doi:10.26443/seismica.v4i2.1494

Automated weighting schemes for DAS data in geophysical inversion: a case study on event location

Emanuele Bozzi * [1,2], **Nicola Piana Agostinetti** [1,3], **Gilberto Saccorotti** [4]

[1]Department of Environmental and Earth Sciences, University of Milano-Bicocca, Milan, Italy, [2]Department of Earth Sciences, University of Pisa, Pisa, Italy, [3]Osservatorio Nazionale Terremoti (ONT), Istituto Nazionale di Geofisica e Vulcanologia (INGV), Rome, Italy, [4]Istituto Nazionale di Geofisica e Vulcanologia (INGV), Pisa, Italy

Author contributions: *Conceptualization*: Emanuele Bozzi, Nicola Piana Agostinetti, Gilberto Saccorotti. *Methodology*: Emanuele Bozzi, Nicola Piana Agostinetti, Gilberto Saccorotti. *Software*: Emanuele Bozzi, Nicola Piana Agostinetti. *Validation*: Emanuele Bozzi, Nicola Piana Agostinetti, Gilberto Saccorotti. *Formal Analysis*: Emanuele Bozzi, Nicola Piana Agostinetti. *Writing - Original draft*: Emanuele Bozzi. *Writing - Review & Editing*: Emanuele Bozzi, Nicola Piana Agostinetti, Gilberto Saccorotti. *Supervision*: Nicola Piana Agostinetti, Gilberto Saccorotti. *Funding acquisition*: Nicola Piana Agostinetti.

Abstract Distributed Acoustic Sensing (DAS) technology offers meter-scale spatial sampling of seismic wavefields, which improves our seismic monitoring capabilities. However, the resulting data volumes often complicate expert-driven analysis, such as weighting measurements in geophysical inverse problems to mitigate the influence of outliers. To address such difficulties in an automated manner, we focus on a representative problem in seismology, source location, and we test a Bayesian weighting procedure applied to P-wave arrival/differential times estimated from DAS waveforms. This approach estimates the Posterior Probability Densities of physical (epicenter) and non-physical (hyperparameters) parameters. The hyperparameters are designed as a set of thresholds and weights that enable the automatic identification of portions of data associated with higher reliability in the inversion. Specifically, the thresholds are values having the same dimension as specific waveform attributes and/or geometrical descriptors. Weights are then the scaling factors for the variances of data that do not meet these thresholds. Consequently, several possible weighting schemes (thresholds and weights) based on signal-to-noise ratios, cross-correlation indices, and interchannel distances are explored in a Bayesian framework. We present synthetic tests and real-data applications that demonstrate the potential of this method as an alternative to a similar approach without data weighting.

Production Editor: Yen Joe Tan
Handling Editor: Wenbo Wu
Copy & Layout Editor: Théa Ragon

Received: October 22, 2024
Accepted: December 7, 2025
Published: December 19, 2025

1 Introduction

Distributed Acoustic Sensing (DAS) technology, from its first applications in the early to mid-2010s, is nowadays an established method for seismic monitoring (Cannon and Aminzadeh, 2013; Daley et al., 2013; Parker et al., 2014; Biondi et al., 2017; Lindsey et al., 2017; Jousset et al., 2018; Zhan, 2020; Wuestefeld et al., 2024). DAS provides meter-scale spatial resolution and an unprecedented full-waveform view of seismic wavefields, enhancing source characterization and subsurface imaging (Li et al., 2023a; Lanza et al., 2024; Noe et al., 2026; Biondi et al., 2023; Miller et al., 2025; Strumia et al., 2024). However, the resulting data volumes not only stress our storage capabilities (Seguí et al., 2025) but often complicate efficient data processing and expert identification of outliers affecting source parameter estimation.

Among seismological monitoring tasks, event location, which is one of the oldest problems in seismology, remains an active area of research, with a particular focus on phase picking and the management of data uncertainties (Lay and Wallace, 1995; Aki and Richards, 2002). In fact, time picks are influenced by both experimental uncertainties, such as the accuracy of the picker for a specific seismic phase, and the forward model imprecision, which affects the accuracy of travel time predictions. Such uncertainties are routinely addressed using subjective, though refined, data processing in the source location inverse problem. For instance, arrival times from an automated picker are routinely weighted using thresholds based on Signal-to-Noise Ratio (SNR), allowing experts to assign greater importance to measurements where the signal is better distinguished from background noise. Alternatively, when the geometry permits, array processing techniques can exploit phase coherence to obtain differential times (Rost and Thomas, 2002). In such a case, differential times are inverted and often weighted based on, e.g., waveform similarity, using the Maximum values of the Cross Correlation (MCC) functions. Differential times estimated from cross-correlation are considered more reliable with higher MCC values. Consequently, subjective choices are made to identify appropriate threshold values in the MCC data space.

DAS technology has emerged as a valuable distributed sensor for earthquake location, as it can provide new data from non-conventional experimental setups (e.g., glaciers, deep oceans) or add valuable observations closer to the target hypocenter (Walter et al., 2020; Ajo-

*Corresponding author: emanuele.bozzi@dst.unipi.it, e.bozzi3@campus.unimib.it, emanuele.bozzi96@gmail.com

Franklin et al., 2019; Li et al., 2023b; Piana Agostinetti et al., 2022; Lellouch et al., 2020; Nishimura et al., 2021; Bocchini et al., 2025; Tilmann et al., 2024; Lior et al., 2021; Hudson et al., 2024; Fichtner et al., 2022). Given its unprecedented meter-scale spatial sampling, automatic picking methods have been developed to address the impossibility of manually picking seismic phases recorded at the typically thousands of DAS channels (Xiao et al., 2025; Zhu et al., 2023; Latorre et al., 2025). Such arrival times can be incorporated into existing monitoring pipelines to improve source location accuracy (Obermann et al., 2022; Baillet et al., 2025; Biondi et al., 2025). Moreover, DAS has been successfully exploited as a receiver antenna to constrain the source of a variety of signals (Klaasen et al., 2021; Biagioli et al., 2024). Despite such evident advantages, the resulting volumes of data may expose the event location process to the influence of unidentified, meaningless data (e.g., arrival times), such as noise transients, which cause spurious automatic triggers, often manually identified and treated as outliers in conventional monitoring procedures. A possible solution is the automatic selection of good-quality DAS channels, and thus of the associated arrival times (Rodriguez et al., 2025; Bozzi et al., 2025). Alternatively, if the goal is to exploit the full seismic-wavefield resolution provided by fiber-optic sensors, automatic data weighting can be considered to mitigate potential bias from noisy channels in source location estimation. In this framework, data selection is particularly important for reducing the amount of data and supporting real-time applications, whereas weighting is more critical for offline, more refined seismic analysis. Here, the focus is on the latter application, as data selection has already been covered in previous studies.

While expert (manual) data weighting is feasible for a limited number of stations, it becomes more challenging with DAS sensors, given the large number of available observations (DAS channels). Additionally, DAS waveforms are often affected by lower SNRs and strong signal spatial amplitude/coherence variations due to directional sensitivity ground coupling (Celli et al., 2024; Bozzi et al., 2024) and site effects (Yang et al., 2022; Pecci et al., 2025). As a result, data are potentially influenced by additional and typically complex noise sources, making standard data processing less effective and pushing the adoption of less conventional methods (Trabattoni et al., 2023; Porras et al., 2024; Pascucci et al., 2025; Lapins et al., 2024; Isken et al., 2022), while also increasing the difficulty of subjective weighting.

To address the complexity and manage the redundancy of DAS data, here we test a Bayesian automatic weighting procedure, focusing on the event location inverse problem. Compared to expert-defined and often subjective weighting schemes, our approach identifies data families (i.e., portions of the complete set of data) that are considered more reliable in the inversion, through a probabilistic exploration of the data space. The method is based on a hierarchical Markov Chain Monte Carlo (MCMC) algorithm, thus estimating the Posterior Probability Distribution (PPD) of both physical and non-physical (hyperparameters) parameters (Mosegaard and Tarantola, 2002; Malinverno and Briggs, 2004; Tarantola, 2005). It works by producing a modified, i.e., weighted, version of the covariance matrix, computed from a set of hyperparameters (Piana Agostinetti et al., 2023). Different weighting schemes are accepted or rejected according to the Metropolis rule (Sambridge and Mosegaard, 2002), thus it is guided by the agreement between the observed data and the model predictions. Here, we define the hyperparameters as a set of "thresholds" and "weights". The thresholds are values with the same dimensions as specific waveform attributes, such as SNR and MCC, and identify families of data within the data space. The weights are values that scale the entries of the data covariance matrix, and thus the likelihood function, corresponding to the data families not meeting the thresholds.

We test a DAS data weighting procedure tailored to either P-wave arrival or differential times, considering these waveform attributes: a) SNR, b) MCC, and c) INTER-channel DISTance (INTER-DIST). The weighting scheme for P-wave arrival times is defined by a single threshold on SNR and an associated weight: data below the threshold are automatically weighted in the inversion. For P-wave differential times, we develop models defined by two sets of thresholds, one on MCC and one on the INTER-DIST, and associated weights: data below the MCC threshold and above the INTER-DIST threshold are weighted in the inversion.

Results demonstrate the robustness of the automatic weighting algorithm for both P-wave arrival times and differential times in synthetic tests. Additionally, applications to real data indicate improvements in event location accuracy while maintaining comparable computational costs relative to a more standard unweighted approach.

2 Data

To test the algorithm, we exploit earthquake signals of tectonic and volcano-tectonic origin, recorded by two distinct DAS deployments: a) a terrestrial cable running along a civil route close to Azuma volcano (ID. AZ-V) in Japan (Figure 1a) and b) an ocean bottom cable northeast of the Gran-Canaria (ID. G-C) Island (Figure 1b) (Ugalde et al., 2023). Tabs. S1 and S2 (see Supplementary Materials) summarize the DAS acquisition parameters and the analyzed earthquake signals, respectively.

First, P-wave arrival and differential times are estimated using a conventional STA/LTA trigger (Allen, 1982) and a multichannel cross-correlation approach, respectively. The data spaces have different dimensions: given Q available DAS channels, P-wave arrival times are a subset S of Q (i.e., the channels triggered by STA/LTA), while P-wave differential times are $(S * (S - 1))/2$.

SNR, MCC, and INTER-DIST waveform attributes are estimated at the triggered channels within a window centered on the P-wave arrival times, or at linearly interpolated arrival times when a few channels are missing (not triggered). When inverting P-wave arrival times, SNR is the waveform attribute used to determine the threshold for weighting the data (hyperparameter). In-

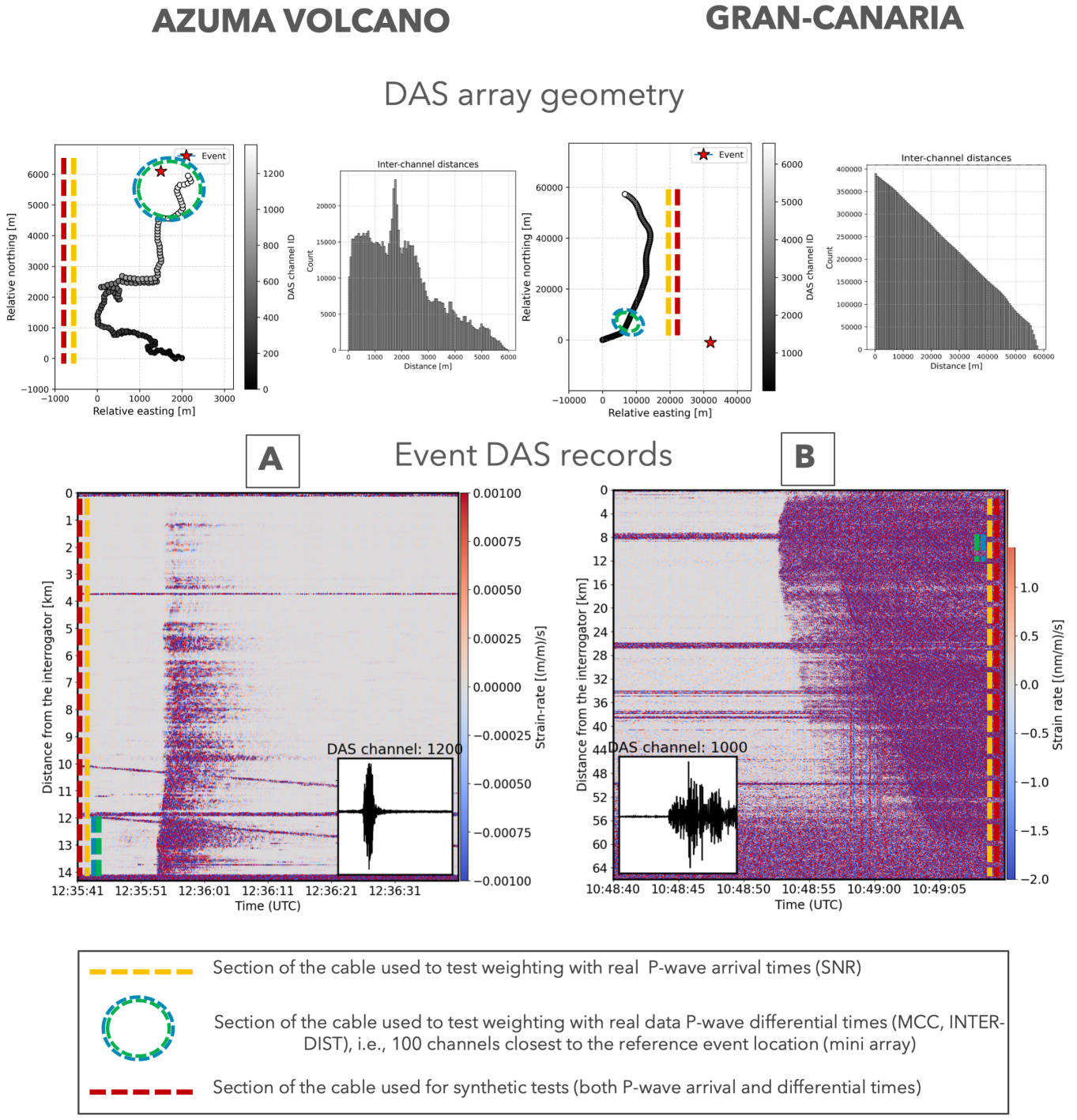

Figure 1 Datasets used in the study. Panels (a) and (b) for AZ-V and G-C deployments illustrate (i) the DAS array-event geometrical relations, with the distribution of the inter-channel distances used for weighting data in the inversion, and (ii) the event recordings (adopted filtering described in Tab. S2 of the Supplementary Materials). DAS data from the full length of cable deployments are used for all synthetic (red dotted line) and real-data tests (yellow dotted line) using P-wave arrival-time-based weighting. For real-data tests using P-wave differential time weighting schemes, the focus is instead on subsections of 100 high-SNR channels closer to earthquake sources (green-blue dotted line).

stead, when inverting P-wave differential times, MCC and INTER-DIST are the waveform attributes. Tabs. S3 and S4 (see Supplementary Materials) provide examples of data and associated attributes used in the algorithm described in this work.

3 Methods

3.1 Computation of P-wave arrival/differential times and waveform attributes

P-wave arrival times are estimated using a conventional STA/LTA triggering (Allen, 1982). Then, SNRs at the triggered channels are computed using the power ratios of the signal before and after the identified P-wave arrival time (Eq. 1). Tailored time windows are defined to isolate the section of the seismogram corresponding to the P-wave arrival time, thus depending on the event duration. In this work, as very local or local events are used, time windows are set to two seconds. For a DAS channel i triggered by a defined STA/LTA threshold:

$$\mathbf{SNR}_i = 10 \log_{10} \left(\frac{\overline{s_i^2}}{\overline{n_i^2}} \right) \tag{1}$$

where $\overline{s_i^2}$ and $\overline{n_i^2}$ are the mean squared amplitudes of the signal and noise windows, respectively. SNR is expressed in decibels.

P-wave differential times are estimated using signal cross-correlation for all triggered DAS channel pairs, evaluated over time windows starting at the estimated P-wave arrival times. When using data from the entire length of the cable (synthetic tests), we spatially downsample by stacking a few contiguous waveforms (10 for AZ-V and 50 for G-C) to keep the number of channel pairs manageable for the algorithm. Cross-correlations computed at channels lacking a P-wave arrival time, and thus relying on interpolated estimates, may be affected by a misestimate of the true arrival time. However, because the purpose of the automatic weighting is to reduce the influence of outliers in large DAS datasets, we choose to retain a simple and broadly applicable workflow that maximizes the number of usable P-wave differential times, rather than adopt alternative, possibly more accurate, methods that may be less generalizable. Nevertheless, we also focus on high-SNR and high-coherency sections of the cable and avoid spatial downsampling so that the algorithm can be evaluated with DAS data at its full channel density. For such an application, despite the limited azimuthal coverage, the resulting abundance of P-wave differential times is expected to sufficiently constrain source locations. For two DAS channels with strain/strain-rate waveforms U_i and U_j, the cross-correlation is computed as shown in Eq. 2:

$$CC_{i,j}(k) = \frac{\sum_{n=0}^{N-1} U_i(n)\, U_j(n+k)}{\sqrt{\sum_{n=0}^{N-1} U_i^2(n)} \sqrt{\sum_{n=0}^{N-1} U_j^2(n)}} \tag{2}$$

where N is the length of the selected time window around the P-wave arrival time, and k is the time lag. $\mathrm{CC}_{i,j}(k)$ is normalized to the product of the norms of U_i and U_j, thus providing a value of 1 for autocorrelation and 0 for uncorrelated signals.

The MCC values and the P-wave differential times τ are, respectively, the maximum value of the cross-correlation function (Eq. 3) and its corresponding temporal lag (Eq. 4).

$$MCC_{i,j} = \max_k CC_{i,j}(k) \tag{3}$$

$$\tau_{i,j} = \arg\max_k CC_{i,j}(k) \cdot dt \tag{4}$$

where dt is the sampling rate.

3.2 Automated DAS data weighting

3.2.1 Hyperparameters to define the weighting scheme

To develop the automated weighting algorithm for P-wave arrival and differential times, we follow the method presented in Piana Agostinetti et al. (2023). Our starting point is an existing hierarchical MCMC event locator (Riva et al., 2024), which, in addition to sampling physical parameters corresponding to the source location, samples a non-physical parameter for an independent data error estimation. Event locations are estimated within a homogeneous velocity model, where the P-wave velocity (V_P) can be jointly sampled with the source location parameters (easting, northing, depth, origin time). In this work, however, we fix the P-wave velocity and event depths to focus on epicentral locations and to avoid possible trade-offs among event depth, P-wave velocity, and origin time that arise from inverting P-phase-only-derived arrival/differential times. In Riva et al. (2024), the hyperparameter for P-wave arrival times (H_1 in this work) is treated as a value sampled within the Markov chains, alongside the other physical parameters, and it scales the data error standard deviation. In other words, it provides an independent estimate of the "uncertainty" associated with the observed data and the assumed forward model, since the hyperparameter adjusts the original, user-defined value (Malinverno and Briggs, 2004). Here, the novelty is the presence of additional hyperparameters (H_2, H_3, H_4, H_5) that define *where* and *how* the data are weighted for the event location. These hyperparameters include: a) "thresholds" based on waveform attributes (such as SNR) to identify a subset of data with greater importance in the inversion, and b) "weights" to adjust the error covariance matrix in the likelihood function. Despite the presence of new hyperparameters to be sampled, the overall computational costs remain similar to those of the original algorithm that does not weight the data (Riva et al., 2024) for applications involving P-wave arrival times: with just a few minutes (Intel i5-11320H, 3.20GHz, 8 logical CPUs, 2 threads, RAM 15 GB, GPU Intel Corporation TigerLake-LP GT2 – Iris Xe Graphics) the user can estimate a location considering a relevant number of Monte Carlo models ($> 10^5$). Applying the method to P-wave differential times naturally requires additional processing time due to the larger data volume involved.

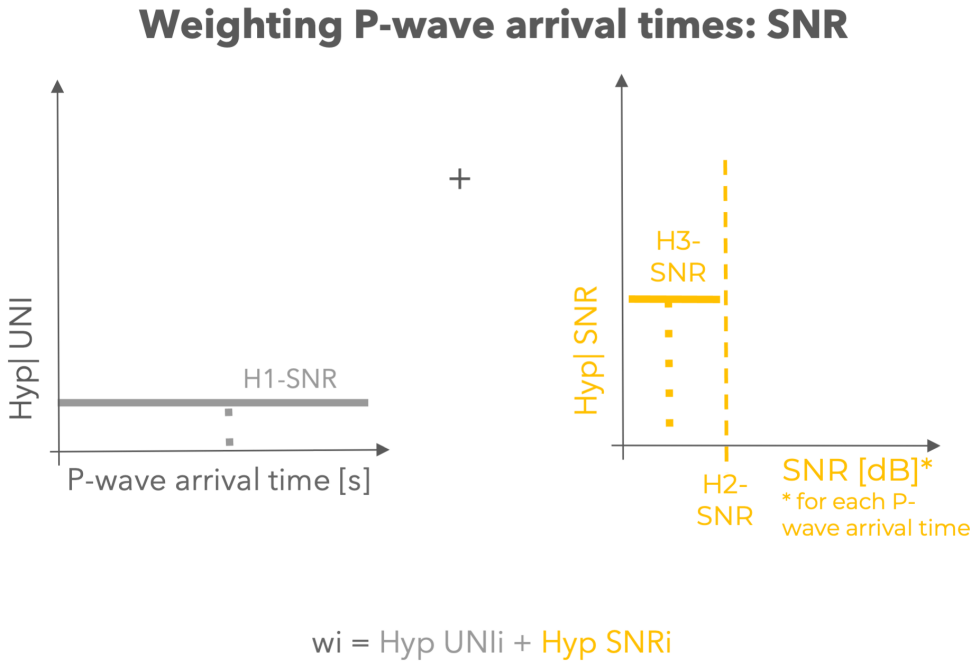

Figure 2 Weighting scheme for P-wave arrival times using SNR attributes. H_1 represents the hyperparameter that uniformly multiplies the diagonal covariance matrix, as in (Riva et al., 2024). The weighting scheme is represented by $H_{2(SNR)}$ and $H_{3(SNR)}$. $H_{2(SNR)}$ is the threshold sampled on the SNR, while the associated weight ($H_{3(SNR)}$) scales the covariance matrix entries for data below $H_{2(SNR)}$. The total factor scaling each covariance matrix entry w_s is therefore the combination of the original uniform scaling (Hyp_UNI_i in the figure) and the attribute-based weighting (Hyp_SNR_i in the figure).

Data weighting is achieved by scaling the entries of the diagonal error covariance matrix for the set of data that do not satisfy the Bayesian-sampled threshold in the form of a non-physical parameter. Specifically, the adopted weighting schemes, implemented in two versions of the algorithm for application to either P-wave arrival times or differential times, determine whether data are weighted above (exceeding the threshold) or below (falling short of the threshold) the threshold. Consequently, data that exceed (falling short) the sampled threshold are penalized in the event location process by multiplying the original, uniform entries of the diagonal covariance matrix (COV_{orig}), defined a priori by the user and already coherently scaled by the H_1 hyperparameter (Riva et al., 2024), by an additional factor. The modified covariance matrix, also diagonal (COV_{weight}), is then computed as follows (Eq. 5):

$$\mathbf{COV}_{\text{weight}} = \mathbf{W}^{-1} \cdot \mathbf{COV}_{\text{orig}} \cdot \mathbf{W}^{-1} \tag{5}$$

where $\mathbf{W}$ is a diagonal matrix built from the sampled set of weights $\mathbf{w}_s$, defined as

$$\mathbf{W} = \text{diag}\left(10^{-\mathbf{w}_s}\right).$$

$\mathbf{COV}_{\text{weight}}$ enters the likelihood function, and the weights are sampled according to the Metropolis rule (Sambridge and Mosegaard, 2002), as are the thresholds and physical parameters. More details of the general hierarchical MCMC algorithm can be found in (Piana Agostinetti et al., 2023).

3.2.2 Adopted weighting schemes

Two versions of the algorithm are developed to account for automated Bayesian weighting of P-wave arrival and differential times.

P-wave arrival times The weighting scheme for absolute arrival times is based on two hyperparameters that depend on the SNR and is defined as follows: a threshold ($H_{2(SNR)}$) on the SNR and a weight ($H_{3(SNR)}$) which, together with the uniform (un-weighting) error estimate ($H_{1(SNR)}$) is multiplied by the a priori-defined and equal variances of the data having SNR lower than $H_{2(SNR)}$ (Eq. 6). $w_{s,SNR}$ is thus an array of weights with the same dimension as the data.

$$\begin{cases} w_{s,SNR} = H_{3(SNR)} & \text{if } SNR_s < H_{2(SNR)}, \\ w_{s,SNR} = 0 & \text{if } SNR_s > H_{2(SNR)}, \end{cases} \tag{6}$$

As a result, the final multiplicative factor of the error covariance matrix is given by the uniform hyperparameter, $H_{1(SNR)}$, and the weighting, non-uniform hyperparameter, H_3 (Figure 2).

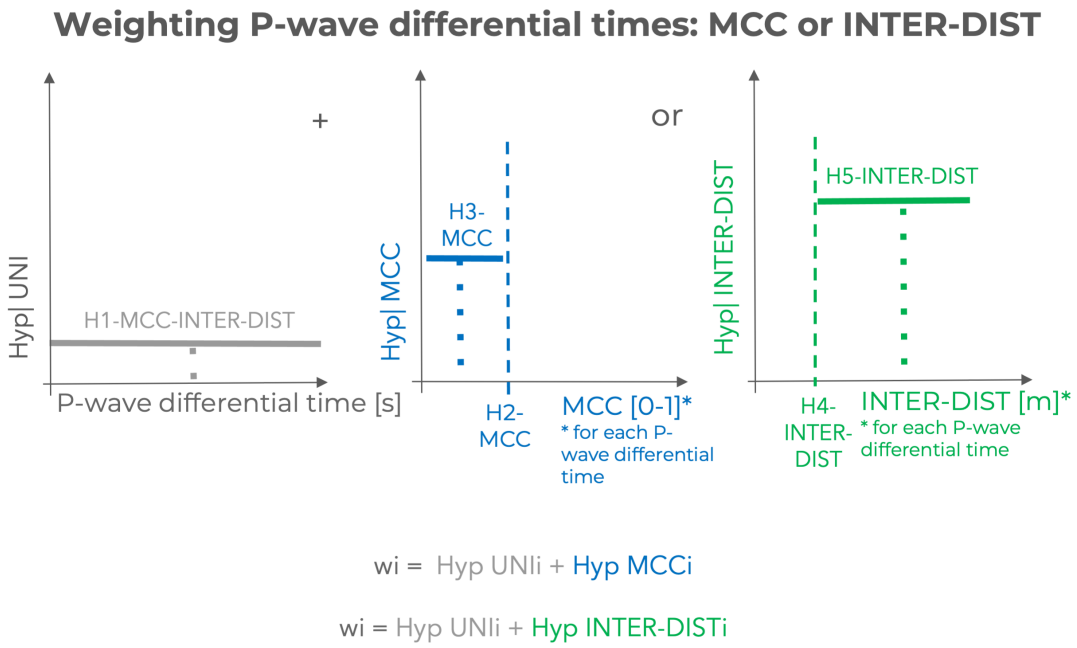

Figure 3 Weighting scheme for P-wave differential times using MCC and INTER-DIST attributes. As in Figure 2, $H_{1(MCC,INTER-DIST)}$ represents the hyperparameter that uniformly multiplies the diagonal covariance matrix, as in (Riva et al., 2024). The weighting scheme is presented here with two alternatives: MCC or INTER-DIST. $H_{2(MCC)}$ and $H_{4(INTER-DIST)}$ are the thresholds on MCC and INTER-DIST, respectively, with the associated weights ($H_{3(MCC)}$ and $H_{5(INTER-DIST)}$) scaling the covariance entries for data below $H_{2(MCC)}$ and above $H_{4(INTER-DIST)}$. The total factor scaling each covariance matrix entry w_i is therefore the combination of the original uniform scaling (Hyp_UNI_i) and the MCC or INTER-DIST weighting (Hyp_MCC_i and $Hyp\_INTER-DIST_i$).

P-wave differential times The weighting scheme for P-wave differential times is based on four hyperparameters that depend on MCC and INTER-DIST and are defined as follows:

- A threshold ($H_{2(MCC)}$) on the MCC and a weight ($H_{3(MCC)}$) which is multiplied by the variances of the data having an MCC value lower than $H_{2(MCC)}$;
- A threshold ($H_{4(INTER-DIST)}$) on the INTER-DIST and a weight ($H_{5(INTER-DIST)}$) which is multiplied by the variances of the data having an INTER-DIST higher than $H_{4(INTER-DIST)}$.

While the algorithm can, in principle, estimate all four hyperparameters simultaneously, we used MCC and INTER-DIST weighting independently to avoid potential trade-offs among the hyperparameters. Moreover, while the MCC typically decreases with increasing inter-channel distance (Menke et al., 1990), in a previous study, we found that the expected decay is often more complex in DAS data (Bozzi et al., 2024). Accordingly, the tests are conducted using two distinct schemes: $H_{2(MCC)}$ and $H_{3(MCC)}$, or $H_{4(INTER-DIST)}$ and $H_{5(INTER-DIST)}$. The final weight w_s for each is defined by the two sub-weights ($w_{s,MCC}$ or $w_{s,INTER-DIST}$), depending on the application (Eq. 7):

$$\begin{cases} w_{s,MCC} = H_{3(MCC)}, & \text{if MCC}_s < H_{2(MCC)}, \\ w_{s,MCC} = 0, & \text{if MCC}_s > H_{2(MCC)}, \\ w_{s,INTER-DIST} = H_{5(INTER-DIST)}, & \text{if INTER-DIST}_s > H_{4(INTER-DIST)}, \\ w_{s,INTER-DIST} = 0, & \text{if INTER-DIST}_s < H_{4(INTER-DIST)}, \end{cases} \tag{7}$$

As a result, the final multiplicative factor of the error covariance matrix is determined by the uniform hyperparameter H_1 along with the appropriate non-uniform weighting hyperparameter: $H_{3(\text{MCC})}$ when the MCC data space is explored, or $H_{5(\text{INTER-DIST})}$ when the INTER-DIST space is explored (Figure 3).

Figure 2 and 3 provide an overview of the weighting schemes tested in this work. Overall, the information required for the algorithm differs depending on the data type: (a) P-wave arrival times and SNR at each triggered DAS channel, or (b) P-wave differential times, MCC, and INTER-DIST values at a set of DAS channel pairs.

3.3 Tests

The automatic DAS data weighting algorithm is tested on synthetic data, part of which is shown in the text (Figs. 2 to 7) and part in the supplementary material (Figs. S1 to S15). Then, two earthquake recordings are used as a first

application to real data (Figs. 8 and 9 and S16 to S18). Synthetic tests are conducted by predicting P-wave arrival and differential times from a known source in the model space, with Gaussian noise added according to the chosen hyperparameters, mimicking the weighting schemes. More specifically, we first select the threshold (hyperparameter) over waveform attributes, thereby determining the portions of the data contaminated with Gaussian noise. The noise standard deviation is then defined by the corresponding weight (hyperparameter). This procedure is consistent with the computation of the weighting matrix (scaling the covariance matrix) W (Eq. 5).

For synthetic tests, we use the geometrical information from AZ-V and G-C deployments, along with the associated earthquake recordings, to compute the corresponding waveform attributes (see Figures 1a and 1b). These attributes are then paired with a chosen set of hyperparameters to contaminate a portion of the synthetic traveltimes with Gaussian noise. Tab. 5 (see Supplementary Materials) summarizes the selected hyperparameter values for each synthetic test. We synthetically invert for easting (X), northing (Y) and origin time (T_0), and the set of hyperparameters ($H_{1(SNR)}$ with $H_{2(SNR)}$ and $H_{3(SNR)}$ or $H_{1(MCC)}$ with $H_{2(MCC)}$ and $H_{3(MCC)}$/$H_{1(INTER-DIST)}$ with $H_{4(INTER-DIST)}$ and $H_{5(INTER-DIST)}$). Thus, we fix the depth and model velocity to avoid trade-offs and to focus on the effects of the weighting hyperparameters on the epicentral location, since the P-wave is the only seismic phase considered. For synthetic P-wave arrival times, we use the entire cable geometry for both AZ-V and G-C. For differential times, however, we either use stacked waveforms from the full cable, providing a spatially downsampled version of the original channels, or we use full-resolution channel data, but only for selected cable sections. This strategy limits computational costs given the large number of possible channel pairs. Thus, while the stacked waveforms ensure azimuthal coverage, the original traces focus on sections closer to the epicenter, where signal quality is generally higher.

For the applications to real DAS data, we use two local earthquakes recorded by the AZ-V and G-C deployments. While in the synthetic tests, the objective is to recover the chosen epicenter and hyperparameters, for the real-data tests, our goals are to: (i) determine whether the method retrieves meaningful (e.g., not zero, very-low, very-high) hyperparameter thresholds on waveform attributes, and (ii) evaluate whether the automatic weighting procedure improves event-location accuracy relative to the unweighted approach (Riva et al., 2024), via comparison with catalog locations. The reference locations are known from a local dense network of seismometers (Nishimura et al., 2021) (AZ-V) and from the catalog of the local monitoring agency, that is, the Instituto Geográfico Nacional of Spain (G-C) (Instituto Geográfico Nacional (IGN), 2022). For consistency among the unweighted and weighted cases, we fix the model velocity to 2500 m/s and 5000 m/s for AZ-V and G-C, respectively. The depth is fixed at –800 m for AZ-V and –10,000 m for G-C. To weight the data, we set flat and wide prior probability distributions for the physical parameters and hyperparameters (see Tab. S6 of the Supplementary Materials). Thus, the algorithm samples from a broad range of possible combinations of SNR, MCC, and INTER-DIST thresholds together with their associated weights. In principle, we also allow for inverse weighting by setting very wide prior distributions, including negative values, a further test of the algorithm's ability to retrieve meaningful thresholds and weights.

4 Results

4.1 Synthetic tests

We present results from synthetic tests using automatic weighting schemes based on P-wave arrival and differential times for both the AZ-V and G-C deployments.

4.1.1 P-wave arrival times

Figure 4 shows the source location and hyperparameters PPDs for a synthetic test of the automatic weighting algorithm using P-wave arrival times for the AZ-V deployment. Additional tests are provided in the supplementary material (Figs. S1 to S4), which, through the exploration of different source positions and weighting schemes, support the presented results and the robustness of the interpretations.

The estimated model parameters and hyperparameters (Figures 4a and 4b) are generally consistent with the true values used to generate the synthetic data. Specifically, the true location falls within the 90% confidence bound of the possible solutions. The model parameter PPDs exhibit clear bell-shaped distributions. The PPDs for the hyperparameters $H_{1(SNR)}$ and $H_{3(SNR)}$ are also bell-shaped, while the PPD for $H_{2(SNR)}$ has a more box-like shape (Figure 4c).

Similar observations are made for the inversion of synthetic P-wave arrival times from the G-C deployment, whose results are illustrated in Figure 5 and Figure S5. Wider epicentral location PPDs are observed, as expected, given the quasi-rectilinear shape of the cable (Figure 5c).

Ten Markov chains, each with 1×10^5 samples, are used to sample the PPDs in these two tests. The runtime is less than one minute on the same machine mentioned above.

4.1.2 P-wave differential times

Figure 6 shows the results from the test with synthetic P-wave differential times for the AZ-V deployment. Additional tests are provided in the supplementary material (Figs. S6 to S15).

As in the synthetic tests with P-wave arrival times, the estimated model parameters (Figures 6a and 6b) and hyperparameters (Figures 6c and 6d) are generally consistent with the true values used to generate the synthetic data. However, in this specific weighting scheme, the PPDs for the hyperparameters $H_{1(MCC)}$ and $H_{3(MCC)}$ deviate from a Gaussian distribution, exhibiting spike-shaped distributions for both MCC (Figure 6c) and INTER-DIST schemes (Figure 6d). Addi-

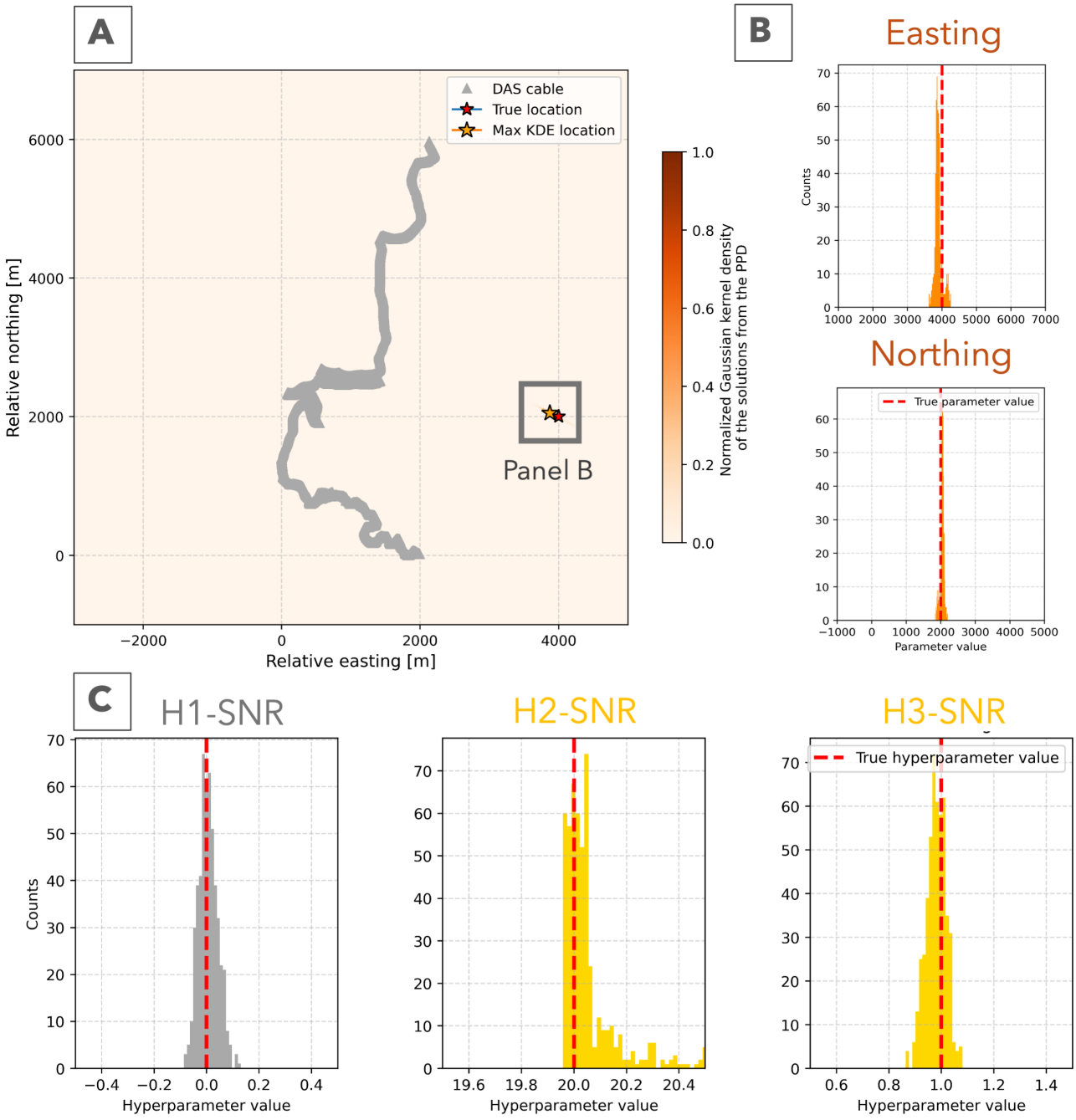

Figure 4 Synthetic test for P-wave arrival times weighted using SNR attributes and AZ-V fiber layout. (a) and (b) PPDs for easting and northing showing the solution density (left) and corresponding histograms (right), given the very narrow distribution of samples around the true epicenter (red dashed vertical line); (c) PPDs for the hyperparameters $H_{1(SNR)}$, $H_{2(SNR)}$ and $H_{3(SNR)}$ compared with their true values (red dashed lines).

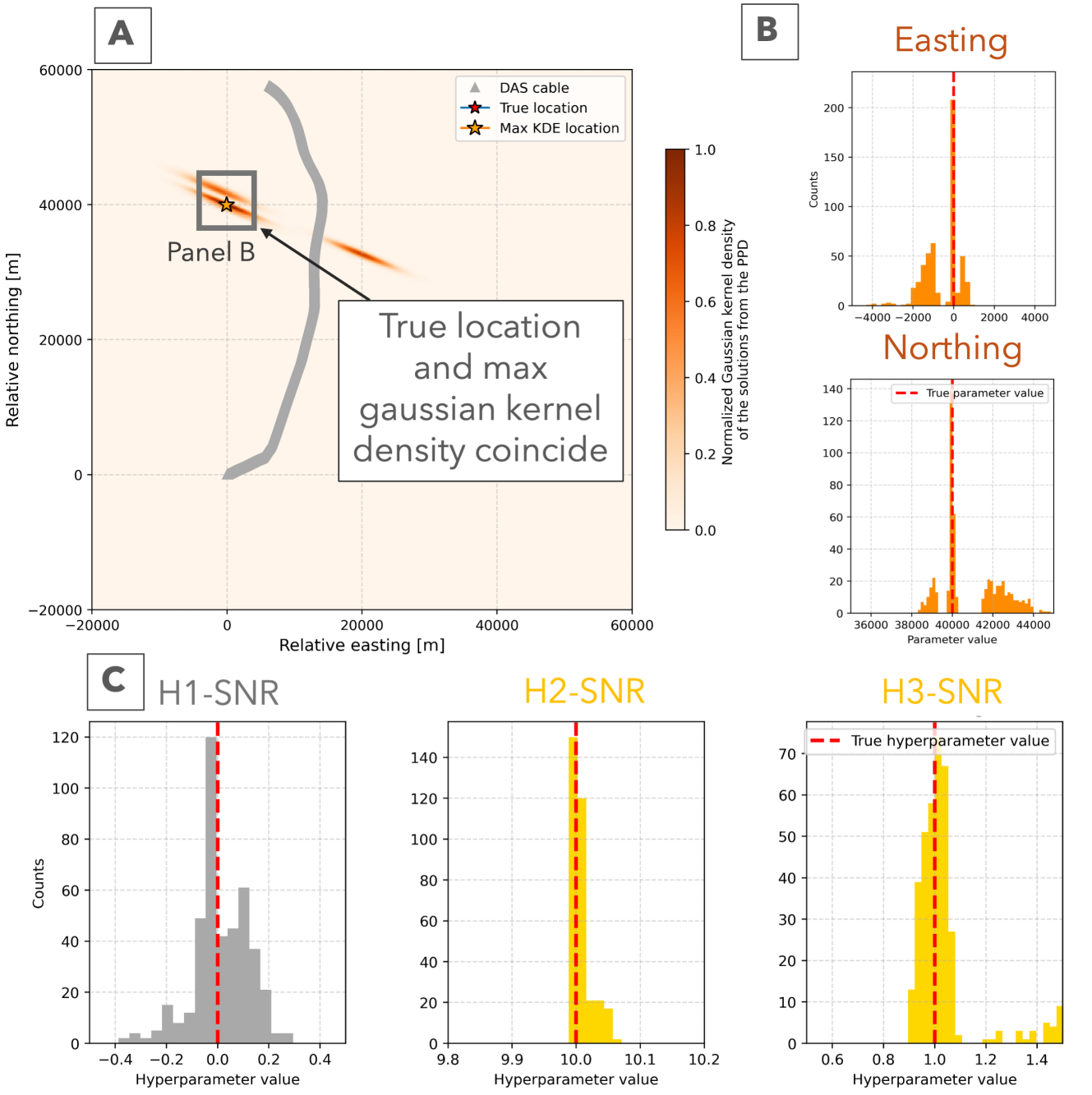

Figure 5 Synthetic test for P-wave arrival times weighted using SNR attributes and G-C fiber layout. (a) and (b) PPDs for easting and northing showing the solution density (left) and corresponding histograms (right), given the very narrow distribution of samples around the true epicenter (red dashed vertical line), with the true location coinciding with the maximum-density solution (orange star) in (a); PPDs for the hyperparameters $H_{1(SNR)}$, $H_{2(SNR)}$ and $H_{3(SNR)}$ compared with their true values (red dashed lines).

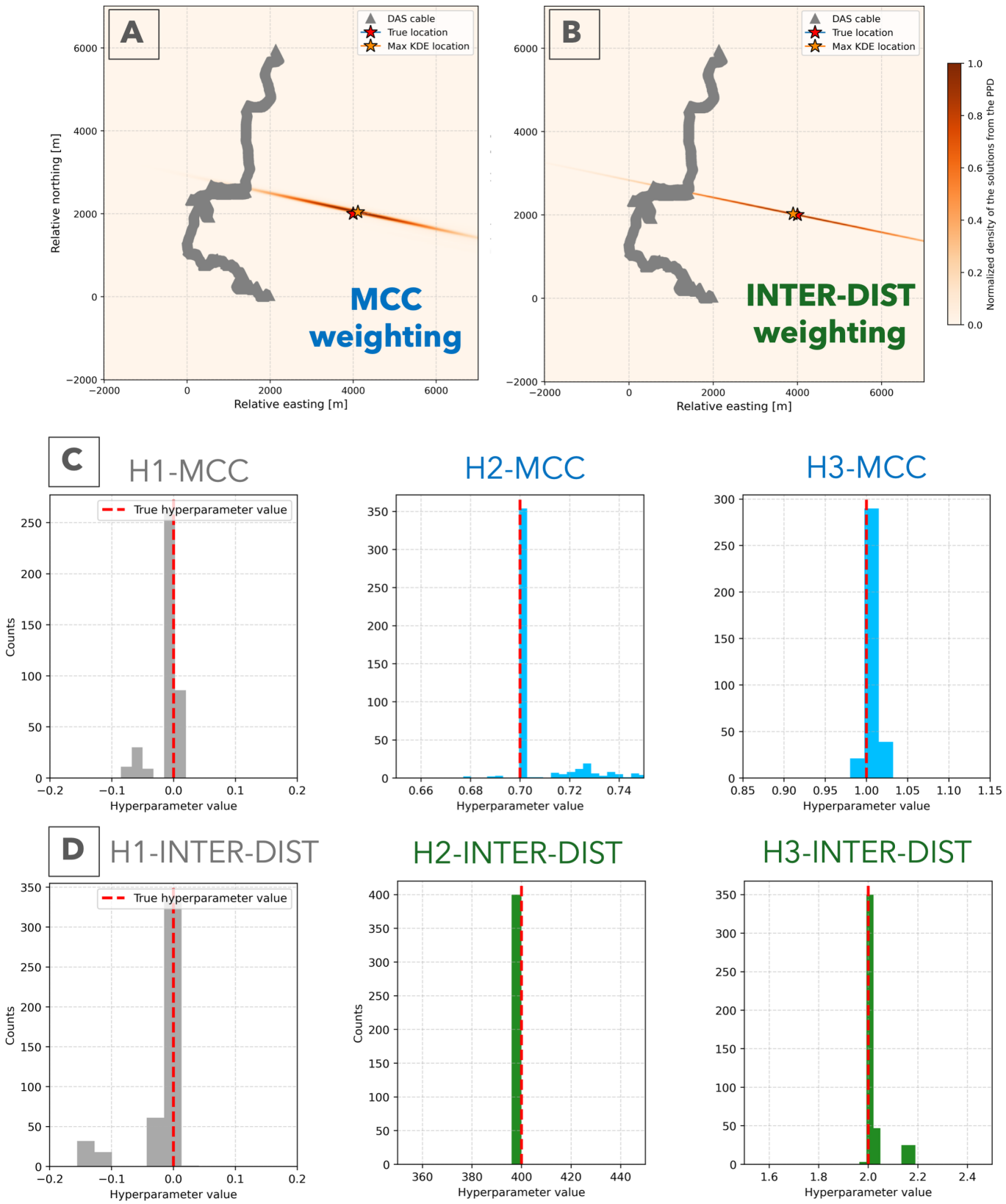

Figure 6 Synthetic test for P-wave differential times weighted using MCC and INTER-DIST attributes with the AZ-V fiber layout. (a) and (b) PPDs for easting and northing showing the solution densities for MCC and INTER-DIST weighting possibilities, respectively. As in Figure 4, the maximum-density solution values are plotted as orange stars and compared with the true solution (red star); (c) PPDs for the hyperparameters when MCC attributes are used for data weighting, compared with their true values (red dashed lines); (d) PPDs for the hyperparameters when INTER-DIST attributes are used for data weighting, compared with their true values (red dashed lines). An initial trade-off between $H_{1(INTER-DIST)}$ and $H_{3(INTER-DIST)}$ is observed, with secondary peaks.

tionally, a secondary peak is observed for INTER-DIST weighting, with corresponding weights that, when combined, almost reproduce the true value (given the true H_1 value equal to zero).

Figure 7 presents the results of the test for synthetic P-wave differential times from the G-C deployment, with additional tests shown in Figs. S11 to S15. Given the quasi-rectilinear geometry of the cable, ambiguities in the solution are expected. Nevertheless, the estimated model parameters (Figures 7a and 7b) are consistent with the true epicenter. The hyperparameters (Figures 7c and 7d) are consistent with the true values, although their distributions deviate from a Gaussian shape, as in the AZ-V case. In this case as well, a secondary peak is present for $H_{1(MCC)}$ and $H_{3(MCC)}$ with MCC-based weighting.

Similar computational costs are observed compared to the P-wave arrival-time case, due to the dimensionality reduction achieved through channel stacking.

4.2 Applications to real data

We present results from the application of the automatic weighting schemes based on P-wave arrival and differential times to real DAS data from both the AZ-V and G-C deployments.

4.2.1 P-wave arrival times

Figure 8 shows locations obtained with the weighting schemes, which are either comparable to the unweighted solutions (Riva et al., 2024) (Figure 8c) or closer to the reference epicenters (Figure 8d). We used the distance between the peak density of the solutions and the reference epicentral solutions as a metric for accuracy. Moreover, the PPDs for the SNR threshold hyperparameters (H_2) are centered on values far from model boundaries: about 10 and 4 dB for AZ-V and G-C, respectively (Figures 8e and 8f). Indeed, the identified thresholds effectively separate data with clearly lower SNRs from higher-quality data (Figs. 8a and 8b). In other words, data with associated SNRs below 10 dB for AZ-V and 4 dB for G-C are treated by the algorithm as less reliable in the inversion.

4.2.2 P-wave differential times

Figure 9 shows the automatic weighting scheme based on P-wave differential times applied to subsections of 100 channels. Additionally, results using data from the full cable length are shown in Figs. S16 and S17. We omit these latter results from the main text because of the poor event locations in both the weighted and unweighted cases, with high H_1 values indicating large data uncertainties in this configuration (Piana Agostinetti et al., 2023).

Overall, the results for cable subsections show very similar unweighted and weighted epicentral solutions (Figures 9c and 9d). The PPDs for the MCC and INTER-DIST threshold hyperparameters ($H_{2(MCC)}$ and $H_{4(INTER-DIST)}$) are centered on values far from model boundaries: about 0.67 and 0.64 for MCC, and 90 m and 700 m for INTER-DIST, for AZ-V and G-C, respectively (Figures 9e and 9f). In other words, P-wave differential times with MCC values below thresholds of 0.67 for AZ-V and 0.64 for G-C are treated by the algorithm as less reliable in the inversion (see also Figure 9b). Similarly, P-wave differential times with INTER-DIST values above thresholds of 90 m for AZ-V and 700 m for G-C are treated as less reliable (see also Figure 9b). Nevertheless, all identified weights ($H_{2(MCC)}$ and $H_{4(INTER-DIST)}$) are very small compared with H_1, as shown in Figure S18. Moreover, as observed in some of the synthetic tests, the distributions for $H_{3(MCC)}$ and $H_{5(INTER-DIST)}$ are closer to a Gaussian shape when compared with those for $H_{2(MCC)}$ and $H_{4(INTER-DIST)}$.

5 Discussion

We present the testing of a hierarchical MCMC method that automatically weights DAS data for epicentral location. Synthetic tests are conducted using both P-wave arrival and differential times for two cable geometries: one complex, with subsections of different orientations in a volcanic area, and one quasi-rectilinear in a submarine context. Results highlight the method's robustness to variations in source–cable geometry and to different threshold and weight configurations, which are here sampled as hyperparameters of the MCMC algorithm.

Applications of the method to real data show that automatic weighting allows the recovery of event locations at least comparable to those obtained with unweighted inversions (Riva et al., 2024) (Figure 8). Notably, the use of a homogeneous velocity model is not ideal for the complex installation contexts of the two DAS arrays. Therefore, the resulting event locations are not expected to be more accurate than those in the reference catalog, which were obtained using 1D velocity models. However, since the purpose of this study is to evaluate the relative impact of the automatic weighting scheme compared to an unweighted location, the adoption of a homogeneous model is justified, as it enables a consistent comparison with the reference catalog. For the weighting scheme based on SNR and P-wave arrival times, we observe improvements in event locations, with maxima of the epicentral parameters' PPDs closer to the reference locations. Additionally, for both P-wave arrival and differential times, the algorithm consistently identifies realistic SNR, MCC, and INTER-DIST thresholds in real-data tests, even when the prior model space is wide and allows for opposite weighting (i.e., higher weights for less accurate observations). In these cases, it automatically favors high-SNR, high-MCC, and low-INTER-DIST observations as the most accurate for inversion, while down-weighting less trustworthy data, without external constraints. Since the computational cost is comparable to that of the unweighted algorithm and the resulting locations are at least as accurate, the presented MCMC-based weighting method represents an opportunity to potentially obtain more accurate epicentral estimates.

The assumptions of this probabilistic weighting, including a wide prior model space for hyperparameters,

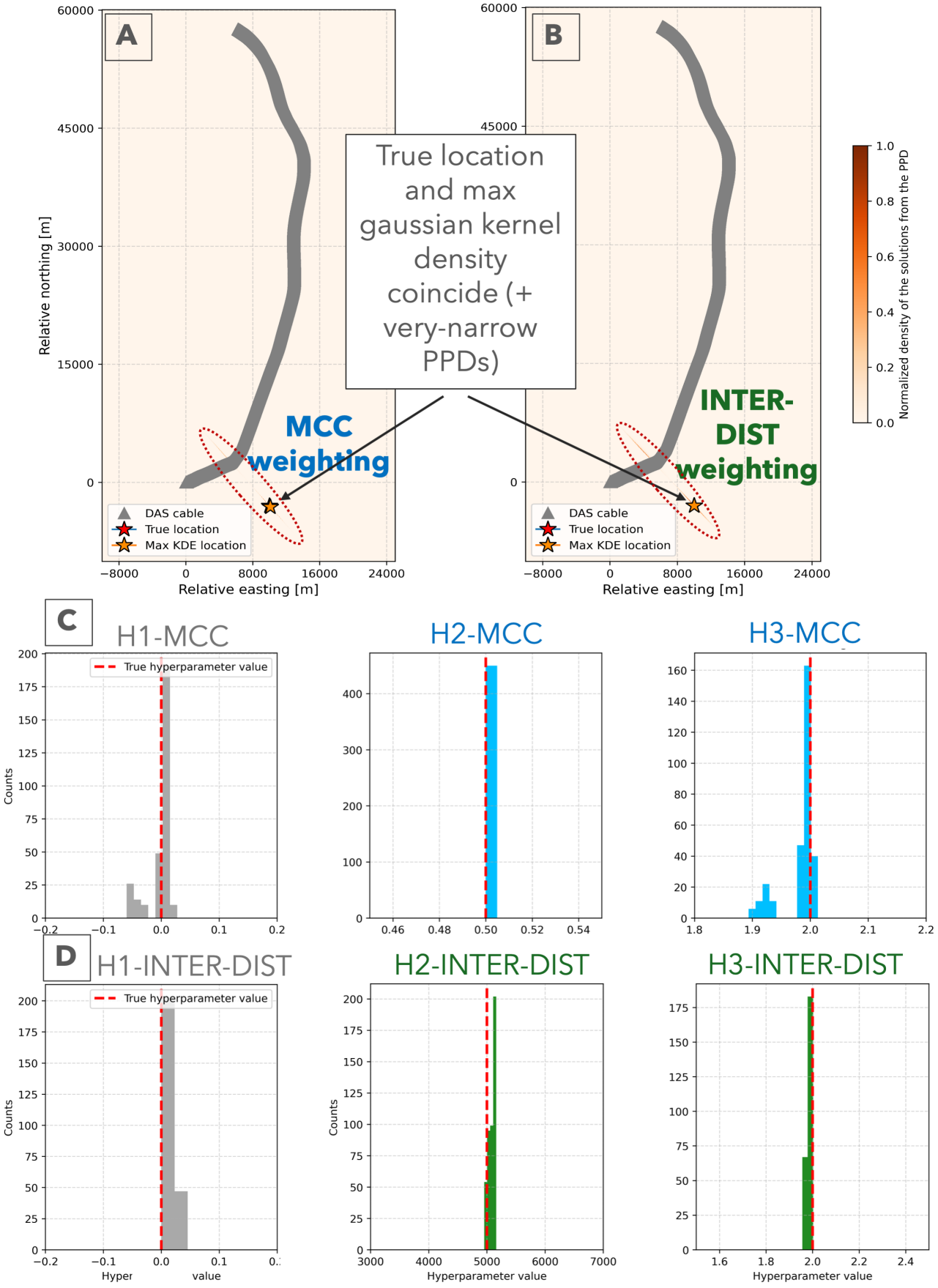

Figure 7 Synthetic test for P-wave differential times weighted using MCC and INTER-DIST attributes with the G-C fiber layout. (a) and (b) PPDs for easting and northing showing the solution densities for MCC and INTER-DIST weighting possibilities, respectively. The maximum-density solution is plotted as an orange star and compared with the true solution (red star); (c) PPDs for the hyperparameters when MCC attributes are used for data weighting, compared with their true values (red dashed lines). An initial trade-off between $H_1(MCC)$ and $H_3(MCC)$ is observed, with secondary peaks; (d) PPDs for the hyperparameters when INTER-DIST attributes are used for data weighting, compared with their true values (red dashed lines).

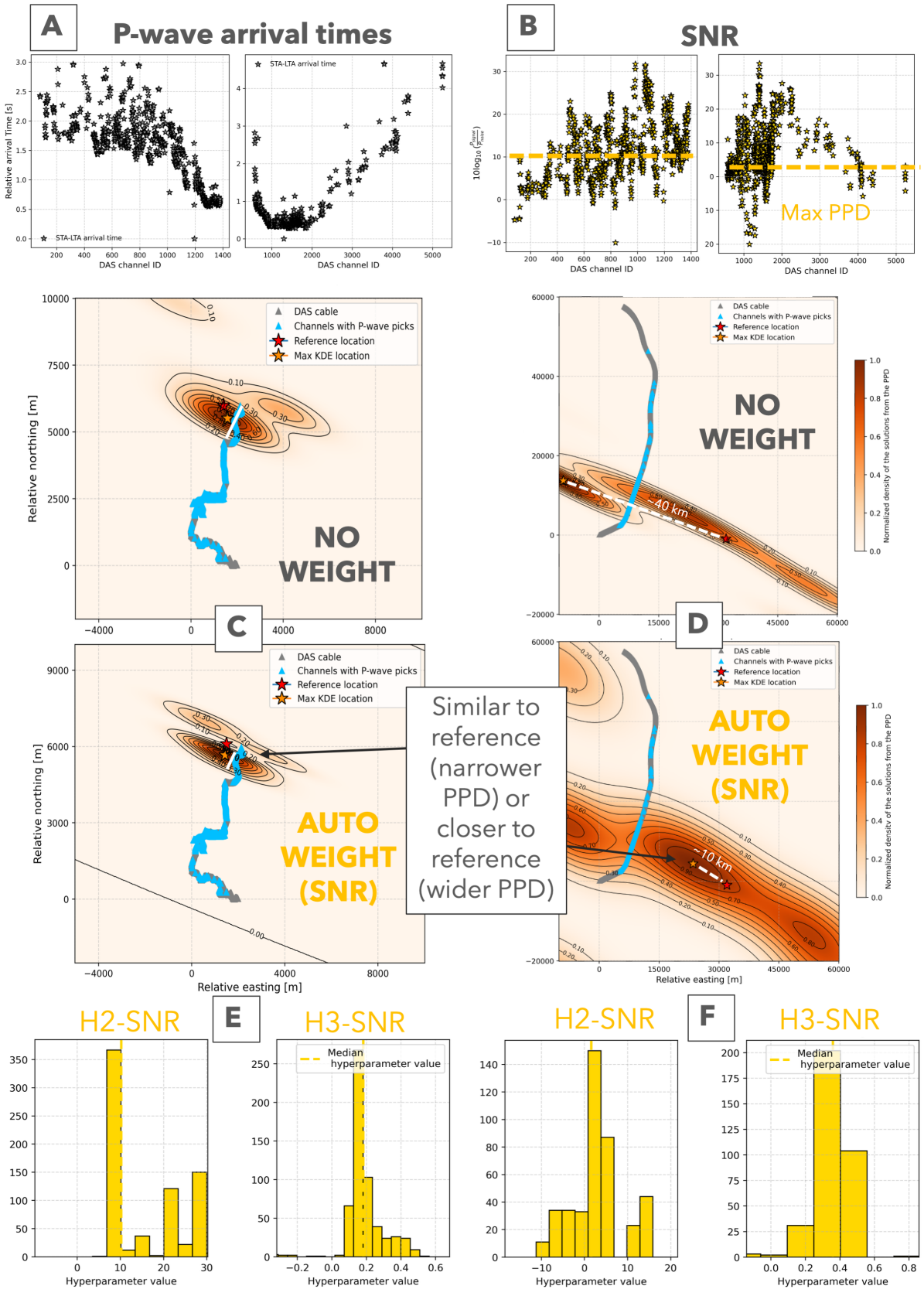

Figure 8 Real-data test for P-wave arrival times weighted using SNR attributes and the AZ-V and G-C fiber layouts. (a) and (b) Estimated P-wave arrival times obtained with STA/LTA and their associated SNR values. The yellow dotted line represents the identified SNR thresholds, automatically classifying data as "less reliable" in the inversion (below the threshold) or "more reliable" (above the threshold); (c) (AZ-V) and (d) (G-C) PPDs for epicentral location showing solution densities for the unweighted solution (Riva et al., 2024) and the automatic weighting presented here. The maximum density of the solution is shown as an orange star for comparison with the reference solution (red star). The distance between the peak density and the reference solution, used as a metric for accuracy, is shown with a white dotted line. DAS channels with an associated estimated P-wave arrival time are plotted in light blue; (e) (AZ-V) and (f) (G-C) PPDs for the hyperparameters $H_{2(SNR)}$ and $H_{3(SNR)}$.

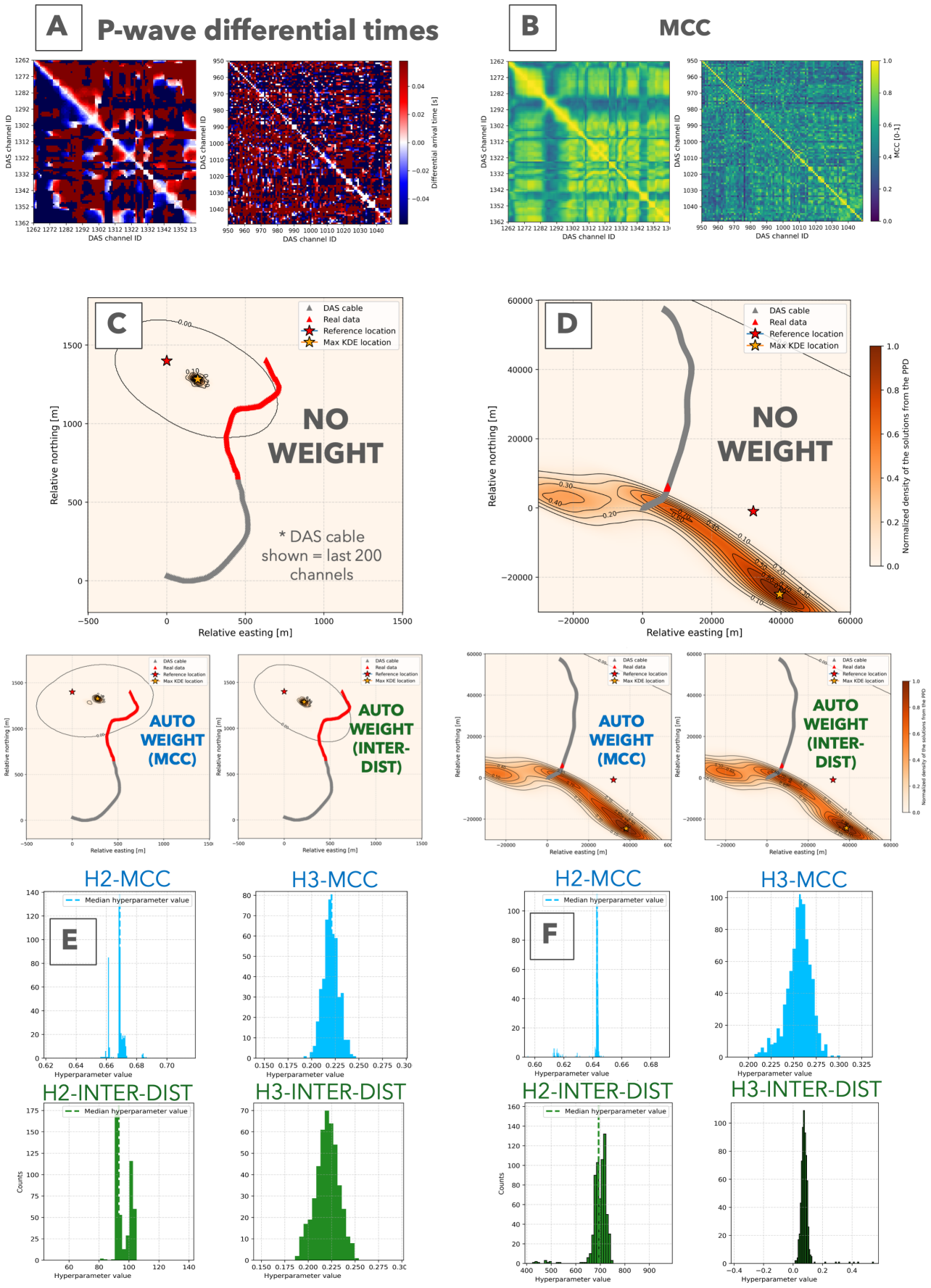

Figure 9 Real-data test for P-wave differential times weighted using MCC and INTER-DIST attributes and the AZ-V and G-C fiber layouts. (a) and (b) Estimated P-wave differential times for the channels highlighted in red in (c) and (d), along with their associated MCC values; (c) (AZ-V) and (d) (G-C) PPDs for epicentral location showing solution densities for the unweighted solution (Riva et al., 2024) and the two automatic weighting cases presented here (MCC and INTER-DIST). The maximum density of each solution is shown as an orange star for comparison with the reference solution (red star); (e) (AZ-V) and (f) (G-C) PPDs for the hyperparameters $H_{2(MCC)}$, $H_{3(MCC)}$, $H_{4(INTER\text{-}DIST)}$, and $H_{5(INTER\text{-}DIST)}$.

are appropriate for the specific characteristics of DAS technology. Indeed, compared to point-like conventional sensors, DAS sensors and their data are affected by additional factors such as axial sensitivity, poorly constrained site conditions, and cable coupling. These factors lead to strong spatio-temporal variability in DAS signal attributes, such as SNR and waveform coherence, making subjective data weighting difficult either because of the data volume or the complexity of modeling DAS signal properties. The proposed probabilistic framework explores a large number of potential weighting configurations without prior assumptions, relying exclusively on waveform attributes such as SNR, MCC, or INTER-DIST. This approach minimizes subjective intervention while identifying less reliable observations through systematic data-space exploration.

In principle, additional user-defined attributes, such as cable–ground coupling (Celli et al., 2024) or directional sensitivity (Martin et al., 2017), could also be incorporated. Moreover, as new DAS quality metrics are emerging (Rodriguez et al., 2025), these can readily serve as input to the algorithm. Indeed, the method remains flexible and effective as long as selected attributes are quantifiable within a defined range of values. We therefore encourage tests by the broader seismological community.

Most difficulties arise when locating real events using P-wave differential times, likely due to complex signal coherence and challenges in tracking the P phase along the entire cable length (van den Ende and Ampuero, 2021; Bozzi et al., 2024). Our tests confirm that event locations derived from differential times using channel pairs along the full cable are mislocated (Figs S16 and S17), near the limits of the model space, even when automatic weighting is applied. Conversely, when using P-wave differential times retrieved from high-SNR cable sections, the estimated locations are closer to the reference catalog. Still, when applying automatic weighting, the algorithm assigns very low weights, suggesting difficulties in isolating clear outlier patterns within the data space. Overall, tests on real data using P-wave differential times reveal limitations likely unrelated to the weighting scheme, confirmed to work with synthetic tests, but instead to the highly complex data-space structure.

From a computational perspective, the use of P-wave differential times substantially increases the number of data points to be inverted, as each channel is cross-correlated with all others, compared with the arrival-time case. The algorithm remains fast for absolute P-wave arrival times, achieving converged PPDs within seconds even for thousands of DAS channels. However, computational costs increase approximately quadratically when using differential times, as the number of observations grows from N to $N(N-1)/2$.

Because automated processing is essential for data-intensive technologies like DAS, the proposed method can support their integration into conventional seismic networks (Biondi et al., 2025; Baillet et al., 2025). Its main advantage lies in mitigating the influence of unreliable DAS data, ensuring that such data are not weighted equally with the more precise, often manually verified, measurements from traditional instruments. The weighting framework currently operates on P-wave arrival and differential times, with a justified focus on P-waves due to the difficulty of automatically identifying S-phases in DAS data using energy-based methods such as STA/LTA. However, as new AI-based algorithms (Zhu et al., 2023) improve coherent P- and S-phase picking in DAS, extending the weighting framework to include S-phases represents a natural next step.

6 Conclusions

The algorithm presented here represents a step toward the automatic weighting of DAS P-wave arrival and differential times for source location. Weighting schemes are defined through thresholds and weights applied to measured waveform attributes, which are sampled as non-physical parameters (hyperparameters) in a Bayesian framework that scales the covariance matrix of the inverse problem. This allows the algorithm to automatically distinguish between portions of data, more or less reliable, during the inversion.

Synthetic tests demonstrate that the identified weighting schemes successfully recover both the epicentral parameters and the hyperparameters. We show that automatic data weighting applied to real datasets can improve epicentral location accuracy compared to unweighted solutions for the application involving P-wave arrival times weighted using SNR attributes. On the other hand, no significant improvements are observed when applying automatic P-wave–differential-time–based weighting schemes using either the MCC or INTER-DIST features. Since the unweighted and weighted solutions are obtained in comparable runtime and with at least similar accuracy, the proposed algorithm provides a valid and efficient alternative.

The method is fully data-driven and therefore well-suited for instruments that generate high data density, such as DAS, where subjective data treatment is particularly challenging. Although this study focused on attributes such as SNR and MCC for weighting, the algorithm remains flexible and can incorporate other properties, provided they are quantifiable within a defined range of values.

Acknowledgements

We thank Takeshi Nishimura, Professor at Tohoku University (Japan), for sharing the data and metadata from the AZ-V DAS deployment.

We thank Arantza Ugalde, Senior Researcher from the Institut de Ciències del Mar, CSIC (Spain), for the support on the G-C dataset. Moreover, we thank CANALINK-Canarias submarine link S.L. (http://www.canalink.tel), part of the Instituto Tecnológico y de Energías Renovables (ITER) and Cabildo de Tenerife, owners of the Canalink fiber optic cable. Funding for the acquisition of the CANDAS2 dataset was provided by the European Union NextGenerationEU/PRTR Program projects PSI (ref. PLEC2021-007875) and TREMORS (ref. CPP2021-008869).

This paper and the related research activities have been conducted during the PhD course in "Chemical, Geological and Environmental Sciences" at the University of Milano-Bicocca. (https://www.disat.unimib.it/it/ricerca/dottorato-scienze-chimiche-geologiche-e-ambientali).

Data and code availability

Data and metadata of the study can be found in:

1. (Nishimura et al., 2021) and a Zenodo repository https://doi.org/10.5281/zenodo.8177033 for the AZ-V dataset;
2. (Ugalde et al., 2023) for the G-C dataset.

The Hierarchical MCMC code used as a base to develop the weighting schemes based on hyperparameters is described in detail in (Riva et al., 2024). Instead, the two versions of the algorithm to reproduce the results shown in this work are available at the following Gitlab repository (versions 1.4 and 1.5): https://gitlab.com/rjMCMC-1D-geophysics/1d-loc-MCMC/-/tree/main/REPOSITORY?ref_type=heads.

Competing interests

The authors have no competing interests.

References

Ajo-Franklin, J. B., Dou, S., Lindsey, N. J., Monga, I., Tracy, C., Robertson, M., Rodriguez Tribaldos, V., Ulrich, C., Freifeld, B., Daley, T., and Li, X. Distributed acoustic sensing using dark fiber for near-surface characterization and broadband seismic event detection. *Scientific reports*, 9(1328):1–14, 2019. doi: 10.1038/s41598-018-36675-8.

Aki, K. and Richards, P. G. *Quantitative seismology*. University Science Books, 2002.

Allen, R. Automatic phase pickers: Their present use and future prospects. *Bulletin of the Seismological Society of America*, 72 (6B):S225–S242, 1982. doi: 10.1785/BSSA07206B0225.

Baillet, M., Rivet, D., Trabattoni, A., Cheze, J., Peix, F., Ambrois, D., van den Ende, M., Vernet, C., Strumia, C., Potin, B., et al. Automatic earthquake catalogs from a permanent DAS offshore network. *Journal of Geophysical Research: Solid Earth*, 130(10): e2025JB031565, 2025.

Biagioli, F., Métaxian, J.-P., Stutzmann, E., Ripepe, M., Bernard, P., Trabattoni, A., Longo, R., and Bouin, M.-P. Array analysis of seismo-volcanic activity with distributed acoustic sensing. *Geophysical Journal International*, 236(1):607–620, 2024. doi: 10.1093/gji/ggad427.

Biondi, B., Martin, E., Cole, S., Karrenbach, M., and Lindsey, N. Earthquakes analysis using data recorded by the Stanford DAS array. In *SEG International Exposition and Annual Meeting*, pages SEG–2017. SEG, 2017.

Biondi, E., Zhu, W., Li, J., Williams, E. F., and Zhan, Z. An upper-crust lid over the Long Valley magma chamber. *Science Advances*, 9(42):eadi9878, 2023. doi: 10.1126/sciadv.adi9878.

Biondi, E., Tepp, G., Yu, E., Saunders, J. K., Yartsev, V., Black, M., Watkins, M., Bhaskaran, A., Bhadha, R., Zhan, Z., et al. Real-time processing of distributed acoustic sensing data for earthquake monitoring operations. *arXiv preprint arXiv:2505.24077*, 2025.

Bocchini, G. M., Bozzi, E., Roth, M. P., Gaviano, S., Pascucci, G., Grigoli, F., Biondi, E., Sokos, E., and Harrington, R. M. Earthquake Catalog and Continuous Waveforms from a Two-Week Distributed Acoustic Sensing experiment on Kefalonia Island, Greece. *Authorea Preprints*, 2025.

Bozzi, E., Saccorotti, G., Agostinetti, P. N., Becerril, C., Fichtner, A., Klaasen, S., Nishimura, T., Shen, J., Walter, F., and Zhu, T. Complex spatial distribution of onset amplitude and waveform correlation: case studies from different DAS experiments. *Bulletin of Geophysics and Oceanography*, 2024. doi: 10.4430/bgo00458.

Bozzi, E., Pascucci, G., Rapagnani, G., Bocchini, G. M., Harrington, R., Ugalde, A., Saccorotti, G., and Grigoli, F. Near real-time channel selection for Distributed Acoustic Sensing technology, 2025. https://arxiv.org/abs/2512.08845.

Cannon, R. and Aminzadeh, F. Distributed acoustic sensing: State of the art. In *SPE Digital Energy Conference and Exhibition*, pages SPE–163688. SPE, 2013. doi: 10.2118/163688-MS.

Celli, N. L., Bean, C. J., and O'Brien, G. S. Full-waveform simulation of DAS records, response and cable-ground coupling. *Geophysical Journal International*, 236(1):659–674, 2024. doi: 10.1093/gji/ggad449.

Daley, T. M., Freifeld, B. M., Ajo-Franklin, J., Dou, S., Pevzner, R., Shulakova, V., Kashikar, S., Miller, D. E., Goetz, J., Henninges, J., et al. Field testing of fiber-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring. *The Leading Edge*, 32(6):699–706, 2013.

Fichtner, A., Klaasen, S., Thrastarson, S., Çubuk-Sabuncu, Y., Paitz, P., and Jónsdóttir, K. Fiber-Optic Observation of Volcanic Tremor through Floating Ice Sheet Resonance. *The Seismic Record*, 2(3):148–155, 2022. doi: 10.1785/0320220010.

Hudson, T., Klaasen, S., Fontaine, O., Zunino, A., van Meulebrouck, S., Walter, F., Jonsdottir, K., and Fichtner, A. Towards a generic fibreoptic earthquake detection and location algorithm for arbitrary fibre geometries and hybrid fibre-seismometer networks. Technical report, Copernicus Meetings, 2024. doi: 10.5194/egusphere-gc12-fibreoptic-26.

Instituto Geográfico Nacional (IGN). Spanish Seismic Catalogue. https://doi.org/10.7419/162.03.2022, 2022. doi: 10.7419/162.03.2022. Accessed: 2025-08-27.

Isken, M. P., Vasyura-Bathke, H., Dahm, T., and Heimann, S. Denoising distributed acoustic sensing data using an adaptive frequency–wavenumber filter. *Geophysical Journal International*, 231(2):944–949, 2022. doi: 10.1093/gji/ggac229.

Jousset, P., Reinsch, T., Ryberg, T., Blanck, H., Clarke, A., Aghayev, R., Hersir, G. P., Henninges, J., Weber, M., and Krawczyk, C. M. Dynamic strain determination using fibre-optic cables allows imaging of seismological and structural features. *Nature communications*, 9(1):1–11, 2018. doi: 10.1038/s41467-018-04860-y.

Klaasen, S., Paitz, P., Lindner, N., Dettmer, J., and Fichtner, A. Distributed acoustic sensing in volcano-glacial environments: Mount Meager, British Columbia. *Journal of Geophysical Research: Solid Earth*, 126(11), 2021. doi: 10.1029/2021JB022358.

Lanza, F., Tuinstra, K. B., Rinaldi, A. P., Tribaldos, V. R., Hopp, C., Bethmann, F., Dyer, B., Meier, P., Serbeto, F., Hertrich, M., et al. A Prospective Evaluation of Borehole Distributed Acoustic Sensing Technologies for Detecting and Locating Out-of-Network Microseismicity. *Distributed Acoustic Sensing in Borehole Geophysics*, pages 311–329, 2024.

Lapins, S., Butcher, A., Kendall, J.-M., Hudson, T. S., Stork, A. L., Werner, M. J., Gunning, J., and Brisbourne, A. M. DAS-N2N: machine learning distributed acoustic sensing (DAS) signal denoising without clean data. *Geophysical Journal International*, 236 (2):1026–1041, 2024. doi: 10.1093/gji/ggad460.

Latorre, H., Ventosa, S., Ugalde, A., Villaseñor, A., Bartolomé, R.,

and Ranero, C. R. KVP: a multiscale kurtosis approach for seismic phase picking. *Geophysical Journal International*, 241(3): 1922–1934, 2025.

Lay, T. and Wallace, T. C. *Modern global seismology*. Elsevier, 1995.

Lellouch, A., Lindsey, N. J., Ellsworth, W. L., and Biondi, B. L. Comparison between distributed acoustic sensing and geophones: Downhole microseismic monitoring of the FORGE geothermal experiment. *Seismological Society of America*, 91(6):3256–3268, 2020. doi: 10.1785/0220200149.

Li, J., Kim, T., Lapusta, N., Biondi, E., and Zhan, Z. The break of earthquake asperities imaged by distributed acoustic sensing. *Nature*, 620(7975):800–806, 2023a.

Li, J., Zhu, W., Biondi, E., and Zhan, Z. Earthquake focal mechanisms with distributed acoustic sensing. *Nature Communications*, 14(4181), 2023b. doi: 10.1038/s41467-023-39639-3.

Lindsey, N. J., Martin, E. R., Dreger, D. S., Freifeld, B., Cole, S., James, S. R., Biondi, B. L., and Ajo-Franklin, J. B. Fiber-optic network observations of earthquake wavefields. *Geophysical Research Letters*, 44(23):11–792, 2017. doi: 10.1002/2017GL075722.

Lior, I., Sladen, A., Rivet, D., Ampuero, J.-P., Hello, Y., Becerril, C., Martins, H. F., Lamare, P., Jestin, C., Tsagkli, S., and Markou, C. On the detection capabilities of underwater distributed acoustic sensing. *Journal of Geophysical Research: Solid Earth*, 126(3): e2020JB020925, 2021. doi: 10.1029/2020JB020925.

Malinverno, A. and Briggs, V. A. Expanded uncertainty quantification in inverse problems: Hierarchical Bayes and empirical Bayes. *Geophysics*, 69(4):1005–1016, 2004. doi: 10.1190/1.1778243.

Martin, E., Biondi, B., Cole, S., and Karrenbach, M. Overview of the Stanford DAS Array-1 (SDASA-1), 2017.

Menke, W., Lerner-Lam, A. L., Dubendorff, B., and Pacheco, J. Polarization and coherence of 5 to 30 Hz seismic wave fields at a hard-rock site and their relevance to velocity heterogeneities in the crust. *Bulletin of the Seismological Society of America*, 80(2): 430–449, 1990.

Miller, M. S., Townend, J., and Lai, V. H. The South Island seismology at the speed of light experiment (SISSLE): Distributed acoustic sensing across and along the Alpine Fault, south Westland, New Zealand. *Seismological Research Letters*, 96(3): 2065–2078, 2025.

Mosegaard, K. and Tarantola, A. Probabilistic approach to inverse problems. *International Geophysics Series*, 81(A):237–268, 2002. doi: 10.1016/S0074-6142(02)80219-4.

Nishimura, T., Emoto, K., Nakahara, H., Miura, S., Yamamoto, M., Sugimura, S., Ishikawa, A., and Kimura, T. Source location of volcanic earthquakes and subsurface characterization using fiber-optic cable and distributed acoustic sensing system. *Scientific reports*, 11(1):1–12, 2021. doi: 10.1038/s41598-021-85621-8.

Noe, S., Tuinstra, K. B., Klaasen, S., Krischer, L., and Fichtner, A. Theoretical background for full-waveform inversion with distributed acoustic sensing and integrated strain sensing. *Geophysical Journal International*, 244(1):ggaf406, 2026.

Obermann, A., Sánchez-Pastor, P., Wu, S.-M., Wollin, C., Baird, A. F., Isken, M. P., Clinton, J., Goertz-Allmann, B. P., Dahm, T., Wuestefeld, A., and ... Wiemer, S. Combined Large-N Seismic Arrays and DAS Fiber Optic Cables across the Hengill Geothermal Field, Iceland. *Seismological Society of America*, 93(5):2498–2514, 2022. doi: 10.1785/0220220073.

Parker, T., Shatalin, S., and Farhadiroushan, M. Distributed Acoustic Sensing–a new tool for seismic applications. *first break*, 32 (2), 2014. doi: 10.3997/1365-2397.2013034.

Pascucci, G., Gaviano, S., Pozzoli, A., and Grigoli, F. Signal Enhancement of Distributed Acoustic Sensing data using a Spectral Subtraction-based Approach. *Seismological Research Letters*, 2025. doi: 10.1785/0220250105.

Pecci, D., Cesca, S., Niemz, P., Pankow, K., Carelli, G., and Grigoli, F. Noise Analysis of Distributed Acoustic Sensing (DAS) Data in Borehole Installations. *Seismological Research Letters*, 96(3): 1703–1718, 2025.

Piana Agostinetti, N., Villa, A., and Saccorotti, G. Distributed acoustic sensing as a tool for subsurface mapping and seismic event monitoring: a proof of concept. *Solid Earth*, 13(2):449–468, 2022. doi: 10.5194/se-13-449-2022, 2022.

Piana Agostinetti, N., Malinverno, A., Bodin, T., Dahner, C., Dineva, S., and Kissling, E. Weighing Geophysical Data With Trans-Dimensional Algorithms: An Earthquake Location Case Study. *Geophysical Research Letters*, 50(22):e2023GL102983, 2023. doi: 10.1029/2023GL102983.

Porras, J., Pecci, D., Bocchini, G. M., Gaviano, S., De Solda, M., Tuinstra, K., Lanza, F., Tognarelli, A., Stucchi, E., and Grigoli, F. A semblance-based microseismic event detector for DAS data. *Geophysical Journal International*, 236(3):1716–1727, 2024. doi: 10.1093/gji/ggae016.

Riva, F., Agostinetti, N. P., Marzorati, S., and Horan, C. The microseismicity of Co. Donegal (Ireland): Defining baseline seismicity in a region of slow lithospheric deformation. *Terra Nova*, 36(1): 62–76, 2024. doi: 10.1111/ter.12691.

Rodriguez, T., Seguí, A., Ugalde, A., Armas, M. C., Latorre, H., Ventosa, S., and Monfret, T. Automatic Identification of High-Quality Channels in Distributed Acoustic Sensing Through Implementation of a Channel Quality Index. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 18, 19869–19883*, 2025. doi: 10.1109/JSTARS.2025.3593330.

Rost, S. and Thomas, C. Array seismology: Methods and applications. *Reviews of geophysics*, 40(3):2–1, 2002. doi: 10.1029/2000RG000100.

Sambridge, M. and Mosegaard, K. Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40(3):3–1, 2002. doi: 10.1029/2000RG000089.

Seguí, A., Ugalde, A., Fichtner, A., Ventosa, S., and Morros, J. R. DASPack: Controlled Data Compression for Distributed Acoustic Sensing. *Geophysical Journal International*, page ggaf397, 2025.

Strumia, C., Trabattoni, A., Supino, M., Baillet, M., Rivet, D., and Festa, G. Sensing optical fibers for earthquake source characterization using raw DAS records. *Journal of Geophysical Research: Solid Earth*, 129(1):e2023JB027860, 2024.

Tarantola, A. *Inverse problem theory and methods for model parameter estimation*. SIAM, 2005.

Tilmann, F., Atherton, C., Asero, C., Evangelidis, C., Charalampakis, M., Landrø, M., Rondenay, S., Ottemöller, L., Maji, V., Corela, C., and the SUBMERSE WP3 team, . Towards continuous fibre-optic monitoring in the oceans with submarine telecommunications cables–the SUBMERSE project. Technical report, Copernicus Meetings, 2024. doi: 10.5194/egusphere-gc12-fibreoptic-31.

Trabattoni, A., Biagioli, F., Strumia, C., van den Ende, M., Scotto di Uccio, F., Festa, G., Rivet, D., Sladen, A., Ampuero, J. P., Métaxian, J.-P., and Strutzamann, E. From strain to displacement: using deformation to enhance distributed acoustic sensing applications. *Geophysical Journal International*, 235(3):2372–2384, 2023. doi: 10.1093/gji/ggad365.

Ugalde, A., Latorre, H., Vidal, P., Martins, H. F., Martin-Lopez, S., and Gonzalez-Herraez, M. Canary Islands seismic monitoring with Distributed Acoustic Sensing (CANDAS) [Dataset], 2023. doi: 10.7914/73k1-1369.

van den Ende, M. and Ampuero, J.-P. Evaluating seismic beamforming capabilities of distributed acoustic sensing arrays. *Solid Earth*, 12(4):915–934, 2021. doi: 10.5194/se-12-915-2021, 2021.

Walter, F., Gräff, D., Lindner, F., Paitz, P., Köpfli, M., Chmiel, M., and Fichtner, A. Distributed acoustic sensing of microseismic sources and wave propagation in glaciated terrain. *Nature communications*, 11(1):1–10, 2020. doi: 10.1190/geo2019-0752.1.

Wuestefeld, A., Spica, Z. J., Aderhold, K., Huang, H.-H., Ma, K.-F., Lai, V. H., Miller, M., Urmantseva, L., Zapf, D., Bowden, D. C., and ... Wang, H. The Global DAS Month of February 2023. *Seismological Research Letters*, 95(3):1569–1577, 2024. doi: 10.1785/0220230180.

Xiao, H., Tilmann, F., van den Ende, M., Rivet, D., Loureiro, A., Tsuji, T., and Ugalde, A. DeepDAS: An Earthquake Phase Picker for Submarine Distributed Acoustic Sensing Data. In *EGU General Assembly Conference Abstracts*, pages EGU25–9380, 2025.

Yang, Y., Atterholt, J. W., Shen, Z., Muir, J. B., Williams, E. F., and Zhan, Z. Sub-kilometer correlation between near-surface structure and ground motion measured with distributed acoustic sensing. *Geophysical Research Letters*, 49(1):e2021GL096503, 2022. doi: 10.1029/2021GL096503.

Zhan, Z. Distributed acoustic sensing turns fiber-optic cables into sensitive seismic antennas. *Seismological Research Letters*, 91 (1):1–15, 2020. doi: 10.1785/0220190112.

Zhu, W., Biondi, E., Li, J., Yin, J., Ross, Z. E., and Zhan, Z. Seismic arrival-time picking on distributed acoustic sensing data using semi-supervised learning. *Nature Communication*, 14(8192), 2023. doi: 10.1038/s41467-023-43355-3.