Source Properties of the Induced M_L 0.0–1.8 Earthquakes from Local Beamforming and Backprojection in the Helsinki Area, Southern Finland

Bo Li[*1], Alice-Agnes Gabriel[2,3], and Gregor Hillers[4]

Abstract

Seismic arrays constrain local wave propagation that can be used to infer earthquake source characteristics. Array processing is routinely used to infer detailed earthquake properties of intermediate and large events. However, the source properties of microseismicity often remain elusive. In this study, we use high signal-to-noise ratio seismograms of 204 M_L 0.0–1.8 earthquakes induced by the 6 km deep 2018 Espoo/Helsinki geothermal stimulation to evaluate the performance and capabilities of beamforming and backprojection array methods. Using accurate travel-time-based event locations as a reference, we first show that miniarray beamforming is sensitive to medium heterogeneities and requires calibration to mitigate local systematic slowness biases. A catalog-based calibration significantly improves our multiarray beam raytracing estimates of source locations. Second, the application of the backprojection technique using *P*-wave signals with sufficient azimuthal coverage yields hypocenter estimates with generally good horizontal but poor vertical resolution. The short local source–receiver distances result in incomplete separation of *P*- and *S*-wave arrivals during backprojection. Numerical tests show that the relatively large *S*-wave amplitudes can influence coherent *P*-wave stacks, resulting in large location errors. Our combined *P*- and *S*-wave backprojection approach mitigates the influence of the large *S*-wave amplitude and improves the depth resolution significantly. The average depth offset to the reference catalog locations reduces from ≥ 1.4 km to ~91 m. Third, 3D numerical simulations demonstrate that backprojection swimming patterns are not merely processing or configuration artifacts. We show that the swimming patterns correlate with and can resolve the source focal mechanism when the azimuthal wavefield sampling is sufficiently complete. Our work demonstrates that the backprojection techniques can help to better constrain important properties of local-scale microseismicity.

Cite this article as Li, B., A.-A. Gabriel, and G. Hillers (2024). Source Properties of the Induced M_L 0.0–1.8 Earthquakes from Local Beamforming and Backprojection in the Helsinki Area, Southern Finland, *Seismol. Res. Lett.* **96**, 111–129, doi: 10.1785/0220240122.

Supplemental Material

Introduction

Geothermal energy and its enhanced geothermal systems (EGSs) can help address the global energy crisis that includes the need to reduce CO_2 emissions (Giovanni *et al.*, 2005; Cuenot *et al.*, 2015; Garcia *et al.*, 2016; Buijze *et al.*, 2019; Rathnaweera *et al.*, 2020). The fluid injection during an EGS stimulation induces or triggers earthquakes, which can, in the worst case, lead to the suspension or termination of a project (Giardini, 2009; Diehl *et al.*, 2017; Seithel *et al.*, 2019). Monitoring microseismicity is thus essential to support reservoir management and operation and to control the seismicity evolution for hazard mitigation. However, these seismic sources are usually difficult to characterize due to their small magnitudes and the associated low signal-to-noise ratio (SNR) seismograms. The heterogeneous and often ill-constrained 3D subsurface structure dissipates seismic information. Borehole sensors can yield better quality signals of small-magnitude events but are typically less convenient and economic to deploy compared to surface seismic arrays (Kwiatek *et al.*, 2019). Here, we utilize surface seismic array data that are typically employed to

1. Physical Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, https://orcid.org/0000-0002-3548-5496 (BL); 2. Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, U.S.A., https://orcid.org/0000-0003-0112-8412 (A-AG); 3. Department of Earth and Environmental Sciences, Ludwig-Maximilians-University Munich, Munich, Germany; 4. Institute of Seismology, University of Helsinki, Helsinki, Finland, https://orcid.org/0000-0003-2341-1892 (GH)

*Corresponding author: bo.li.3@kaust.edu.sa

improve the SNR of small event seismograms (Douglas, 2002) to detect, locate, and characterize induced seismic sources.

Array beamforming analyzes the differential travel times of an assumed planar wavefront across an array to infer local wave propagation properties (Birtill and Whiteway, 1965; Weichert, 1975; Krüger *et al.*, 1993; Rost and Thomas, 2002). Beamforming has been applied to detect and discriminate seismic events to study the behavior of scattered waves reflected off structural heterogeneities and discontinuities and to resolve the Earth structure (e.g., Wright, 1972; Hedlin *et al.*, 1991; Krüger *et al.*, 1993). More recent applications of small aperture array data demonstrate the possibility to derive wavefield gradients and rotational motion for the complementary subsurface characterization (Suryanto *et al.*, 2006; Langston, 2007; Taylor *et al.*, 2021) and an improved detection capability of low-amplitude signals associated with small earthquakes, tremors, and low-frequency earthquakes (Li and Ghosh, 2017; Verdon *et al.*, 2017; Meng and Ben-Zion, 2018). However, the sensitivity of the target high-frequency wavefield to structural heterogeneity tends to increase the uncertainty of associated event location estimates (e.g., Schweitzer, 2001; Schulte-Pelkum *et al.*, 2003).

Backprojection array processing (Kiser and Ishii, 2017) forms a signal beam to image the earthquake rupture process in sliding time windows (Ishii *et al.*, 2005; Meng, Ampuero, Sladen, and Rendon, 2012; Yin *et al.*, 2016; Li and Ghosh, 2016; Li *et al.*, 2022; Zhang *et al.*, 2023). These applications target the direct *P* waveforms, and large-aperture teleseismic arrays are ideally used to minimize the interference from later arriving body and surface waves. Teleseismic backprojection-based event detection (Kiser and Ishii, 2013; Li, 2019) can potentially be applied to earthquakes with magnitudes as small as *M* 3.5, and local configurations can lower this threshold significantly (Chambers *et al.*, 2010; Vlček *et al.*, 2016; Beskardes *et al.*, 2018). Microseismicity applications typically require signal enhancement (Inbal *et al.*, 2015) and systematic detection quality assessment (Yang *et al.*, 2021). However, such improvements are difficult to assess in the absence of well-located and well-characterized reference events.

During 49 days from 4 June 2018 to 22 July 2018, the St1 Deep Heat Company stimulated a geothermal reservoir in the Espoo/Helsinki region, southern Finland (Kwiatek *et al.*, 2019). Thousands of microearthquakes with a maximum magnitude M_L 1.8 were induced in a compact region at around 6 km depth in a structurally comparatively simple cratonic environment in response to the injection of ∼18,000 m^3 fresh water (Kwiatek *et al.*, 2019). The absence of a dissipating sedimentary layer results in high-SNR seismograms recorded by a diverse local network, including six surface seismic miniarrays and several surface and borehole stand-alone stations (Fig. 1; Ader *et al.*, 2020; Hillers *et al.*, 2020). Here, we work with the seismograms and catalog data of 204 earthquakes in the M_L 0.0–1.8 range (Taylor *et al.*, 2021) that were obtained using standard location procedures based on manually picked *P*- and *S*- wave arrival times (Hillers *et al.*, 2020). The reference catalog locations are characterized by a 20 m average uncertainty and errors of 11, 11, and 14 m in longitude, latitude, and depth, respectively (Hillers *et al.*, 2020).

We use the high-quality seismograms of these small-magnitude events to test the performance and resolution power of the beamforming and backprojection array methods for locating and characterizing small earthquakes induced by this geothermal stimulation. The first focus of this study is to evaluate the event location performance of these two staple array-processing techniques, analyze their limitations, and explore strategies for improvement using the reference catalog observations. Using data in the 2–35 Hz frequency band, we show that the local slowness bias is significant despite the local configuration and the simple subsurface structure. We mitigate this using a calibration approach that is based on the available reference locations. The short source–receiver distance and high-wave speeds affect the backprojection results because of the short *S*-minus-*P* times. This prompts us to apply a combined *P*- and *S*-wave stack, which significantly improves the vertical resolution of the associated hypocenter estimates.

The second key contribution of this study is the resolution of small earthquake source properties from backprojection swimming patterns, which are well-known artifacts characterized by a systematic time-dependent drift of the stacked energy and the resulting backprojected location (Meng, Ampuero, Luo, *et al.*, 2012). Supported by 3D numerical experiments, we demonstrate a link between earthquake focal mechanisms and swimming patterns that have long been assumed to be artifacts. Our results suggest the new research avenues using local array observations for improved characterization of the properties and underlying physics of microseismicity.

Beamforming

Data and methods

Beamforming assumes a plane wavefront propagating across an array. It uses the slowness vector composed of the horizontal slowness and back azimuth to quantify the differential travel times at each array station relative to the array center (Rost and Thomas, 2002). We use the three-component data from three square-patch of miniarrays deployed at the sites Seurasaari (SS), Elfvik (EV), and Toppelund (TL) located to the east, north–west, and south–west of the injection site, respectively, to perform beamforming (Fig. 1a). The arrays consist of nominally 25 sensors, and the aperture and interstation distances are 150 and 25 m, respectively.

We filter the seismograms between 2 and 35 Hz and perform beamforming independently for each miniarray and each component. We perform a grid-search beamforming in the slowness domain and calculate the beam power at each slowness associated with the *P* or *S* wave (Appendix A). The grid element with the maximum beam power indicates the local wave propagation direction relative to the array center. We

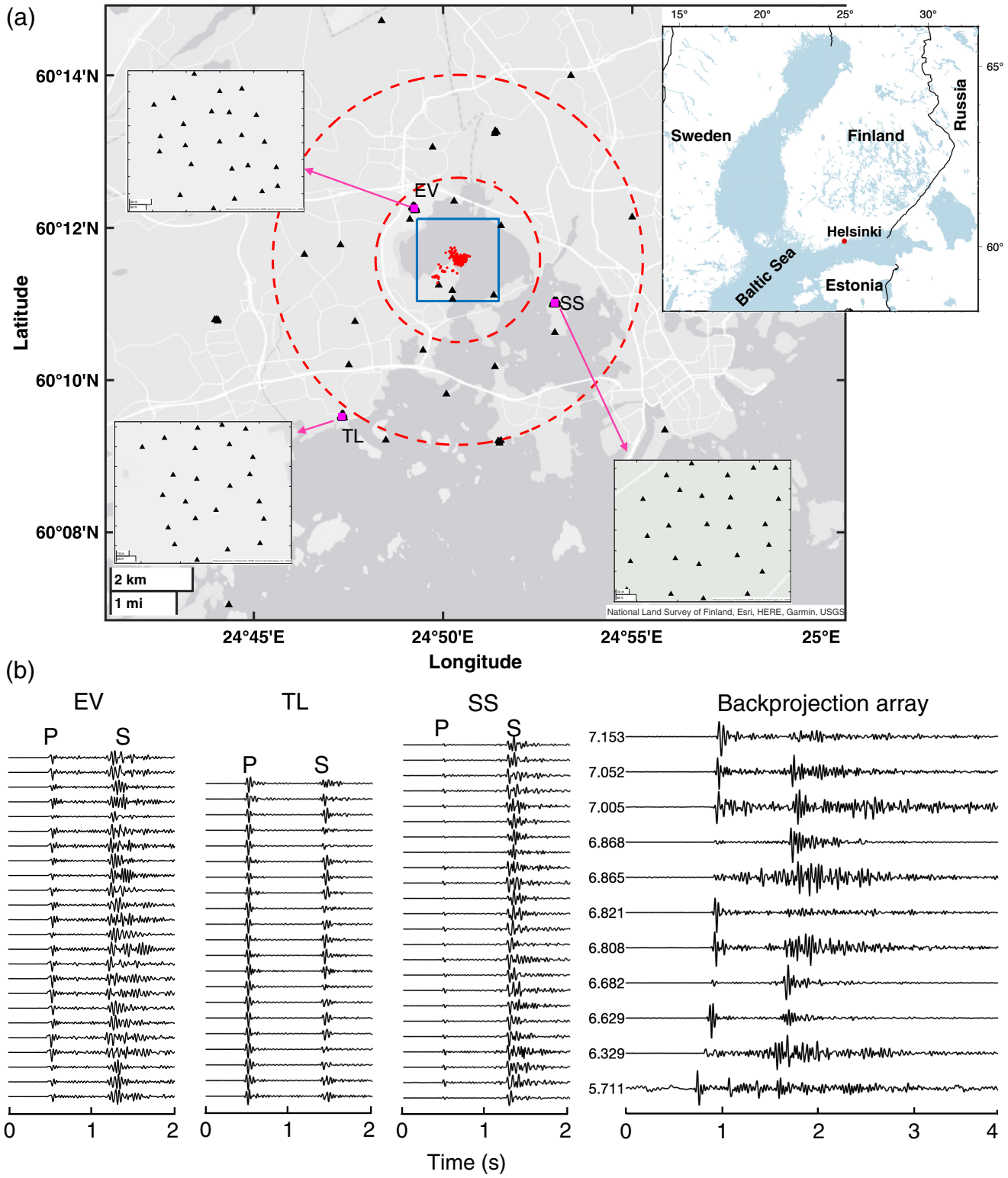
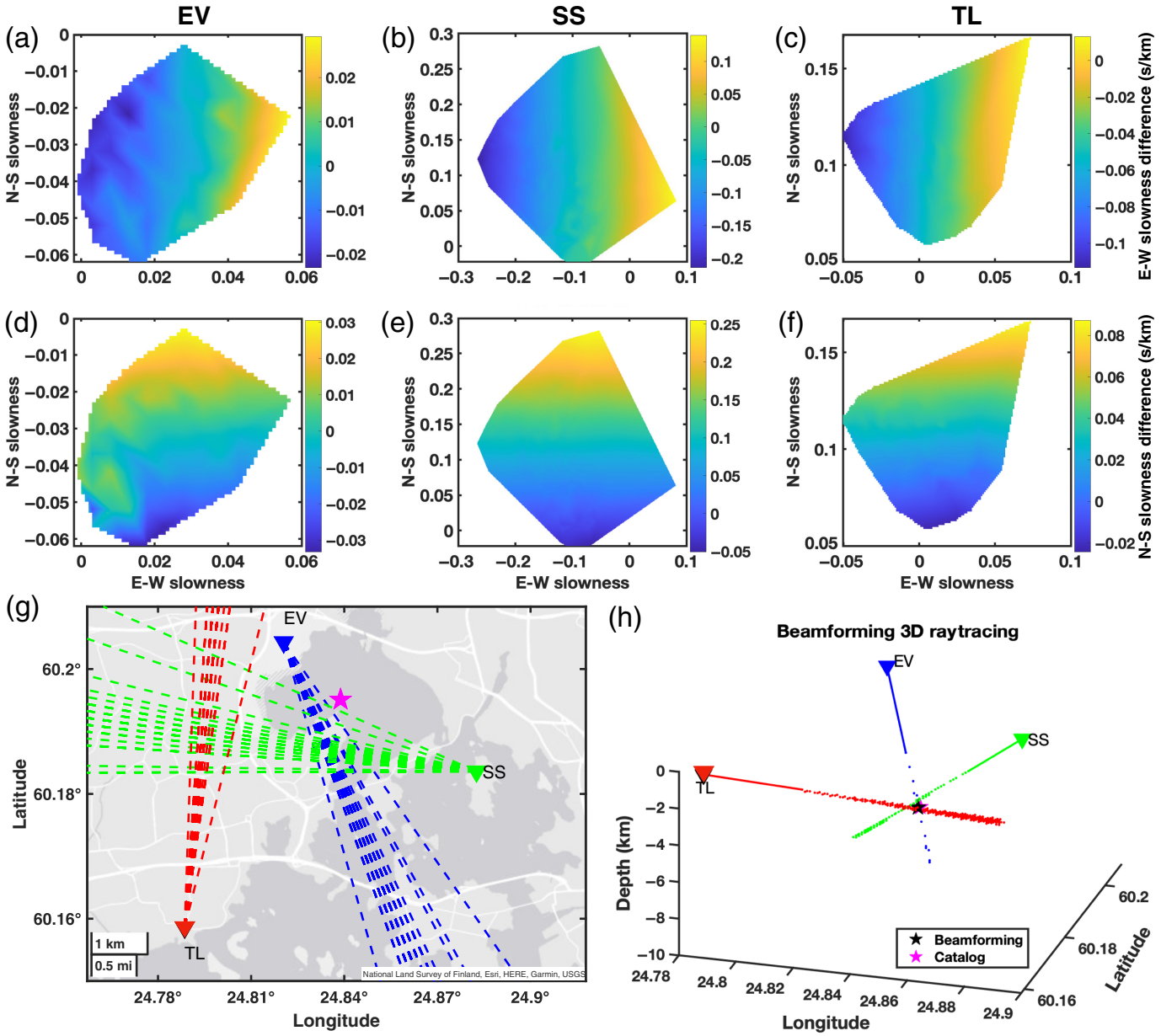


Figure 1. (a) Map of the Helsinki study region with the station distribution. The inset map in the upper right shows the regional setting. The red dots in the center of the main map represent the catalog earthquake locations. Black triangles indicate individual seismic stations. The pink squares indicate the three miniarrays, EV, SS, and TL, that are used for beamforming. The three 100 m scale insets show the configuration of these miniarrays. The two dashed circles indicate the 2–4.5 km distance range relative to the event cluster. Our backprojection array consists of stations between these two dashed circles. The blue box frames the backprojection target region. (b) 2–35 Hz filtered high signal-to-noise ratio vertical-component seismograms for the 21 July 2018 M_L 1.0 event, recorded at the three beamforming miniarrays and the backprojection array. The backprojection stations are organized by hypocentral distance in kilometers indicated to the left of each waveform. EV, Elfvik; SS, Seurasaari; TL, Toppelund. The color version of this figure is available only in the electronic edition.



use the bootstrap method (Efron, 1992) to quantify the slowness uncertainty at each array (Appendix A). We estimate the event epicenter or hypocenter by performing 2D or 3D raytracing of the slowness vectors from each array center, utilizing a 1D homogeneous velocity model with $V_P = 6.2$ and $V_S = 3.62$ km/s that is used for regional earthquake analysis (Kortström *et al.*, 2018).

Beamforming results and calibration

For each miniarray, we observe a systematic slowness bias in the beamforming results (Fig. 2a–f). These array specific patterns are controlled by unaccounted for structural variations that influence the assumed plane-wave propagation and that lead to erroneous but systematic back-azimuth variations away from the source direction. As a result, the back-azimuth rays from each array center do not converge and fail to properly constrain source locations (Fig. 2g). This result agrees with previous propagation direction estimates based on the same array data (Hillers *et al.*, 2020; Taylor *et al.*, 2021) and is likely due to

Figure 2. Vertical-component *P*-wave beamforming results. (a–f) The slowness calibration functions are the average slowness difference between the beamforming results and the catalog location-based slowness estimates. (a,d) Calibration functions in the east–west (E-W) and north–south (N-S) direction for the EV miniarray. Panels (b,e), (c,f) are the corresponding patterns for the SS and TL arrays. (g) Beamforming-based back-azimuth raytracing before calibration. The pink star marks the location of the target M_L 1.0 event. The dashed lines show the bootstrap raytracing results. (h) Beamforming-based 3D raytracing for the same event using the calibrated *P*-wave slowness. The black and pink stars mark the beamforming and catalog location. The color version of this figure is available only in the electronic edition.

medium heterogeneity that is not accounted for by the velocity model (Shearer *et al.*, 2023), including the effects of spatially variable near-surface properties of weathered bedrock (Hillers *et al.*, 2020; Eulenfeld *et al.*, 2023).

We use the reference catalog locations to correct for the slowness bias. For this, we randomly choose 200 of the 204 event locations covering the compact earthquake source region to establish a robust calibration function. For each miniarray, we perform beamforming to obtain a slowness estimate for each selected event. These estimates are averaged across bins of 0.005 s/km width along the east–west and north–south directions. The target slowness calibration function is the average difference relative to the theoretical slowness based on the catalog locations and the homogeneous velocity model. Figure 2a–f shows the resulting calibration function of the *P*-wave slowness bias pattern for the three miniarrays. Color indicates the average difference compared to the catalog-based slowness prediction. The similar apertures 144, 119, and 135 m for the SS, TL, and EV array, respectively, are unlikely to control the different patterns and varying amplitudes of the calibration functions. The EV results have the smallest bias with the maximum 0.03 s/km difference in both east–west and north–south directions. The SS array shows the highest deviation up to 0.25 s/km. The slowness bias of the SS array is less than 0.07 s/km for the beamforming results using the horizontal components (Figs. S1, S2, available in the supplemental material to this article). The average back azimuth of EV and SS mostly points to the same source quadrant, whereas some TL results show more significant variations for the estimated source–receiver directions.

Again, for demonstration purposes, we can use these calibration functions to correct the bias in the observed beamforming slowness of the remaining four events. Figure 2g shows the map view of the bootstrap solutions of the slowness back-azimuth raytracing before calibration, for example, 21 July 2018 M_L 1.0 event. Clearly, the solutions do not converge on the catalog location (pink star). After calibration, the back-azimuth raytracing from three array centers converges much better around the reference solution, even in 3D (Fig. 2h). For this example event, the multiarray *P*-wave beamforming locates the source 87 and 110 m off the catalog location in the horizontal and depth direction, respectively. Corresponding estimates based on horizontal component *S*-wave data yield a 233 m horizontal and 220 m vertical offset compared to the catalog location using the east–west component (Fig. S1) and a 200 m horizontal and 30 m vertical offset using the north–south component (Fig. S2). Calibration results of the other three events are collected in Figure S3.

For these examples, we use the very large fraction of the database—200 out of 204 events—to maximize the data coverage for robust calibration functions. To study the stability of the calibration function constructed from a smaller database, we randomly deselect data from bins that contain many events. However, we maintain the number of bins and hence the spatial coverage of the reservoir. The results for the EV array using *P* waves suggest that 50% of the data are sufficient to estimate a stable calibration function (Fig. S4). This indicates that the deselected data have a very similar slowness offset compared to the average estimated from the almost complete dataset. Using 50% of the data for the *P*-wave beamforming calibration function results in an average horizontal offset of the other half of the dataset of 139 m and a vertical offset of 151 m (Fig. S5). These values are insensitive to the choice of the deselected data. The calibration functions diverge using yet smaller data subsets, for example, using 33% of the data and one sample per bin (Fig. S4).

Backprojection

The backprojection method shares the signal time shifting and stacking concept with beamforming. In contrast to beamforming, which discretizes the slowness space to estimate local plane wave propagation, backprojection discretizes the space of the location of energy radiation to calculate a corresponding suite of travel times of the curved wavefront, here for our 1D homogeneous velocity model, across an array or network of sensors, similar to matched field processing (Corciulo *et al.*, 2012; Gradon *et al.*, 2019). Together with the observed coherent wavefront data, it is then possible to successively estimate the time and location of the energy excitation during a large earthquake (Ishii *et al.*, 2005; Krüger and Ohrnberger, 2005). For our hypocenter search and source focal mechanism study, we analyze data and synthetic waveforms.

Data analysis

In addition to the miniarrays, the network consists of 12 shallow borehole stations (Kwiatek *et al.*, 2019) with 500 Hz sampling rate and five broadband sensors with 250 Hz sampling rate (Hillers *et al.*, 2020). These are located within a 7 km radius around the stimulation wellhead. For the backprojection approach, we refer to the "local array" configuration, which includes the single borehole and surface stations and one sensor from each miniarray (Fig. 1a). The horizontal orientation of the borehole sensors is not calibrated, and only vertical-component data are used for backprojection. The target source volume for our data analysis has a lateral extension of 2.8 km × 2.2 km (Fig. 1a) with a 55 m discretization and a vertical extension that ranges from 3 to 8 km depth with a 50 m interval. For our hypocenter location search, we use a moving time window of 0.3 s and a timestep equal to the sampling interval (Appendix B). Time windows that include the target body-wave energy has significantly enhanced backprojection beam power compared to the background level. The maximum beam energy associated with a grid search cell then indicates the source location for each time window.

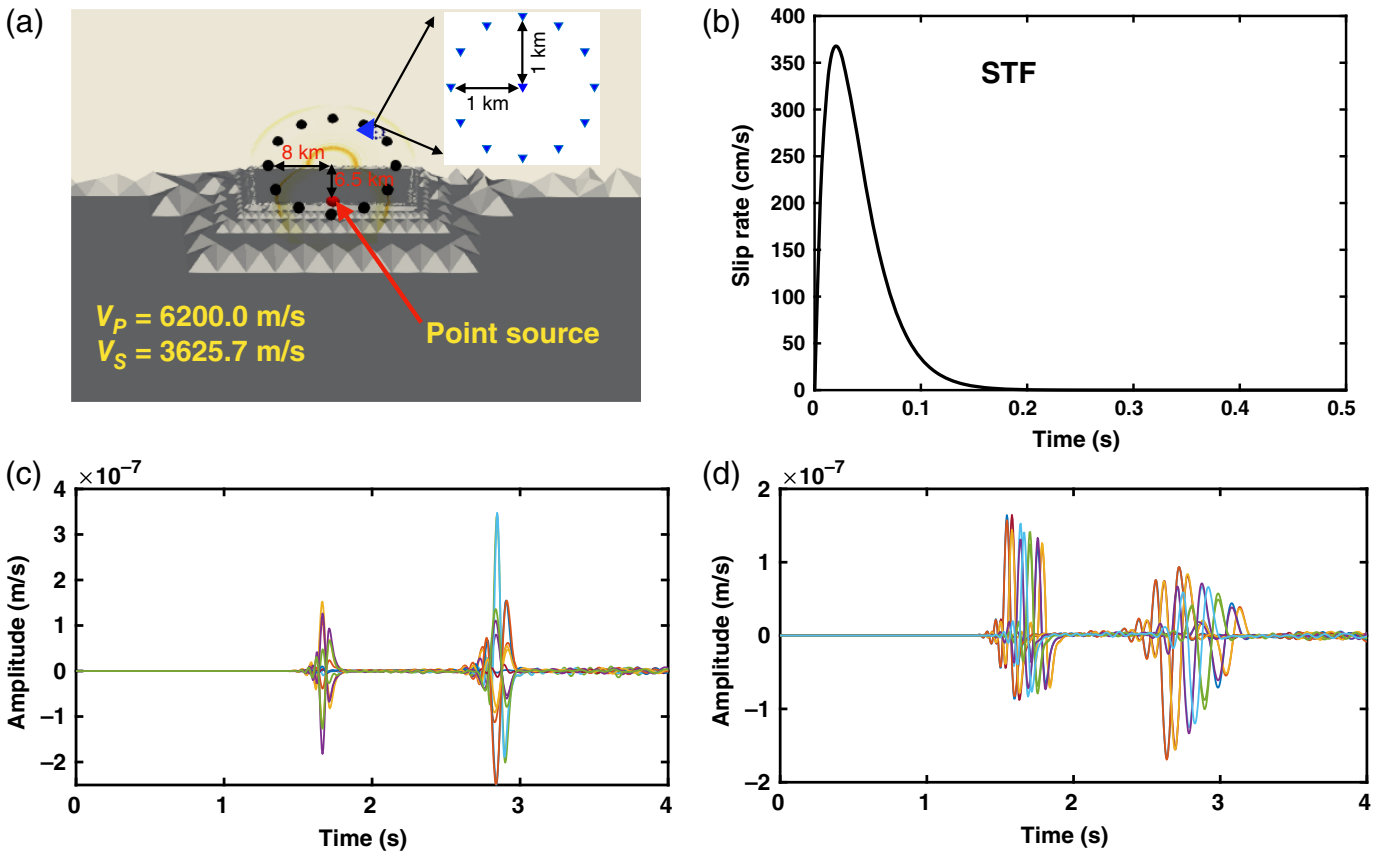


Figure 3. Model setup for the synthetic point-source experiments. (a) Point-source simulation configuration, with a view of the clipped model domain around the source and arrays. The red circle in the center represents the point source. Black circles indicate the full azimuth coverage array (FACA) at an 8 km epicentral distance. The array has an azimuth range from 30° to 360°, with a 30° interval. The circular array configuration of FACA is distorted in the figure due to the view angle. The triangles denote the limited azimuth coverage array (LACA) at 8 km epicentral distance and 45° azimuth relative to the point source. The inset shows the array configuration with an aperture of 2 km. The grayscale indicates static refinement in computational mesh element size. (b) The prescribed source time function (STF) of the point source. (c) Example FACA waveforms for a point source with strike 0°, dip 90°, and rake 45°. Colors distinguish waveforms at different stations. (d) As in panel (c) but for stations of the LACA. The color version of this figure is available only in the electronic edition.

Numerical tests

We perform numerical experiments (Appendix C) using SeisSol (Breuer *et al.*, 2014; Uphoff *et al.*, 2017; Krenz *et al.*, 2021) to investigate the governing mechanism and solution of the obtained large location offsets (see the Backprojection location estimates section) and to analyze the potential link between the observed swimming patterns, focal mechanisms, and variable source-array configurations (see The backprojection "swimming artifact" section). A moment tensor point source is located at 6.5 km depth and at the center of the domain with size $150 \times 150 \times 100$ km^3 (Fig. 3a). A refined mesh is used for the central region to resolve frequencies up to 4.5 Hz. We set up two synthetic arrays. The model array referred to as full azimuth coverage array (FACA) consists of 12 stations that are distributed at constant 8 km epicentral distance and with 30° equidistant spacing. The array referred to as limited azimuth coverage array (LACA) consists of 13 stations that are evenly distributed in a 1 km radius circle that is centered 8 km away from the epicenter in a 45° direction. We use the same homogeneous velocity model that we use for the beamforming and backprojection data analysis. We apply backprojection to waveforms associated with 3024 point sources with systematically varying focal mechanisms. For all cases, we use the same asymmetric source time function with the duration of 0.15 s (Fig. 3b). Figure 3c,d shows the synthetic waveforms at the FACA and LACA for a point source with strike 0°, dip 90°, and rake 45°.

Backprojection location estimates

***P*-wave data analysis results.** To enhance the spatial resolution of our *P*-wave backprojection location estimates, we use catalog locations and apply cross correlation to correct the waveform polarity before stacking (Appendix B). This polarity correction approach results in a higher resolution compared to various alternative methods that have been suggested to mitigate the polarity ambiguity (e.g., Inbal *et al.*, 2015; Vlček *et al.*, 2016; Beskardes *et al.*, 2018), including the waveform envelope

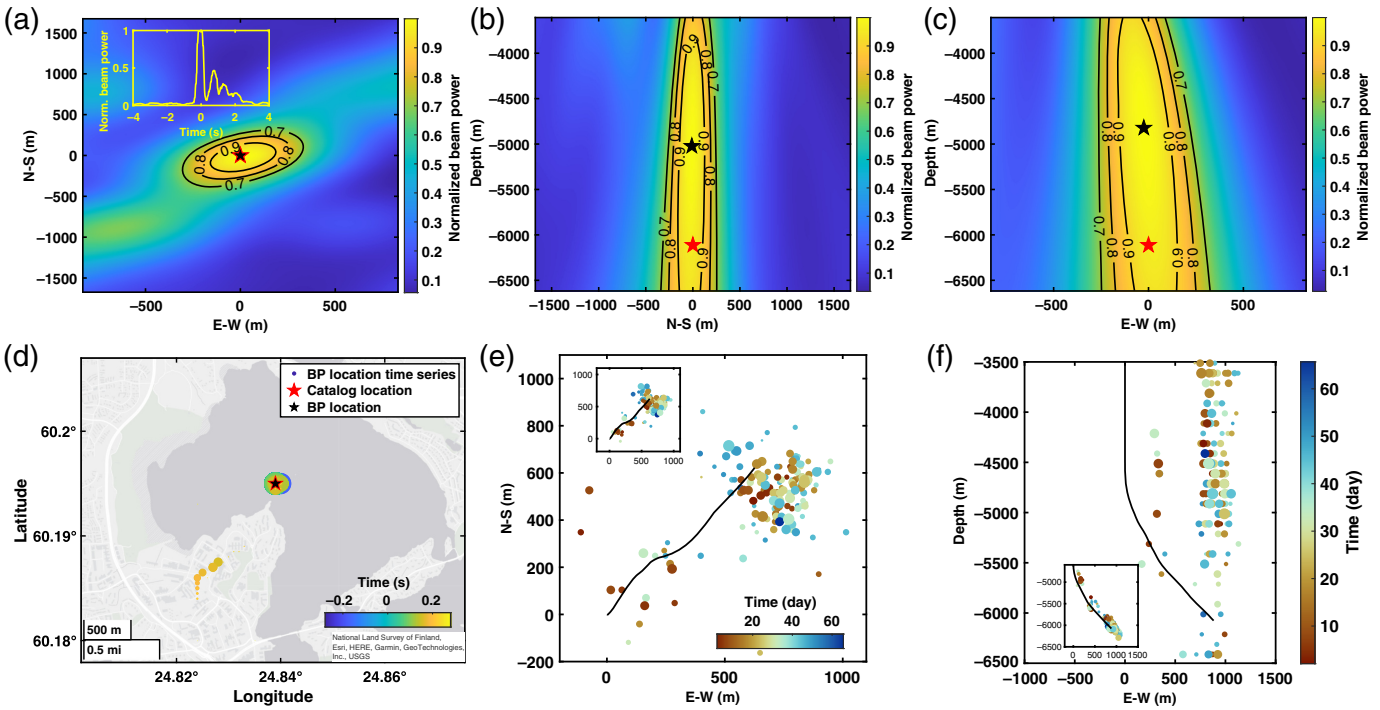


Figure 4. *P*-wave backprojection example using the same M_L 1.0 event as in Figure 2. (a) Horizontal plane backprojection image of the relative energy radiation at the event catalog depth and at the time of peak beam power. The black and red stars mark the backprojection based on the peak beam power and the catalog location. The black contours indicate 70%, 80%, and 90% of the peak beam power. The inset shows the beam power evolution with time. Panels (b) and (c) show the backprojection depth resolution at the peak beam power timestep. Backprojection image of the relative energy radiation in the (b) north–south–vertical plane at the event longitude and in the (c) east–west–vertical plane at the event latitude. (d) *P*-wave backprojection results for this M_L 1.0 event. (e) *P*-wave backprojection epicenters of the 204 catalog events. The black line denotes the borehole trajectory. The inset shows the catalog locations. (f) *P*-wave backprojection hypocenter location estimates. The inset shows the catalog locations. The color version of this figure is available only in the electronic edition.

that we use in the following sections. By enhancing the resolution, we can show that the observed intermittent significant location offsets are not controlled by the choice of a low-resolution waveform function. Figure 4a–c shows the horizontal and vertical resolution of the *P*-wave backprojection of the example M_L 1.0 event using the local array consisting of single sensors. It demonstrates good horizontal resolution, and the backprojection epicenter indicated by the peak beam power (Fig. 4a,d) is in good agreement with the catalog location. The poor vertical resolution is indicated by the offsets on the order of 1 km toward the surface compared to the reference (Fig. 4b,c). The average horizontal offset between backprojection and catalog locations is 97.5 m (Fig. 4e). This value is influenced by a few outliers, for which the backprojection locations differ significantly from the catalog estimates. In contrast, the poor depth distribution is typical for backprojection solutions (Fig. 4f) and is comparable to other travel-time-based location methods. The backprojection tends to locate the sources shallower than the reference, with a significant average offset of ~1.4 km. When using the waveform envelope, the average depth offset increases to ~3 km (Appendix B).

The effect of *S* waves. Our analysis of the synthetic waveforms using the FACA shows that the *S* waves can strongly affect the *P*-wave backprojection results and thus contribute to the obtained large location errors. Comparing the results using the FACA for two point sources with the same strike and dip but two different rake angles of 45° and 75° (Fig. 5) shows that the 45° rake model generates much larger *S*-wave amplitudes at most model stations (Fig. 5a). Importantly, in comparison to coherent *P*-wave-only stacks at the proper source location, this can lead to inaccurate location estimates because of the stronger beams associated with biasing large-amplitude *S* waves (Fig. 5b). In contrast, the rake 75° model yields *P*- and *S*-wave time window stacks with comparable beam power (Fig. 5c). In this case, backprojection images sources of radiation that seem to migrate in the northeast direction, but with peak beam power located at the actual synthetic source location (Fig. 5d). We conclude that these *S*-wave effects can at least partially explain some of the *P*-wave-based locations in Figure 4e that have a low consistency with the reference locations.

Backprojection with *P* and *S* waves. The synthetic tests show that not fully separated *P*- and *S*-wave arrival signals can

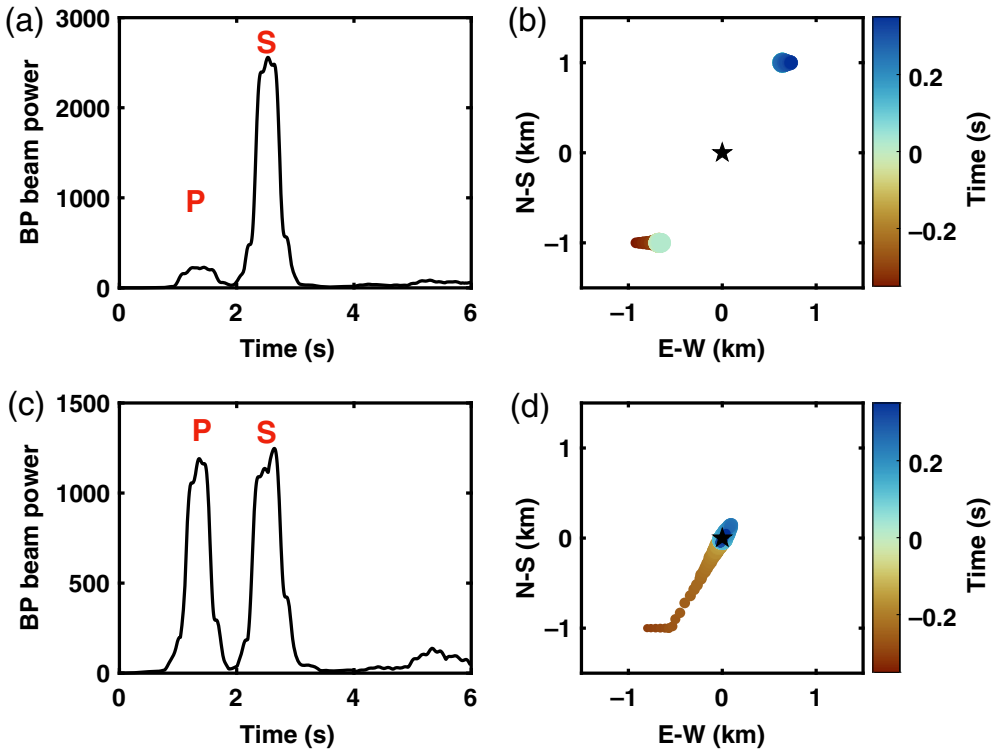


Figure 5. Synthetic backprojection results with FACA demonstrate the *S*-wave effect on the location estimates. (a,b) Backprojection results for a point source with strike 30°, dip 90°, and rake 45°. (c,d) Backprojection results for a point source with strike 30°, dip 90°, and rake 75°. Panels (a) and (c) show beam power time series. The *P*- and *S*-wave arrivals are indicated. Panels (b) and (d) show the evolution of the backprojected source location estimate with time. The black star marks the synthetic source location. The backprojection grid search is performed in a 2 × 2 km^2 area. The color version of this figure is available only in the electronic edition.

yield a large location error, especially when the *S*-wave amplitudes are larger at most stations. To avoid location bias due to the peak stacked energy associated with a mixed *P*- and *S*-wave stack, we combine the *P* and *S* waves and perform backprojection by, first, stacking the theoretical *P*- and *S*-waves time windows separately and, second, using the summed stack to represent the beam energy at each source grid. The combined *P*- and *S*-wave stack can be expressed as

$$s^2_{j-psbp}(t,t+t_w) = \int_t^{t+t_w}\left(\sum_{i=1}^{N} w_i u_i(t+\tau_p^{ij})\right)^2 \times \mathrm{d}t^2, + \int_t^{t+t_w}\left(\sum_{i=1}^{N} w_i u_i(t+\tau_s^{ij})\right)^2 \times \mathrm{d}t^2, \quad (1)$$

in which $s^2_{j-psbp}(t,t+t_w)$ is the obtained power for the source grid j in the time window $(t,t+t_w)$, with $t_w = 0.3$ s the backprojection stack time window length. The signal envelope of the station i is u_i, N is the number of stations, w_i is the normalization factor, defined as the maximum value in u_i, τ_p^{ij} and τ_s^{ij} are, respectively, the *P*- and *S*-wave travel-time estimates from the source j to a receiver i, and dt is the sampling rate.

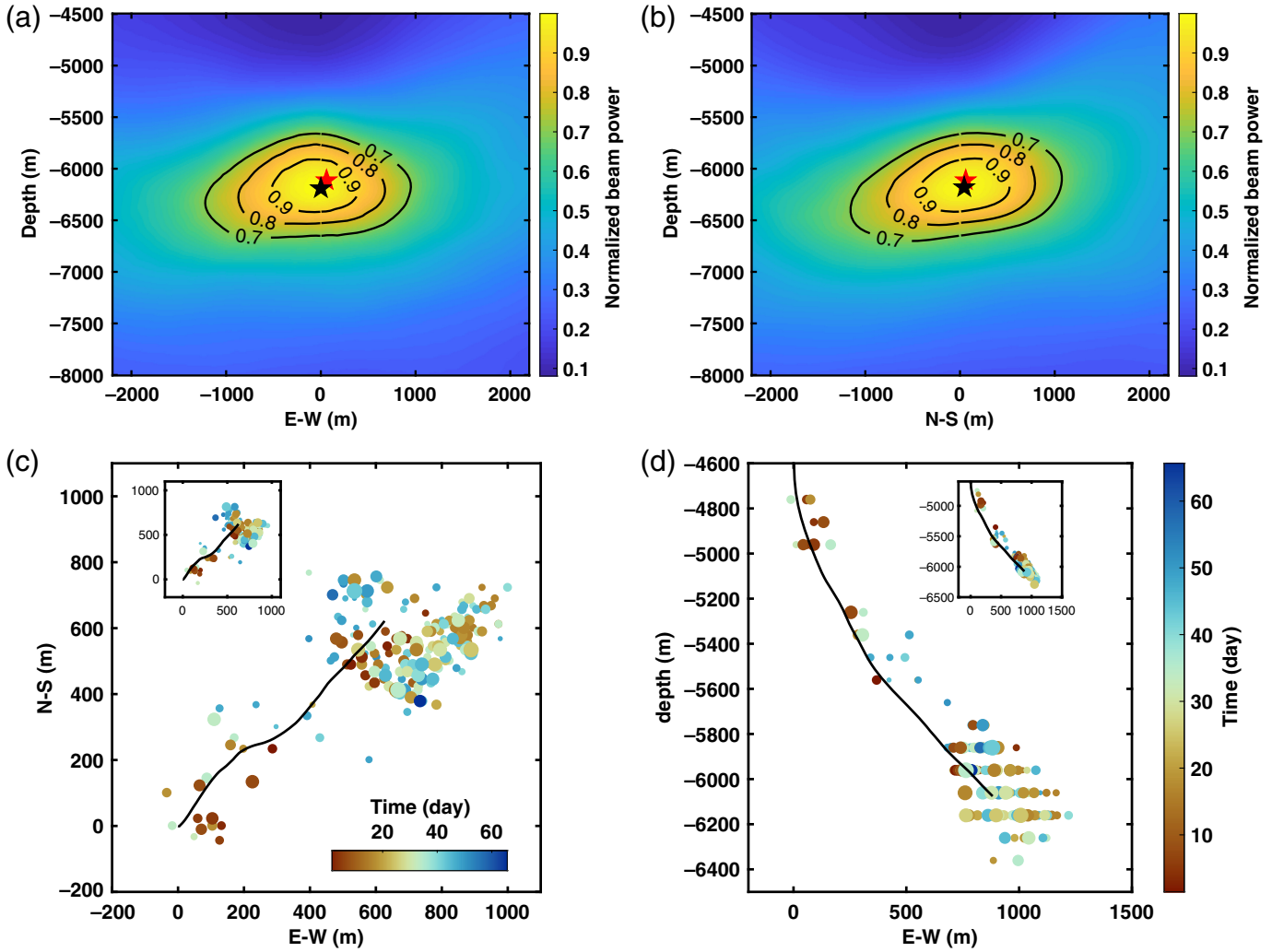
As before, the timestep and location associated with the peak beam power through all time windows represent the event time and location.

Both *P* and *S* waves stack coherently only at the source location. This results in the largest summed *P*- and *S*-wave beam energy across all grid points, independent of the relative *P*- and *S*-wave amplitudes. In addition, the depth-dependent travel-time difference between the *P*- and *S*-wave arrivals can also improve the depth resolution. To test this, we apply the combined *P*- and *S*-waves backprojection to the 3024 synthetic focal mechanisms associated with one event location, using a grid-search interval of 50 m in all three directions. The average (maximum) offset between the backprojection locations and true source locations are 0.6 m (50 m), 25 m (50 m), and 58 m (150 m) for the east–west, north–south, and vertical directions, respectively (Fig. S6). This indicates that the proposed procedure effectively locates the hypocenter across the full range of tested source mechanisms.

Figure 6 shows the results of the combined *P*- and *S*-wave backprojection approach applied to the records of the 204 catalog events. Compared to the spatially variable beamforming calibration functions, we observe a laterally homogeneous average error or offset between the backprojection estimates and the catalog locations (Fig. S7). Hence, we can use the mean offset between these solutions to correct for the systematic average error associated with the velocity model uncertainty and the incomplete array configuration. In comparison with the *P*-wave-only estimates (Fig. 4), the resulting calibrated locations show an improved consistency with the travel-time-based catalog locations (Fig. 6). The calibrated *P*- and *S*-wave backprojection locations have an average difference of ~74.9 m in horizontal and ~91.3 m in vertical direction to the reference estimates.

The backprojection "swimming artifact"

Our *P*-wave backprojection results are characterized by a migration of the obtained location estimate (Fig. 7). These so-called "swimming artifacts" (Koper *et al.*, 2012; Meng,



Ampuero, Luo, *et al.*, 2012) or "shooting star artifacts" (Beskardes *et al.*, 2018) are the manifestation of the array response function of a transient signal that sweeps across an array. These artifacts result from the beam main lobe migration and exist even for a stationary radiation source. Previous observations of the swimming pattern have been attributed to the array's limited azimuthal coverage relative to the source. However, we observe swimming migration using an array with good azimuthal coverage, as demonstrated by the time-lapse backprojection results of two M_L 1.0 and 1.1 events (Fig. 7a,b), for which the sources appear to move along a linear trajectory towards southwest or northwest across the catalog source location. Shirzad *et al.* (2020) propose that 3D backprojection images of an M_w 4.1 event can constrain its source focal mechanism. Following this observation, we explore if the horizontal swimming patterns observed here correlate with the radiation pattern and the event focal mechanism.

The swimming effect in synthetic data. We use numerical experiments to systematically analyze the effects of source focal mechanisms and source-array configurations on backprojection swimming patterns. Backprojection results

Figure 6. Combined *P*- and *S*-wave backprojection results. Panels (a,b) show the depth resolution example using the same M_L 1.0 event as in Figure 2. Backprojection image of the relative energy radiation in the (a) east–west–vertical plane at the event latitude and in the (b) north–south–vertical plane at the event longitude at the timestep of the peak beam power. The black and red stars mark the backprojection and catalog location. The black contours indicate percentages of the peak beam power. *P*- and *S*-wave backprojection (c) epicenters and (d) hypocenters of the catalog events. The insets show the corresponding catalog locations. The black line denotes the borehole trajectory. The circle size is proportional to the magnitude. The color version of this figure is available only in the electronic edition.

(Fig. 8) of two point sources with focal mechanisms that have a different 0° and 90° strike but the same 80° dip and 15° rake show that the FACA data yield different and source-dependent migration patterns, whereas the LACA data always result in a swimming pattern along the source-array direction. This demonstrates that both the source focal mechanisms and the source-array configurations affect the observed swimming patterns. Genuine source effects can only be resolved with the good azimuthal FACA-type coverage, and the relative LACA

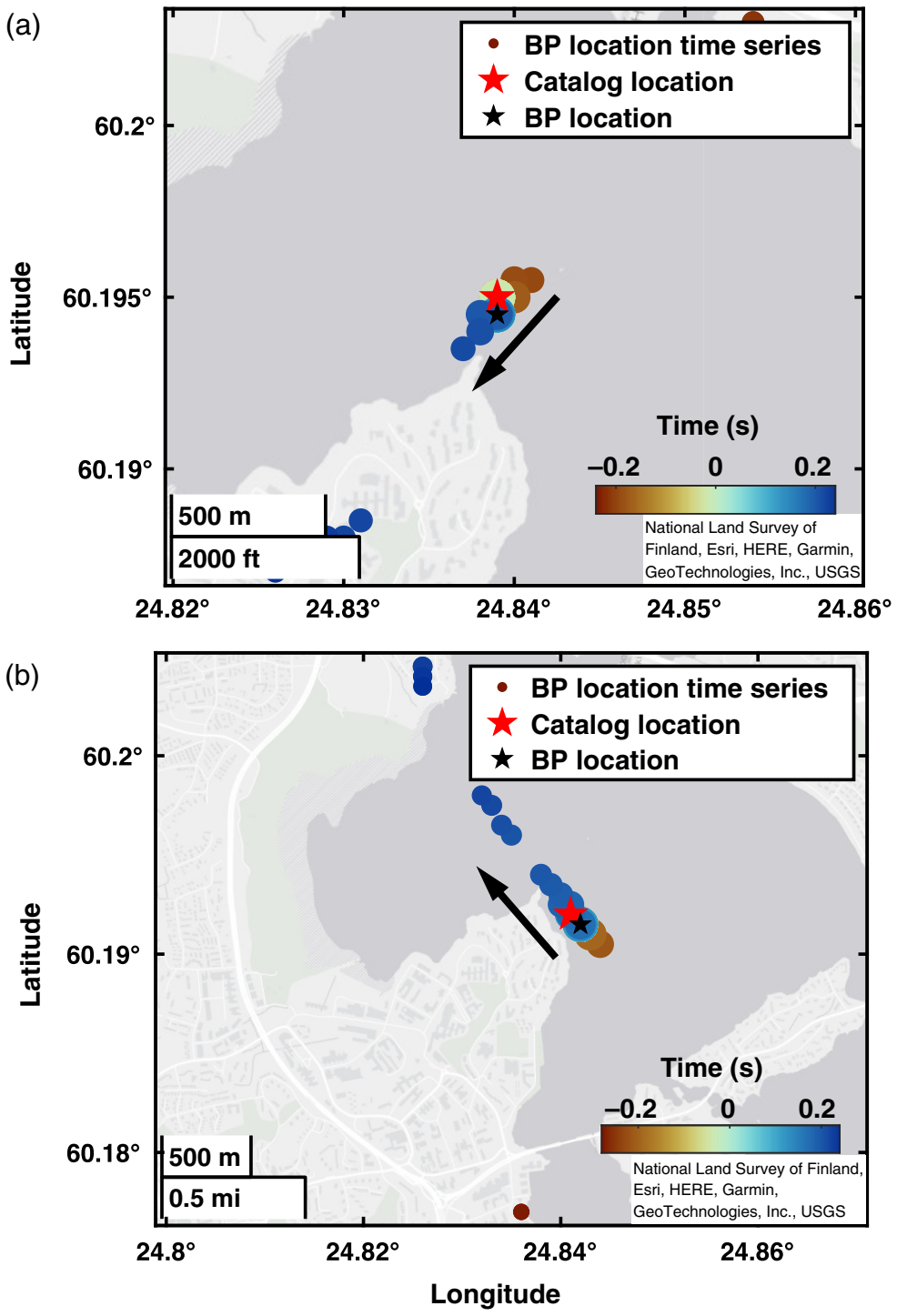
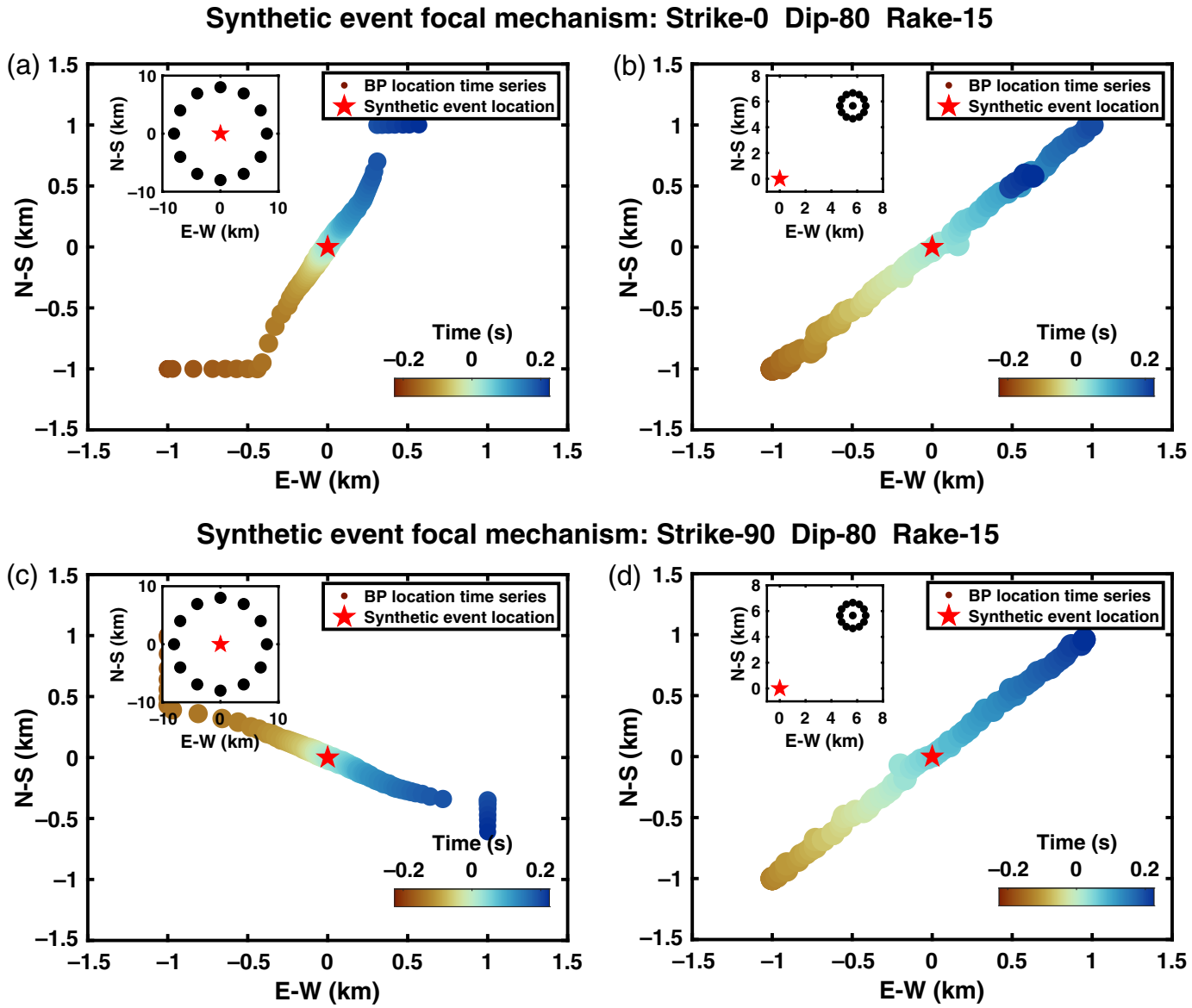


Figure 7. *P*-wave backprojection swimming pattern examples of two catalog events. (a) Backprojection results of the example M_L 1.0 event shows migration toward the southwest direction. The circle color indicates the sequence of backprojection locations, and the circle size is proportional to the backprojection energy. (b) Backprojection results of an M_L 1.1 event show migration toward the northwest direction. The color version of this figure is available only in the electronic edition.

position controls the obtained spurious swimming direction. Importantly, the dip and rake angles also affect the radiation patterns and thus the swimming patterns. To investigate the full relation between source focal mechanisms and backprojection swimming patterns, we perform *P*- and *S*-wave-only backprojection.

Figure 9a,b collects the azimuthal swimming direction using the FACA for all 3024 simulated point sources. Our results show that the swimming directions dependent systematically on the source focal mechanism, but the patterns differ for *P* and *S* waves. For certain focal mechanisms for which the dip is 90° (not shown), and for which the rake takes the values 0°, 90°, 180°, and 270°, the radiation pattern is symmetric across the source, resulting in no swimming patterns. We randomly generate three point sources with parameter values that are not included in the 3024 database (Fig. 9c, gray–yellow focal mechanism plots) and backproject the associated synthetic waveforms to estimate their *P*- and *S*-wave swimming directions. To compare, we first interpolate the results from the swimming direction database in Figure 9a,b to a refined interval from 5° to 1°. We then use it to estimate the three focal mechanisms based on their migration directions. The different *P*- and *S*-wave sensitivities together fully constrain the focal mechanisms, and the obtained minimum residual solutions (Fig. 9c, gray–red focal mechanism plots) are excellent estimates of the true focal mechanisms in all three examples. We apply this approach to 500 additional point sources with random focal mechanisms. The backprojection inferred focal mechanisms show a highly consistent match (Fig. S8). The median Kagan angle (Kagan, 1991), which describes the spatial rotation angle between two "focal mechanism plot" solutions, for our configuration is 4.88°. This indicates backprojection swimming patterns have important implications for constraining the source



properties of small-magnitude events, for which it is challenging to reliably estimate focal mechanisms.

Before we use the swimming pattern to resolve the focal mechanisms of the induced earthquakes, we investigate how different levels of noise affect the backprojection migration direction. For this analysis, we add random noise to our synthetic records at each station, where the ratio of peak noise amplitude to peak signal amplitude varies between 0.1 and 0.9 in 0.1 increments. For each noise level, we run 20 realizations and compare the swimming direction with the noise-free swimming direction. We fix the source focal mechanism as strike 332°, dip 41°, and rake 155° (Fig. 9c, the second case). The obtained deviation of the swimming direction correlates with the noise level (Fig. S9). When the noise level is 0.2 or smaller, the swimming direction can be estimated with sufficient confidence. The direction estimates are uncertain when the noise level is larger than 0.5, at which level we infer a greater than 50% chance that the deviation is larger than 10°.

Figure 8. Array configuration effects on the swimming patterns of synthetic backprojection experiments. (a) Backprojection results for a point source with focal mechanism strike 0°, dip 80°, and rake 15°, using the FACA. (b) Backprojection results of the same synthetic event using the LACA. Results in panels (c) and (d) correspond to panels (a) and (b) for a focal mechanism with strike 90°, dip 80°, and rake 15°, respectively. The color version of this figure is available only in the electronic edition.

The swimming pattern constrains induced event properties. Our numerical tests demonstrate that the swimming directions in backprojection results can constrain focal mechanisms. To further investigate this relation, we apply this approach to observed waveforms. Hillers *et al.* (2020) use *P*-wave polarities, *S*-wave-to-*P*-wave amplitude ratios, take-off angles, and station azimuths to resolve double-couple focal mechanisms of induced events. The obtained solutions are corroborated by complementary moment tensor analyses

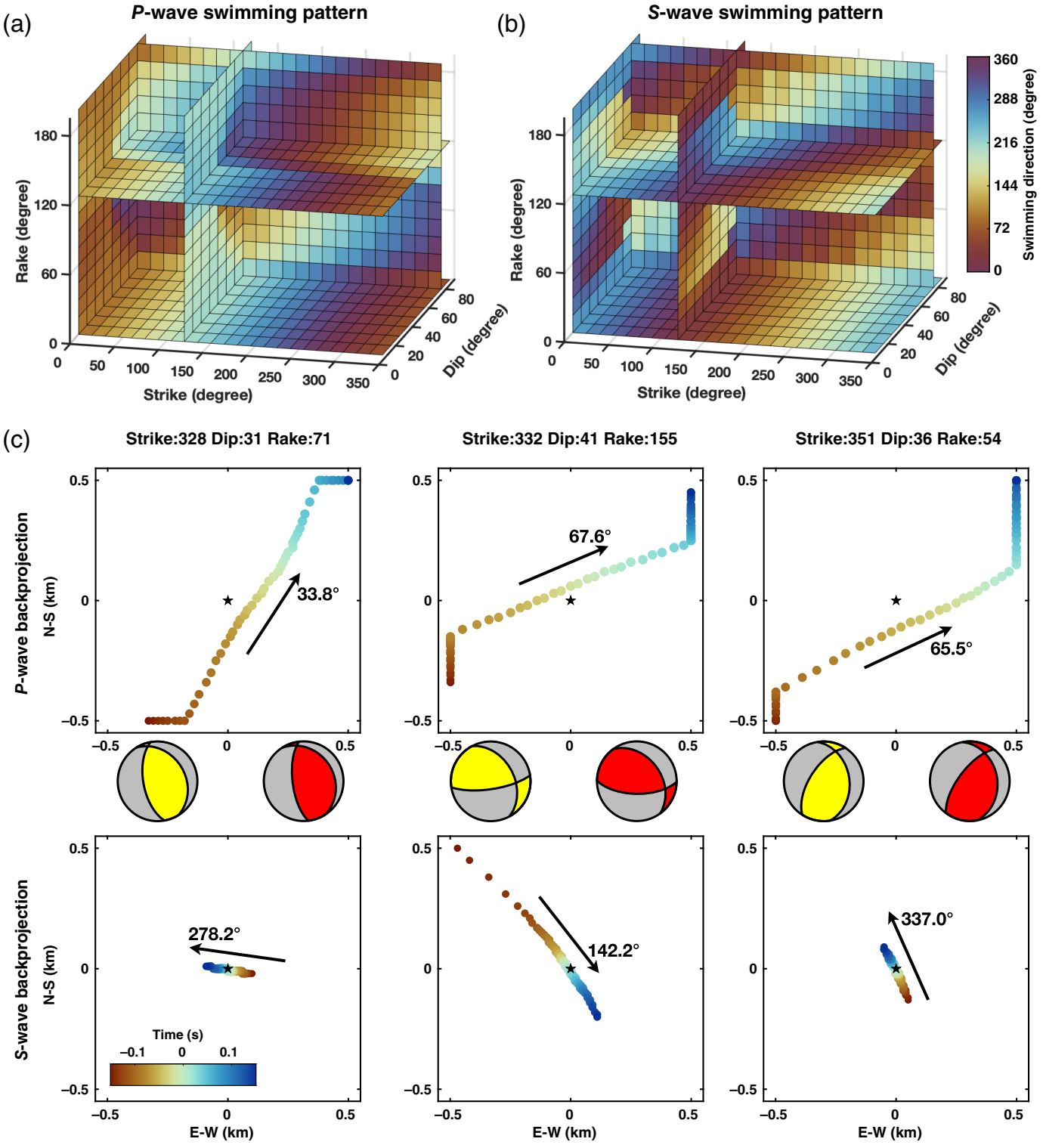


Figure 9. Numerical experiment results of backprojection swimming patterns and the resolution of synthetic source focal mechanisms. Panels (a) and (b) show the *P*- and *S*-wave swimming patterns of 3024 synthetic point-source simulations for various focal mechanisms obtained with the FACA configuration. (c) We use the obtained swimming patterns to infer the focal mechanisms of three random point sources. The black arrows denote the swimming direction for the *P*- and *S*-wave backprojection. The direction relative to the north is indicated at the arrowhead. The gray–yellow and gray–red focal mechanism plots represent the actual and inferred source focal mechanisms, respectively. The color version of this figure is available only in the electronic edition.

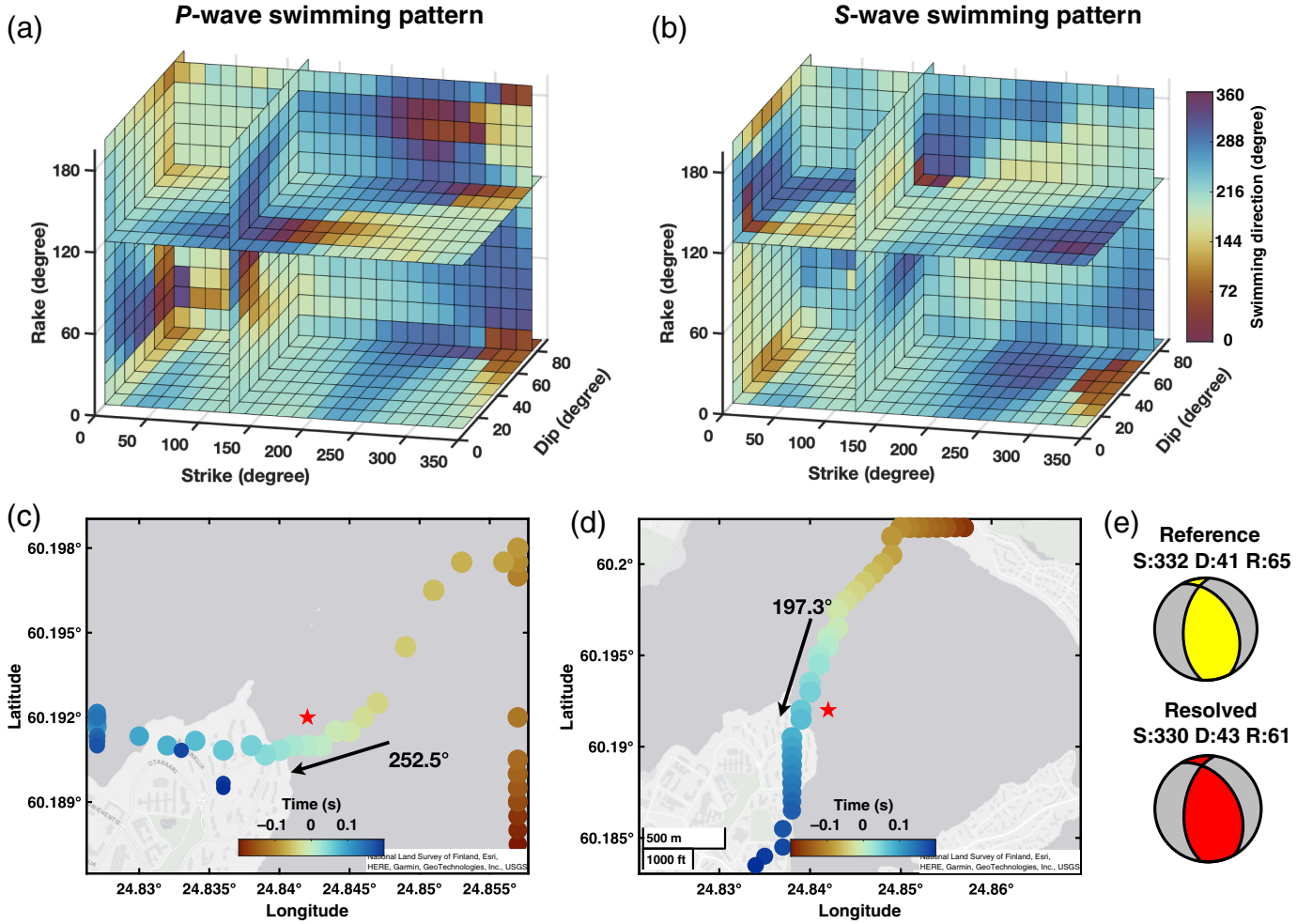


Figure 10. Synthetic swimming patterns can resolve the catalog source focal mechanism. Panels (a) and (b) are the *P*- and *S*-wave backprojection swimming patterns of 3024 synthetic point simulations for various focal mechanisms, using the 11 stations shown in Figure S9. Panels (c) and (d) show the *P*- and *S*-wave swimming direction obtained for the largest induced M_L 1.8 earthquake, using the same station configuration. (e) The inferred focal mechanism (red focal mechanism plot) is in good agreement with the reference focal mechanism (yellow focal mechanism plot). The color version of this figure is available only in the electronic edition.

(Leonhardt *et al.*, 2021; Rintamäki *et al.*, 2023). We now implement the field station configuration (Fig. 1a) for our numerical point-source simulations (Fig. S10). We use the 11 stations in the 2–4.5 km epicentral distance range relative to the mean event location to perform the backprojection for the synthetic and recorded data. This narrow distance range avoids potentially biasing variations in the duration of the body wave phases and results in optimal waveform coherence. Following the aforeintroduced procedure, we first build the database that contains the relations between the synthetic swimming pattern and the input focal mechanisms. It is then used to constrain the focal mechanisms of the induced events based on their backprojection swimming directions. Figure 10a,b shows the synthetic swimming patterns for the *P* and *S* wave, respectively. The comparison to the more homogeneous FACA results in Figure 9a,b demonstrates the effect of the array configuration on the backprojection swimming patterns. The 11 unevenly distributed stations provide good azimuthal coverage, with a maximum gap of 98° in the northeast direction. Using the same array configuration, the backprojection of the largest induced M_L 1.8 event data exhibits a swimming direction of 252.5° and 197.3° for the *P* and *S* wave, respectively (Fig. 10c,d), which corresponds to a point source with strike 330°, dip 43°, and rake 61°. The inferred focal mechanism is in very good agreement with the estimated focal mechanism of strike 332°, dip 41°, and rake 65° obtained by Hillers *et al.* (2020).

Discussion

In seismic array beamforming applications, two types of medium heterogeneity can bias the event location estimates based on raytracing along a great-circle path that uses the array observations of the local propagation direction. Large-scale heterogeneity, compared to the wavelength, can lead to bent rays, which can be accounted for using a sufficiently accurate

velocity model and raytracing. However, small-scale heterogeneity, in particular in the vicinity or below the array, can distort the assumed planarity of the wave front. This leads to the poor estimates of the source–receiver direction irrespective of the general knowledge of the velocity structure, and it hence degrades estimates of the source location using raytracing (Berteussen, 1976; Schweitzer, 2001; Schulte-Pelkum *et al.*, 2003). Approaches to correct or mitigate this bias include slowness and azimuth station corrections (Bondár *et al.*, 1999), empirical matched field processing (Harris and Kvaerna, 2010), the limited sensor-pair correlation (Gibbons *et al.*, 2018), and source-location-based station corrections (Shearer *et al.*, 2023). To locate induced microseismicity, we here use highly coherent miniarray waveforms without signal polarity variation. Miniarrays have been effectively used to detect low-amplitude signals (e.g., Li and Ghosh, 2017; Verdon *et al.*, 2017; Meng and Ben-Zion, 2018).

Our beamforming of 204 induced events in the M_L 0.0–1.8 range using 2–35 Hz data from three ∼25-station miniarrays at about 2–5 km epicentral distance shows that slowness estimates are strongly affected by medium heterogeneity even in the structurally relatively simple environment of the southern Fennoscandian Shield. Using the same data, Taylor *et al.* (2021) observe similar deviations between great circle and local propagation directions estimated from array-derived rotational motion. Our calibration in this study demonstrates that a catalog of well-located events that cover the target region can be used to significantly reduce the systematic slowness bias and improve 3D location estimates using beamforming-based raytracing.

The good azimuthal coverage of the local array supports our backprojection application. The resulting location estimates are generally less sensitive to the medium heterogeneity compared to the beamforming-based raytracing. For backprojection, fluctuations associated with unaccounted for heterogeneity are well mitigated by the stacking of signals observed across a wide azimuthal range. In our study, *P*-wave backprojection shows good horizontal resolution but comparably poor depth resolution. For teleseismic events, combining multiple seismic phases and reflected phases can improve the depth resolution in backprojection (Kiser *et al.*, 2011; Yagi *et al.*, 2012). In our local study setup, we similarly show that a combined *P*- and *S*-wave analysis significantly improves the depth resolution for overall well-constrained hypocenter estimates.

Swimming artifacts are commonly observed in teleseismic (Meng *et al.*, 2011; Koper *et al.*, 2012) and local backprojection (Beskardes *et al.*, 2018) applications. An azimuthally incomplete source-array configuration governs the associated horizontal and vertical swimming patterns (Ishii *et al.*, 2007; Xu *et al.*, 2009; Beskardes *et al.*, 2018). Our analysis of synthetic waveforms using such a typical limited azimuth array demonstrates the relation between the obtained swimming direction and the back azimuth. More importantly, however, our systematic synthetic backprojection analysis shows that the swimming pattern is not an artifact associated with the array response function but contains information about the source mechanism for azimuthally well-distributed stations.

The *P*- or *S*-wave-based swimming artifacts are individually not sufficient to uniquely constrain all three focal mechanism angles strike, dip, and rake. For example, focal mechanisms with a strike/dip/rake of 45°/60°/45°, 75°/70°/75°, 210°/70°/120°, 300°/10°/150°, and 315°/30°/165° all yield similar *P*-wave swimming directions; however, they uniquely differ in their *S*-wave patterns. In a successful first application of this connection, we resolve the catalog earthquake focal mechanism. This shows that the backprojection of data from a good azimuthal distribution is an effective tool to monitor and characterize microseismicity at local scales.

The short distances and high wavespeeds in our local study imply that the *S*-wave window cannot always be fully separated from the *P*-wave window in backprojection applications. The radiation pattern, velocity structure, and source-array distance govern the relative *P*- and *S*-wave amplitudes, and *S*-wave signals that are interfering with the *P*-wave stack across the array can lead to false detections and erroneous location estimates. When the *S*-wave amplitude exceeds the *P*-wave amplitude at most array stations, the stacking of weighted waveforms results in an erroneously increased beam power controlled by the incoherent *S*-wave energy. Beskardes *et al.* (2018) show that when the *S* wave is well sampled by a properly designed array, the short-time-average to long-time-average ratio and kurtosis processing can reduce the sensitivity of backprojection to *S* waves. Here, we show that combined *P* and *S* waves backprojection can avoid strong *S*-wave effects and, therefore, improve the depth resolution, even when stations in a limited distance range are used.

Location estimates obtained with calibrated beamforming and combined *P*- and *S*-wave backprojection show an average 3D offset of ∼140 and ∼118 m compared to the reference catalog locations, respectively. These reference locations themselves differ from the relocated industrial catalog that employs deep borehole string data and a refined 1D model (Kwiatek *et al.*, 2019). Our reported offsets may be further reduced using a better velocity model together with refined search grids or in combination with other optimization strategies.

Conclusions

We investigate the performance and limitations of the two widely applied beamforming and backprojection array processing techniques using high SNR seismic records of 204 catalog M_L 0.0–1.8 earthquakes induced by the ∼6 km deep 2018 Espoo/Helsinki geothermal stimulation experiment. Our local-scale results demonstrate that a calibration function constructed from catalog reference data is required to mitigate the systematic slowness bias associated with medium

heterogeneity and to improve 2D and 3D location estimates for both *P*- and *S*-wave beamforming. *P*-wave backprojection results using polarity corrected waveform data obtained with azimuthally well-distributed single stations show very good horizontal but comparatively poor depth resolution. The average offset to the catalog locations are ~97.5 m in the horizontal direction and ~1.4 km in the vertical direction. A large *S*/*P* amplitude ratio can lead to backprojection locations away from the travel-time-based reference catalog sources. We establish that a *P* and *S* wave combined backprojection approach can mitigate the *S*-wave effects. The combined approach significantly reduces the average depth offsets to the reference locations. Using waveform envelope data the average offset is reduced from ~3 km to ~91 m. Numerical experiments and data processing of the largest induced event demonstrate that the backprojection swimming pattern is not merely an artifact of the array response function; if the backprojection array has good azimuthal coverage, it can constrain earthquake focal mechanism parameters. These results have important implications for monitoring and characterizing the abundance of small-magnitude seismicity around EGSs and in natural environments using seismic array techniques.

Data and Resources

Waveform data of the induced earthquakes are available from Vuorinen *et al.* (2023). A supplementary file to Taylor *et al.* (2021) (doi: 10.1029/2020GL090403) contains information about the 204 events used in this study. We use the software package SeisSol (https://www.seissol.org) to simulate the point-source wavefields. SeisSol is an open-source and freely available at https://github.com/SeisSol/SeisSol. Documentation for downloading, compiling, and using SeisSol is available at https://seissol.readthedocs.io. The point-source synthetic simulation data can be accessed through Zenodo at doi: 10.5281/zenodo.7541076. All websites were last accessed in May 2024. The supplemental material includes the calibration function stability test, beamforming results using horizontal components, calibrated beamforming locations for the other three events, backprojection results using *P*-wave envelope data, a backprojection synthetic noise test, array configurations, synthetic location offsets, Kagan angle results for the focal mechanism solutions of 500 random synthetic sources, and systematic location offsets to the reference catalog using combined *P* and *S* waves.

Declaration of Competing Interests

The authors acknowledge that there are no conflicts of interest recorded.

Acknowledgments

The authors thank Peter Shearer for his constructive comments. The 2018 temporary deployment was supported by the Geophysical Instrument Pool Potsdam under Grant Number 201802-ORS. This work received funding from the European Union's (EU) Horizon 2020 research and innovation program (Truly Extended Earthquake Rupture (TEAR) project funded by the European Research Council (ERC) Starting Grant Agreement Number 852992) and the Horizon Europe (DT-Geo Grant Agreement Number 101058129, Geo-INQUIRE Grant Agreement Number 101058518, and ChEESE-2P Grant Agreement Number 101093038), the King Abdullah University of Science and Technology (KAUST, Grant Number BAS/1/1339-01-01), the National Science Foundation (NSF Grant Numbers EAR-2121568, EAR-2225286, OAC-2311208, and OAC-2139536), the National Aeronautics and Space Administration (NASA; Grant Number 80NSSC20K0495), and the Research Council of Finland (Funding Decision Numbers 337913, 338075, and 339670). The authors acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu, last accessed May, 2022, project pr63qo) for funding this project by providing computing time on the Gauss Centre for Supercomputing (GCS) Supercomputer SuperMUC-NG at Leibniz Supercomputing Centre (www.lrz.de). Computing resources were also provided by the Institute of Geophysics of LMU Munich (Oeser *et al.*, 2006). The authors thank two anonymous reviewers for comments that helped to improve the article.

References

Ader, T., M. Chendorain, M. Free, T. Saarno, P. Heikkinen, P. E. Malin, P. Leary, G. Kwiatek, G. Dresen, F. Bluemle, *et al.* (2020). Design and implementation of a traffic light system for deep geothermal well stimulation in Finland, *J. Seismol.* **24,** no. 5, 991–1014.

Berteussen, K. (1976). The origin of slowness and azimuth anomalies at large arrays, *Bull. Seismol. Soc. Am.* **66,** no. 3, 719–741.

Beskardes, G., J. Hole, K. Wang, M. Michaelides, Q. Wu, M. Chapman, K. Davenport, L. Brown, and D. Quiros (2018). A comparison of earthquake backprojection imaging methods for dense local arrays, *Geophys. J. Int.* **212,** no. 3, 1986–2002.

Birtill, J., and F. Whiteway (1965). The application of phased arrays to the analysis of seisbody waves, *Phil. Trans. Roy. Soc. London* **258,** no. 1091, 421–493.

Bondár, I., R. G. North, and G. Beall (1999). Teleseismic slowness-azimuth station corrections for the International Monitoring System seismic network, *Bull. Seismol. Soc. Am.* **89,** no. 4, 989–1003.

Breuer, A., A. Heinecke, S. Rettenberger, M. Bader, A.-A. Gabriel, and C. Pelties (2014). Sustained petascale performance of seismic simulations with SeisSol on SuperMUC, *International Supercomputing Conference*, Springer, 1–18.

Buijze, L., L. van Bijsterveldt, H. Cremer, B. Paap, H. Veldkamp, B. B. Wassing, J.-D. VanWees, G. C. van Yperen, J. H. Heege ter, and B. Jaarsma (2019). Review of induced seismicity in geothermal systems worldwide and implications for geothermal systems in the Netherlands, *Neth. J. Geosci.* **98,** e13 , doi: 10.1017/njg.2019.6.

Chambers, K., J.-M. Kendall, S. Brandsberg-Dahl, and J. Rueda (2010). Testing the ability of surface arrays to monitor microseismic activity, *Geophys. Prospect.* **58,** no. 5, 821–830.

Corciulo, M., P. Roux, M. Campillo, D. Dubucq, and W. A. Kuperman (2012). Multiscale matched-field processing for noise-source localization in exploration geophysics, *Geophysics* **77,** no. 5, KS33–KS41, doi: 10.1190/geo2011-0438.1.

Cuenot, N., J. Scheiber, W. Moeckes, and A. Genter (2015). Evolution of the natural radioactivity on the Soultz-sous-Forêt EGS power plant and implication for radiation protection, *Proc. of the World Geothermal Congress*, Melbourne, Australia.

Diehl, T., T. Kraft, E. Kissling, and S. Wiemer (2017). The induced earthquake sequence related to the St. Gallen deep geothermal

project (Switzerland): Fault reactivation and fluid inter actions imaged by microseismicity, *J. Geophys. Res.* **122,** no. 9, 7272–7290.

Douglas, A. (2002). Seismometer arrays - Their use in earthquake and test ban seismology, in *International Handbook of Earthquake & Engineering Seismology, Part A*, W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger (Editors), Academic Press, London, England, 357–367, doi: 10.1016/S0074-6142(02)80226-1.

Efron, B. (1992). Bootstrap methods: Another look at the jackknife, in *Breakthroughs in Statistics*, S. Kotz and N. L. Johnson (Editors), Springer, New York, 569–593, doi: 10.1007/978-1-4612-4380-9_41.

Eulenfeld, T., G. Hillers, T. A. Vuorinen, and U. Wegler (2023). Induced earthquake source parameters, attenuation, and site effects from waveform envelopes in the Fennoscandian Shield, *J. Geophys. Res.* doi: 10.1029/2022JB025162.

Folch, A., C. Abril, M. Afanasiev, G. Amati, M. Bader, R. M. Badia, H. B. Bayraktar, S. Barsotti, R. Basili, F. Bernardi, *et al.* (2023). The EU Center of Excellence for Exascale in Solid Earth (ChEESE): Implementation, results, and roadmap for the second phase, *Future Gener. Comput. Syst.* **146,** 47–61.

Garcia, J., C. Hartline, M. Walters, M. Wright, J. Rutqvist, P. F. Dobson, and P. Jeanne (2016). The Northwest Geysers EGS demonstration project, California: Part 1: Characterization and reservoir response to injection, *Geothermics* **63,** 97–119.

Giardini, D. (2009). Geothermal quake risks must be faced, *Nature* **462,** no. 7275, 848–849.

Gibbons, S. J., S. P. Näsholm, E. Ruigrok, and T. Kværna (2018). Improving slowness estimate stability and visualization using limited sensor pair correlation on seismic arrays, *Geophys. J. Int.* **213,** no. 1, 447–460.

Giovanni, B., C. Guido, and F. Adolfo (2005). Characteristics of geothermal fields in Italy, *Giornale di Geologia Applicata* **1,** 247–254.

Gradon, C., L. Moreau, P. Roux, and Y. Ben-Zion (2019). Analysis of surface and seismic sources in dense array data with match field processing and Markov chain Monte Carlo sampling, *Geophys. J. Int.* **218,** no. 2, 1044–1056, doi: 10.1093/gji/ggz224.

Harris, D. B., and T. Kvaerna (2010). Superresolution with seismic arrays using empirical matched field processing, *Geophys. J. Int.* **182,** no. 3, 1455–1477.

Hedlin, M. A., J. B. Minster, and J. A. Orcutt (1991). Beam-stack imaging using a small aperture array, *Geophys. Res. Lett.* **18,** no. 9, 1771–1774.

Hillers, G., T. Vuorinen, M. R. Uski, J. T. Kortström, P. B. Mäntyniemi, T. Tiira, P. E. Malin, and T. Saarno (2020). The 2018 geothermal reservoir stimulation in Espoo/Helsinki, Southern Finland: Seismic network anatomy and data features, *Seismol. Res. Lett.* **91,** no. 2A, 770–786.

Inbal, A., R. W. Clayton, and J.-P. Ampuero (2015). Imaging widespread seismicity at midlower crustal depths beneath Long Beach, CA, with a dense seismic array: Evidence for a depth dependent earthquake size distribution, *Geophys. Res. Lett.* **42,** no. 15, 6314–6323.

Ishii, M., P. M. Shearer, H. Houston, and J. E. Vidale (2005). Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array. *Nature* **435,** no. 7044, 933–936.

Ishii, M., P. M. Shearer, H. Houston, and J. E. Vidale (2007). Teleseismic P wave imaging of the 26 December 2004 Sumatra-Andaman and 28 March 2005 Sumatra earthquake ruptures using the Hi-net array, *J. Geophys. Res.* **112,** no. B11, doi: 10.1029/2006JB004700.

Kagan, Y. (1991). 3-D rotation of double-couple earthquake sources, *Geophys. J. Int.* **106,** no. 3, 709–716.

Kiser, E., and M. Ishii (2013). Hidden aftershocks of the 2011 Mw 9.0 Tohoku, Japan earthquake imaged with the backprojection method, *J. Geophys. Res.* **118,** no. 10, 5564–5576.

Kiser, E., and M. Ishii (2017). Back-projection imaging of earthquakes, *Annu. Rev. Earth Planet. Sci.* **45,** 271–299.

Kiser, E., M. Ishii, C. H. Langmuir, P. Shearer, and H. Hirose (2011). Insights into the mechanism of intermediate-depth earthquakes from source properties as imaged by back projection of multiple seismic phases, *J. Geophys. Res.* **116,** no. B6, doi: 10.1029/2010JB007831.

Koper, K. D., A. R. Hutko, T. Lay, and O. Sufri (2012). Imaging short-period seismic radiation from the 27 February 2010 Chile (Mw 8.8) earthquake by back-projection of P, PP, and PKIKP waves, *J. Geophys. Res.* **117,** no. B2, doi: 10.1029/2011JB008576.

Kortström, J., M. Uski, and K. Oinonen (2018). The Finnish National Seismic Network, *Summ. Bull. Int. Seismol. Cent.*, **52,** 41–52, doi: 10.31905/59QRNANC.

Krenz, L., C. Uphoff, T. Ulrich, A.-A. Gabriel, L. S. Abrahams, E. M. Dunham, and M. Bader (2021). 3D acoustic-elastic coupling with gravity: The dynamics of the 2018 Palu, Su earthquake and tsunami, *Proc. of the International Conf. for High Performance Computing, Networking, Storage and Analysis*, ACM, New York, 14–19 November 2021.

Krüger, F., and M. Ohrnberger (2005). Tracking the rupture of the Mw9.3 Sumatra earthquake over 1,150 km at teleseismic distance, *Nature* **435,** no. 7044, 937–939.

Krüger, F., M. Weber, F. Scherbaum, and J. Schlittenhardt (1993). Double beam analysis of anomalies in the core-mantle boundary region, *Geophys. Res. Lett.* **20,** no. 14, 1475–1478.

Kwiatek, G., T. Saarno, T. Ader, F. Bluemle, M. Bohnhoff, M. Chendorain, G. Dresen, P. Heikkinen, I. Kukkonen, P. Leary, *et al.* (2019). Controlling fluid-induced seismicity during a 6.1-km-deep geothermal stimulation in Finland, *Sci. Adv.* **5,** no. 5, eaav7224 , doi: 10.1785/0220190253.

Langston, C. A. (2007). Wave gradiometry in two dimensions, *Bull. Seismol. Soc. Am.* **97,** no. 2, 401–416, doi: 10.1785/0120060138.

Leonhardt, M., G. Kwiatek, P. Martínez-Garzón, M. Bohnhoff, T. Saarno, P. Heikkinen, and G. Dresen (2021). Seismicity during and after stimulation of a 6.1 km deep enhanced geothermal system in Helsinki, Finland, *Solid Earth* **12,** no. 3, 581–594, doi: 10.5194/se-12-581-2021.

Li, B. (2019). A broad spectrum of fault behaviors in fast and slow earthquakes, University of California, Riverside, California, available at https://escholarship.org/uc/item/4149s6nr (last accessed August 2024).

Li, B., and A. Ghosh (2016). Imaging rupture process of the 2015 M_w 8.3 Illapel earthquake using the US Seismic array, *Pure Appl. Geophys.*, **173,** 2245–2255, doi: 10.1007/s00024-016-1323-y.

Li, B., and A. Ghosh (2017). Near-continuous tremor and low-frequency earthquake activities in the Alaska-Aleutian subduction zone revealed by a mini seismic array, *Geophys. Res. Lett.* **44,** no. 11, 5427–5435.

Li, B., B. Wu, H. Bao, D. D. Oglesby, A. Ghosh, A.-A. Gabriel, L. Meng, and R. Chu (2022). Rupture heterogeneity and directivity effects in back-projection analysis, *J. Geophys. Res.* **127,** no. 3, e2021JB022663 , doi: 10.1029/2021JB022663.

Mai, P. M., T. Aspiotis, T. A. Aquib, E. V. Cano, D. Castro-Cruz, A. Espindola-Carmona, B. Li, X. Li, J. Liu, R. Matrau, *et al.* (2023). The destructive earthquake doublet of 6 February 2023 in South-Central Türkiye and northwestern Syria: Initial observations and analyses, *Seism. Record* **3,** no. 2, 105–115.

Meng, H., and Y. Ben-Zion (2018). Detection of small earthquakes with dense array data: Example from the San Jacinto fault zone, southern California, *Geophys. J. Int.* **212,** no. 1, 442–457.

Meng, L., J.-P. Ampuero, Y. Luo, W. Wu, and S. Ni (2012). Mitigating artifacts in back projection source imaging with implications for frequency-dependent properties of the Tohoku-Oki earthquake, *Earth Planets Space* **64,** no. 12, 1101–1109.

Meng, L., J.-P. Ampuero, A. Sladen, and H. Rendon (2012). High-resolution backprojection at regional distance: Application to the Haiti M7.0 earthquake and comparisons with finite source studies, *J. Geophys. Res.* **117,** no. B4, doi: 10.1029/2011JB008702.

Meng, L., A. Inbal, and J.-P. Ampuero (2011). A window into the complexity of the dynamic rupture of the 2011 Mw 9 Tohoku-Oki earthquake, *Geophys. Res. Lett.* **38,** no. 7, doi: 10.1029/2011GL048118.

Oeser, J., H.-P. Bunge, and M. Mohr (2006). Cluster design in the earth sciences Tethys, *International Conf. on High Performance Computing and Communications*, Vol. 640, Springer, 31–40.

Rathnaweera, T. D., W. Wu, Y. Ji, and R. P. Gamage (2020). Understanding injection-induced seismicity in enhanced geothermal systems: From the coupled thermo-hydro-mechanical chemical process to anthropogenic earthquake prediction, *Earth Sci. Rev.* **205,** 103182.

Rintamäki, A., G. Hillers, S. Heimann, T. Dahm, and A. Korja (2023). Centroid full moment tensor analysis reveals fluid channels opened by induced seismicity at EGS, Helsinki region, southern Finland, *EGU General Assembly Conference Abstracts*, EGU23–12756 pp., doi: 10.5194/egusphere-egu23-12756.

Rost, S., and C. Thomas (2002). Array seismology: Methods and applications, *Rev. Geophys.* **40,** no. 3, 2–1.

Schulte-Pelkum, V., F. L. Vernon, and J. Eakins (2003). Large teleseismic P wavefront deflections observed with broadband arrays, *Bull. Seismol. Soc. Am.* **93,** no. 2, 747–756.

Schweitzer, J. (2001). Slowness corrections—One way to improve IDC products, *Pure Appl. Geophys.* **158,** no. 1, 375–396.

Seithel, R., E. Gaucher, B. Mueller, U. Steiner, and T. Kohl (2019). Probability of fault reactivation in the Bavarian Molasse Basin, *Geothermics* **82,** 81–90.

Shearer, P. M., H. Meng, and W. Fan (2023). Earthquake detection using a nodal array on the San Jacinto fault in California: Evidence for high foreshock rates preceding many events, *J. Geophys. Res.* **128,** no. 3, e2022JB025279 , doi: 10.1029/2022JB025279.

Shirzad, T., S. Lasocki, and B. Orlecka-Sikora (2020). Tracking the development of seismic fracture network by considering the fault rupture method, *EGU General Assembly Conference Abstracts*, 13,079 pp.

Suryanto, W., H. Igel, J. Wassermann, A. Cochard, B. Schuberth, D. Vollmer, F. Scherbaum, U. Schreiber, and A. Velikoseltsev (2006). First comparison of array-derived rotational ground motions with direct ring laser measurements, *Bull. Seismol. Soc. Am.* **96,** no. 6, 2059–2071, doi: 10.1785/0120060004.

Taufiqurrahman, T., A.-A. Gabriel, D. Li, T. Ulrich, B. Li, S. Carena, A. Verdecchia, and F. Gallovič (2023). Dynamics, interactions and delays of the 2019 Ridgecrest rupture sequence, *Nature* **618,** 308–315.

Taylor, G., G. Hillers, and T. Vuorinen (2021). Using array-derived rotational motion to obtain local wave propagation properties from earthquakes induced by the 2018 geothermal stimulation in Finland, *Geophys. Res. Lett.* **48,** no. 6, e2020GL090403 , doi: 10.1029/2020GL090403.

Uphoff, C., S. Rettenberger, M. Bader, E. H. Madden, T. Ulrich, S. Wollherr, and A.-A. Gabriel (2017). Extreme scale multi-physics simulations of the tsunamigenic 2004 Sumatra megathrust earthquake, *Proc. of the International Conf. for High Performance Computing, Networking, Storage and Analysis*, 1–16.

Verdon, J. P., J.-M. Kendall, S. P. Hicks, and P. Hill (2017). Using beamforming to maximise the detection capability of small, sparse seismometer arrays deployed to monitor oil field activities, *Geophys. Prospect.* **65,** no. 6, 1582–1596.

Vlček, J., T. Fischer, and J. Vilhelm (2016). Back-projection stacking of P-and S-waves to determine location and focal mechanism of microseismic events recorded by a surface array, *Geophys. Prospect.* **64,** no. 6, 1428–1440.

Vuorinen, T. A., T. Veikkolainen, K. Oinonen, J. Hällsten, M. Uski, T. Heikkilä, A. Rintamäki, A. Voutilainen, I. Salmenperä, J. Kortström, *et al.* (2023). ISUH waveform, time and location data products from stimulations of deep geothermal wells in Espoo in 2020, doi: 10.23729/cdfd937c-37d5-46b0-9c16-f6e0c10bc81f.

Weichert, D. (1975). Reduced false alarm rates in seismic array detection by non-linear beam forming, *Geophys. Res. Lett.* **2,** no. 4, 121–123.

Wright, C. (1972). Array studies of seismic waves arriving between P and PP in the distance range 90 to 115, *Bull. Seismol. Soc. Am.* **62,** no. 1, 385–400.

Xu, Y., K. D. Koper, O. Sufri, L. Zhu, and A. R. Hutko (2009). Rupture imaging of the Mw 7.9 12 May 2008 Wenchuan earthquake from back projection of teleseismic P waves, *Geochem. Geophys. Geosys.* **10,** no. 4, doi: 10.1029/2008GC002335.

Yagi, Y., A. Nakao, and A. Kasahara (2012). Smooth and rapid slip near the Japan Trench during the 2011 Tohoku-Oki earthquake revealed by a hybrid back-projection method, *Earth Planet. Sci. Lett.* **355,** 94–101.

Yang, L., X. Liu, and G. C. Beroza (2021). Revisiting evidence for widespread seismicity in the upper mantle under Los Angeles, *Sci. Adv.* **7,** no. 4, eabf2862 , doi: 10.1126/sciadv.abf2862.

Yin, J., H. Yang, H. Yao, and H. Weng (2016). Coseismic radiation and stress drop during the 2015 Mw 8.3 Illapel, Chile megathrust earthquake, *Geophys. Res. Lett.* **43,** no. 4, 1520–1528.

Zhang, L., Y. Zhou, X. Zhang, A. Zhu, B. Li, S. Wang, S. Liang, C. Jiang, J. Wu, Y. Li, *et al.* (2023). 2022 Mw 6.6 Luding, China, earthquake: A strong continental event illuminating the Moxi seismic gap, *Seismol. Res. Lett.* **94,** 2129–2142.

Appendix A

Beamforming

For event location using beamforming, we perform a grid search in the slowness domain in a polar coordinate system. Beamforming assumes a plane wavefront propagating across the array. It uses the slowness vector with its horizontal

slowness and back azimuth to quantify the differential travel times at each array station relative to the array center (Rost and Thomas, 2002). The correspondingly shifted traces at each station are stacked to yield the beam $s_{j-\mathrm{beam}}$ at the slowness vector grid j,

$$s_{j-\mathrm{beam}} = \sum_{i=1}^{N} u_i(t + \boldsymbol{r}_j \cdot \boldsymbol{u}_i), \tag{A1}$$

in which N is the number of array stations, u_i is the seismogram at station i, $\boldsymbol{r}_i$ is the distance vector of the station i relative to the array center, $\boldsymbol{u}_j$ is the horizontal slowness vector at the slowness grid element j, and the dot product $\boldsymbol{r}_j \cdot \boldsymbol{u}_i$ is the corresponding time shift. The beam power or beam energy at each grid element j is computed as an integral of $s^2_{j-\mathrm{beam}}$ over a time window around the P or S wave. The slowness vector grid element with the maximum beam energy is the solution that indicates the local wave propagation direction relative to the array center. The array center of each miniarray is the origin of the slowness domain coordinate system. The domain is gridded using a 0.01 s/km slowness interval from 0 to 0.3 s/km and a 2° azimuth interval from 0° to 360°. The high signal-to-noise ratio (SNR) allows us to process the data in a wide frequency range between 2 and 35 Hz. We perform beamforming independently for each miniarray. We use vertical-component P-wave data and independent S waveforms from the two horizontal components.

We use the bootstrap method (Efron, 1992) to estimate the slowness uncertainty at each array. For this, we iterate the aforementioned analysis using all but one station of the array. Ideally, the back azimuth of each slowness vector will point to the source direction from the array center. In applications, it typically requires correction for the structural heterogeneity along the travel paths. We calibrate the slowness bias following the method discussed in the Beamforming results and calibration section. We then use the calibrated slowness vector to locate the source. For this, we select a volume that extends the catalog limits by 500 m in the east–west and the north–south directions and by 2 km in the vertical direction. We discretize the volume into cells with an interval of 50 m and calculate the theoretical slowness vector of each grid cell relative to each array center, using a 1D homogeneous velocity model, with $V_P = 6.2$ and $V_S = 3.62$ km/s (Kortström *et al.*, 2018). These values from the uppermost 15 km thick layer of a regional velocity model were also used for the catalog locations (Hillers *et al.*, 2020). Then, we estimate the event location by ray tracing the slowness vector from each array center, with the uncertainty from the bootstrap estimates (Fig. 2). Because of the residuals of the calibration function, the traced rays may not intersect at one source grid. We use the cell location with the minimum summed distance to the three rays as the location estimate.

Appendix B

Backprojection

When using backprojection to image the rupture propagation of large earthquakes, the known hypocenter location is used to correct the signal polarity of the array recordings (e.g., Mai *et al.*, 2023; Taufiqurrahman *et al.*, 2023). For unknown sources, functions such as the envelope, short-term averaging over long-term averaging, and kurtosis have been applied to mitigate the polarity ambiguity in local backprojection (e.g., Inbal *et al.*, 2015; Vlček *et al.*, 2016; Beskardes *et al.*, 2018). However, these strategies negatively affect backprojection resolution, event magnitude estimates, and the detection threshold (Beskardes *et al.*, 2018). We test the P-wave envelope backprojection and obtain a poor depth resolution. Most results indicate the largest beam power at the shallowest depth of the grid-search domain, which is 3 km in our study (Fig. S11). This results in an average depth offset of ∼3 km to the catalog location. To gain a more comprehensive understanding of the limitations and performance of backprojection, we used reference catalog locations to find the P-wave arrival time and performed cross correlation around the P phase to correct the polarity in the *P*-wave Data Analysis Results section. Waveform envelope is used in the other sections. We implement the backprojection method using a moving time window of 0.3 s length and a 0.004 s timestep for the catalog data and 0.005 s timestep for the synthetic data. For each time window, we compute a linear stack across the array (Kiser and Ishii, 2017):

$$s_{j-\mathrm{bp}}(t) = \sum_{i=1}^{N} w^{ij} u_i(t + \tau^{ij}), \tag{B1}$$

in which $s_{j-\mathrm{bp}}(t)$ is the stacked time series of N stations that corresponds to the target grid point j, w^{ij} is a weighting factor that is used to correct the polarity and to normalize $u_i(t)$, $u_i(t)$ is the seismogram waveform or envelope at station i, and τ^{ij} is the travel time from source j to station i. The backprojection stacked beam power or energy at time t is computed as an integral of $s^2_{j-\mathrm{bp}}(t)$ over a time window from t to $t + t_w$, with $t_w = 0.3$ s.

To avoid potentially problematic variations in the duration of the body phases and for optimal waveform coherence, we limit the analysis to data from azimuthally well-distributed local array stations at short epicentral distances between 2 and 4.5 km.

Our earthquake catalog (Hillers *et al.*, 2020) lists induced earthquakes between 24.8286° and 24.8450° east and between 60.1880° and 60.1957° north. We extend the lateral dimension of the backprojection search source volume around these limits from 24.82° to 24.86° east and from 60.18° to 60.205° north, which corresponds to an area of 2.8 km × 2.2 km (Fig. 1). We divide this region into cells with 0.0005° and 0.001° spacing in

latitude and longitude, respectively. This leads to the square grid elements with edge lengths of about 55 m. To estimate 3D locations, we consider a depth range between 3 and 8 km depth with 50 m spacing. We perform a grid search and compute the relative radiation strength from each source cell by stacking the waveforms of all stations. We shift the moveouts by synthetic travel times obtained from the homogeneous velocity model (Kortström *et al.*, 2018). We use a moving time window during the location search, and the maximum stacked energy across all cells indicates the source location for a given time window. The backprojection stacked beam power significantly exceeds the background level in time windows that contain coherent wave arrivals.

Appendix C

Numerical simulation

We use SeisSol (Breuer *et al.*, 2014; Uphoff *et al.*, 2017; Krenz *et al.*, 2021), a flagship code in the EU Center of Excellence for Exascale in Solid Earth (ChEESE) project (Folch *et al.*, 2023), to perform 3D point-source seismic wave propagation simulations with varying focal mechanisms and to generate synthetic waveforms at two model arrays for backprojection analysis. SeisSol solutions feature high-order accuracy in space and time. We choose polynomial basis functions of the order $p = 3$, leading to a fourth-order accurate scheme. The full domain is $150 \times 150 \times 100$ km^3. The spatial discretization of the statically adaptive, fully unstructured tetrahedral mesh (Fig. 3a) is refined to 200 m in the central $20 \times 20 \times 8$ km^3 target high-resolution area, which covers the source and array region. This can resolve frequencies up to 4.5 Hz. The computational cost of each point-source simulation is ∼130 central processing unit hours for 8 s simulation time.

To estimate the swimming direction of the synthetic backprojection results, we use the direction between two points before and after the backprojection location "swims" across the epicenter. We first find the time at which the beam power is maximized, and the associated grid point is generally indicative of the synthetic source location. The two points are then the backprojection locations associated with 10 time increments before and after the timing of the peak beam power. A larger temporal average does not improve the estimates of the linear swimming direction across the epicenter, and it can be influenced by backprojection location at the boundary of the target region (Fig. 5). For the synthetic data, we use the signal envelope to avoid the polarity ambiguity and investigate the solution to improve the source location.

Manuscript received 27 March 2024
Published online 4 September 2024