Geophysical Research Letters

Supporting Information for

Local beamforming and back-projection of induced earthquakes in Helsinki, southern Finland

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

1. Department of Earth and Environmental Sciences, Ludwig-Maximilians-Universität München, Munich, Germany
2. Physical Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
3. Scripps Institution of Oceanography, UC San Diego, La Jolla, CA, USA
4. Institute of Seismology, University of Helsinki, Helsinki, Finland

Contents of this file

Introduction

The supporting information presented here includes the text about detailed description of the beamforming (Section 6.1) and back-projection array (Section 6.2) methods, and the synthetic experiments (Section 6.3). It also includes seven figures which present additional information or results to support the paper.

6.1. Beamforming

For event location using beamforming, we perform a grid search in the slowness domain in a polar coordinate system. Beamforming assumes a plane wave front propagating across the array, and uses the slowness vector with its elements horizontal slowness and back azimuth to quantify the differential travel times at each array station relative to the array center (Rost & Thomas, 2002). The correspondingly shifted traces at each station are stacked to yield the beam s_{j-beam} at the slowness vector grid j

$$s_{j-beam} = \sum_{i=1}^{N} u_i\left(t + \boldsymbol{r_i} \cdot \boldsymbol{u_j}\right),$$

where N is the number of array stations, u_i is the seismogram at station i, $\boldsymbol{r_i}$ is the distance vector of 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 $\boldsymbol{r_i} \cdot \boldsymbol{u_j}$ 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-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 mini array is the origin of the slowness domain coordinate system. The domain is divided into grids with a 0.01 s/km slowness interval from 0 s/km to 0.3 s/km, and a 2° azimuth interval from $0°$to $360°$. The high signal-to-noise ratio allows us to process the data in a wide frequency range between 2 Hz and 35 Hz. We perform beamforming independently for each mini array. 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 above analysis using all but one stations in the array. Ideally, the back-azimuth of each slowness vector will point to the source direction from the array center, but in practice, it typically requires correction for the structural heterogeneity along the travel paths. Thus, we calibrate the slowness bias following the calibration method mentioned in Section 3.1. Next, we use the calibrated slowness vector to locate the source. we then select a volume by extending 500 m further in the East-West and in the North-South direction, and 2 km in the vertical direction from the catalog boundary locations. 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 $v_P = 6.2$ km/s and $v_S = 3.62$ km/s (Kortström et al., 2018). These values from the uppermost 15 km thick layer of a regional 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 (Figure 2). Due to the residuals of the calibration function and the limited refinement of the searching cell size, the slowness ray traces may not intersect at one source grid, and we use the cell location with the minimum summed distance to the three traces as the location estimate.

6.2. Back-projection

The back-projection method has the signal time shifting and stacking in common with beamforming. In contrast to beamforming, however, which discretizes the slowness space to estimate local plane wave propagation, back-projection discretizes the space of the location of energy radiation, and then calculates a corresponding suite of travel times for a given velocity model across an array or network of sensors. In case of a dense array, the curvature of the wavefront is essential and explicitly taken into account. Together with the observed coherent wave front data it is then possible to estimate time

and location of successive radiation sources during a large earthquake (Ishii et al., 2005; Krüger & Ohrnberger, 2005). According to our earthquake catalog (Hillers et al., 2020), the induced earthquakes are located between 24.8286° and 24.8450° East, and between 60.1880° and 60.1957° North. Hence, we define the lateral extension of the search source volume from 24.82° to 24.86° East and from 60.18° to 60.205° North, which corresponds to a 2.8 km by 2.2 km area (Figure 1). We divide this region into cells with 0.0005° and 0.001° spacing in latitude and longitude, respectively. This leads to square grid elements with edge length of about 55 m. For 3D location, we add the depth dimension from 3 km to 8 km depth with 50 m spacing. We perform a grid search and the relative radiation strength from each source cell is computed by stacking the waveform of all stations after shifting the moveout by the synthetic travel times obtained with the above described homogeneous velocity model (Kortström et al., 2018). We implement the back-projection method using a moving time window of 0.3 s length and a 0.05 s time step, and for each time window we compute a linear stack across the array (Kiser & Ishii, 2017)

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

where $s_{j-bp}(t)$ is the stacked time series of N stations, corresponding to the target grid point j, $u_i(t)$ is the seismogram at station i, τ^{ij} is the travel time from source j to station i, and w^{ij} is a weighting factor that is used to adjust the polarity and normalize $u_i(t)$. The back-projection beam power or energy at time t is computed as an integral of $s^2_{j-bp}(t)$ over a time window from t to $t + t_w$, which is 0.3 s in our study. The maximum beam energy across all cells indicates the estimated source location in a given time window. In the time window where the source signal is included, the back-projection beam power is significantly larger compared to the background level. To avoid potentially problematic large 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 with narrow epicentral distance between 2 km and 4.5 km.

6.3. Back-projection synthetic experiments

We use SeisSol (https://github.com/SeisSol/SeisSol) to perform point source seismic wave propagation simulations with varying focal mechanisms, and generate synthetic waveforms at two model arrays for back-projection analysis. SeisSol solutions feature high-order accuracy in space and time. We choose polynomial basis functions of order $p = 3$ leading to a fourth-order accurate scheme. A moment tensor point source is located at the center of the domain and at 6.5 km depth (Figure S1a). We use the same homogeneous velocity model as for the beamforming and back-projection analysis. 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 2 km diameter circle that is centered 8 km away from the epicenter in a $45°$ direction. The full domain is $150 \times 150 \times 100\ km^3$. The spatial

discretization of the statically adaptive, fully unstructured tetrahedral mesh (Figure S1a) is refined to 200~m in the central $20 \times 20 \times 8\ km^3$ target high-resolution area, which completely covers the source and array region. This can resolve frequencies up to 4.5 Hz. For all 3024 point sources, we use the same asymmetric source time function with a duration of 0.15 s (Figure S1b). Figures S1c and S1d show the synthetic waveforms at the FACA and LACA arrays, for an exemplary point source with strike $0°$, dip $90°$, and rake $45°$.

To estimate the swimming direction of the synthetic back-projection results, we use the direction of two points before and after the back-projection location ``swim'' across the source point. We first find the time step with the peak beam power from the time window analysis. The peak beam power is generally associated with the synthetic source location. Then we identify two points, the average locations of 10 time steps of the back-projection location estimate before and after the peak beam power point. A larger time step average will not improve the estimates of the linear swimming direction across the source point, but may introduce some bias by including points at the boundary of the target region (Figure 4c). The azimuth from the average point before the peak beam power to the average point after the peak beam power represents the swimming direction.

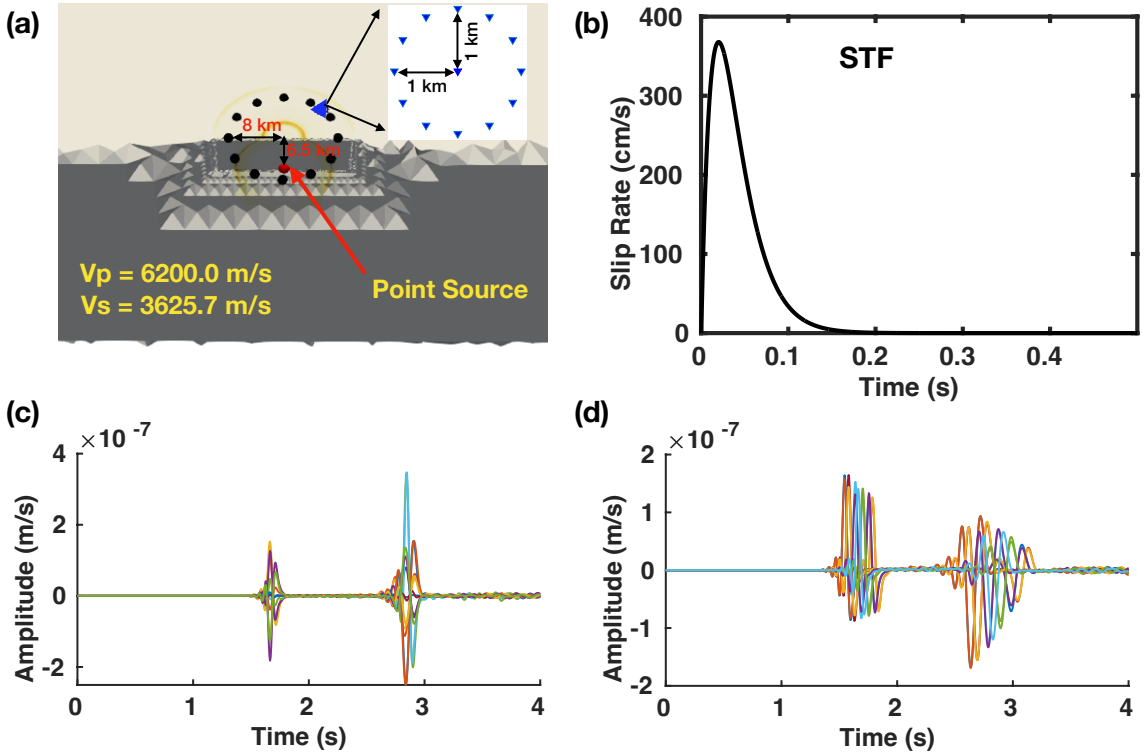

Figure S1. Model setup for the point source synthetic experiments. (a) Point source simulation configuration, with a view of the clipped model domain around the source and arrays. The red circle in the center represents the point source. Black circles indicate

the full azimuth coverage array (FACA) at 8 km epicentral distance. It has an azimuth range from 30° to 360°, with a 30° interval. 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. (b) The source time function (STF) of the point source. (c) Example waveforms of the FACA array for the point source with strike 0°, dip 90°, and rake 45°. (d) As in (c) but for stations of the LACA array.

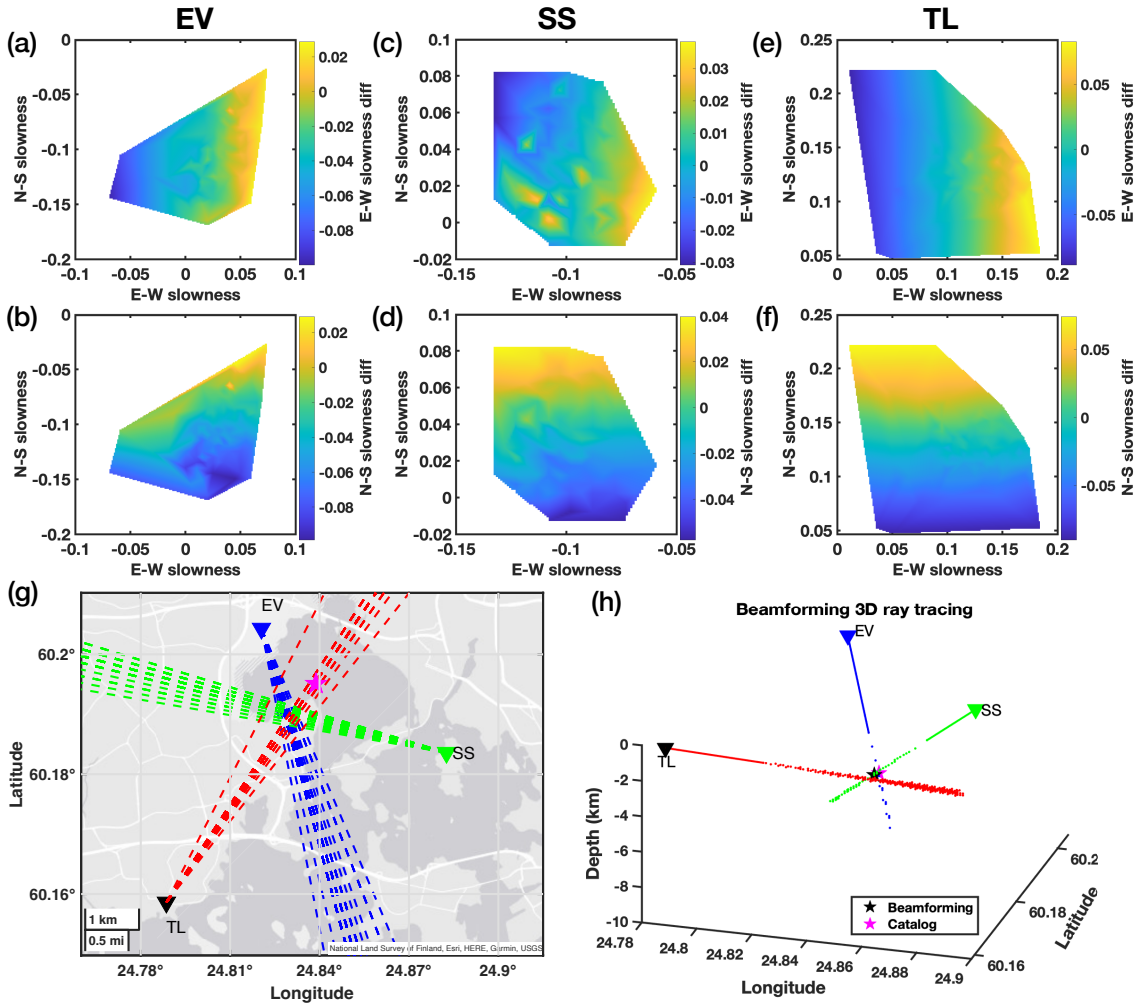

Figure S2. E-W component S wave beamforming results. Panels (a)-(f) show slowness calibration functions or average slowness differences between the beamforming results and the catalog location based slowness estimates. (a)-(b) Calibration functions in the East-West (E-W) and North-South (N-S) direction for the EV mini array. Panels (c)-(d) and (e)-(f) show the corresponding patterns for the SS and TL arrays. (g) Beamforming based back-azimuth ray tracing before calibration. The pink star marks the location of a M_L1.0 event. The dashed lines show the ray tracing results from the bootstrap beamforming. (h)

Beamforming based 3D ray tracing for the same event using the calibrated P wave slowness. The black and pink star mark the beamforming and catalog location.

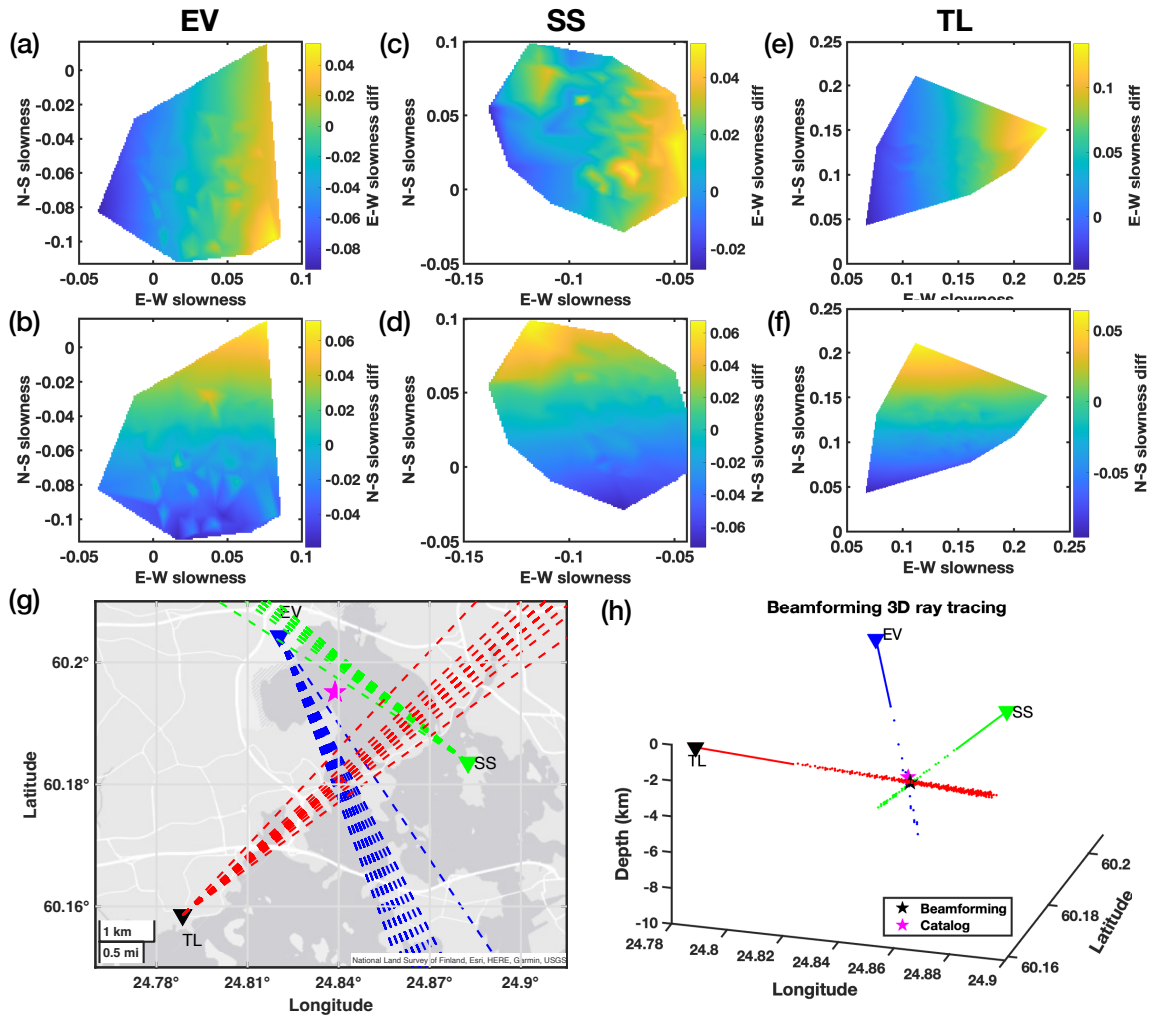

Figure S3. As Figure S2 for N-S component S wave data from the same M_{L}1.0 event.

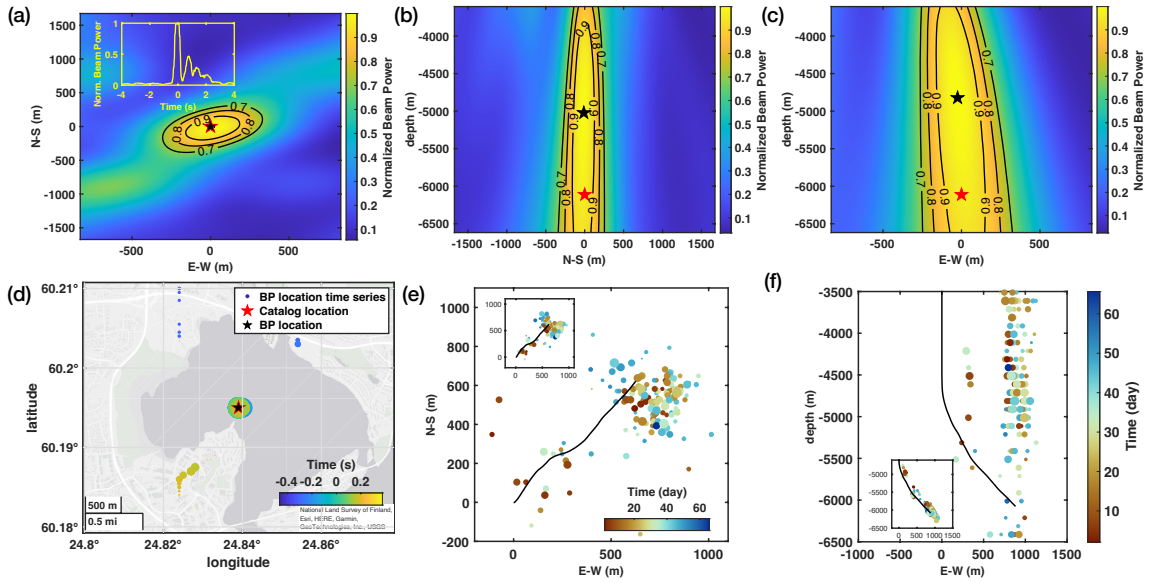

Figure S4. Back-projection example using the same M_L1.0 event as in Figure 2. (a) Horizontal plane back-projection image of the relative energy radiation at the event depth and at the time of peak beam power. The black and red stars mark the back-projection and catalog location. The black contours indicate 70%, 80%, and 90% of the peak beam power. The inset shows the beam power evolution with time. Panels (b) and (c) show the back-projection depth resolution at the peak beam power time step. Back-projection 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 back-projection results of this M_L1.0 event. (e) P wave back-projection epicenters of the catalog events. The black line denotes the borehole trajectory. The inset shows the catalog locations. (f) P wave back-projection hypocenter location estimates. The inset shows the catalog locations.

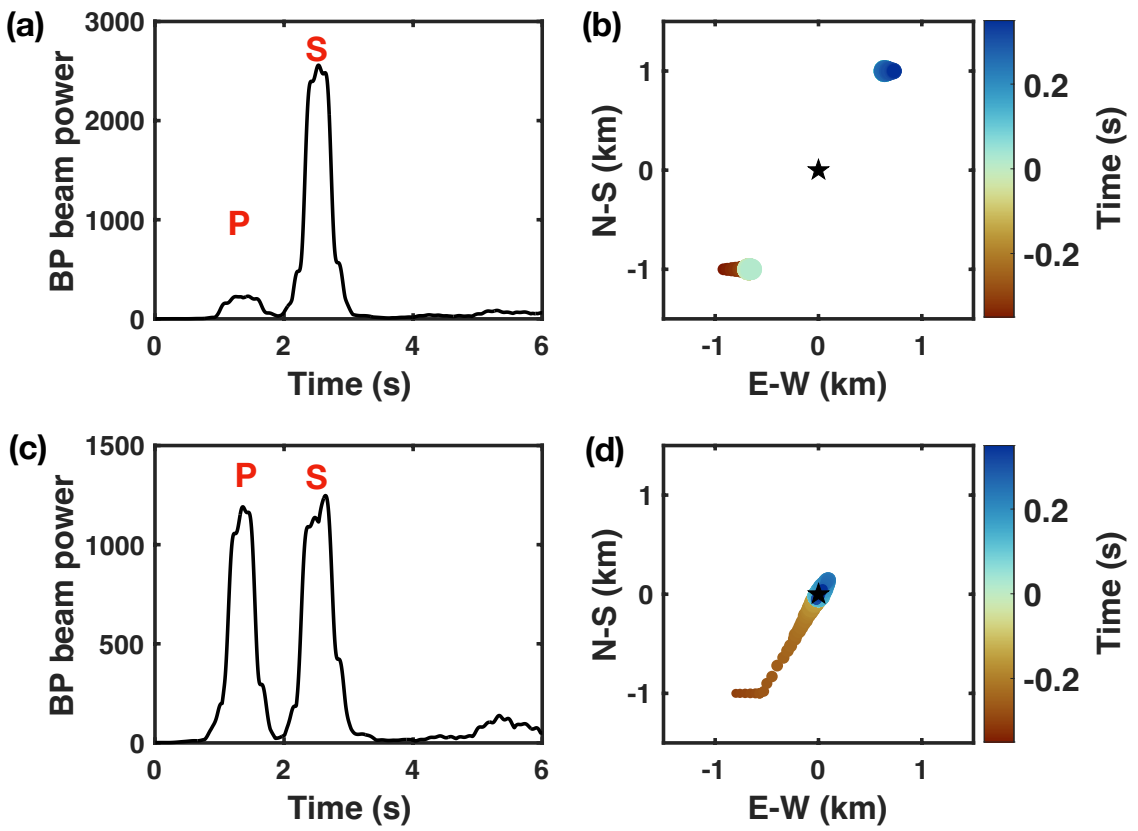

Figure S5. Synthetic back-projection results demonstrating the S wave effect on back-projection locations. (a)-(b) Back-projection results for a point source with strike $30°$, dip $90°$, and rake $45°$. (c)-(d) Back-projection results for a point source with strike $30°$, dip $90°$, and rake $75°$. Panels (a) and (c) show beam power time series. The P and S wave time window is indicated. Panels (b) and (d) show the back-projected source location evolution with time. The black star marks the synthetic source location. The back-projection grid search is performed in a square 4 km^2 area.

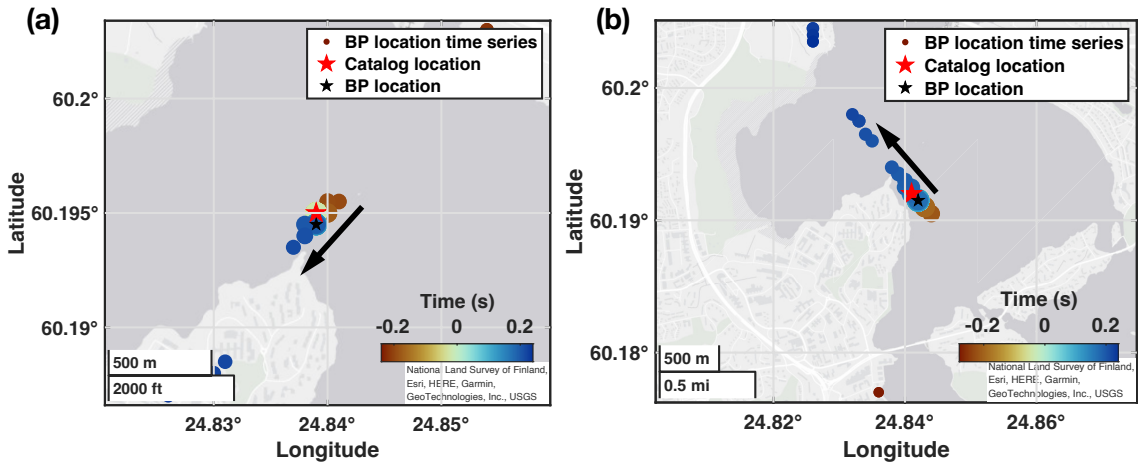

Figure S6. Back-projection swimming pattern examples of two catalog events. (a) Back-projection results of the example M_L1.0 event show migration towards the southwest direction. The circle color indicates the temporal pattern of back-projection locations, and the circle size is proportional to the back-projection energy. (b) Back-projection results of a M_L1.1 event show migration towards the northwest direction.

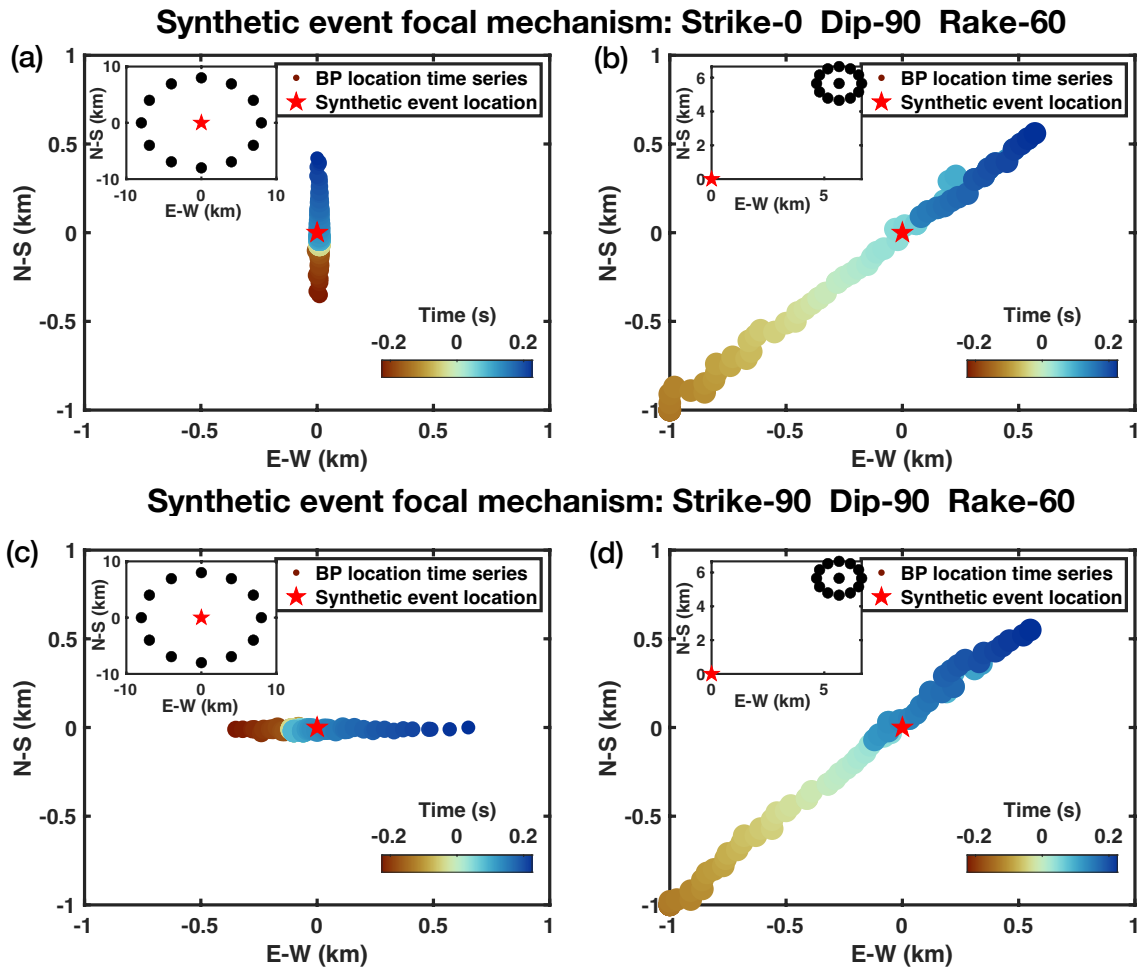

Figure S7. Swimming patterns of synthetic back-projection experiments. (a) Back-projection results of a synthetic event, with focal mechanism strike 0°, dip 90°, and rake 60°, using the full azimuth coverage array (FACA). (b) Back-projection results of the same synthetic event using the limited azimuth coverage array (LACA). Results in panels (c) and (d) correspond to (a) and (b) for a focal mechanism strike strike 90°, dip 90°, and rake 60°.