# 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.

**Key Points:**

- Calibration of systematic beamforming slowness bias improves location estimates of small induced events.
- Combined $P$ and $S$ wave back-projection can produce high resolution 3D source locations.
- Numerical experiments show that back-projection swimming patterns correlate with source mechanisms.

Corresponding author: Bo Li, bli@geophysik.uni-muenchen.de

## Abstract

Seismic 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 $M_{\mathrm{L}}0.0{-}1.8$ earthquakes induced by the 6 km deep 2018 Espoo/Helsinki geothermal stimulation to evaluate the capabilities of beamforming and back-projection array methods. We show that mini array beamforming is sensitive to medium heterogeneity and requires calibration to mitigate systematic slowness biases. A combined $P$ and $S$ wave back-projection approach significantly improves depth resolution, reducing offsets to catalogue locations from ∼1.4 km to ∼91 m. Supported by numerical experiments, we demonstrate that back-projection swimming patterns can constrain focal mechanisms. Our results imply that back-projection of data collected over a wide azimuthal range can be used to monitor and characterize local-scale microseismicity, whereas beamforming calibration requires independently obtained reference observations.


## Plain Language Summary

Small magnitude events generally produce low signal-to-noise ratio (SNR) recordings at a limited number of seismic instruments, which makes them challenging to analyze with traditional seismic methods. Researchers have developed array methods that use records from seismic arrays consisting of tens of stations to improve the SNR and derived source location estimates, but it is difficult to evaluate the performance. The 2018 Helsinki geothermal stimulation in a competent bedrock environment produced high SNR array records of induced microseismicity. In this article, using the high SNR waveforms together with independently obtained reference locations, we show that both the beamforming and the back-projection method can adequately resolve small event locations. For this, we explore a range of novel and previously applied processing schemes to improve the location accuracy, which includes beamforming calibration, and a combined $P$ and $S$ wave back-projection algorithm. In addition to the seismic data analysis, we perform numerical experiments of wave propagation. The results demonstrate that the so-called swimming pattern, which so far has been interpreted to be an artifact associated with the observational configuration, does contain physically meaningful information and can thus be analyzed to constrain properties of the earthquake source mechanisms.

## 1 Introduction

Geothermal energy and with it Enhanced Geothermal Systems (EGSs) can help address the world energy crisis that includes the need to reduce $CO_2$ emissions (e.g., Giovanni et al., 2005; Cuenot et al., 2015; Garcia et al., 2016; Buijze et al., 2019; Rathnaweera et al., 2020). The injection during an EGS stimulation induces or triggers earthquakes, which can ultimately lead to the suspension or termination of a project (e.g., 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. Borehole sensors can yield better quality signals of small magnitude events, but are typically less convenient and cost-efficient to deploy compared to surface seismic arrays. Here we utilize phased surface array instrument data that can also improve the signal-to-noise ratio (SNR) of small event seismograms (Douglas, 2002) to detect, locate, and characterize induced seismic sources.

Array beamforming analyzes the differential travel times of a planar wave front across an array to infer local wave propagation properties (Birtill & Whiteway, 1965; Weichert, 1975; Krüger et al., 1993; Rost & Thomas, 2002). Beamforming has been applied to detect and discriminate events, to study the migration of scattered or reflected waves, and to resolve earth structure (e.g., Wright, 1972; Hedlin et al., 1991; Krüger et al., 1993). More recent applications of small aperture array data demonstrate the improved detection capability of low amplitude signals associated with small earthquakes, tremors, and low frequency earthquakes (e.g., Li & Ghosh, 2017b; Verdon et al., 2017; H. Meng & Ben-Zion, 2018). However, the sensitivity of the high frequency wave field to structural heterogeneity tends to increase the uncertainty of location estimates (e.g. Schweitzer, 2001; Schulte-Pelkum et al., 2003).

Back-projection array processing (Kiser & Ishii, 2017) forms a signal beam to image the earthquake rupture process in sliding time windows (e.g., Ishii et al., 2005; L. Meng, Ampuero, Sladen, & Rendon, 2012; Yin et al., 2016; Li & Ghosh, 2017a; Li et al., 2022). These applications target the arrival of the direct $P$ wave, and ideally large-aperture teleseismic arrays are used to minimize interference from later arriving body and surface waves. Teleseismic back-projection based event detection (Kiser & Ishii, 2013; Li, 2019) can potentially be applied to earthquakes with magnitudes as small as $M3.5$, whereas local configurations can lower this threshold significantly (e.g. Chambers et al., 2010; Vlček et al.,

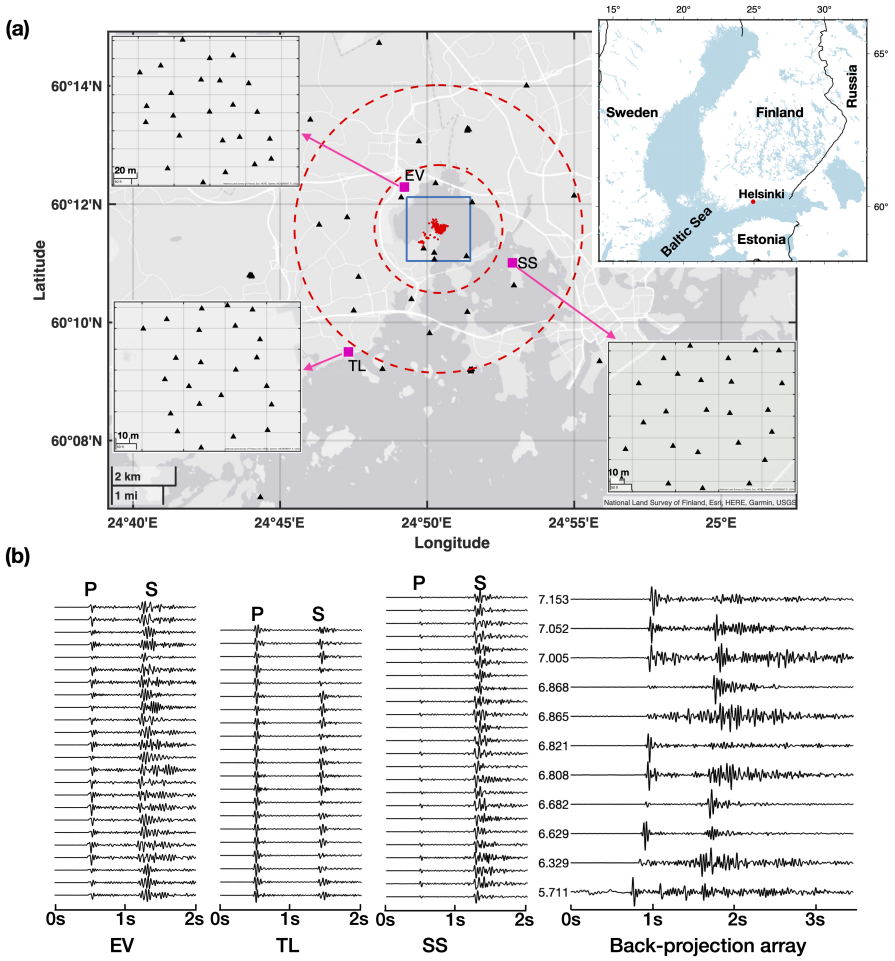


**Figure 1.** (a) Map of the Helsinki study region, larger scale in the top right inset, and station distribution. The red dots represent the catalog earthquake locations. Each black triangle is one seismic station. The pink squares indicate the three mini arrays EV, SS, and TL, that are used for beamforming. The three top left and bottom insets show the configuration of the mini arrays. Stations between the two dashed circles that indicate the 2−4.5 km distance range relative to the event cluster are used for back-projection. The blue box is the back-projection target region. (b) 2−35 Hz filtered high signal-to-noise ratio vertical component seismograms for a 21 July 2018 $M_L$1.0 event, recorded at the three beamforming mini arrays, and at the back-projection array that consists of stations between the two dashed circles in (a). The back-projection stations are sorted by hypocentral distance in kilometers indicated to the left of each waveform.

2016; Beskardes et al., 2018). Microseismicity applications typically require signal enhancement (Inbal et al., 2015) and systematic detection quality assessment (Yang et al., 2021), improvements that are, however, difficult to assess in the absence of well-located reference events.

Here we evaluate the capabilities of the beamforming and back-projection methods for locating and characterizing small earthquakes induced by a geothermal stimulation. 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 $M_{\mathrm{L}}1.8$ were induced in a compact region at around 6 km depth in response to the injection of $\sim$18,000 $\mathrm{m}^3$ fresh water (Kwiatek et al., 2019). The absence of a dissipating sedimentary layer results in high SNR seismograms recorded by a local network consisting of several mini arrays, and stand-alone surface and borehole stations (Figure 1). We use cataloged earthquake data of 204 earthquakes in the $M_{\mathrm{L}}0.0-M_{\mathrm{L}}1.8$ 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).

In this study, we quantify a mini array dependent systematic slowness bias in the associated beamforming patterns in the target 2−35 Hz frequency band, and to calibrate the responses for improved consistency between reference and beamforming based source location estimates (Section 3.1). $P$ wave back-projection using the local array data can generally resolve event epicenters well, but has poor depth resolution and can be strongly affected by the relatively large $S$ wave. We show that combined $P$ and $S$ wave data processing can avoid strong $S$ wave interference and significantly improves the depth resolution (Section 3.2). In Section 3.3 we present results from a data and synthetic waveform analysis that demonstrate the link between properties of source focal mechanisms and back-projection swimming patterns, which have long been assumed to be artifacts.

## 2 Data and Methods

### 2.1 Data

A diverse network including six seismic surface mini arrays, and several surface and borehole stand-alone stations was deployed during the 2018 Helsinki geothermal stim-

ulation (Ader et al., 2020; Hillers et al., 2020). Here we use three component data from 12 borehole stations sampling with 500 Hz, from 5 250 Hz sampling broadband sensors, and from 100 geophones with a 400 Hz sampling rate (Figure 1a) located within a ∼7 km radius around the stimulation wellhead. The horizontal orientation of the borehole sensors is not calibrated and only vertical components are used in back-projection. For the beamforming approach we use data from three square patch or mini arrays deployed at the sites Seurasaari (SS), Elfvik (EV), and Toppelund (TL). They consist of nominally 25 sensors, and the aperture and interstation distance are ∼150 m and ∼25 m. For the back-projection approach we refer to the local array configuration which includes the single borehole and surface stations, and one sensor from each mini array. The techniques' event location estimates are assessed against the locations of 204 manually processed catalog events. The formal error of the reference locations is 20 m, and in longitude, latitude, and depth the error is 11 m, 11 m, and 14 m (Hillers et al., 2020). The velocity model used for these locations and in our analysis consists of the parameters of the 15 km thick topmost layer of the regional model applied in routine analysis (Kortström et al., 2018). It reflects the first-order structural homogeneity of the Fennoscandian Shield environment which results in the observed clean waveforms with ∼1 s $S-P$ times (Figure 1b).

### 2.2 Beamforming

Beamforming assumes a plane wave front propagating across the array (Rost & Thomas, 2002). We perform a grid search in the slowness domain and calculate the beam power at each slowness associated with the $P$ or $S$ wave (Supplementary Section 6.1). The grid element with the maximum beam power indicates the local wave propagation direction relative to the array center. We filter the seismograms between 2−35 Hz and perform beamforming independently for each mini array and for each component. We use the bootstrap method (Efron, 1992) to quantify the slowness uncertainty at each array, and we estimate the event epicenter or hypocenter by 2D or 3D ray tracing the slowness vectors from each array center.

### 2.3 Back-projection

The back-projection method has the signal time shifting and stacking in common with beamforming, but calculates the travel times for a curved wave front based on a given velocity model, and applies a grid search in the source volume around potential

event locations (Ishii et al., 2005; Krüger & Ohrnberger, 2005). In our implementation (Supplementary Section 6.2), the target source volume has a lateral extension of 2.8 km by 2.2 km (Figure 1a) with a 55 m discretization, and a vertical extension that ranges from 3 km to 8 km depth with a 50 m interval. We use a moving time window in the location search, and the maximum beam energy across all cells indicates the source location for a given time window. In the time window where the source signal is included, the back-projection beam power would be much higher than the background level.

Compared to back-projection of large earthquakes using teleseismic waveforms, local array back-projection needs to address two issues. One, uncorrected signal polarity due to unknown source locations, which affects the stack efficiency. Two, not fully separated $P$ and $S$ wave arrival signals, which could lead to false locations, especially when the $S$ wave has larger amplitudes in the majority of the array stations. In addition, local earthquake catalogs require higher accuracy of source locations than global catalogs. One of our motivations here is to improve the poor depth resolution in back-projection.

To account for the aforementioned requirements, we use the signal envelope, and combine both $P$ and $S$ waves in back-projection to improve the source location accuracy and depth resolution. For each time step $t$ that an event could start at, we stack the waveform around the theoretical $P$ and $S$ arrival time window estimated with the velocity model

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

where $s^2_{j-psbp}(t, t+t_w)$ is the stacked beam power for the source grid $j$ in the time window $(t, t+t_w)$, with $t_w$=0.3 s the back-projection stack time window length. The signal envelope of station $i$ is $u_i$, $dt$ is the sampling rate, $w_i$ is the normalization factor for station $i$, $\tau_p^{ij}$ and $\tau_s^{ij}$ are the $P$ and $S$ wave travel time estimates from the source $j$ to a receiver $i$, and $N$ is number of stations. The time and location associated peak value solution represent the event time and location.

### 2.4 Back-projection synthetic experiments

To investigate the potential link between the observed swimming pattern of back-projection results and event focal mechanisms, and the effect of source-array configura-

tions, we perform 3D point source wave propagation simulations using SeisSol (Breuer et al., 2014; Uphoff et al., 2017; Krenz et al., 2021). We simulate the seismic wave field up to ∼4.5 Hz in a volume domain consisting of ∼2.17 million tetrahedral elements. Our mesh is statically refined to an element edge length of 200 m in a 20×20×8 $km^3$ subdomain (Figure S1a) and we use a fourth order accurate scheme in space and time. We vary the strike and rake from 0° to 345° in 15° steps, the dip from 10° to 90° in 10° steps, and the rake from 0° to 195° in 15° steps, which results in 3024 simulations with different focal mechanisms (Supplementary Section 6.3). The computational cost of each point source simulation is ∼130 CPU Hours for 8 seconds simulation time. We use two synthetic surface arrays (Figure S1), the full azimuth coverage array (FACA) and limited azimuth coverage array (LACA), to perform back-projection with synthetic waveforms. More details are given in Supplementary Section 6.3.

## 3 Results

### 3.1 Beamforming with mini arrays

For each mini array, we observe a systematic slowness bias in the beamforming results (Figures 2a-f), which leads to erroneous back-azimuth estimates away from the source direction. As a result, the back-azimuth traces from each array center do not converge and fail to properly constrain source locations (Figure 2g). This is compatible with previous propagation direction estimates based on the same array data (Hillers et al., 2020; Taylor et al., 2021), and is likely due to structural heterogeneity in the medium not accounted for by the velocity model, including effects of a shallow, few tens of meters thick, low velocity layer (Hillers et al., 2020). We use the reference locations to correct for the slowness bias. We first choose 200 of the 204 event locations to demonstrate the estimation of a robust calibration function that covers the compact earthquake source region, and then apply this calibration function to the remaining four events. For each mini array, we first perform beamforming to obtain a slowness estimate of each selected event. These estimates are averaged across bins of 0.005 s/km width along the East-West and North-South direction, and the average difference relative to the theoretical slowness based on catalog locations and the velocity model is the target slowness calibration vector or function.

Figures 2a-f show the resulting calibration function of the $P$ wave slowness bias pattern for the three mini arrays, with color-indicated differences from the catalog based slowness. The similar apertures 144 m, 119 m, 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 based results have the smallest bias with a maximum 0.03 s/km bias in both East-West and North-South directions. The SS array shows the highest deviation of up to 0.25 s/km. The slowness bias of SS array significantly reduces to less than 0.07 s/km in the beamforming results using the horizontal components (Figures S2, S3). The back-azimuth of EV and SS mostly points to the same quadrant where the source is, whereas some TL results directs to another quadrant, resulting in a far away partial convergence.

Again for demonstration, 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 ray tracing before calibration for an example 21 July 2018 $M_{\mathrm{L}}$1.0 event. Clearly, the solutions do not converge on the pink star catalog location. After calibration, the back-azimuth ray tracing from three array centers converges much better around the reference solution, even in 3D (Figure 2h). For this example event, the multi-array $P$ wave beamforming locates the source 87 m and 110 m off the catalog location in the horizontal and depth direction, respectively. Corresponding results based on horizontal component $S$ wave data shows 233 m horizontal and 220 m vertical off the catalog using the East-West component (Figure S2), and 200 m horizontal and 30 m vertical off the catalog using the North-South component (Figure S3).

### 3.2 Back-projection using the local array

#### *3.2.1 Back-projection using $P$ waves*

The epicenter estimate of the example $M_{\mathrm{L}}$1.0 event obtained with $P$ wave back-projection using the local array consisting of stand-alone sensors is highly accurate (Figure S4a). Source locations are generally characterized by good horizontal (Figure S4a) but poor vertical (Figures S4b, c) resolution. The typical horizontal offset between back-projection and catalog locations (Figure S4e) is better than the 97.5 m average offset between individual solutions that is dominated by few poor back-projection results of small-

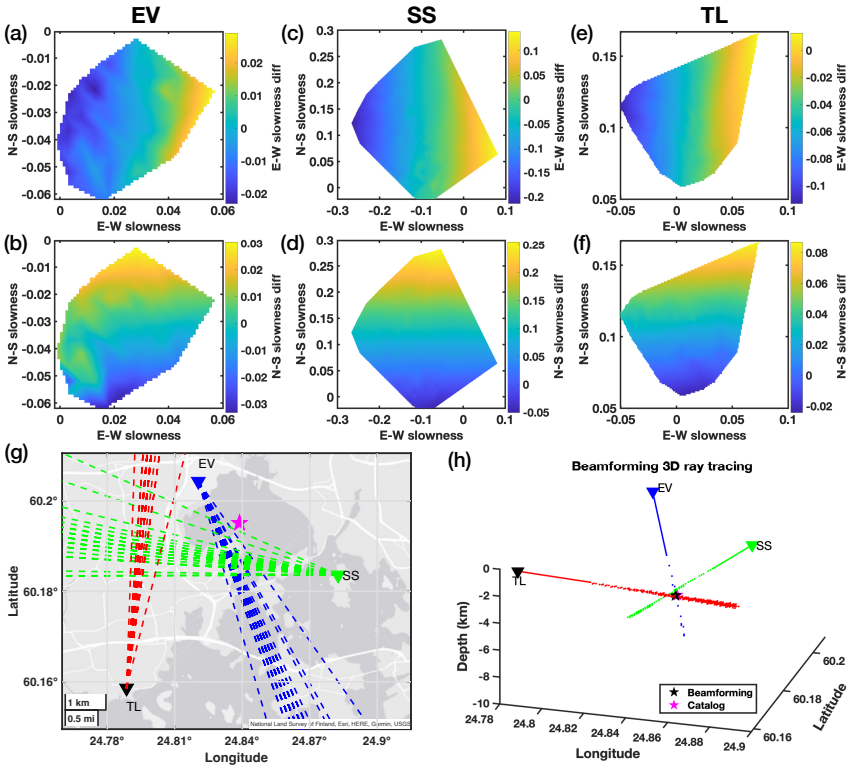


**Figure 2.** Vertical component $P$ wave beamforming results. (a)−(f) Slowness calibration functions or average slowness differences between the beamforming results and the catalog location based slowness estimates. (a)−(b) Calibration functions in the East-West (E-W) and North-South (N-S) direction for the EV mini array. (c)−(d), (e)−(f) 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_L$1.0 event. The dashed lines show the bootstrap ray tracing results. (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.

magnitude events. In contrast, back-projection location cross sections at the peak beam power time step (Figures S4b, c) together with the obtained location distribution (Figure S4f) indicate an overall poor depth resolution. The back-projection tends to locate the sources more shallow compared to the reference, with a significant average offset of ∼1.4 km.

#### *3.2.2 Strong $S$ wave effects*

Analysis of the synthetic SeisSol waveform data shows that $S$ waves can have a strong effect on back-projection results and can lead to large location errors. Comparing the results for two point sources with the same strike and dip but two different rake angles of 45° and 75° (Figure S5) shows that the 45° rake model generates much larger $S$ wave amplitudes at most model stations (Figure S5a). Importantly, this influences the coherent stack of $P$ waves and results in poor location estimates (Figure S5b). In contrast, the rake 75° model yields $P$ and $S$ wave time window stacks that have comparable beam power (Figure S5c). In this case, back-projection images radiation sources that seem to migrate in the northeast direction, but with peak beam power located at the synthetic source location (Figure S5d). This $S$ wave effect also explains some of the $P$ wave based locations in Figure S4e that show low consistency with the catalog reference solutions.

#### *3.2.3 Back-projection with $P$ and $S$ waves*

To improve the poor depth resolution (Figures S4b, c), we combine $P$ and $S$ wave information using equation 1. The combination of $P$ and $S$ waves also mitigates the disadvantageous $S$ wave effect on the back-projection results. Our combined approach leads to much improved agreement between synthetic and estimated point source locations (Figure 3) and improved depth-constraints. Figure 3d collects the 204 3D event locations using this approach. Using the mean offset between the back-projection and catalog locations of 10 random events to correct the average system error due to velocity model uncertainty, we quantify that the calibrated back-projection locations have an average difference of ∼74.9 m in horizontal and ∼91.3 m in vertical direction to the reference estimates.

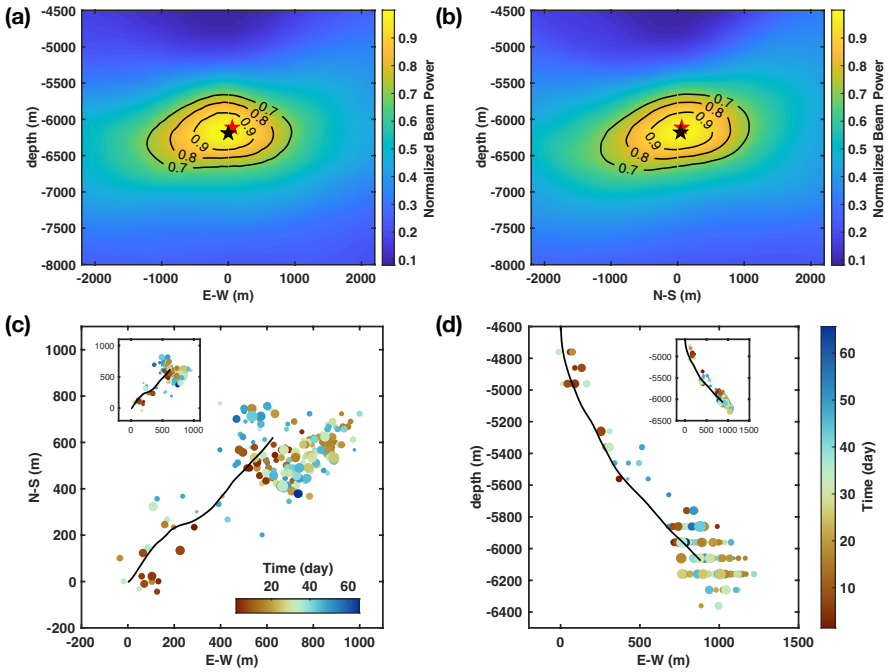


**Figure 3.** $P$ and $S$ wave back-projection results. Panels (a)−(b) show the depth resolution example using the same $M_{\mathrm{L}}1.0$ event as in Figure 2. Back-projection image of the relative energy radiation in the (a) East-West-Vertical plane at the event latitude, and in the (b) North-South-Vertical plane at the event longitude, at the time step of the peak beam power. The black and red stars mark the back-projection and catalog location. The black contours indicate percentages of the peak beam power. $P$ and $S$ wave back-projection (c) epicenters and (d) hypocenters of the catalog events. The insets show the catalog locations. The black line denotes the borehole trajectory. The circle size is proportional to magnitude.

### 3.3 The back-projection swimming artifacts and focal mechanisms

Our back-projection application produces so-called swimming migrations with time. This swimming artifact (Koper et al., 2012; L. Meng, Ampuero, Luo, et al., 2012) or shooting star artifact (Beskardes et al., 2018) is the manifestation of the array response function of a non-stationary signal, which decays as a function of time, across the array. Previous observations of the swimming migration have been attributed to limited azimuthal coverage of the array relative to the source. Here, however, we observe swimming migration using an array with good azimuthal coverage, as shown in the time lapse back-projection results of two $M_{\mathrm{L}}1.0$ and $M_{\mathrm{L}}1.1$ events in Figures S6a and S6b, where the sources appear to move along a linear trajectory towards southwest or northwest, each across the catalog source location. We note that Shirzad et al. (2020) propose that 3D back-projection images of a larger $M_{\mathrm{w}}4.1$ event can constrain its source focal mechanism.

Using synthetic numerical experiments we are able to systematically analyze the effects of source focal mechanisms and source-array configurations on back-projection swimming patterns. Back-projection results of two point sources with focal mechanisms (Figure S7) that differ in their 0° and 90° strike, but have the same dip 90° and rake 60°, show that the full azimuth coverage array (FACA) data yield source-dependent migration patterns along the strike orientation, whereas the limited azimuth coverage array (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, and that source effects can only be resolved with good azimuthal FACA-type coverage.

Of course, the dip and rake angles can also affect radiation and thus the swimming pattern. To investigate the full relation between source focal mechanisms and back-projection swimming patterns, we perform $P$ wave only and $S$ wave only back-projection. Figures 4a and 4b collect the azimuthal swimming direction using the FACA array for all simulated 3024 point sources (Supplementary Section 6.3). The results show that the swimming directions vary systematically with the change of the source focal mechanism, but the pattern differs for $P$ and $S$ waves. Note that for certain focal mechanisms with dip 90°, and rake 0°, 90°, 180° and 270°, the radiation pattern is symmetric across the source, resulting in no swimming patterns. We randomly generate three point sources (Figure 4c, gray-yellow beach balls) and back-project the associated synthetic waveforms to estimate

their $P$ and $S$ wave swimming directions. We compare these estimates to interpolated results from the swimming direction data base in Figures 4a and 4b. The different $P$ and $S$ wave sensitivities together fully constrain the focal mechanisms, and the obtained minimum residual solutions (Figure 4c, gray-red beach balls) are excellent estimates of the true focal mechanisms used for all three point sources. This is an important implication for constraining the source properties of small magnitude events for which focal mechanisms are challenging to estimate reliably.

## 4 Discussion

In beamforming applications, the systematic slowness bias associated with medium heterogeneity and in particular with heterogeneity directly below the array (Berteussen, 1976; Schweitzer, 2001; Schulte-Pelkum et al., 2003) can lead to poor estimates of the source-receiver direction and hence of the source location. Earlier attempts to correct or mitigate this bias include the Slowness and Azimuth Station Corrections (Bondár et al., 1999), empirical matched field processing (Harris & Kvaerna, 2010), and the limited sensor-pair correlation (Gibbons et al., 2018). For locating induced microseismicity we here use highly coherent mini array waveforms without signal polarity variation, that have also shown strong capability in detecting low amplitude signals (e.g., Li & Ghosh, 2017b; Verdon et al., 2017; H. Meng & Ben-Zion, 2018). Beamforming 204 induced events in the $M_{\mathrm{L}}0.0{-}M_{\mathrm{L}}1.8$ range using 2−35 Hz data from three mini arrays 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. The calibration in this study demonstrates that a well located catalog that densely spreads over the target region, can significantly reduce the systematic slowness bias and improve 3D location estimates using beamforming-based ray-tracing.

The good azimuthal coverage of the local array supports our back-projection application. The resulting location estimates are generally less sensitive to the medium heterogeneity. In our study, $P$ wave back-projection 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 back-projection (Kiser et al., 2011; Yagi et al., 2012). In our local situation we similarly show that a combined

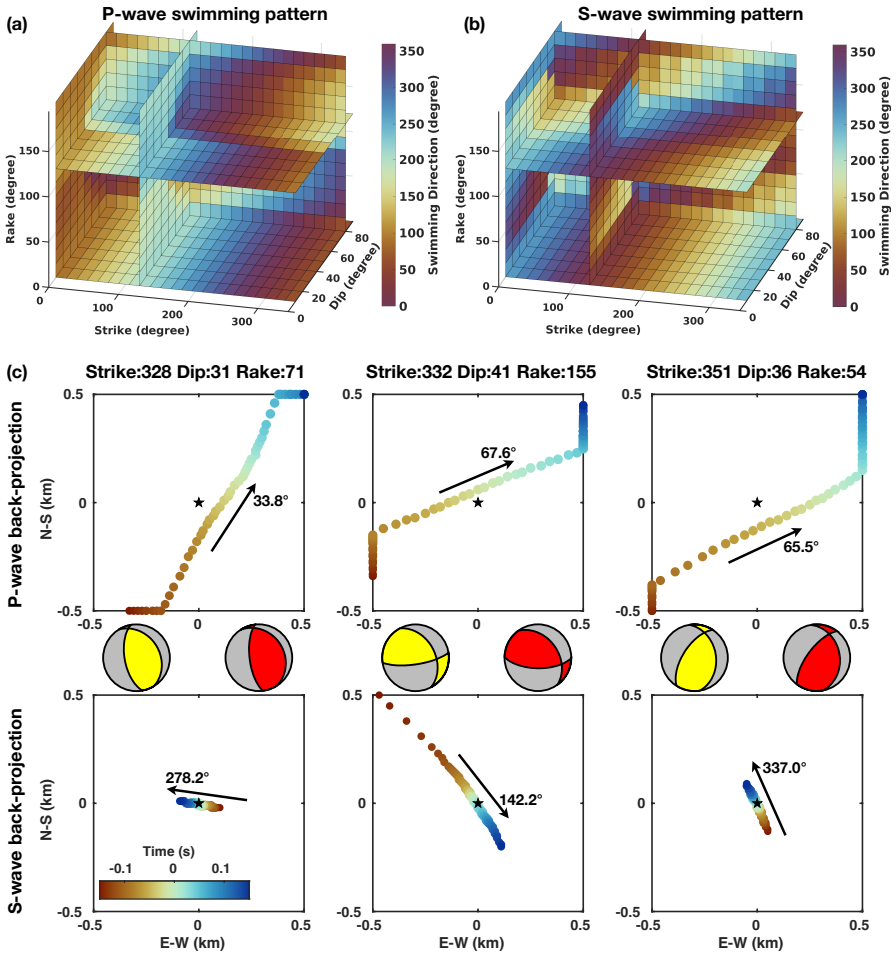


**Figure 4.** Synthetic experiment results of back-projection swimming patterns and the application to resolve source focal mechanisms. Panels (a) and (b) show the $P$ and $S$ wave swimming patterns of 3024 synthetic point simulations for various focal mechanisms, using the full azimuth coverage array (FACA). (c) We use the established swimming patterns to infer focal mechanisms of three random point sources. The black arrows denote the swimming direction for the $P$ and $S$ wave back-projection, with the direction relative to the north labeled at the arrowhead. The gray-yellow and gray-red beach-balls represent the actual and inferred source focal mechanism, respectively.

$P$ and $S$ wave analysis significantly improves the depth resolution for overall well constrained hypocenter estimates.

Swimming artifacts are commonly observed in teleseismic (Koper et al., 2012; L. Meng et al., 2011) and local back-projection (Beskardes et al., 2018). The respective source-array configurations govern 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 a limited azimuth array demonstrates clear towards-array migration (Figure S7). However, our systematic synthetic back-projection results using a full azimuth coverage array imply that the swimming patterns are not just an artifact of the array response function but contain information about the source properties. The $P$ wave (Figure 4a) or $S$ wave (Figure 4b) based swimming artifacts are individually not sufficient to uniquely constrain all three parameters, 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) yield similar $P$ wave swimming directions, however, they uniquely differ in their $S$ wave patterns (Supplementary Section 6.3).

The short distances and high wave speeds in our local study imply that the $S$ wave window can not always be fully separated from the $P$ wave window. If $S$ wave signals are overprinting the preferred $P$ wave stack across the array, this can lead to false detections and erroneous location estimates (Figure S5) when grid searching across the full back-projection target region at the same time step. The radiation pattern, velocity structure and source-array distance governs the relative $P$ and $S$ wave amplitudes. When the $S$ wave amplitude exceeds the $P$ wave amplitude at most array stations, stacking normalized waveforms results in incorrectly increased beam power controlled by incoherent $S$ waves. 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 back-projection to $S$ waves. Here we show that combined $P$ and $S$ waves back-projection (Equation 1) can avoid strong $S$ wave effects and therefore improves 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 back-projection show an average 3D offset of $\sim$140 m and $\sim$118 m compared to the reference catalog locations (Section 2.1). 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). Thus, our here reported offsets may be further reduced by using a better velocity model together with refined search grids.

## 5 Conclusions

We investigate the capabilities of the two widely applied beamforming and back-projection array techniques using high-quality seismic data from 204 $M_{\mathrm{L}}0.0-M_{\mathrm{L}}1.8$ 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 3D location estimates, in both $P$ and $S$ wave beamforming. $P$ wave back-projection results obtained with an azimuthally well distributed single station configuration show very good horizontal but comparatively poor depth resolution. A strong $S/P$ amplitude ratio can lead to back-projection locations away from the true sources. We establish that a $P$ and $S$ wave combined back-projection approach can mitigate strong phase effects. This significantly reduces the average depth offsets to the reference catalogue locations from $\sim$1.4 km to $\sim$91 m. Numerical experiments demonstrate that the back-projection swimming pattern is not just an artifact of the array response function, but it can constrain earthquake focal mechanism parameters if the back-projection array has a good azimuthal coverage. In summary, our results have important implications for monitoring and characterizing the abundance of small-magnitude seismicity around Enhanced Geothermal Systems and in natural environments using seismic array techniques.

## Acknowledgments

The 2018 temporary deployment was supported by the Geophysical Instrument Pool Potsdam under grant 201802. This work received funding from the European Union's Horizon 2020 research and innovation programme (TEAR ERC Starting grant agreement No. 852992) and Horizon Europe (DT-Geo grant agreement No. 101058129, Geo-Inquire grant agreement No. 101058518, ChEESE-2P grant agreement No. 101093038), and King Abdullah University of Science and Technology (KAUST, Grant BAS/1/1339-01-01). The authors acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu, project pr63qo) for funding this project by providing computing time on the GCS Su-

percomputer 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).

**Open Research**

Waveform data from the induced earthquakes are available from Vuorinen et al. (2023) (https://doi.org/10.23729/39cfac4f-4d0d-4fb4-83dc-6f67e8ba8dce). A supplementary file to Taylor et al. (2021) (`https://doi.org/10.1029/2020GL090403`) contains information about the 204 events used in this study. We use the open-source software package SeisSol (`https://www.seissol.org`) to simulate the point source wave fields. SeisSol is open-source and freely available from `https://github.com/SeisSol/SeisSol`. Documentation for downloading, compiling, and using SeisSol is provided at `https://seissol.readthedocs.io`. We provide a quick start docker container and jupyter training notebooks at `https://github.com/SeisSol/Training`. Example problems and benchmark model configuration files are provided at `https://github.com/SeisSol/Examples`. The point source synthetic simulation data can be accessed through Zenodo at `https://doi.org/10.5281/zenodo.7541076`.

## References

Ader, T., Chendorain, M., Free, M., Saarno, T., Heikkinen, P., Malin, P. E., ... Vuorinen, T. (2020). Design and implementation of a traffic light system for deep geothermal well stimulation in Finland. *Journal of Seismology*, *24*(5), 991–1014.

Berteussen, K. (1976). The origin of slowness and azimuth anomalies at large arrays. *Bulletin of the Seismological Society of America*, *66*(3), 719–741.

Beskardes, G., Hole, J., Wang, K., Michaelides, M., Wu, Q., Chapman, M., ... Quiros, D. (2018). A comparison of earthquake backprojection imaging methods for dense local arrays. *Geophysical Journal International*, *212*(3), 1986–2002.

Birtill, J., & Whiteway, F. (1965). The application of phased arrays to the analysis of seismic body waves. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *258*(1091), 421–493.

Bondár, I., North, R. G., & Beall, G. (1999). Teleseismic slowness-azimuth station corrections for the International Monitoring System seismic network. *Bulletin of the Seismological Society of America*, *89*(4), 989–1003.

Breuer, A., Heinecke, A., Rettenberger, S., Bader, M., Gabriel, A.-A., & Pelties, C.

(2014). Sustained petascale performance of seismic simulations with SeisSol on SuperMUC. In *International Supercomputing Conference* (pp. 1–18).

Buijze, L., van Bijsterveldt, L., Cremer, H., Paap, B., Veldkamp, H., Wassing, B. B., ... Jaarsma, B. (2019). Review of induced seismicity in geothermal systems worldwide and implications for geothermal systems in the Netherlands. *Netherlands Journal of Geosciences*, *98*.

Chambers, K., Kendall, J.-M., Brandsberg-Dahl, S., & Rueda, J. (2010). Testing the ability of surface arrays to monitor microseismic activity. *Geophysical Prospecting*, *58*(5), 821–830.

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

Diehl, T., Kraft, T., Kissling, E., & Wiemer, S. (2017). The induced earthquake sequence related to the St. Gallen deep geothermal project (Switzerland): Fault reactivation and fluid interactions imaged by microseismicity. *Journal of Geophysical Research: Solid Earth*, *122*(9), 7272–7290.

Douglas, A. (2002). Seismometer arrays - their use in earthquake and test ban seismology. *International Handbook of Earthquake & Engineering Seismology, Part A*, 357.

Efron, B. (1992). Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics* (pp. 569–593). Springer.

Garcia, J., Hartline, C., Walters, M., Wright, M., Rutqvist, J., Dobson, P. F., & Jeanne, P. (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*(7275), 848–849.

Gibbons, S. J., Näsholm, S. P., Ruigrok, E., & Kværna, T. (2018). Improving slowness estimate stability and visualization using limited sensor pair correlation on seismic arrays. *Geophysical Journal International*, *213*(1), 447–460.

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

Harris, D. B., & Kvaerna, T. (2010). Superresolution with seismic arrays using empirical matched field processing. *Geophysical Journal International*, *182*(3), 1455–1477.

Hedlin, M. A., Minster, J. B., & Orcutt, J. A. (1991). Beam-stack imaging using a small aperture array. *Geophysical Research Letters*, *18*(9), 1771–1774.

Hillers, G., T. Vuorinen, T. A., Uski, M. R., Kortström, J. T., Mäntyniemi, P. B., Tiira, T., ... Saarno, T. (2020). The 2018 geothermal reservoir stimulation in Espoo/Helsinki, Southern Finland: Seismic network anatomy and data features. *Seismological Research Letters*, *91*(2A), 770–786.

Inbal, A., Clayton, R. W., & Ampuero, J.-P. (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. *Geophysical Research Letters*, *42*(15), 6314–6323.

Ishii, M., Shearer, P. M., Houston, H., & Vidale, J. E. (2005). Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array. *Nature*, *435*(7044), 933–936.

Ishii, M., Shearer, P. M., Houston, H., & Vidale, J. E. (2007). Teleseismic P wave imaging of the 26 December 2004 Sumatra-Andaman and 28 March 2005 Sumatra earthquake ruptures using the Hi-net array. *Journal of Geophysical Research: Solid Earth*, *112*(B11).

Kiser, E., & Ishii, M. (2013). Hidden aftershocks of the 2011 Mw 9.0 Tohoku, Japan earthquake imaged with the backprojection method. *Journal of Geophysical Research: Solid Earth*, *118*(10), 5564–5576.

Kiser, E., & Ishii, M. (2017). Back-projection imaging of earthquakes. *Annual Review of Earth and Planetary Sciences*, *45*, 271–299.

Kiser, E., Ishii, M., Langmuir, C. H., Shearer, P., & Hirose, H. (2011). Insights into the mechanism of intermediate-depth earthquakes from source properties as imaged by back projection of multiple seismic phases. *Journal of Geophysical Research: Solid Earth*, *116*(B6).

Koper, K. D., Hutko, A. R., Lay, T., & Sufri, O. (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. *Journal of Geophysical Research: Solid Earth*, *117*(B2).

Kortström, J., Uski, M., & Oinonen, K. (2018). The Finnish National Seismic Network. *Summary of the Bulletin of the International Seismological Centre*, *52*(I), 41–52.

Krenz, L., Uphoff, C., Ulrich, T., Gabriel, A.-A., Abrahams, L. S., Dunham, E. M., & Bader, M. (2021). 3D acoustic-elastic coupling with gravity: the dynamics of the 2018 Palu, Sulawesi earthquake and tsunami. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (pp. 1–14).

Krüger, F., & Ohrnberger, M. (2005). Tracking the rupture of the Mw= 9.3 Sumatra earthquake over 1,150 km at teleseismic distance. *Nature*, *435*(7044), 937–939.

Krüger, F., Weber, M., Scherbaum, F., & Schlittenhardt, J. (1993). Double beam analysis of anomalies in the core-mantle boundary region. *Geophysical research letters*, *20*(14), 1475–1478.

Kwiatek, G., Saarno, T., Ader, T., Bluemle, F., Bohnhoff, M., Chendorain, M., ... others (2019). Controlling fluid-induced seismicity during a 6.1-km-deep geothermal stimulation in Finland. *Science Advances*, *5*(5), eaav7224.

Li, B. (2019). *A broad spectrum of fault behaviors in fast and slow earthquakes*. University of California, Riverside.

Li, B., & Ghosh, A. (2017a). Imaging rupture process of the 2015 Mw 8.3 Illapel earthquake using the US Seismic Array. In *The Chile-2015 (Illapel) earthquake and tsunami* (pp. 33–43). Springer.

Li, B., & Ghosh, A. (2017b). Near-continuous tremor and low-frequency earthquake activities in the Alaska-Aleutian subduction zone revealed by a mini seismic array. *Geophysical Research Letters*, *44*(11), 5427–5435.

Li, B., Wu, B., Bao, H., Oglesby, D. D., Ghosh, A., Gabriel, A.-A., ... Chu, R. (2022). Rupture heterogeneity and directivity effects in back-projection analysis. *Journal of Geophysical Research: Solid Earth*, *127*(3), e2021JB022663.

Meng, H., & Ben-Zion, Y. (2018). Detection of small earthquakes with dense array data: Example from the San Jacinto fault zone, southern California. *Geophysical Journal International*, *212*(1), 442–457.

Meng, L., Ampuero, J.-P., Luo, Y., Wu, W., & Ni, S. (2012). Mitigating artifacts in back-projection source imaging with implications for frequency-dependent proper-

ties of the Tohoku-oki earthquake. *Earth, Planets and Space*, *64*(12), 1101–1109.

Meng, L., Ampuero, J.-P., Sladen, A., & Rendon, H. (2012). High-resolution back-projection at regional distance: Application to the haiti M7. 0 earthquake and comparisons with finite source studies. *Journal of Geophysical Research: Solid Earth*, *117*(B4).

Meng, L., Inbal, A., & Ampuero, J.-P. (2011). A window into the complexity of the dynamic rupture of the 2011 Mw 9 Tohoku-oki earthquake. *Geophysical Research Letters*, *38*(7).

Oeser, J., Bunge, H.-P., & Mohr, M. (2006). Cluster design in the Earth Sciences tethys. In *International Conference on High Performance Computing and Communications* (pp. 31–40).

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

Rost, S., & Thomas, C. (2002). Array seismology: Methods and applications. *Reviews of Geophysics*, *40*(3), 2–1.

Schulte-Pelkum, V., Vernon, F. L., & Eakins, J. (2003). Large teleseismic P wavefront deflections observed with broadband arrays. *Bulletin of the Seismological Society of America*, *93*(2), 747–756.

Schweitzer, J. (2001). Slowness corrections—one way to improve IDC products. *Pure and Applied Geophysics*, *158*(1), 375–396.

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

Shirzad, T., Lasocki, S., & Orlecka-Sikora, B. (2020). Tracking the development of seismic fracture network by considering the fault rupture method. In *EGU General Assembly Conference Abstracts* (p. 13079).

Taylor, G., Hillers, G., & Vuorinen, T. (2021). Using array-derived rotational motion to obtain local wave propagation properties from earthquakes induced by the 2018 geothermal stimulation in Finland. *Geophysical Research Letters*, *48*(6), e2020GL090403.

Uphoff, C., Rettenberger, S., Bader, M., Madden, E. H., Ulrich, T., Wollherr, S., & Gabriel, A.-A. (2017). Extreme scale multi-physics simulations of the tsunami-

genic 2004 sumatra megathrust earthquake. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (pp. 1–16).

Verdon, J. P., Kendall, J.-M., Hicks, S. P., & Hill, P. (2017). Using beamforming to maximise the detection capability of small, sparse seismometer arrays deployed to monitor oil field activities. *Geophysical Prospecting*, *65*(6), 1582–1596.

Vlček, J., Fischer, T., & Vilhelm, J. (2016). Back-projection stacking of P-and S-waves to determine location and focal mechanism of microseismic events recorded by a surface array. *Geophysical Prospecting*, *64*(6), 1428–1440.

Vuorinen, T. A., Veikkolainen, T., Oinonen, K., Hällsten, J., Uski, M., Heikkilä, T., ... Tiira, T. (2023, 1). *ISUH waveform, time and location data products from stimulations of deep geothermal wells in espoo in 2018.* `https://doi.org/10.23729/39cfac4f-4d0d-4fb4-83dc-6f67e8ba8dce`.

Weichert, D. (1975). Reduced false alarm rates in seismic array detection by non-linear beamforming. *Geophysical Research Letters*, *2*(4), 121–123.

Wright, C. (1972). Array studies of seismic waves arriving between P and PP in the distance range 90 to 115. *Bulletin of the Seismological Society of America*, *62*(1), 385–400.

Xu, Y., Koper, K. D., Sufri, O., Zhu, L., & Hutko, A. R. (2009). Rupture imaging of the Mw 7.9 12 May 2008 Wenchuan earthquake from back projection of teleseismic P waves. *Geochemistry, Geophysics, Geosystems*, *10*(4).

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

Yang, L., Liu, X., & Beroza, G. C. (2021). Revisiting evidence for widespread seismicity in the upper mantle under Los Angeles. *Science Advances*, *7*(4), eabf2862.

Yin, J., Yang, H., Yao, H., & Weng, H. (2016). Coseismic radiation and stress drop during the 2015 Mw 8.3 Illapel, Chile megathrust earthquake. *Geophysical Research Letters*, *43*(4), 1520–1528.