Earthquake Detectability and Depth Resolution with Dense Arrays in Long Beach, California: Further Evidence for Upper-Mantle Seismicity within a Continental Setting

A. Inbal[1], J.-P. Ampuero[2], R. Clayton[3]

1. Department of Geophysics, Tel Aviv University, Tel Aviv, Israel

2. Université Côte d'Azur, IRD, CNRS, Observatoire de la Côte d'Azur, Géoazur, Valbonne, France

3. Seismological Laboratory, California Institute of Technology, Pasadena, CA, USA

Declaration of Competing Interests

The authors acknowledge there are no conflicts of interest recorded.

Corresponding Author

Asaf Inbal, Department of Geophysics, Tel Aviv University, Tel Aviv, Israel, Email: asafinbal@tauex.tau.ac.il

Abstract

The Newport-Inglewood Fault (NIF) is a slowly-deforming fault cutting through a thin continental crust with a normal geotherm, yet it hosts some of the deepest earthquakes in southern California. The nucleation of deep earthquakes in such a continental setting is not well understood. Moreover, the deep seismogenic zone implies the maximum NIF earthquake magnitude may be larger than expected. Here, we quantify the resolution of the Long-Beach (LB) and the Extended Long-Beach (ELB) dense arrays, used to study deep NIF seismicity. Previous study of the regional catalog and of downward-continued LB array data found NIF seismicity extending into the upper mantle beneath LB. Later studies, which analyzed the ELB raw data, found little evidence for such deep events. To resolve this inconsistency, we quantify the array's microearthquake detectability and resolution power via analysis of pre- and post-downward migrated LB seismograms, and benchmark tests. Downward migration focuses energy onto the source region and de-amplifies the surface noise, thus significantly improving detectability and resolution. The detectability is also improved with the increase in the array-aperture-to-source-depth ratio. The LB array maximum aperture is only 20% larger than the ELB aperture, yet its resolution for deep (>20 km) events is improved by about a factor of two, suggesting that small changes to the array geometry may yield significant improvement to the resolution power. Assuming a constant aperture, we find the LB array maintain resolution with 1% of its sensors used for back-projection. However, the high sensor density is essential for improving the SNR. Analysis of the regional and array-derived NIF catalogs together with newly acquired Moho depths beneath the NIF, suggests mantle seismicity beneath LB is a robust feature of this fault.

Introduction

Seismicity occurring within urban environments is difficult to characterize due to high levels of anthropogenic noise. For example, the Los Angeles (LA) basin, which is the densest population center in southern California, suffers from earthquake detectability that is far lower than the detectability in less-well instrumented regions. Dense array seismology, a methodology which utilizes finely sampled wavefields from closely-spaced seismometer- and smartphone-arrays [*Inbal et al.*, 2015, 2016, 2019; *Yang et al.*, 2021; *Yang and Clayton*, 2023] or fiber optic cables [*Zhan*, 2020; *Lellouch et al.*, 2021], is well suited for signal detection in noisy environments. The main advantage of dense arrays over sparse networks is that dense wavefield sampling may be used suppress noise in data with poor signal-to-noise ratios (SNR). Furthermore, array back-projection may be used to focus incoming signals onto the source region, thereby strongly facilitating their location. However, it is not clear whether dense arrays, which are often deployed in noisy environments, and whose apertures do not exceed a few km, possess sufficient resolution power at seismogenic depths.

We restrict our analysis to geometries in which the potential source lies beneath the array, a situation common in dense array studies [e.g. *Inbal et al.*, 2015, 2016; *Peña Castro et al.*, 2019; *Catchings et al.*, 2020; *Yang et al.*, 2021, 2022; *Yang and Clayton*, 2023]. To investigate dense array source detectability and depth resolution, we consider the Long Beach (LB ; 5200 sensors ; deployed between January and June, 2011) and the Extended Long Beach (ELB ; 2500 sensors ; deployed between January and March, 2012) array datasets. The two arrays were located along adjacent portions of the Newport-Inglewood Fault (NIF), a major fault traversing the LA basin (Figure 1). *Inbal et al.* [2015, 2016, hereafter referred to as I2016 and I2015, respectively] used the LB dataset to compile a catalog for the portion of the NIF in LB, by enhancing the event detectability via sub-array stacking and downward-continuation [*Gazdag*, 1978]. This allowed them to detect abundant seismicity occurring in the lower crust and upper mantle. The depth range was unusual given that, except for a few places, seismicity in southern California is generally confined to the upper 12 km or so [e.g. *Hauksson*, 2011]. Thus, I2016's findings challenged the common understanding regarding the physical mechanisms allowing faulting at depth. Recently, *Yang et al.* [2021, hereafter referred to as Y2021] introduced a new detection scheme which relies on the SNR of the back-projected surface data before and after trace randomization, and applied it to the ELB dataset. Given the proximity between the LB and ELB arrays, the similar deployment intervals, and lack of significant mainshocks during the deployment periods, the gross features of the LB and ELB seismicity catalogs must be quite similar. However, the discrepancy between I2015's and I2016's deep seismicity detection rates to Y2021's deep seismicity detection rates is very significant. Y2021 found that over a three month period 13 lower-crustal and 1 upper-mantle microearthquake had occurred below the ELB array. If we assume the deep seismicity rates below the LB and ELB arrays are similar, then I2016 and I2015 findings imply Y2021 are missing 99% of the deep earthquakes below ELB. Y2021, on the other end, attribute the high detection rates in I2015's and I2016's catalogs to artifacts introduced in the back-projection procedure. Since the width of the NIF seismogenic zone determines the maximum magnitude the fault can produce, and since the physics governing the nucleation of upper-mantle earthquakes on continental transform faults is not well understood, it is important to resolve these discrepancies. Additionally, a growing number of studies rely on the detection of weak signals in back-projected seismic array data, underscoring the importance of quantifying the array resolution and the robustness of the detection scheme. In this study, we reexamine the NIF seismic catalogs along with newly acquired Moho depths in the LB area [*Clayton*, 2020]. We review the detection schemes of I2015, I2016, and Y2021, and test what fraction of events detected by I2015 and I2016 might have been missed by Y2021. Then, we assess the discriminative power and depth resolution of dense arrays by using seismograms of deep NIF earthquakes and a set of synthetic tests. In light of these results, we confirm the findings of I2015 and I2016.

Observing Deep Earthquakes on the LB array

To motivate the discussion on dense-array resolution, we begin by reviewing the available seismicity catalogs and Moho depths, and the detections schemes of I2015, I2016, and Y2021. The distribution of seismicity along the NIF obtained from the LB array, and from the regional Southern California Seismic Network (SCSN), together with newly acquired Moho depths [*Clayton*, 2020] are shown in Figure 2. As was previously suggested by I2015 and I2016, many of the events in the LB section of the NIF occur in the lower crust, and some events occur in the upper mantle (Figure 2a). The frequency-magnitude distribution in the LB back-projection-based catalog is complete down to about $M = -1$. After adjusting for the area and time-window of the LB array deployment, the frequency-magnitude distribution of earthquakes occurring above 15 km depth in I2015 and I2016 catalogs nicely extrapolates to the frequency-magnitude distribution in the SCSN catalog, which is complete above M$\sim$2 [*Inbal et al.*, 2015]. Note that the NIF intersects the Moho at about 17 km depth (Figure 2a). Thus, we think the LB array-derived catalog, which contains widespread lower-crustal seismicity observed over a 6-month period, but whose magnitude of completeness is about three units lower than the SCSN catalog completeness magnitude, reflects the long-term behavior of the NIF and not some transient behavior. Additionally, note that both the SCSN and the LB array-derived catalog depths increase along the A-A' profile, and that this trend in consistent with the increase in Moho depth along the same profile. Accounting for the combined uncertainty on the Moho and source depths suggests the deepest events in the SCSN catalog are well within the upper mantle.

An adjacent cross-section located below the ELB array is shown in Figure 2b. The focal depth distribution for the ELB section is skewed towards depth larger than 10 km, in disagreement with the distribution of Y2021, which mostly consists of events occurring in the upper 10 km (see Fig. 3 in Y2021). Given *Clayton* [2020]'s Moho depths, Y2021's catalog contains 13 lower-crustal earthquakes and one or two upper-mantle earthquakes. Although the SCSN and QTM catalogs do not contain mantle earthquakes below ELB, the ELB focal depths are skewed towards values larger than the depth distribution observed along seismically active fault sections cutting through thin-crustal zones in southern California. Thus, similar to the LB section of the NIF, the regional catalogs suggest the ELB section also hosts earthquakes occurring at depths that are larger than the ones expected given the local geotherm and strain rates (see also discussion in I2016).

We find that the detection rate of earthquakes occurring at a depth between 12 and 20 km is somewhat lower in the QTM catalog than the in the SCSN catalog. This is likely due to the low number of available templates and the poor SNR conditions typical for the SCSN stations in the greater LA area. For example, a recent study found that a $M \approx 1$ NIF and a $M \approx 1$ off-shore earthquake, showing SNR>1 on a number of stations located within 30 to 100 km from the epicenter, were missing from the SCSN and QTM catalogs [*Inbal et al.*, 2023]. Because the urban noise amplitude generally decays more rapidly with distance than the earthquake signal, the SNR of $M \leq 1$ NIF events is sometimes higher on stations located outside the LA basin than on near-epicentral stations. This may cause traditional or template-based detection schemes to miss some events, since those schemes rely primarily on phase arrivals observed on near-epicentral stations.

Unlike traditional network detection techniques, dense array analysis enhances the SNR by beamforming (i.e. delay-and-sum) the array's seismograms. Assuming the noise recorded by the array is uncorrelated between the array's sensors, this procedure improves the SNR by a factor proportional to $\sqrt{N}$, where N is the number of sensors in the array [e.g. *Rost and Thomas*, 2002]. If the target area lies beneath the array, and if a detailed velocity model is available, then further SNR improvement can be obtained by wavefield extrapolation using downward-continuation [*Gazdag*, 1978], which enhances near-vertical signals impinging on the array. We discuss the improvement in source-depth resolution due to downward-continuation in the section Spatial Resolution Analysis. Our experiments showed that strong LB surface noise sources were resilient to beamforming. Those sources showed as local maxima in the back-projected LB array images, making it difficult to discriminate between earthquake and noise signals. To improve the SNR of the LB data, I2015 and I2016 downward-continued them according to the following steps. I2015 and I2016 first stacked the array data over small sub-arrays, each of which consisting of 5 sensors, and then interpolated the sub-array-averaged data onto a regular grid. The interpolated data were Fourier transformed and then downward-continued to a depth of 5 km by applying a set of phase shifts whose magnitude was computed based on the local velocity model, frequency and wavenumber content (see I2015 and I2016 for further details). These steps significantly improved the SNR. Figure 3 illustrates the SNR improvement obtained by interpolating and downward-continuing an LB wavefield containing a signal from an earthquake which occurred 17 km beneath the array. The data were filtered between 2 and 8 Hz before they were downward continued. Visual inspection of these images discloses significant SNR improvements leading to enhancements the facilitate the location procedure. Due to scattered strong noise sources, the amplitudes of the 2 to 8 Hz filtered surface data (Figure 3a) are not well-correlated with the epicentral location. Applying plain-stack (i.e. setting the inter-sub-array time-lags to zero) increases the amplitude of vertically propagating energy due to the deep source relative to horizontally propagating energy due to shallow sources, effectively suppressing isolated surface noise sources. The interpolation also removes some of the effects caused by isolated noise sources, each recorded by a few sensors (Figure 3b). Downward-continuation assumes the wavefield is composed only of vertically propagating energy, which is useful for removing surface waves, and for focusing vertical energy onto deep sources. These effects are clearly demonstrated in Figure 3c, which shows that most of the isolated surface noise-sources were de-focused, whereas energy from the deep earthquake is focused onto the source.

Following downward-continuation, I2015 and I2016 enveloped the data and back-projected them onto the volume beneath the LB array. The statistical attributes of the back-projected image maxima were analyzed. I2015 and I2016 found that the statistical distribution of the post-downward continued back-projection images containing newly identified tectonic sources was significantly different from the one associated with back-projection images of post-downward continued non-tectonic sources. The former follows a power-law distribution, while the latter follows a Gumbel distribution. That separation facilitated the discrimination stage. I2015 and I2016 declared a detection if the maximum amplitude of the back-projection image exceeded 5 times the Median Absolute Deviation (MAD) of the amplitude of the back-projection images around the detection time. Using this detection threshold and the cumulative probabilities of the signal and noise back-projection images, I2015 found the false detection rate to be 2×10^{-3} per night.

Y2021 took a different approach for discriminating coherent seismic sources from noise sources in dense array recordings, which they refer to as Trace Randomization (TR). To test for the presence of a tectonic signal, the TR scheme spatially redistributes envelopes of the array seismograms by assigning them random positions within the array. The TR-detection criteria is based on the degree of back-projected energy reduction due to the randomization, derived from the ratio between the pre- and post-randomized maximal back-projected energy amplitudes as:

$$R = 1 - \frac{E^{post}}{E^{pre}}, \tag{1}$$

where E^{pre} and E^{post} are the pre- and post-TR maximal energy levels, respectively. Neglecting random uncorrelated noise fields which occasionally give rise to $E^{post} > E^{pre}$, Y2021 proposed an R-based detection criteria, applied to windows with $E^{pre} > 5 \times MAD(E^{pre})$ around the detection time. According to that scheme, uncorrelated noise sources should exhibit $R \sim 0$, while coherent tectonic sources should exhibit $R \sim 1$. Thus, the statistical properties of a distribution of R-values computed over multiple time windows, would allow one to discriminate between deep, temporally-isolated coherent sources to shallow uncorrelated noise sources common in continuous urban dense array data.

Given that Y2021 found only a few deep NIF earthquakes, it is instructive to characterize the LB and ELB array's capacity for detecting small-magnitude events in the pre-downward continued data. We do that by employing the TR scheme on LB array data containing signals from deep earthquakes occurring along the NIF. Many of the NIF earthquakes, which are located directly beneath the LB and ELB arrays, exhibit poor surface SNR. Some of the events, however, may be identified on the filtered pre-downward-continued array data. An example is shown in Figure 4, which presents LB array data containing 5 earthquakes recorded during March 2011, whose magnitudes were between 0 and 0.2, and whose focal depths were found to lie between

15 and 20 km. The top row shows the amplitudes of ground-velocity envelopes, computed by filtering the seismograms between 2 and 10 Hz, squaring, and smoothing using a 0.1 s running median window. The traces are ordered with respect to the hypocentral distance obtained by I2016. For each trace we compute the P-wave train SNR by taking the ratio between the mean energy in a 2 s window around the P-wave arrival to the mean in the 6 s preceding the event. Panels a to e show the amplitudes for traces with SNR>1, totaling about 40% of the array's recordings. The seismic arrivals are clearly observed between 33 and 38 s in each of the record sections (see also Figures S1-S5). The panels on the bottom row in Figure 4 show the distribution of the SNR as a function of the sensor location. Note that in a few cases (e.g. panel f and i), the epicenter is located near a cluster of high SNR traces. However, the surface detection pattern is generally not well correlated with the epicentral location, which complicates the detection procedure. For the earthquakes shown in Figure 4, the array-averaged SNR are between 1.02 and 1.06, within a few percent of the median SNR of LB events occurring below 15 km. Thus these 5 events represent the SNR conditions of many of the deep earthquakes in the LB catalog.

We used the relation in Equation 1 to compute the R-values for the time windows containing the arrivals in the seismograms shown in Figure 4a-e, and found that R varies between 0.01 and 0.2 for these five events. Y2021 state that time-windows they associated with noise had $R >0.2735$, which is considered here as a threshold above which the Y2021 scheme would declare a positive detection. Note that the R^{LB}-values calculated for the events in Figure 4 are lower than the threshold of Y2021 for the ELB dataset. Note also, that because of its smaller aperture, the R^{ELB} associated with arrivals as the ones shown in Figure 4 is expected to be smaller than R^{LB}. The strong presence of noise in the pre-downward continued ELB data and the conservative detection criterion may explain why Y2021's scheme have missed many events below the ELB array.

Synthetic Tests for Characterizing the Effects of the Signal-to-Noise Ratio and Array Aperture on Source Discrimination

To examine why the LB back-projection energy reduction may sometime tend to 1 (i.e. $R = 0$) for seismograms containing tectonic signals correlated among 40% of the array's sensors, and to assess how the R-values are influenced by the array's aperture and SNR levels, we applied a series of tests using two synthetic datasets. In the first set of tests, we generate synthetic seismograms assuming a population of sources whose numbers exponentially decay with depth below 4 km, similar to the source depth distribution in the LB catalog compiled by I2016. For each source, we compute R^{ELB} and R^{LB} for a monochromatic 5-Hz input signal modulated by an envelope whose amplitude decays exponentially with time over a time scale of a few seconds, and which propagates in a uniform velocity model. The spectral content of the synthetic signal is selected based on NIF earthquake seismograms analyzed by I2015 and I2016. We add white noise to the seismograms such that their SNR is smaller than one, similar to urban dense-array datasets. In the second set of tests, we compute R using traces containing uncorrelated random noise. The results are presented in Figure 5. The blue curve in panel a shows the value of R^{ELB} as a function of source depth. Note that R-values are depth-dependent, such that larger values are systematically associated with sources at shallow depths, which implies that an R-based detector may miss deep seismic events. This depth bias is only slightly reduced by increasing the aperture of the array, as shown by the red curve in Figure 5a, which indicates R^{LB} values as a function of source depth. Note that, since $R^{LB} > R^{ELB}$, the TR-based detection statistics obtained for the ELB geometry by Y2021 do not apply straightforwardly to the LB array geometry. Also, the value of R computed for the March 2011 earthquakes shown in Figure 4 is considerably smaller than the synthetic value, which likely reflects the poor SNR conditions (i.e. array-averaged SNR) of the LB array data. However, this does not affect the trend with depth shown in Figure 5a.

The results presented in Figure 5 provide further insights on the importance of the array aperture for source discrimination. That discrimination scheme is most effective for sources associated with a large scatter of the inter-array time delays, a requirement that is met when the array aperture is close to or larger than the source depth. When the aperture-to-source-depth ratio is large, TR is expected to significantly decrease E^{post} relative to E^{pre}, thereby providing a reliable detection statistic. For the LB and ELB arrays, this condition applies to events occurring above approximately 8 km and 12 km, respectively. On the other hand, when the aperture-to-source-depth ratio is much smaller than one, the range of inter-array time delays ("normal moveout") tends to zero, thereby reducing the discriminative power of the array. The discriminative power can be parametrized by the array's time-delay Median Absolute Deviation ($\mathrm{MAD}_{\Delta t}$), the value of which is dependent on the array aperture and source depth, as well as on the SNR and the time delay resolution. In general, $\mathrm{MAD}_{\Delta t}$ decreases with source depth, with faster decrease rates for small-aperture arrays (Figure 6). Thus, for very small arrays or very deep sources, we expect $\mathrm{MAD}_{\Delta t} \rightarrow 0$. The narrow range of time-delays obtained in these situations is expected to yield R-values close to zero, and therefore cause the detector to miss some weak events.

The array's discriminative power is also affected by the SNR. For poor-SNR signals, the ratio E^{post}/E^{pre} can occasionally be significantly smaller than one, which may result in a false detection. To illustrate this effect, we indicate in Figure 5 the R^{ELB}-value reported by Y2021, and the one obtained in this study by the dashed and yellow vertical lines, respectively. Note that

Y2021's R^{ELB}-value was computed using thousands of time windows passing their initial detection criteria, whereas the R^{ELB} reported here is a depth-averaged value computed using only windows containing a coherent synthetic source, yet the two values closely match. Since most of the windows Y2021 used for computing R^{ELB} likely do not contain a tectonic signal, this result suggests the R-based scheme may not be suitable for discriminating deep sources. We find this issue repeats when the approach is applied to data containing random uncorrelated noise. For this type of input, the fluctuations around the mean value of R can be quite large, and are generally dependent on factors such as the sampling interval and the envelope calculation method. For the commonly used nth-root stacking [e.g. *Rost and Thomas*, 2002, with n=3], the average value of R is close to 0, as expected for records containing only uncorrelated random noise. However, after neglecting cases in which $R < 0$, we find that 34% the windows have $0 < R < 0.3$ and 13% of the windows have $0.3 < R < 0.6$ (see dark and light-grey rectangles in panel 5a), within the range of results from tests containing a coherent source (blue curve in Figure 5a). In fact, the range of depths allowing for reliable source discrimination on the pre-downward migrated ELB array is limited to the upper 8 km, since the statistics for deeper sources are not significantly different from the ones associated with a random noise field.

Thus far, we have estimated the detection sensitivity to the array aperture and SNR. Next, we estimate the source depth error by comparing the source depth obtained from back-projecting the LB signal envelopes to the input source depth, after adding white uncorrelated noise. The noise amplitude is uniformly distributed over the range between -0.8 and 0.8 times the maximum envelope amplitude. The results are presented in Figure 5b,c, which shows the distribution of source depth discrepancies and the depth error as a function of input LB source depth. We find that the source depth error is about 2 km, consistent with the results of synthetic tests presented by I2016. In addition, for the range of source-aperture-to-source-depth ratios examined here, we do not find that the depth error correlates with the source depth. This suggests that the dominant factor limiting accurate source depth determination is the array aperture, assuming the sources lie within the array's footprint, and that their signals exceed the noise level. Thus, resolving the depth of earthquakes occurring beneath the array may be obtained by a subset of the array's sensors, given that (1) the source-depth-to-array-aperture ratio is smaller than about 2, and (2) the SNR is larger than 1. We test the validity of this statement by using synthetic tests presented in the next section.

Spatial Resolution Analysis

The results presented in the previous section show that the R-based detector is likely to miss low-SNR signals excited by deep tectonic sources. Because of its smaller aperture, the number of events missed by the ELB array is expected to be larger than the number missed by the LB array. Once a signal has been positively detected, however, its location accuracy is dependent on the SNR and array geometry. Given the LB SNR levels, its vertical location resolution is about 2 km (Figure 5b). To characterize the source imaging resolution, we use Point Spread Functions (PSF), which describe the effect of the imaging system on the imaged object [e.g. *Lecomte et al.*, 2015; *Nakahara and Haney*, 2015]. The degree of source resolution and illumination may be derived from basic principles of ray theory, by considering the density of source-to-array raypaths. In this framework, a well illuminated source is defined as one for which ray paths cover a large fraction of the focal sphere. In an isotropic medium, the wavenumber vector is at any point perpendicular to the wavefront, and thus its orientation and amplitude in the source region may be used to determine the source image spatial resolution. For a source at location j imaged by a station at location i, the local wavenumber vector is defined by the projection of the source

Fourier components onto the local slowness vector [*Lecomte et al.*, 2015]:

$$k_{ij}^{local} = \omega \cdot S_{ij}, \tag{2}$$

where ω represents the angular frequency, and S_{ij} is the local slowness vector, which is parallel to the ray connecting the j'th source with the i'th station. In practice, each frequency component is weighted by the source spectra, and as a result, wideband sources are expected to be better resolved than narrowband sources. The spatial resolution is also dependent on the aperture of the array. Increasing the array aperture will increase the local wavenumber density, which improves the illumination and enhances the imaging resolution. The PSF is obtained from k^{local} after weighting by the source spectra by summing over available source-to-array ray paths, and then taking the inverse spatial Fourier transform. The advantage of this approach is that it allows us to compute PSFs that are independent of the noise, and ensures that the spatial variability of urban noise levels [*Riahi and Gerstoft*, 2015; *Inbal et al.*, 2019] does not affect the resolution estimates.

To quantify the spatial resolution and analyze its dependency on the source depth, we compute the PSF for the LB and ELB array geometries. As input, we use the spectra of the envelope of the 5 Hz exponentially decaying sine function discussed in the previous section. Equation 2 is solved assuming a uniform velocity model of 3.5 km/s, neglecting the effects of scattering on the PSF [*Lecomte et al.*, 2015]. Figure 7 presents the spatial resolution for shallow (10 km; panels a,b) and deep (20 km; panels c,d) sources. In the absence of noise in the input data and velocity model, the only effect reducing the source depth resolution is the limited aperture of the array, which is manifested by the smearing of the PSFs along the depth axis. This affects the ELB and LB array differently, and is most noticeable for sources located below 12 km, for which the vertical resolution of the ELB degrades rapidly with depth. To illustrate this effect, we present in Figure 7e the vertical resolution scale, defined as the vertical extent over which the PSF value decreases down to 80% relative to the maximum PSF value at the focal point. For shallow sources ($<$ 10 km), both arrays can well resolve sources located less than 1 km apart. However, the limited aperture of the ELB array yields images whose resolution power at large depths is reduced relative to the LB array. Events located at depths larger than about 20 km are not well resolved by the ELB array, but may be resolved by the LB array. This effect is an outcome of a modestly wider aperture (both in the NS and in the EW direction ; see Figure 1) of the LB array relative to the ELB array.

We also investigated the effects of downward-continuation [*Gazdag*, 1978] of the wavefield on the vertical resolution. Reducing the vertical separation by wavefield extrapolation has the desired effect of increasing the $\mathrm{MAD}_{\Delta t}$. The direct consequence is a significant increase in the vertical resolution scale. This is illustrated by the dashed curve in Figure 7e, showing the vertical resolution for the LB array after wavefield extrapolation down to 5 km depth. For the deepest events located below 25 km, downward continuation may improve the vertical resolution by as much as 40%. Note that these estimates provide a lower bound on the improvement in the resolution. The SNR may be improved prior to conducting downward continuation by applying plain-stack (i.e. setting the array's time delays equal to zero) of small sub-arrays within the LB array, which tends to de-amplify surface waves generated by shallow sources Figure (3a). Additionally, downward continuation further de-amplifies such arrivals (Figure 3b), and is thus expected to improve the vertical resolution relative to what is shown in Figure 7e.

Recent studies suggest the dramatic increase in the spatial sampling of the seismic wavefield provided by state-of-the-art seismic imaging systems may help improve earthquake detectability and hence refine existing catalogs [*Inbal et al.*, 2019; *Lellouch et al.*, 2021; *Mesimeri et al.*, 2021; *Arrowsmith et al.*, 2022]. For example, *Inbal et al.* [2019] evaluated the earthquake location accuracy achieved by dense noisy smartphone arrays. They found that back-projecting only 0.5% of the available smartphone-derived seismograms in the LA area would allow detection of events with M$\sim$1, approximately one magnitude unit below the catalog magnitude of completeness in that region. This smartphone-user density was required in order to enhance the SNR of smartphone-recorded signals due to M$\sim$1 earthquakes. However, it is not clear what is the minimum density required in order to resolve the location of back-projected signals that standout of the noise level.

Next, we use a bootstrap analysis to assess the sensitivity of location estimates of signals with SNR>1 to the density of the array. To do that, we compute the PSF for the LB configuration by using only 1% of the available LB sensor positions, which we refer to as the sparse array configuration. For each input source depth value, we generate 100 sparse configurations randomly selected from the LB array sensor locations. The results are presented in Figure 7e, which shows the average resolution of the vertical location of the source for the sparse array dataset. Remarkably, we find that the sparse configuration is almost as effective as the dense configuration for resolving earthquake-like signals with SNR>1 located beneath the array. Moreover, we find that the resolution on deep (>15 km) sources obtained by using the sparse configuration exceeds the resolution of the 2500-sensors ELB array for sources lying at this depth range. Thus, an array whose dimensions are comparable to the LB array, but which contains only a small number of sensors, can be used to locate signals excited by deep tectonic events if they exceed the ambient noise level, and occur within the array's footprint. The logic also applies to the local seismic network operating in the LA area, whose inter-sensor distances are of the order of 10 km. Back-projecting signals recorded by this network onto the NIF fault may help obtain robust locations, and reduce the local catalog's magnitude of completeness [*Inbal et al.*, 2023].

Summary

We examine the depth resolution of dense seismic arrays for sources lying beneath the array. We find that the parameter controlling the resolution power is the source-depth-to-array-aperture ratio and the source's bandwidth. The source-array geometry effect on the resolution can be parameterized by the MAD of the inter-array time delay distribution, which is sensitive to modest changes in the aperture. The LB array maximum aperture is only 20% larger than the ELB array maximum aperture, yet its source depth resolution for deep (>20 km) events is improved by about a factor of two (Figure 5), which indicates that small changes to the array geometry may yield significant improvement to the resolution power. In addition, we find that using only 1% of the LB array sensors does not significantly affect the depth resolution of signals with SNR>1, given the sensor subset maintains an aperture close to aperture of the entire array.

We use synthetic tests to evaluate the performance of the TR -based approach of Y2021. We find that this scheme is sensitive to the array aperture, and is expected to detect more shallow-depth events than deep events. This sensitivity also suggests the results obtained by Y2021 for the ELB dataset may not straightforwardly apply to the LB dataset. In addition, the TR-based scheme may sometimes classify a random noise field as a tectonic signal. This is demonstrated in the following manner: if we assume the input source depths are exponentially distributed and truncated below 35 km, and that all time windows contain arrivals from no more than a single earthquake, then we find the mean R^{ELB}-value equal to 0.278 (dashed curve in panel 5b). This value is almost identical the R^{ELB} value computed by Y2021 for noise-dominated time windows.

The factors promoting earthquake nucleation below the seismogenic zone remain poorly resolved. Earthquakes are the result of stick-slip frictional instabilities that occur due to brittle fracture of rock, a behavior that is strongly dependent on the ambient pressure-temperature, lithology, strain rate, and pore pressure. In southern California, the maximum depth of seismicity largely coincides with the 400°C isotherm [*Bonner et al.*, 2003; *Hauksson*, 2011]. That correlation is thought to manifest thermal effects on the rheology, with the deep termination of seismicity corresponding to the onset of plastic yielding in Quartz-rich rocks [e.g. *Scholz*, 2002]. Clusters of deep events are common in thick-crustal, rapidly-deforming regions, where the local isotherm is depressed downwards due to lower-than-average heat-flow [*Bonner et al.*, 2003], or where faults cut through mafic lithology [*Magistrale and Sanders*, 1996; *Magistrale*, 2002], which tend to exhibit brittle behavior at larger depths. The NIF events are an exception to this rule. They represent some of the deepest earthquakes in California, yet they occur on slowly deforming faults cutting through the thinnest crust in California, whose associated heat flow is close to the regional average. Thus, the width of the seismogenic zone along the NIF challenges our understanding of the processes responsible for earthquake rupture. Since the maximum earthquake magnitude for a given fault is a function of its width, the seismicity depth extent also bears strong implications for seismic hazard in the LA urban area.

Data and Resources. The Southern California Earthquake Data Center earthquake catalog is available at the following doi: https://scedc.caltech.edu. The LB seismicity catalog is from *Inbal et al.* [2016], and the ELB seismicity catalog is from *Yang et al.* [2021]. The raw LB data are protected by a license agreement with Signal Hill Petroleum, and will be provided by the authors upon request.

Acknowledgments. We thank three anonymous reviewers for constructive remarks that greatly improved the quality of this manuscript. This research was supported by ISF grant #1802/22.

References

Arrowsmith, S. J., D. T. Trugman, J. MacCarthy, K. J. Bergen, D. Lumley, and M. B. Magnani (2022), Big Data Seismology, *Rev. Geophys.*, *60*(2), e2021RG000,769, doi: https://doi.org/10.1029/2021RG000769.

Bonner, J. L., D. D. Blackwell, and E. T. Herrin (2003), Thermal Constraints on Earthquake Depths in California, *Bull. Seis. Soc. Am.*, *93*(6), 2333–2354, doi:10.1785/0120030041.

Catchings, R. D., M. R. Goldman, J. H. Steidl, J. H. Chan, A. A. Allam, C. J. Criley, Z. Ma, D. S. Langermann, G. J. Huddleston, A. T. McEvilly, D. D. Mongovin, E. M. Berg, and Y. Ben-Zion (2020), Nodal Seismograph Recordings of the 2019 Ridgecrest Earthquake Sequence, *Seismo. Res. Lett.*, *91*(6), 3622–3633, doi:10.1785/0220200203.

Clayton, R. W. (2020), A detailed image of the continent-borderland transition beneath Long Beach, California, *Geophys. J. Int.*, *222*(3), 2102–2107, doi:10.1093/gji/ggaa286.

Gazdag, J. (1978), Wave equation migration with the phase-shift method, *Geophysics*, *13*(7).

Hauksson, E. (2011), Crustal geophysics and seismicity in southern California, *Geophys. J. Int.*, *186*(1), 82–98, doi:10.1111/j.1365-246X.2011.05042.x.

Inbal, A., R. W. Clayton, and J.-P. Ampuero (2015), Imaging widespread seismicity at mid-lower crustal depths beneath Long Beach, CA, with a dense seismic array: Evidence for a depth-dependent earthquake size distribution, *Geophys. Res. Lett.*, *42*(15), 6314–6323, doi: 10.1002/2015GL064942.

Inbal, A., J.-P. Ampuero, and R. W. Clayton (2016), Localized seismic deformation in the upper mantle revealed by dense seismic arrays., *Science*, *354*(6308), 88–92, doi: 10.1126/science.aaf1370.

Inbal, A., Q. Kong, W. Savran, and R. M. Allen (2019), On the Feasibility of Using the Dense MyShake Smartphone Array for Earthquake Location, *Seismo. Res. Lett.*, doi: 10.1785/0220180349.

Inbal, A., T. Cristea-platon, J.-p. Ampuero, G. Hillers, and D. Agnew (2023), Reply to "Comment on 'Sources of Long-Range Anthropogenic Noise in Southern California and Implications for Tectonic Tremor Detection' by Asaf Inbal, Tudor Cristea-Platon, Jean-Paul Ampuero, Gregor Hillers, Duncan Agnew, and Susan E. Hough" by Allie Hutchison, Yijian Zhou, and Abhijit Ghosh, *Bull. Seismo. Soc. Am.*, pp. 1–15, doi:10.1785/0120200007.

Lecomte, I., P. Lubrano-Lavadera, A. Wuestefeld, T. Kaschwich, J. Albaric, and H. N. Gharti (2015), Focusing in migration-based location of weak microseismicity: modelling point-spread function for resolution analyses, in *SEG Technical Program Expanded Abstracts 2015*, pp. 2491–2495, Society of Exploration Geophysicists, doi:10.1190/segam2015-5813282.1.

Lellouch, A., R. Schultz, N. Lindsey, B. Biondi, and W. Ellsworth (2021), Low-Magnitude Seismicity With a Downhole Distributed Acoustic Sensing Array—Examples From the FORGE Geothermal Experiment, *J. Geophys. Res.*, *126*(1), e2020JB020,462, doi: 10.1029/2020JB020462.

Magistrale, H. (2002), Relative contributions of crustal temperature and composition to controlling the depth of earthquakes in Southern California, *Geophys. Res. Lett.*, *29*(10), 87–1–87–4, doi:10.1029/2001GL014375.

Magistrale, H., and C. Sanders (1996), Evidence from precise earthquake hypocenters for segmentation of the San Andreas Fault in San Gorgonio Pass, *J. Geophys. Res.*, *101*(B2), 3031–3044, doi:10.1029/95JB03447.

Mesimeri, M., K. L. Pankow, and J. Rutledge (2021), A Frequency-Domain-Based Algorithm for Detecting Microseismicity Using Dense Surface Seismic Arrays, *Bull. Seismo. Soc. Am.*, *111*(5), 2814–2824, doi:10.1785/0120210062.

Nakahara, H., and M. M. Haney (2015), Point spread functions for earthquake source imaging: an interpretation based on seismic interferometry, *Geophys. J. Int.*, *202*(1), 54–61, doi:10.1093/gji/ggv109.

Peña Castro, A. F., S. L. Dougherty, R. M. Harrington, and E. S. Cochran (2019), Delayed Dynamic Triggering of Disposal-Induced Earthquakes Observed by a Dense Array in Northern Oklahoma, *J. Geophys. Res.*, *124*(4), 3766–3781, doi:10.1029/2018JB017150.

Riahi, N., and P. Gerstoft (2015), The seismic traffic footprint: Tracking trains, aircraft, and cars seismically, *Geophys. Res. Lett.*, *42*(8), 2674–2681, doi:10.1002/2015GL063558.

Ross, Z. E., D. T. Trugman, E. Hauksson, and P. M. Shearer (2019), Searching for hidden earthquakes in Southern California., *Science*, *364*(6442), 767–771, doi:10.1126/science.aaw6888.

Rost, S., and C. Thomas (2002), Array seismology: Methods and applications, *Reviews of Geophysics*, *40*(3), 1008, doi:10.1029/2000RG000100.

Scholz, C. (2002), *The Mechanics of Earthquakes and Faulting*, Cambridge University Press.

Yang, L., X. Liu, and G. C. Beroza (2021), Revisiting evidence for widespread seismicity in the upper mantle under Los Angeles., *Science Adv.*, *7*(4), eabf2862, doi:10.1126/sciadv.abf2862.

Yang, L., X. Liu, W. Zhu, L. Zhao, and G. C. Beroza (2022), Toward improved urban earthquake monitoring through deep-learning-based noise suppression, *Science Adv.*, *8*(15), eabl3564, doi:

10.1126/sciadv.abl3564.

Yang, Y., and R. W. Clayton (2023), Shallow Seismicity in the Long Beach–Seal Beach, California Area, *Seismo. Res. Lett.*, *94*(4), 1948–1956, doi:10.1785/0220220358.

Zhan, Z. (2020), Distributed Acoustic Sensing Turns Fiber-Optic Cables into Sensitive Seismic Antennas, *Seismo. Res. Lett.*, *91*(1), 1–15, doi:10.1785/0220190112.

Addresses

- Asaf Inbal , asafinbal@tauex.tau.ac.il, Department of Geophysics, Tel Aviv University, P.O. Box 39040, Tel Aviv 6997801, Israel
- Jean-Paul Ampuero , ampuero@geoazur.unice.fr, Université Côte d'Azur, IRD, CNRS, Observatoire de la Côte d'Azur, Géoazur, boulevard de l'Observatoire CS 34229 06304 Nice cedex 04, France
- Robert Clayton, clay@gps.caltech.edu, California Institute of Technology, Seismological Laboratory, 1200 E. California Blvd., MS 252-21 Pasadena, California 91125-2100

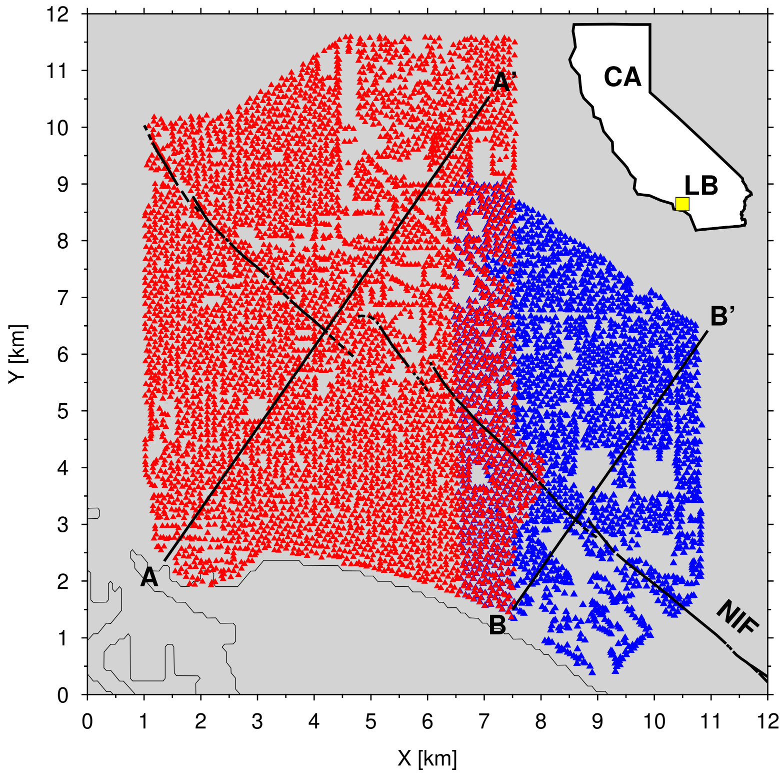

Figure 1. Location map. Red and blue triangles indicate the locations of the LB and ELB array sensors, respectively. The thick black line shows the location of the Newport-Inglewood Fault. Lines A-A' and B-B' refer to depth cross-sections shown in Figure 2. Thin black line marks the coastline. Inset map shows the location of Long Beach within the state of California. Abbreviations: NIF: Newport-Inglewood Fault, LB: Long Beach, CA: California.

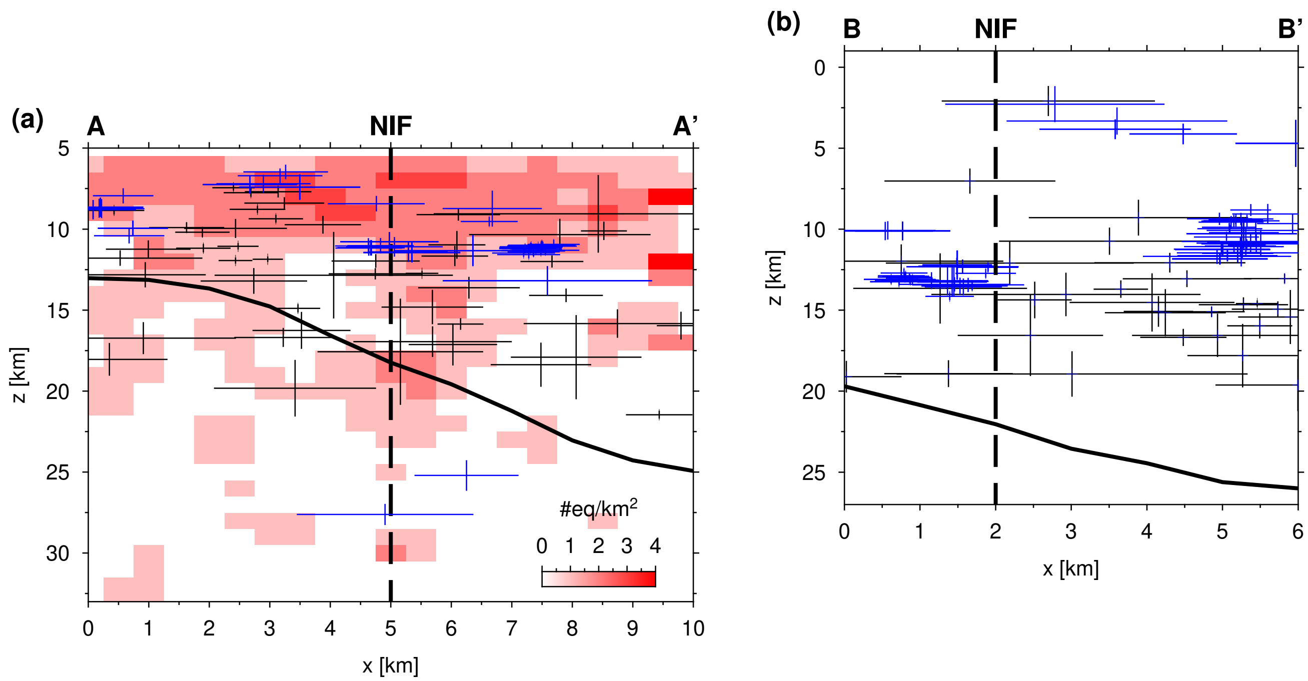

Figure 2. LB seismicity and Moho depth cross-sections. Earthquake densities from the LB array [*Inbal et al.*, 2016] are shown in shades of red. Black and blue crosses indicate the locations of earthquakes found in the regional Southern California Earthquake Center seismicity catalog covering the period between 2005 and 2024, and in the match-filter-based catalog of *Ross et al.* [2019] (QTM) covering the period between 2008 and 2018, respectively. Size of crosses corresponds to the location uncertainty. Solid and dashed curves are for the Moho depth [*Clayton*, 2020], and the Newport-Inglewood Fault, respectively. Red stars in panel b are for the locations in *Yang et al.* [2021]'s ELB catalog. The location of the cross-sections are shown in Figure 1. a. LB cross-section. b. ELB cross-section.

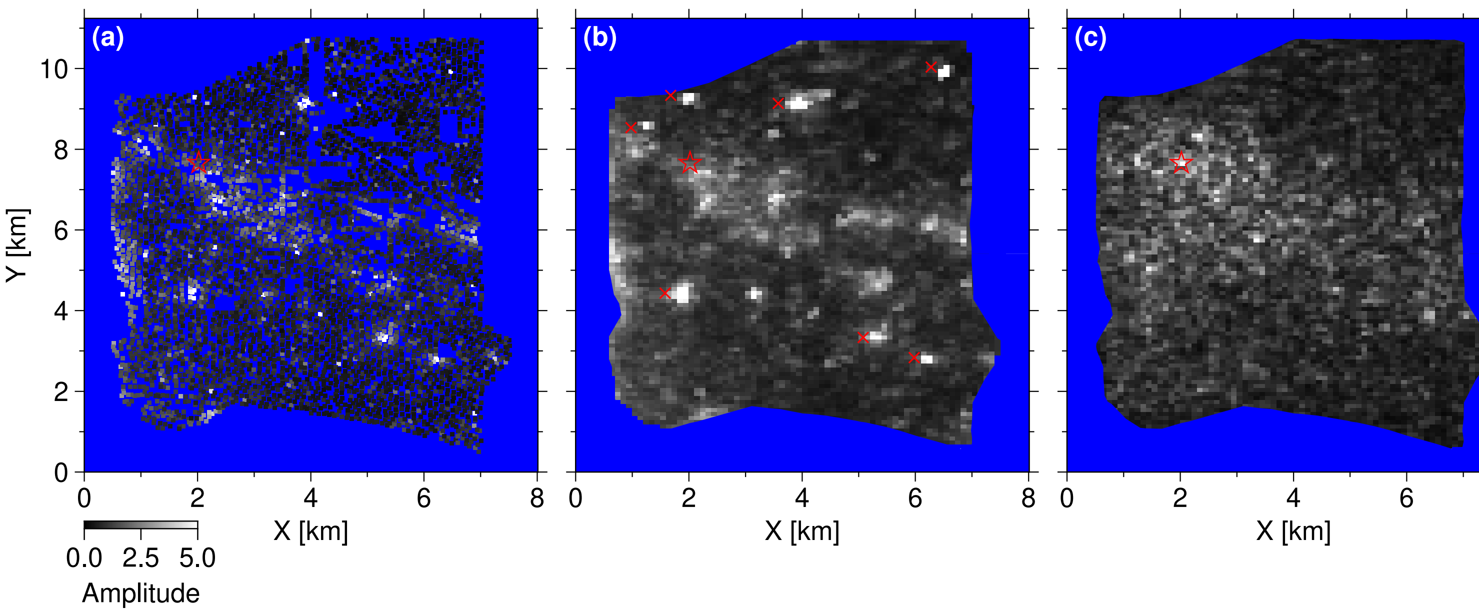

Figure 3. Example of interpolation and downward-continuation of the LB array data. Input data are the LB array amplitudes for the $M0.1$ of March 16, 2011. Event depth was determined by I2016 to be 17 km (see Figure 4a,f). Amplitudes are normalized with respect to the mean array amplitude in each panel. Star indicates the epicentral location. Crosses in panel b highlight strong surface noise sources whose amplitude was decreased by de-focusing during the downward continuation stage. (a) LB array data filtered between 2 and 8 Hz. (b) LB array data after sub-array stacking and interpolation. (c) LB array data after downward continuation to a depth of 5 km.

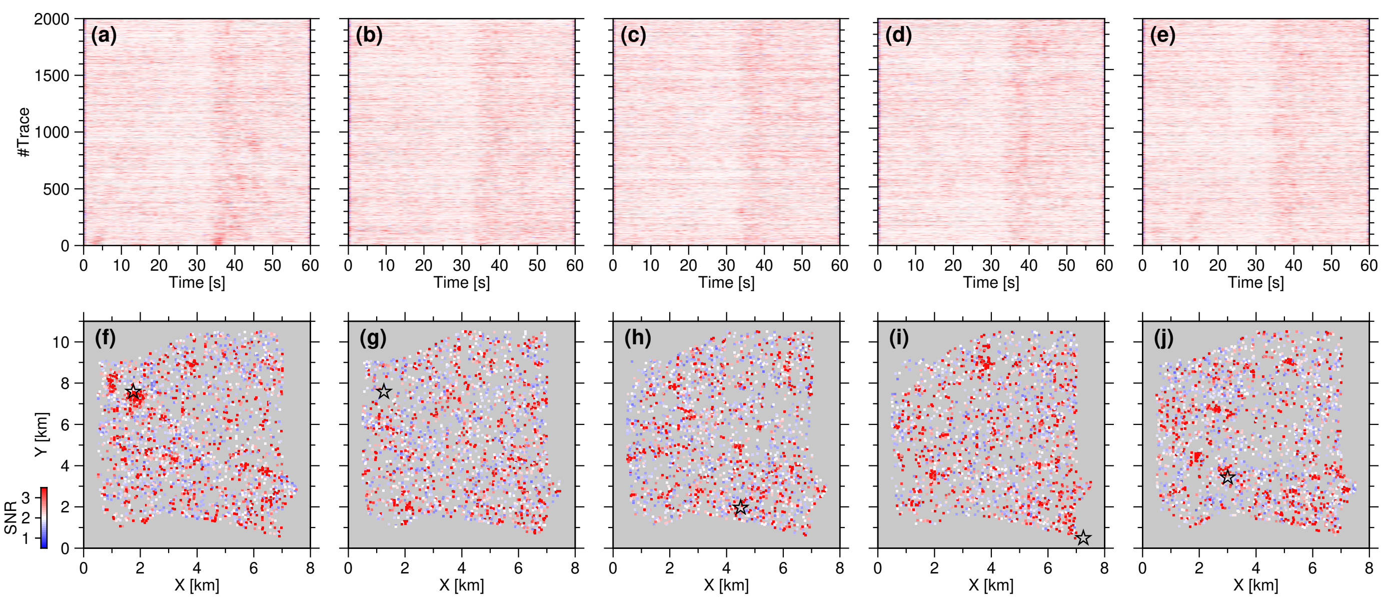

Figure 4. Seismograms recording arrivals from earthquakes occurring during March 2011 beneath the LB array. Top row shows the 2 to 10 Hz envelope amplitudes as a function of time for 2000 traces with SNR>1. Bottom row shows the distribution of the maximal amplitudes relative to the pre-event noise as a function of location. The star indicates the epicentral location. Day of detection, magnitude and depth are as follows: a,f March 16, 2011, $M0.1$, 17 km ; b,g March 18, 2011, $M0.1$, 17 km ; c,h March 8, 2011, $M0.06$, 16 km ; d,i March 15, 2011, $M0.2$, 16 km ; e,j March 5, 2011, $M0.07$, 19 km.

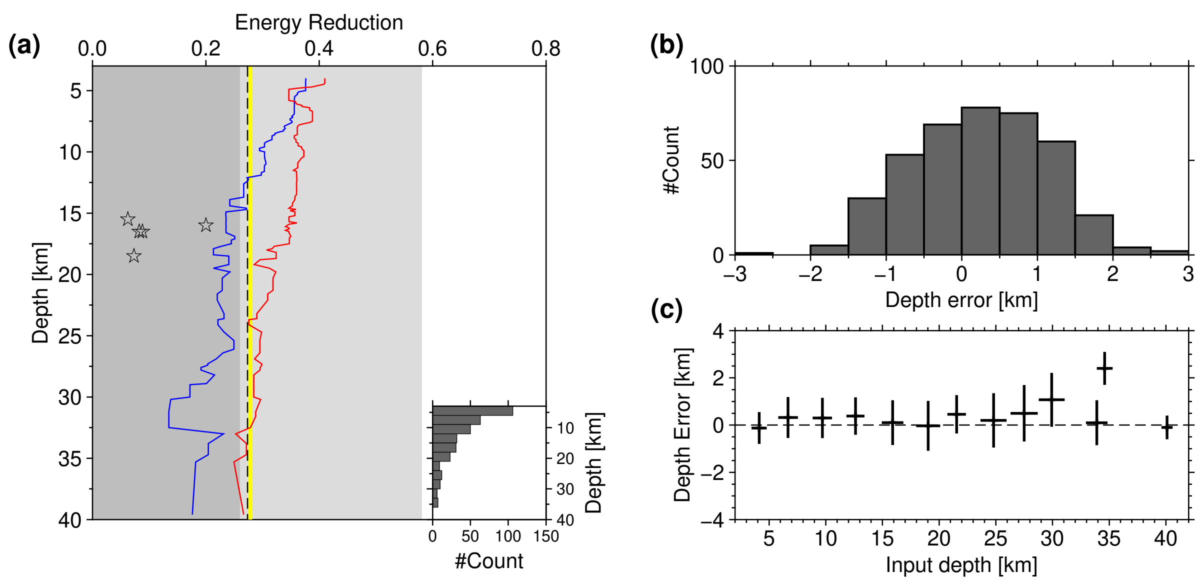

Figure 5. Synthetic tests for source discrimination using Trace Randomization. a. The back-projected energy reduction as a function of the input source depth. Solid lines indicate the level of energy reduction (defined in Equation 1), for synthetic tests in which the input source depths are exponentially distributed (as in the inset histogram), with blue and red colors for the ELB and LB array, respectively. Dashed vertical line indicates the mean back-projected energy reduction for the ELB data reported by *Yang et al.,* 2021, and the yellow line indicates the depth-averaged back-projected energy reduction we obtain for the ELB array. Dark and light grey rectangles indicate the 1- and 2-sigma intervals around the mean stack energy reduction for noise-only input using the ELB array geometry. Stars indicate the energy reduction computed for the 5 NIF earthquakes shown in Figure 4. b. The distribution of source depth error for the LB array. c. The LB source depth error as a function of the source input depth.

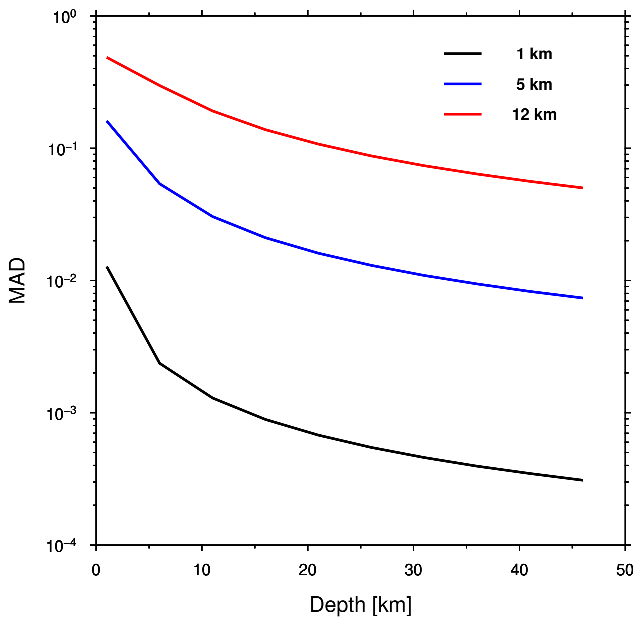

Figure 6. The median absolute deviation of the inter-array time delays as a function of the source depth. Black, blue, and red curves are for 1, 5, and 12 km array apertures, respectively. Travel times are calculated assuming a uniform velocity equal to 3.5 km/s.

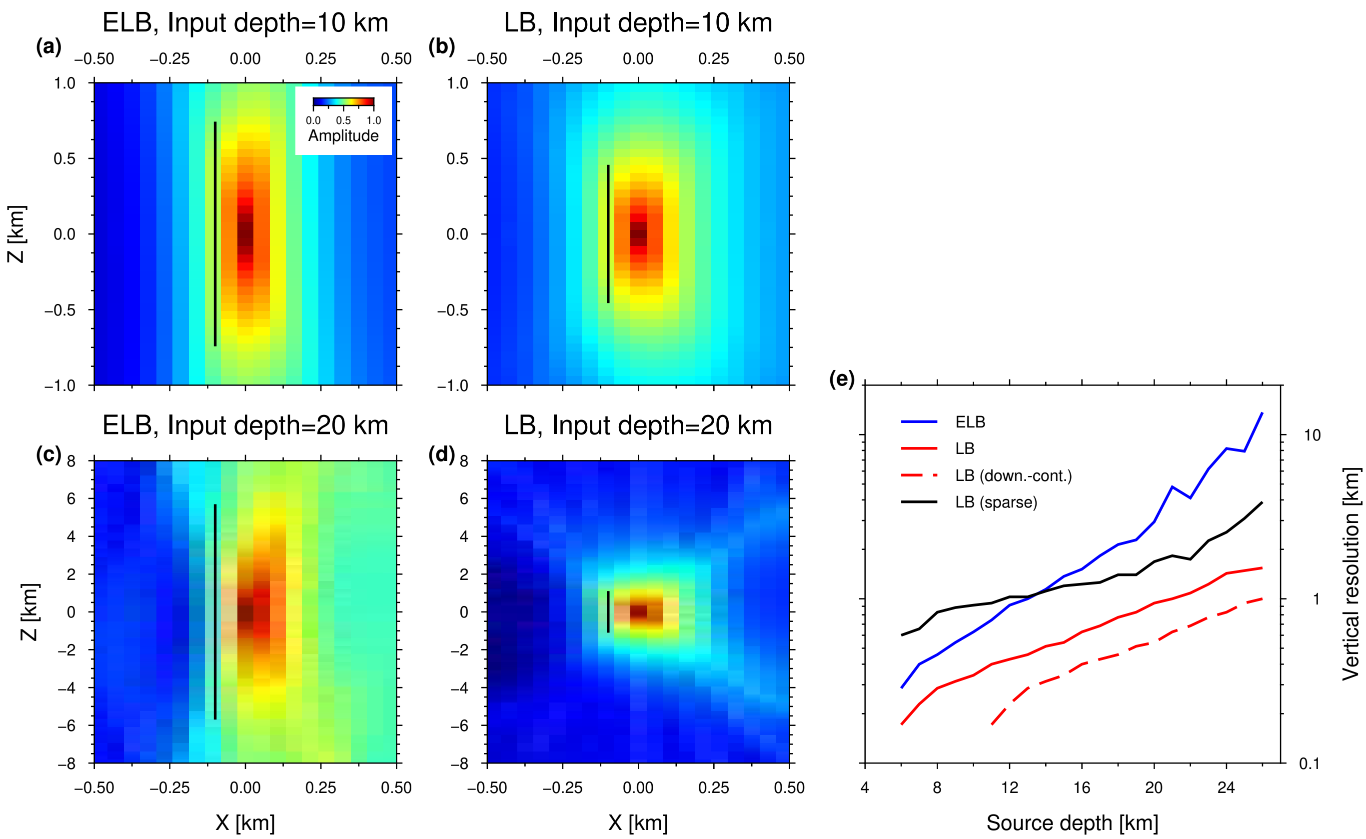

Figure 7. Resolution analysis. a-d. Point spread functions computed for an input source located at depth of 10 km (panels a,b) and 20 km (panels c,d). Vertical lines indicate the vertical resolution, defined as the length scale over which the resolution power decreases down to 80% of the maximum. a,c. ELB array. b,d. LB array. e. The vertical resolution scale as a function of source depth. Blue and red solid curves are for the ELB and LB array, respectively. Dashed red curve indicates the LB array vertical resolution after downward continuation. Black solid curve indicates the vertical resolution obtained using 1% of the LB array sensors.