Supplimentary information

This document contains the supplementary information supporting the article "Frequency-Resolved Seismic Source Localization in Fluvial Settings Using Dense Arrays". It provides a schematic overview of the matched field processing (MFP) workflow used for frequency-resolved source localisation (Text S1), the procedure used to estimate frequency-dependent phase velocities from manually selected coherent flood events (Text S2), and the synthetic framework used to quantify localisation uncertainty as a function of frequency, signal-to-noise ratio, and array geometry (Texts S3).

Text S1. MFP workflow

In figure S1, we provide a schematic overview of the matched field processing (MFP) workflow used throughout this study (Main text, Section 3.1), from raw seismic recordings to frequency-resolved source density maps. The MFP procedure consists of two main steps: first, we estimate the frequency-dependent phase velocity from selected coherent seismic events, and second, we apply these velocities to the continuous dataset for source localisation (see Text S2). In this second step, we transform the observed seismic data into the frequency domain to construct the cross-spectral density matrix, while generating synthetic replica fields for trial source locations and phase velocities. We compare these through the Bartlett processor to obtain the coherence response used to estimate an empirical dispersion relation. We then use the resulting phase velocities in the systematic MFP analysis of the continuous dataset, retain accepted source locations, and summarise them as source-density maps.

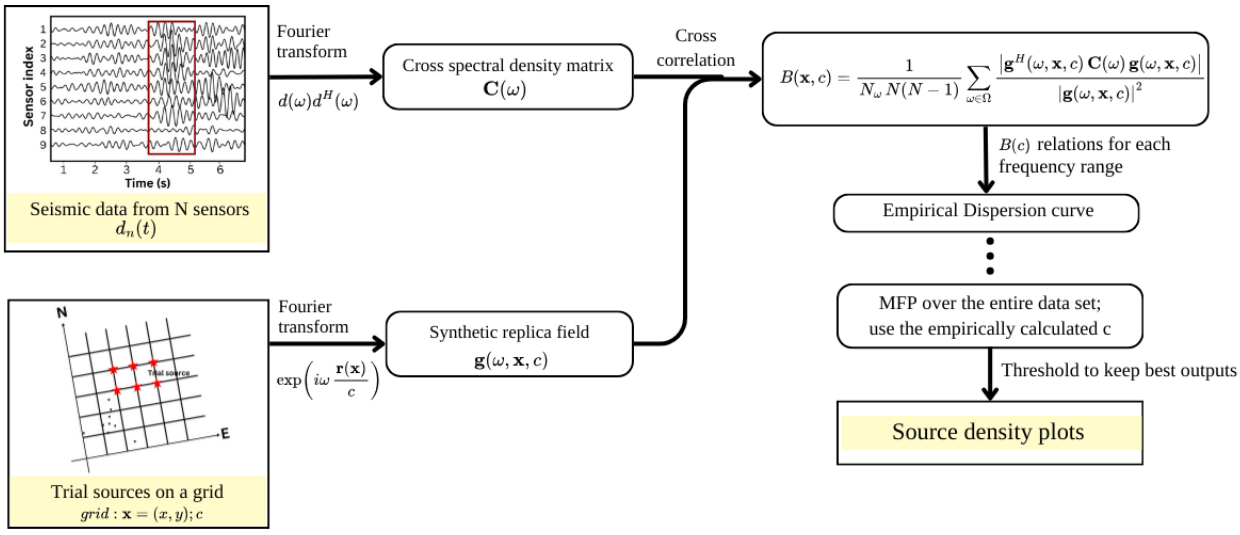


Figure S1. Workflow of the matched field processing (MFP) procedure used in this study.

Text S2. Phase velocity estimation from MFP

We estimate phase velocities using a subset of high-quality seismic events selected from the main flood period on 2 July 2019. More than 40 individual events are manually identified between 17:30 and 23:59 (local time), corresponding to the period of highest discharge and strongest seismic activity. We select events based on visual inspection of the waveforms and beamforming output (Main text, Figure 2b), retaining only signals that exhibit clear temporal coherence across most subarrays.

For each selected event, we apply MFP to short time windows centered on the most coherent portion of the signal. To ensure that no coherent phase information is missed, we analyze three consecutive 3 s windows: one centered

on the event peak, one preceding it by 3 s, and one following it by 3 s. This approach allows us to capture the full coherent wave packet while limiting contamination from unrelated background noise. The MFP is evaluated on a regular spatial grid covering the active channel reach. The grid has a spacing of 5 m and comprises 100×44 trial source locations. Figure S2.1 shows the grid overlaid on the study reach.

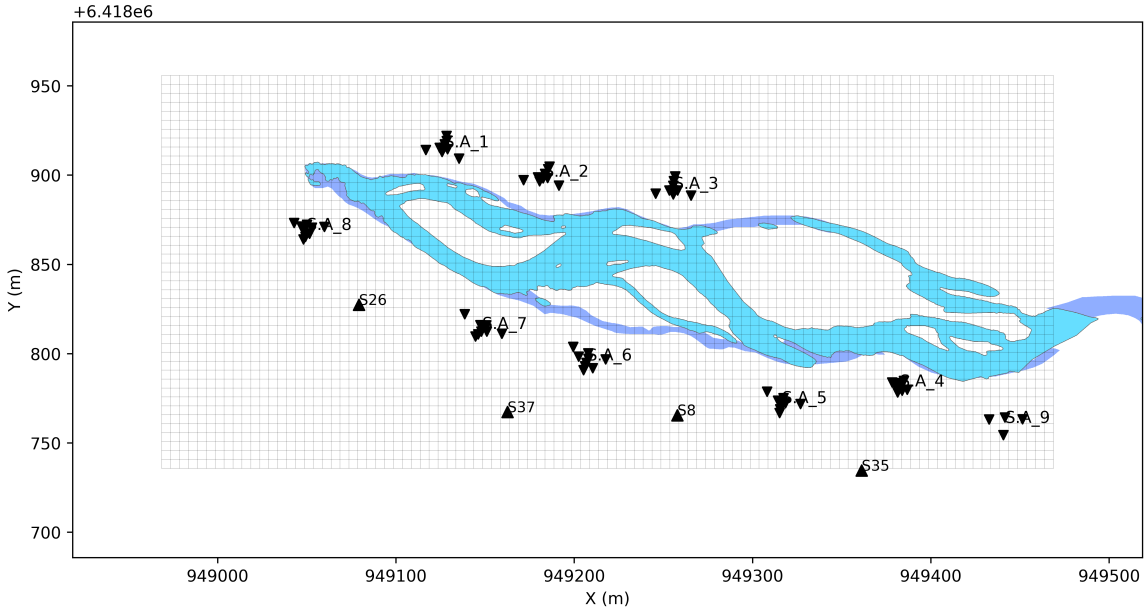


Figure S2.1: Spatial grid used for phase velocity estimation and source localization. The grid spacing is 5 m, resulting in 100×44 trial source locations covering the active channel reach.

To obtain frequency-dependent phase velocities, we use 2 Hz-wide frequency bands and repeat the MFP computation over trial velocities ranging from 10 m s^{-1} to 1000 m s^{-1}, in order to encompass plausible surface-wave velocities in unconsolidated sedimentary and mixed riverbed environments. The maximum Bartlett amplitude over the spatial grid is extracted as a function of trial velocity for every 3 s window, yielding one $B(v)$ curve per window. Figure S2.2 shows these curves for nine representative frequency bands: 4–6, 8–10, 12–14, 18–20, 22–24, 26–28, 30–32, 34–36, and 36–38 Hz. Thin gray lines represent individual windows, while the red curve shows the mean $B(v)$ across all windows and events. The velocity corresponding to the maximum of this mean curve (vertical dashed line) is retained as the representative phase velocity.

The selected events exhibit high beamforming coherence, with Bartlett amplitudes well above the thresholds defined in the main text (Figure 5b). This is expected because the thresholds reported in the main text are derived from the statistical distribution of Bartlett amplitudes over continuous data, whereas the events used here correspond to manually selected, highly coherent signals. Velocity resolution is somewhat poorer in a few frequency bands. Around 8–10 Hz, which lies close to the practical upper limit of the full-network configuration, the mean $B(v)$ curve is broad and the preferred phase velocity is less well constrained than at neighbouring frequencies. A similar decrease in resolution occurs toward the upper end of the investigated range, particularly above $\sim$34 Hz, where wavelengths approach the spatial-aliasing limit of the subarrays and the mean $B(v)$ response becomes flatter and may exhibit secondary local maxima. Despite this increased uncertainty, the selected velocities remain broadly consistent with the overall dispersion trend defined by adjacent frequency bands. We therefore retain the velocity corresponding to the maximum of the mean $B(v)$ curve in each band, while treating the 8–10 Hz and highest-frequency estimates as comparatively less well constrained.

To assess whether uncertainty in the selected phase velocity affects the inferred source distribution, we repeat the 8–10 Hz localization using phase velocities of 440, 540, and 640 m s^{-1}, corresponding to perturbations of ± 100 m s^{-1} around the selected value of 540 m s^{-1} (Figure S2.3). The absolute source density and detailed spatial distribution vary between the three cases, but the dominant source region remains within the same central part of the active channel, with a broadly comparable spatial extent and clustering. The first-order spatial organisation of the inferred sources is therefore not strongly sensitive to phase-velocity variations within the tested range.

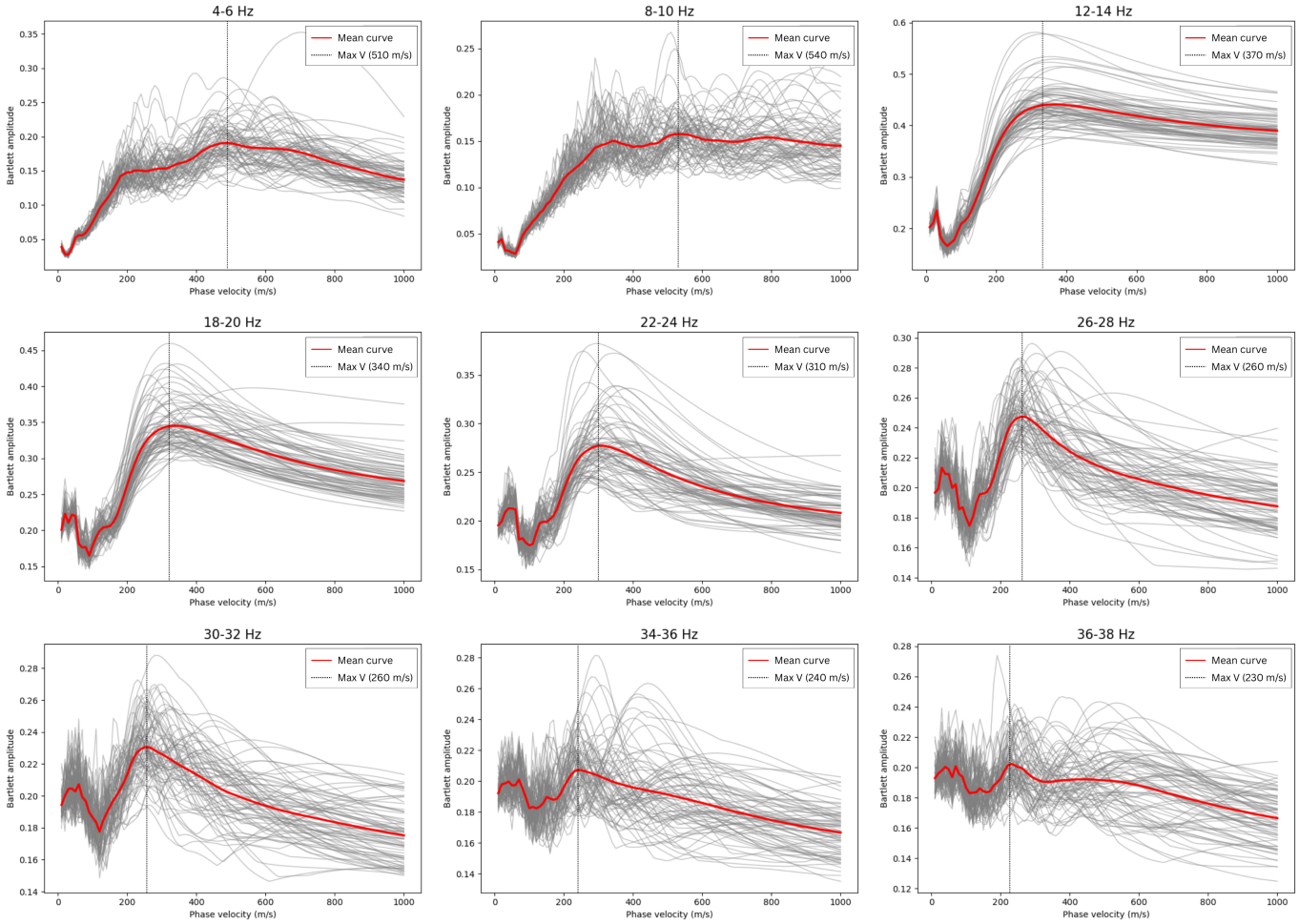


Figure S2.2: Bartlett amplitude as a function of trial phase velocity for nine frequency bands. Thin gray curves correspond to individual 3 s windows from manually selected coherent events during the main flood period (17:30–23:59 on 2 July 2019). The red curve shows the mean $B(v)$ across all windows and events. The dashed black vertical line marks the phase velocity at which the mean $B(v)$ reaches its maximum.

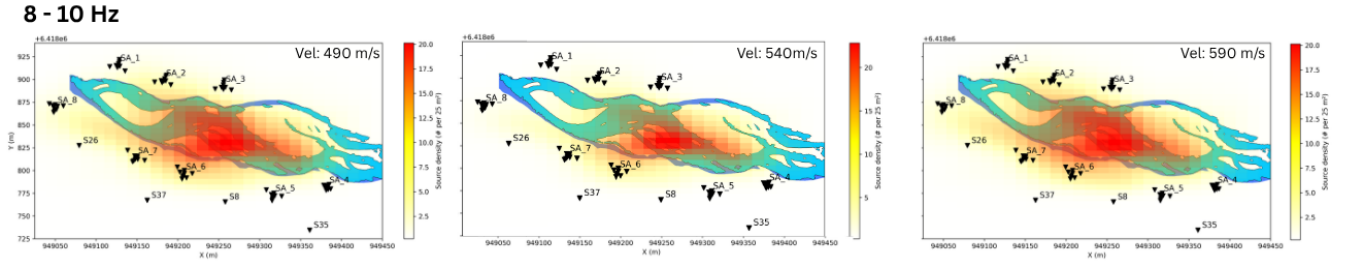


Figure S2.3: Sensitivity of the 8–10 Hz source-density distribution to the phase velocity used for localization. Source-density maps are obtained using phase velocities of (a) 440 m s^{-1}, (b) 540 m s^{-1}, and (c) 640 m s^{-1}. The central value corresponds to the phase velocity selected from the maximum of the mean $B(v)$ curve, while the lower and upper values represent perturbations of ± 100 m s^{-1}.

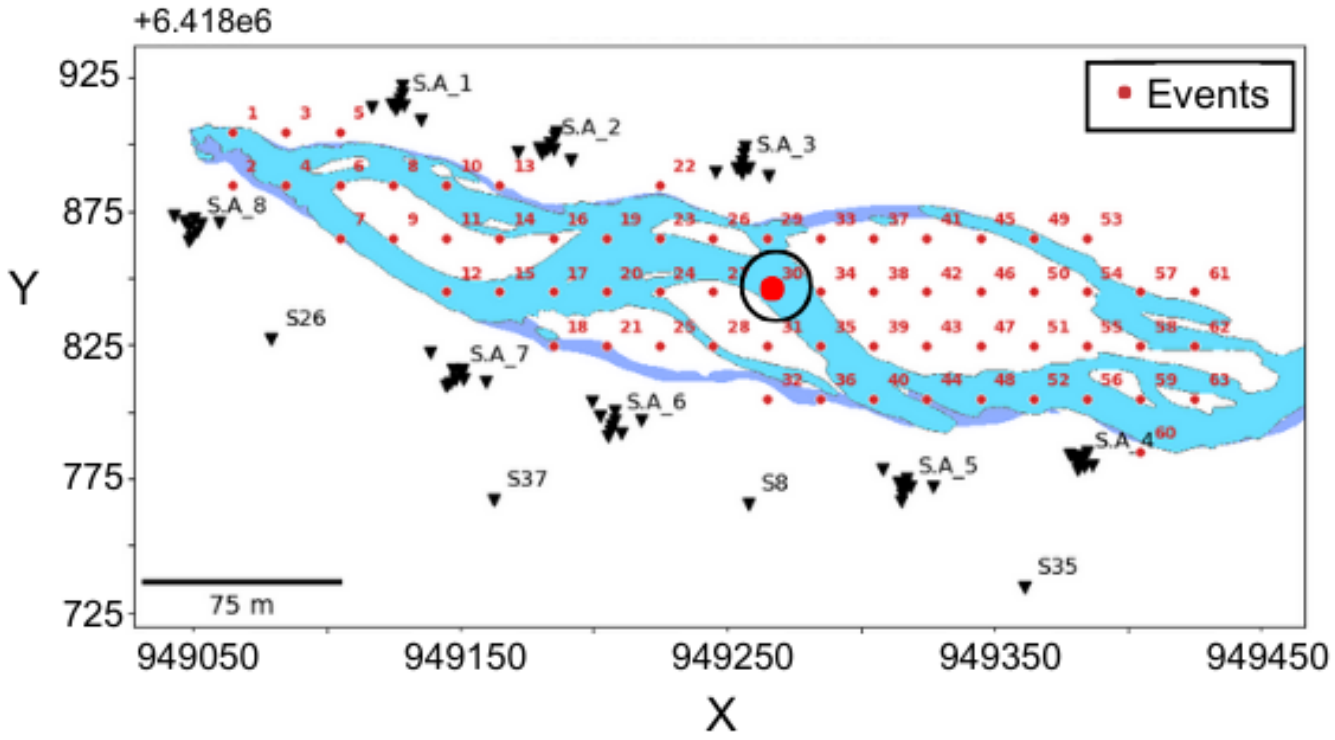


Figure S3.1. Synthetic source grid used for uncertainty estimation. Sources are restricted to the active river channel, where the most robust and coherent seismic sources are observed in the real data. The grid uses the same spatial spacing and coordinate system as the real-data MFP analysis.

Text S3. Uncertainty estimation

S3.1 Synthetic waveforms generation

We generate synthetic Ricker waveforms with a duration of 3 s, defined according to Equation (6) in the main manuscript, at discrete grid points spanning the active river channel of the Séveraisse River. The synthetic source grid is shown in Figure S3.1. Restricting synthetic events to these locations ensures that uncertainty estimates are evaluated only for physically relevant source positions and under array geometries that are representative of the observations. Central frequencies are chosen to match the frequency bands analyzed in the real-data MFP results.

Figure S3.1a shows an example of a noise-free synthetic Ricker wavelet with a central frequency of 12.5 Hz, as recorded by sensors across multiple subarrays. These waveforms illustrate the idealized phase-coherent signal expected in the absence of noise and serve as a reference for assessing the impact of noise on localization performance. To simulate realistic observational conditions, we add spatially uncorrelated Gaussian noise independently to each synthetic waveform. Figure S3.1b shows examples of synthetic waveforms with increasing levels of added noise, illustrating the progressive degradation of phase coherence across the array as signal-to-noise ratio (SNR) decreases.

S3.2 Calibration between Bartlett amplitude and SNR

We apply MFP to the synthetic sources over a range of SNR levels and extract the maximum Bartlett amplitude from each realisation. Repeating this procedure across the event grid and frequency bands provides location- and frequency-dependent calibration curves relating Bartlett amplitude to SNR.

Figure S3.3 shows example calibration curves for synthetic event 30 (Figure S3.1). The curves exhibit a characteristic S-shaped behaviour, with Bartlett amplitudes saturating at high SNR and decreasing rapidly at low SNR. Because this relationship varies with frequency and source position relative to the array geometry, the calibration is performed independently across the synthetic source grid. For example, in the 10–15 Hz band, the real-data threshold of 0.37 corresponds to an estimated SNR of approximately -12.2 dB at event 30. Observed Bartlett amplitudes are then mapped onto the corresponding synthetic calibration curves to estimate the SNR required to reproduce the beamforming output at each grid location (Main text, Figure 7a,b). These inferred SNR values are subsequently used to quantify localisation uncertainty from multiple synthetic noise realisations (Main text, Figure 7c,d).

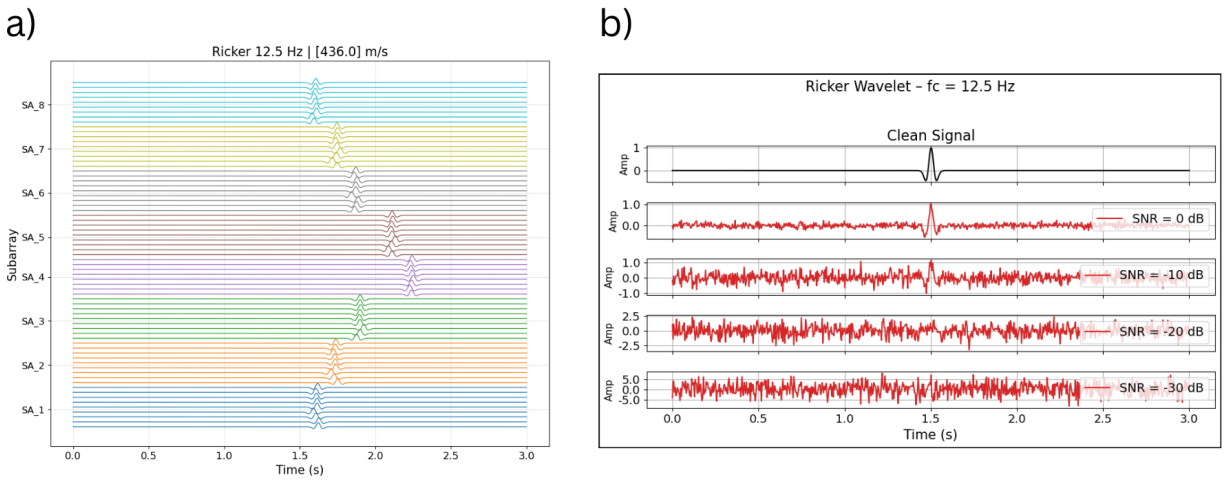


Figure S3.1. a)Noise-free synthetic Ricker wavelet with a central frequency of 12.5 Hz, as recorded by sensors across multiple subarrays. The waveform duration is 3 s and based on Equation (5) in the main manuscript. b)Synthetic Ricker waveforms with increasing levels of added spatially uncorrelated Gaussian noise, illustrating the progressive degradation of signal quality with decreasing signal-to-noise ratio (SNR).

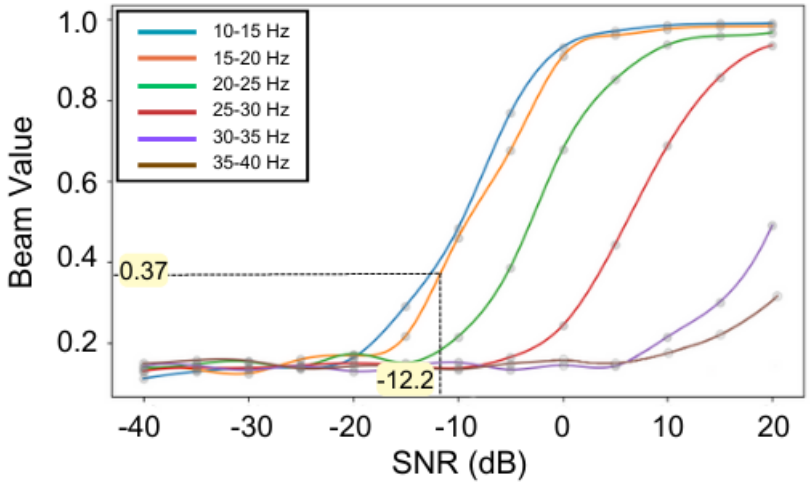


Figure S3.2. Calibration curves relating maximum Bartlett amplitude to signal-to-noise ratio (SNR) for synthetic event 30, shown for multiple frequency bands. For the 10–15 Hz band, a Bartlett amplitude of 0.37 (horizontal dashed line), corresponding to the threshold used in the real data, maps to an estimated SNR of approximately -12.2 dB (vertical dashed line).