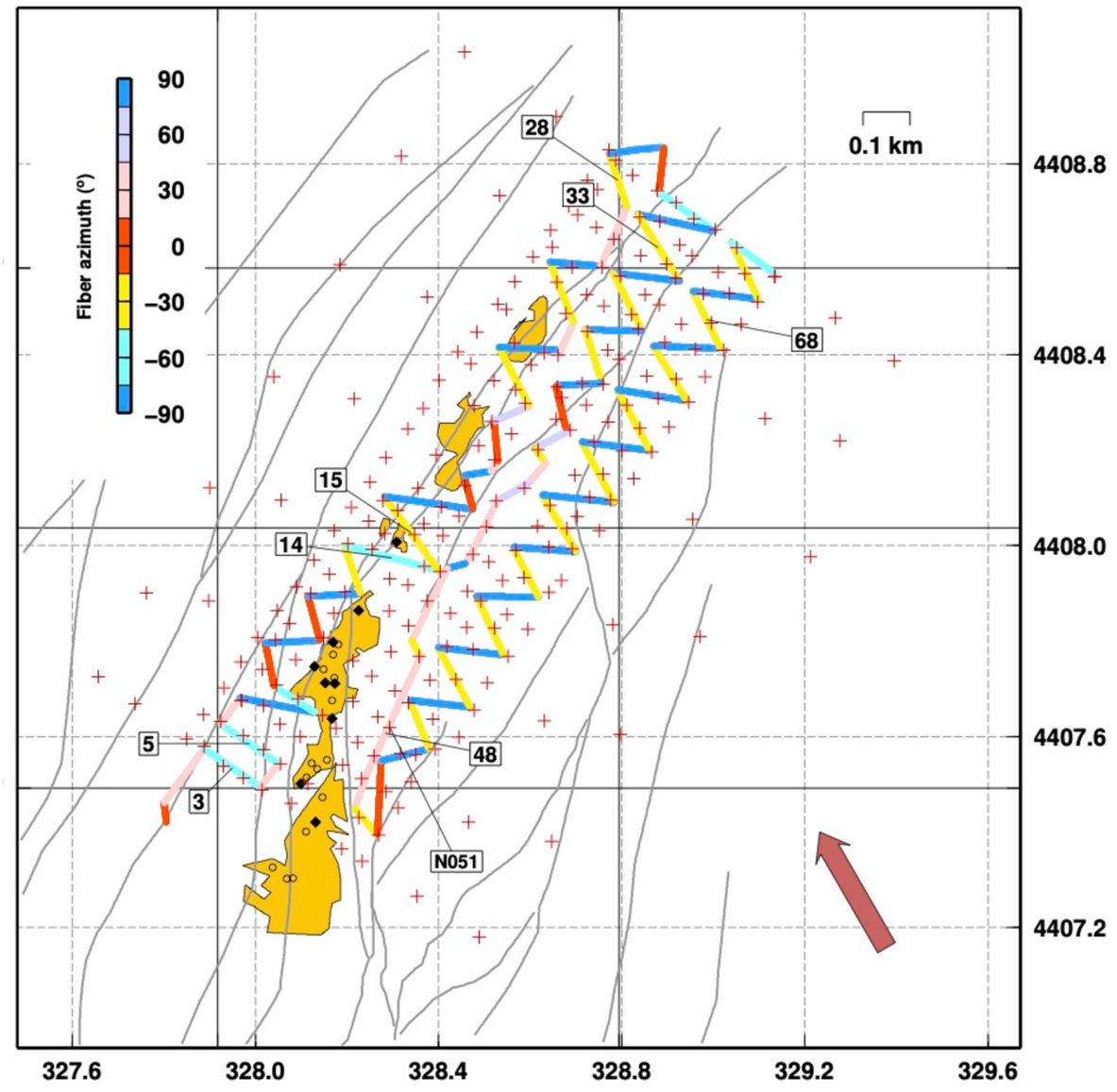
- Automatic picking of **P-wave** on horizontal components (both DAS and geophones) can easily fail
  - EFFECT: mis-labeling of S-wave as P-wave
  - EFFECT: increase uncertainties in seismic event location
- How can we improve event location including uncertainties on P-wave classification?



(Piana Agostinetti et al., Solid Earth, 2022)

TEST-CASE and REFERENCE SOLUTION

- Induced seismic event in Brady geothermal field (Porotomo experiment)
- DATA and METHODS
  - DATA: (0) Manually picked P- and S-wave on geophones vs Automatic picks on (1) Geophones and (2) co-located DAS
  - METHOD: Monte Carlo event location in a homogeneous rock volume (6 physical parameters + 2 hyper-parameters for scaling data uncertainties)



(Piana Agostinetti et al., Solid Earth, 2022)

TEST-CASE and REFERENCE SOLUTION

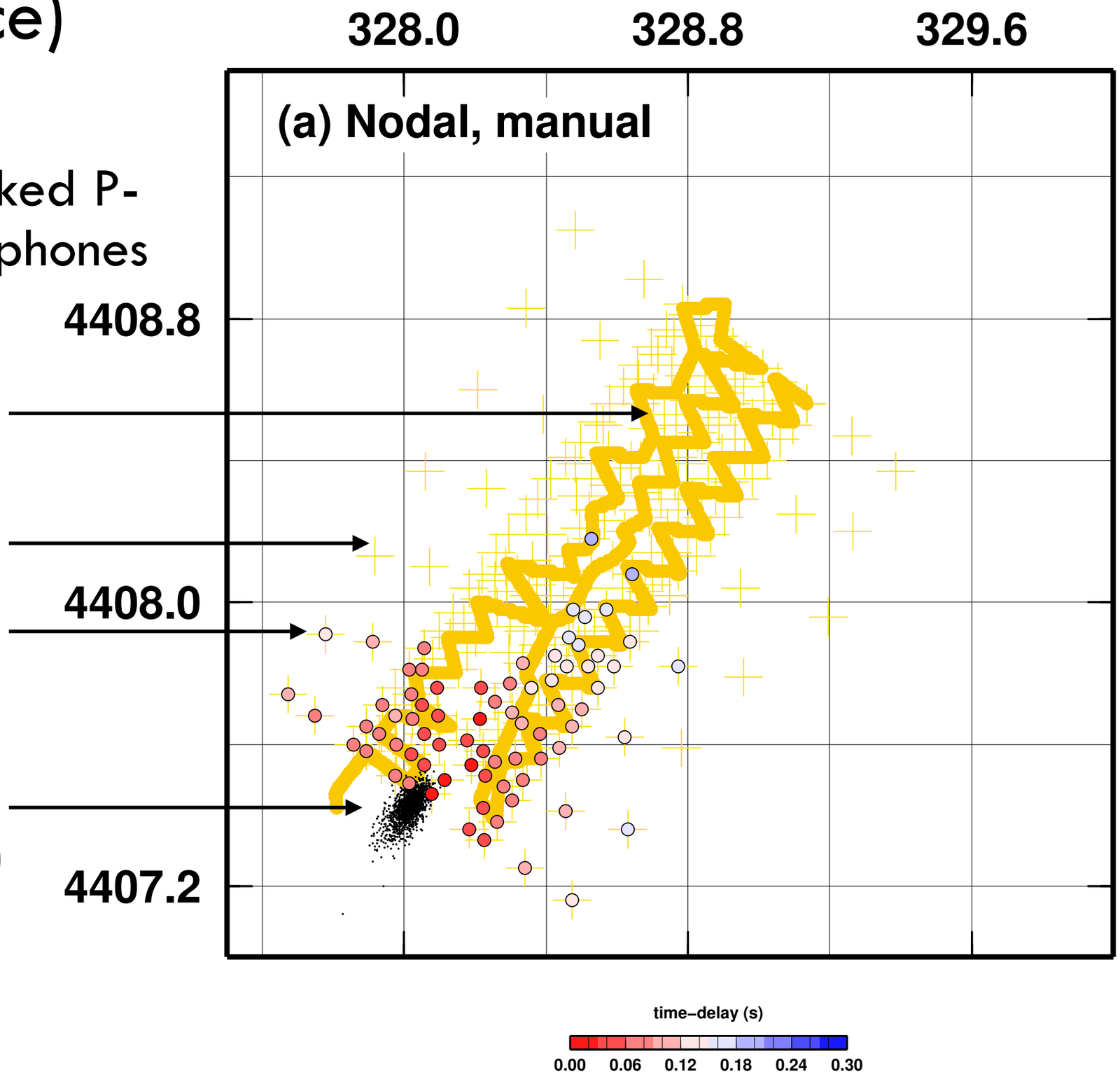
- Precise (reference) location
  - DATA: Manually picked P- and S-wave on geophones

FIBER OPTIC CABLE

GEOPHONE (not used)

GEOPHONE (used)

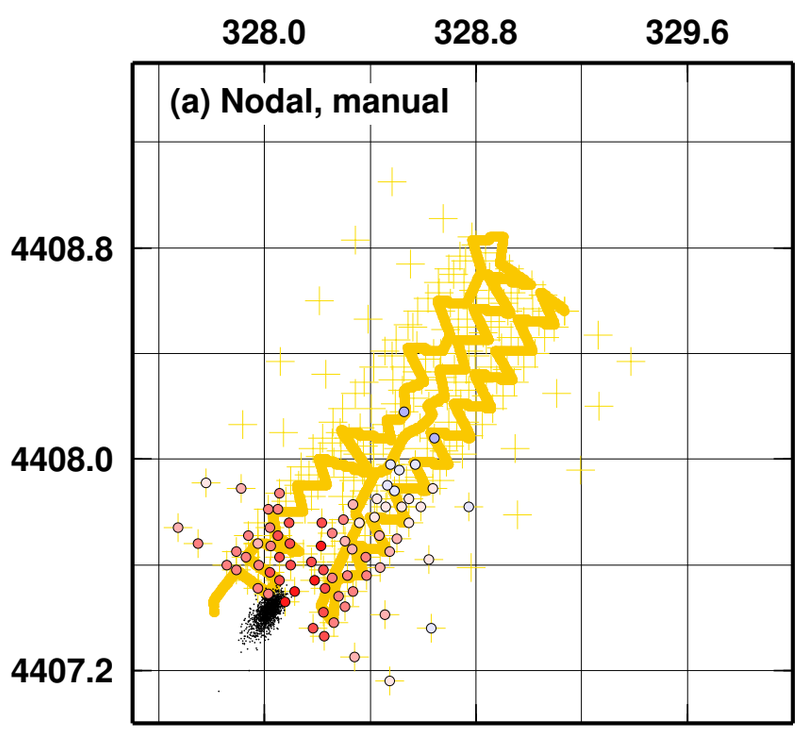
CLOUD OF POINTS: all sampled (potential) epicentres

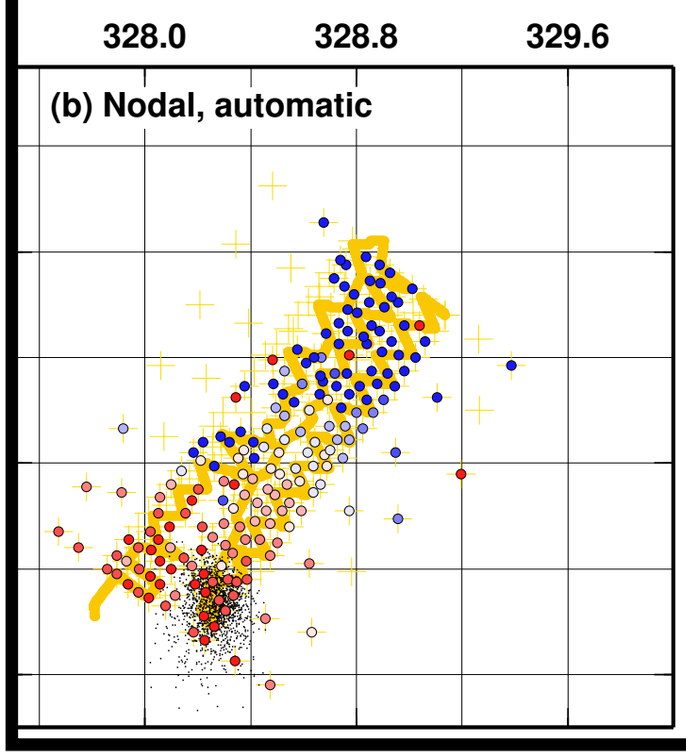


TEST-CASE and REFERENCE SOLUTION

EFFECT OF MIS-LABELLED P-PHASES

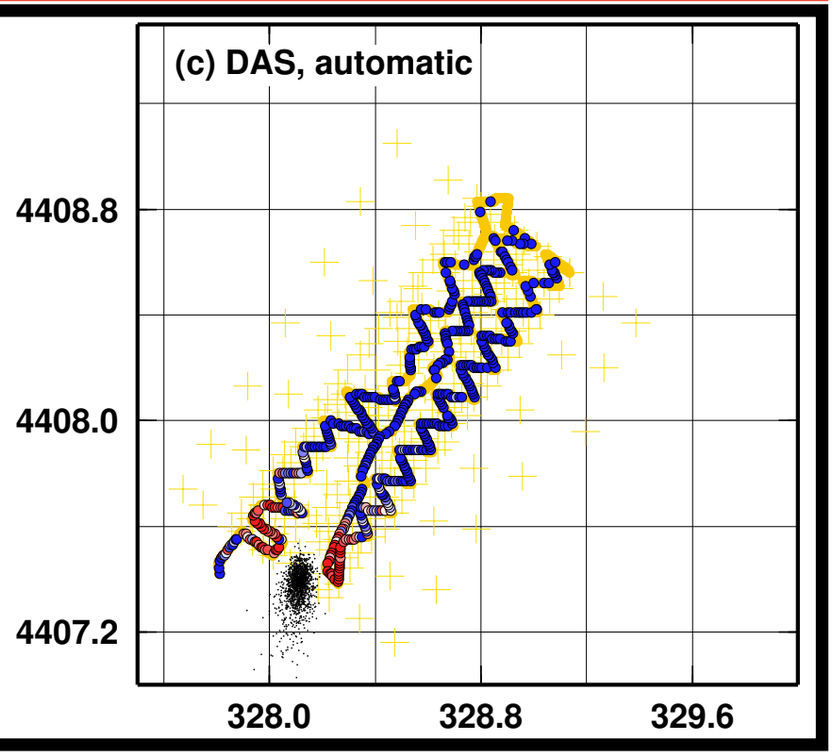
REFERENCE SOLUTION
(GEOPHONES, MANUAL PICKING)

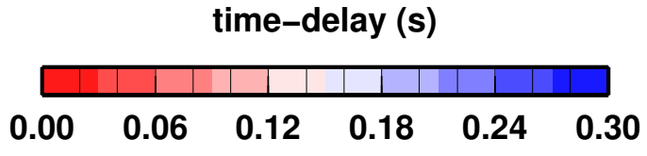




CASE 1:
(GEOPHONES, AUTOMATIC PICKING ON ORIENTED HORIZONTAL COMPONENT)

CASE 2:
(DAS, AUTOMATIC PICKING)





CLOUDS ARE NOT ROBUSTLY OVERLAPPING

TENTATIVE SOLUTIONS

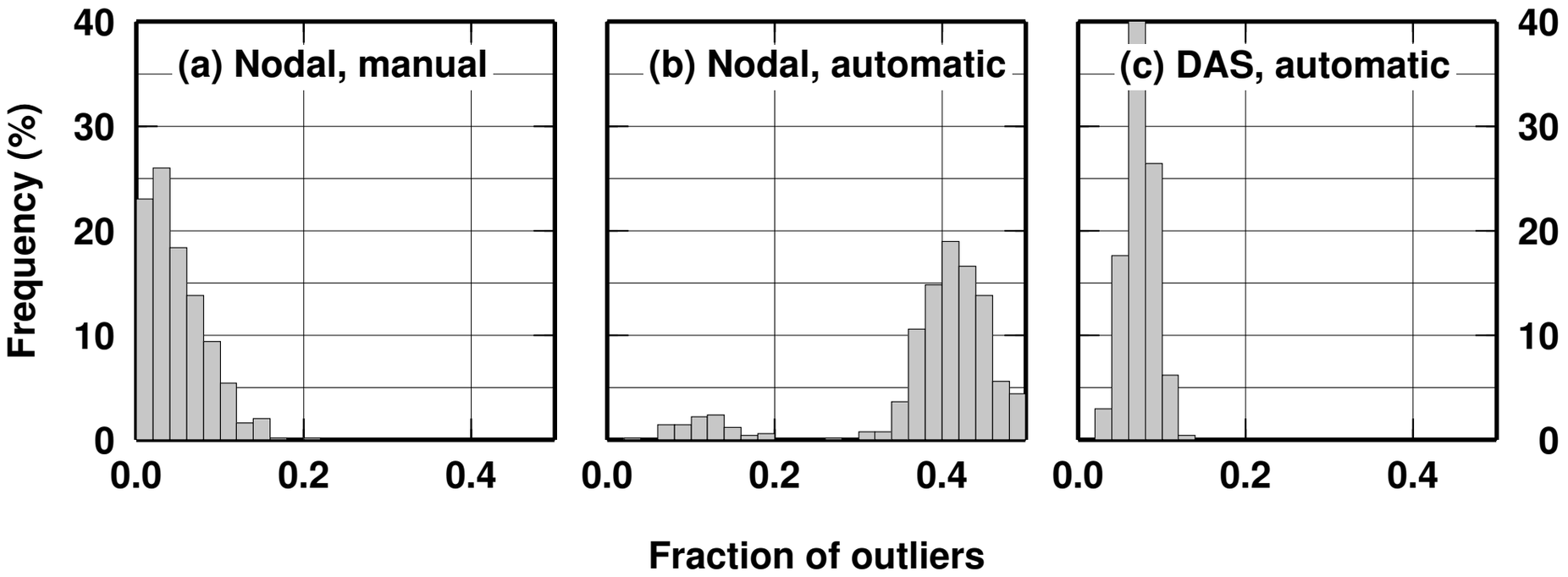
- **Outlier vs "standard" data**
  - A fraction of the data could be outliers
  - Based on: F. Tilmann et al. GJI 2021
- **Probabilistic P- and S- wave classification**
  - Each data-point can probabilistically be a P- or S-phase
  - (NEW)
- **Trans-D data-space exploration**
  - Data variance changes for portions of the fibre optic cable
  - Based on: N. Piana Agostinetti et al. JGR 2021

OUTLIERS vs “standard” DATA: theory

(1) Hypothesis: data-points are all P-wave picks
(2) Hypothesis: some data-points do not follow a standard error statistics and can be labeled as outliers
(3) Based on: Tilmann et al. (2021) "Another look at the treatment of data uncertainty in Markov chain Monte Carlo inversion and other probabilistic methods”

KEY-POINT: data-points are not deterministically defined as “outliers" or “standard" data, only the fraction of outliers/standard data can be defined

PPD of the fraction of outliers for the three data-sets



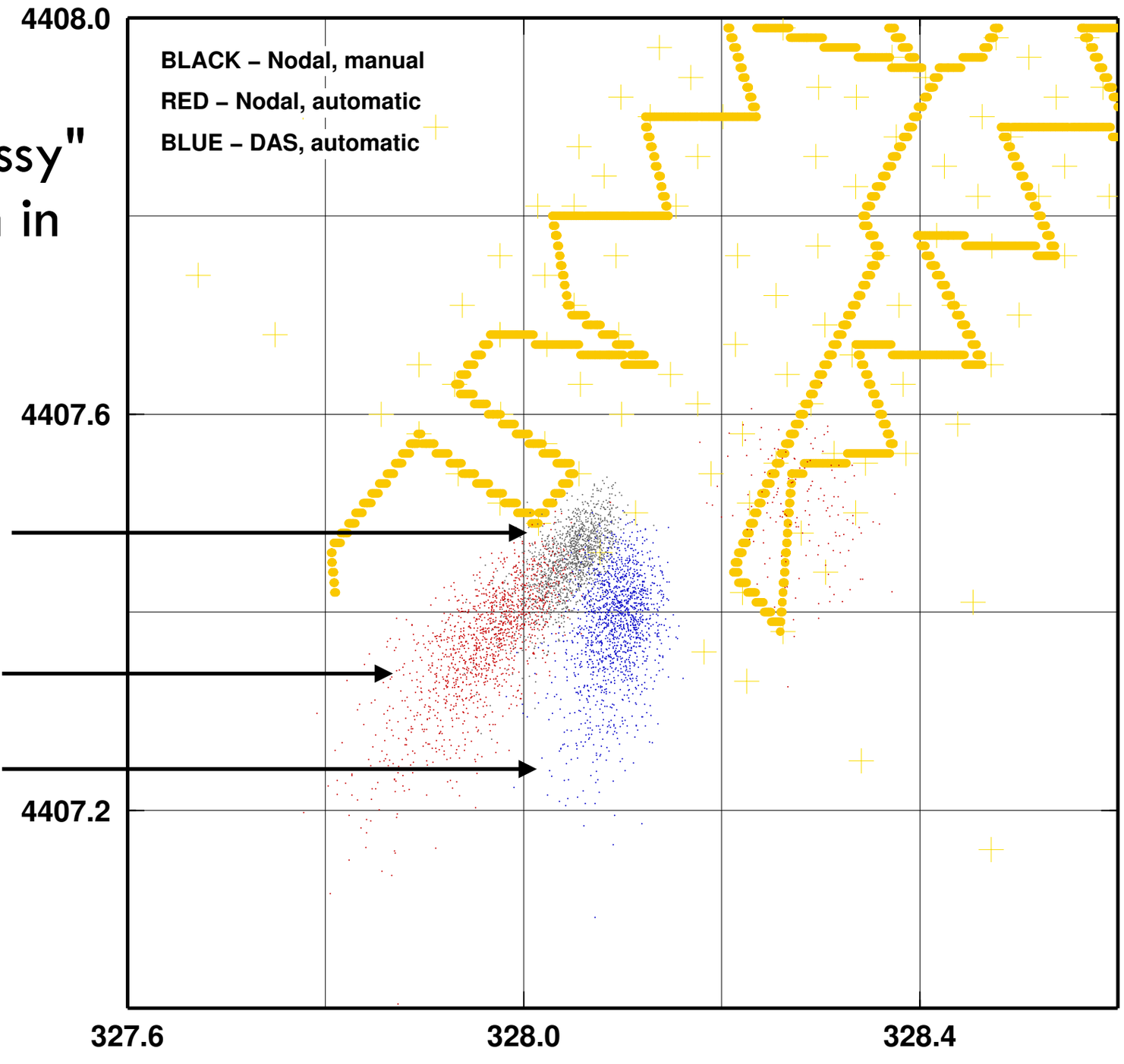
OUTLIERS vs “standard” DATA: results

Take-home message:
It seems to work with "messy" data (no clear separation in two classes, P and S)

BLACK, REFERENCE SOLUTION

RED, CASE 1: NODAL

BLUE, CASE2: DAS

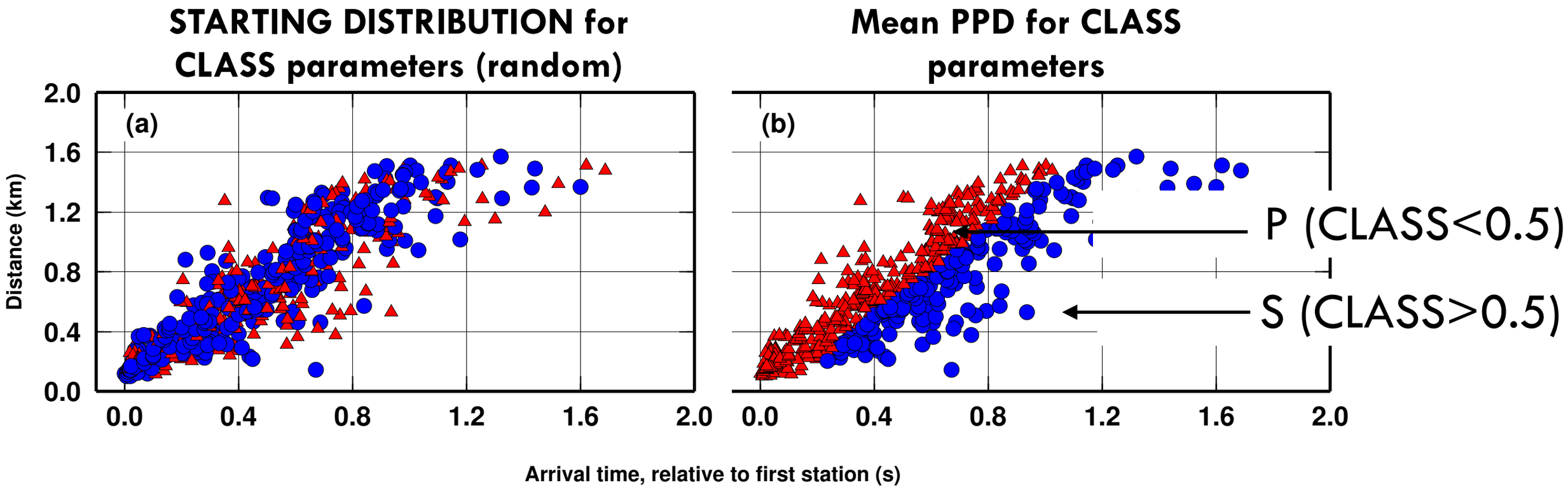


FRACTION OF OUTLIERS DECREASES WITH DATA QUALITY

P- and S- phases classification: theory

(1) Hypthesis: each data-point can be a P-phase pick or a S-phase pick (i.e. can belong to the P CLASS or S CLASS)
(2) Observation: Ps phases are also present in the data
(3) Based on: standard McMC including TWO CLASSES (i.e. one more parameter per data-point, need 10x longer chains with respect to the standard McMC solution)

KEY-POINTS: data-point are not deterministically defined as “P" or “S" data, but we will compute a PPD of their CLASSES



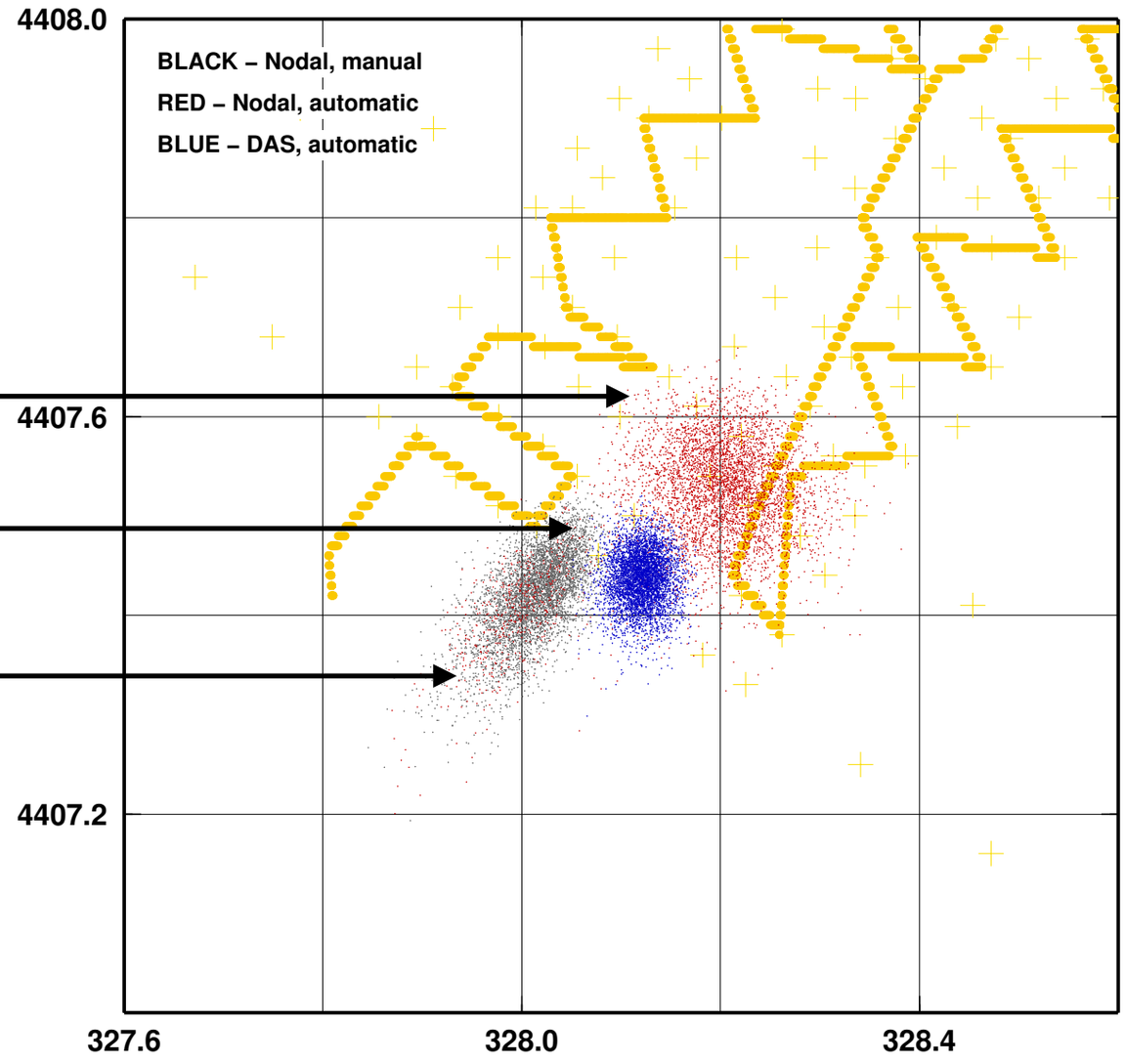
P- and S- phases classification: results

Take-home message:
It seems NOT to work with "messy" data (no clear separation in two classes, P and S)

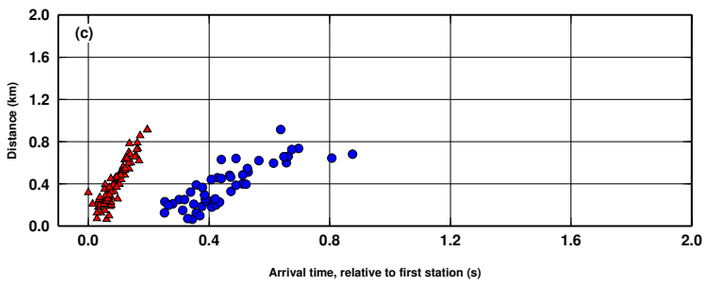
RED, CASE 1: NODAL

BLUE, CASE2: DAS

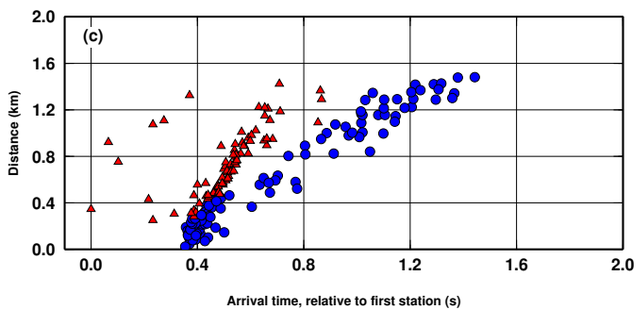
BLACK, REFERENCE SOLUTION



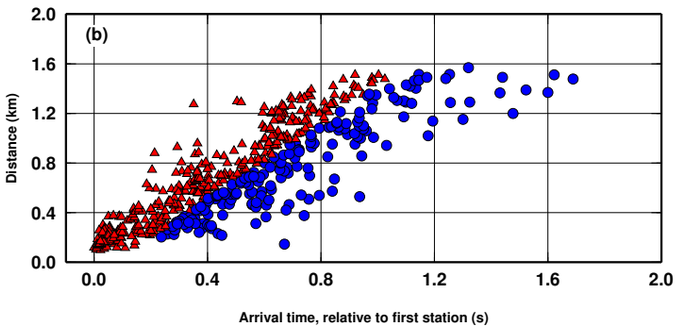
REFERENCE



CASE 1: NODAL



CASE 2: DAS



TRANS-D data-space exploration: theory

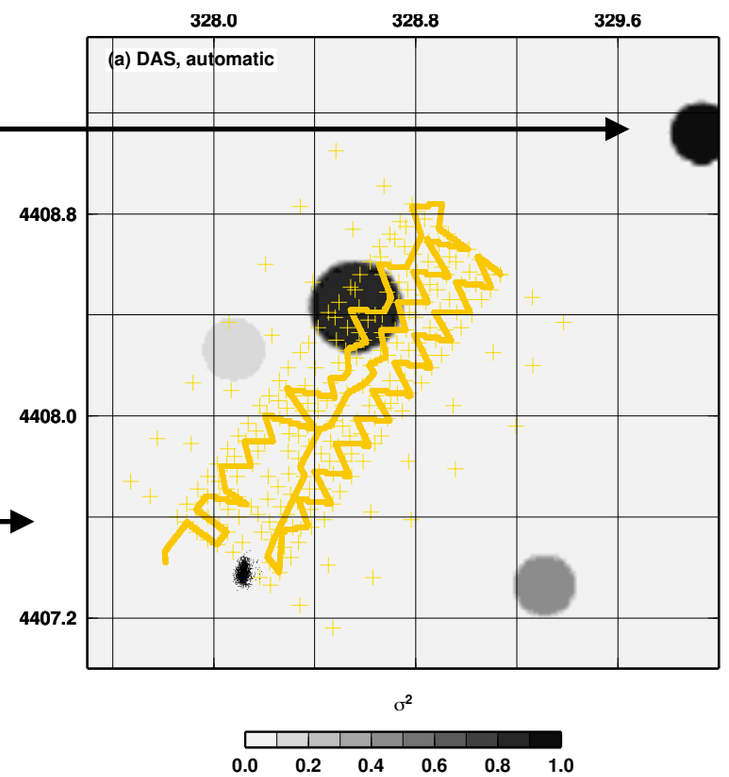
(1) Hypothesis: data-points are all P-wave picks
(2) Method: where hypotheses are violated, data variance is increased to down-weight data importance
(3) Based on: Piana Agostinetti et al. (2021) Exploration of Data Space Through Trans-Dimensional Sampling: A Case Study of 4D Seismics

KEY-POINTS: We can map out "partitions" of data which are un-consistent with our working hypothesis. Covariance matrix is used to down-weight such data-points.

ONE RANDOM MODEL

Data partition: circular area

The number of data partitions is an unknown (trans-D approach)



TRANS-D data-space exploration: results

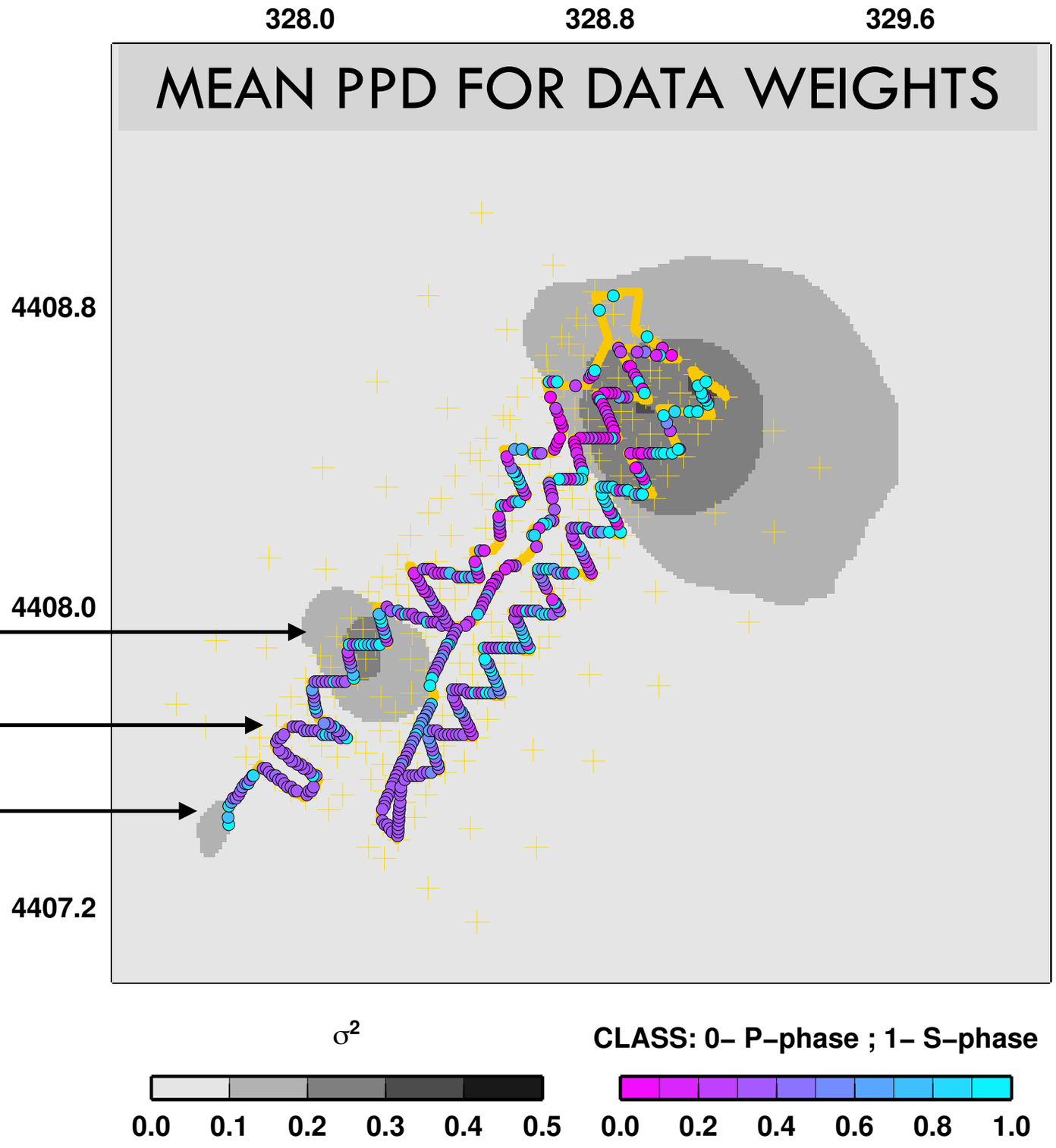
Take-home message:
It maps out the area where most of the S-phases are clustered

DARK GREY, area with inconsistent data

PINK, mostly P-phase

BLUE, mostly S-phase

(P- and S- phases from previous classification)



CONCLUSIONS

- **Outlier vs "standard" data**
  - *Work well on geophone data*, where two classes are NOT so well defined
- **Probabilistic P- and S- wave classification**
  - *Work well on DAS data*, where two classes of data can be easily separated
- **Trans-D data-space exploration**
  - *DAS*: Correctly maps out areas where data do NOT support hypotheses (here, the existence of one P CLASS only)