Detection of small earthquakes with dense array data: example from the San Jacinto fault zone, southern California

Haoran Meng and Yehuda Ben-Zion

Department of Earth Sciences, University of Southern California, Los Angeles, CA 90089–0740, USA. E-mail: haoranme@usc.edu

SUMMARY

We present a technique to detect small earthquakes not included in standard catalogues using data from a dense seismic array. The technique is illustrated with continuous waveforms recorded in a test day by 1108 vertical geophones in a tight array on the San Jacinto fault zone. Waveforms are first stacked without time-shift in nine non-overlapping subarrays to increase the signal-to-noise ratio. The nine envelope functions of the stacked records are then multiplied with each other to suppress signals associated with sources affecting only some of the nine subarrays. Running a short-term moving average/long-term moving average (STA/LTA) detection algorithm on the product leads to 723 triggers in the test day. Using a local *P*-wave velocity model derived for the surface layer from Betsy gunshot data, 5 s long waveforms of all sensors around each STA/LTA trigger are beamformed for various incident directions. Of the 723 triggers, 220 are found to have localized energy sources and 103 of these are confirmed as earthquakes by verifying their observation at 4 or more stations of the regional seismic network. This demonstrates the general validity of the method and allows processing further the validated events using standard techniques. The number of validated events in the test day is >5 times larger than that in the standard catalogue. Using these events as templates can lead to additional detections of many more earthquakes.

Key words: Time-series analysis; Earthquake dynamics; Earthquake source observations; Site effects; Wave propagation.

1 INTRODUCTION

In the last decade, it became a common practice to record seismic waveforms continuously. The continuous recordings, combined with the increasing number of seismometers and array deployments, provide important new opportunities for detecting small events, thereby increasing the ability to resolve seismic processes and structural properties. One highly successful technique is the matched-filter processing (or template matching) that relies on cross-correlating seismograms of events detected by standard techniques with the continuous waveforms (e.g. Gibbons & Ringdal 2006). Wave portions with high similarity to the templates on multiple stations are assumed to reflect smaller events located very close to the event templates. This method enables detection of events with signal-to-noise-ratio (SNR) below 1, and leads to 10–100 times more events than those listed in typical catalogues (e.g. Meng *et al.* 2013; Shelly *et al.* 2016; Ross *et al.* 2017). However, a basic shortcoming of this technique is that it allows only detection of events with hypocentres very close to events already existing in the catalogue.

Array techniques may be used to detect sources of radiation without pre-existing templates by processing moving time windows of continuous waveforms. One example is the matched-field

processing technique (not to be confused with the matched-filter processing), which combines beamforming and backprojection of moving time windows for detection and location of energy sources. This technique was developed in underwater acoustics (Turek & Kuperman 1997) and was recently adapted to geophysical applications (Cros *et al.* 2011; Corciulo *et al.* 2012; Ben-Zion *et al.* 2015). A somewhat related method adopted from medical imaging (Fink 2006) is reverse time migration of waveforms and identification of focusing spots as locations of seismic sources (e.g. Larmat *et al.* 2008; Inbal *et al.* 2015; Nakata & Beroza 2015). An important shortcoming of these techniques is sensitivity to the used velocity model. Also, in most applications, the detected events are not subjected to validation and may correspond to scatterers or other non-genuine seismic sources.

In this paper, we present a simple new technique for detecting small events not included in a seismic catalogue using data of a dense seismic array. The technique is applied to seismic waveforms recorded by a dense array with 1108 vertical geophones (10 Hz) centred on the Clark branch of the San Jacinto fault zone (SJFZ) at the Sage Brush Flat site in the complex trifurcation area southeast of Anza, California (Fig. 1). The array covered an area of about 600 m × 600 m with a core grid consisting of 20 rows perpendicular to and centred on the Clark fault trace. Each row had 50 sensors
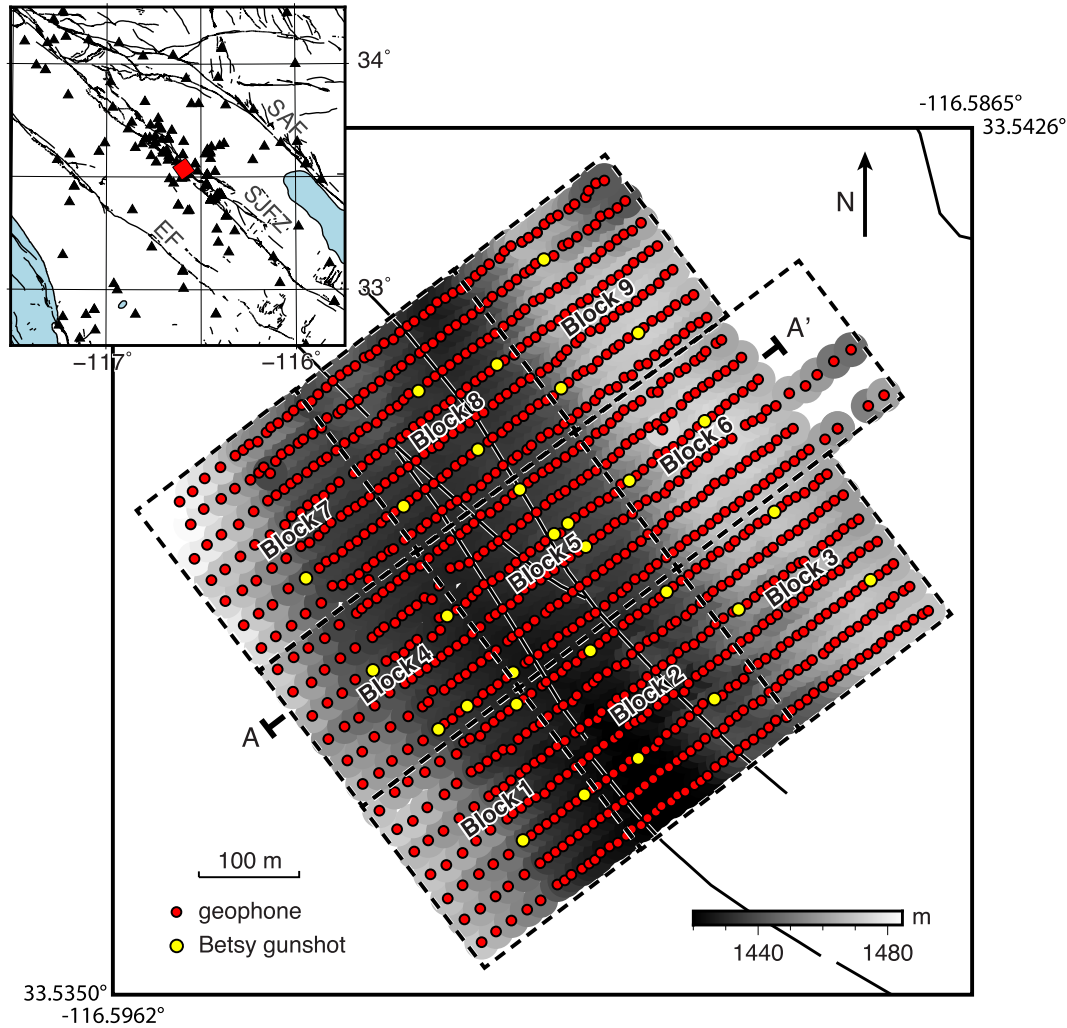
Figure 1. A location map for the dense array providing the data used in this study (Ben-Zion *et al.* 2015). Red circles in the main plot denote vertical geophones and yellow circles mark Betsy gunshots. The background grey colours represent topography. The inset at the top left provides a regional view with the dense array indicated by a red square and regional seismic stations denoted by black triangles.

at a nominal 10 m interstation spacing and the nominal separation between rows was 30 m. The remaining 108 sensors were deployed as extensions to some rows. The array recorded earthquake and noise data continuously at 500 Hz from 2014 May 07 to June 13. Additional controlled source signals were generated by 33 surface Betsy gunshots (yellow circles in Fig. 1) during 2014 June 2 and 3 (Ben-Zion *et al.* 2015).

The basic goal of this paper is methodology development, rather than creating a detailed catalogue of events from all the recorded data. The developed methodology is summarized in a flowchart diagram (Fig. 2) and is illustrated with data of one test day, Julian day 146 (2014 May 26), not associated with Betsy gunshots or field work at the site. The analysis includes a validation step involving searching for seismic signals in data recorded by standard regional stations at times predicted by the detections. The various steps of the analysis procedure are described and demonstrated in the next section. The methodology includes slant stacking and beamforming of waveforms using a detailed local velocity model for the subsurface material derived in the Appendix A. The detected/validated events may be used as templates that can lead to further detections.

2 METHOD AND RESULTS

Stacking the waveforms recorded by the spatially dense array can reduce the random noise. However, strong surface sources, large occasional instrumental glitches and other incoherent sources cannot be suppressed by stacking the data across the entire array. The presented methodology has several steps to resolve this problem and focus on signals likely generated by sources below the surface (Fig. 2). We first divide the array of 1108 sensors into nine non-overlapping subarrays covering each an area of about 200 m × 200 m (Fig. 1). The choice of subarrays is motivated by the fact that amplitudes of waveforms generated (in a different day) by the Betsy gunshots decay by 2 orders of magnitude of more during a propagation distance of 100 m (Fig. S1). The raw waveforms within each subarray are stacked without time-shift to increase the SNR of coherent signals in the subarray. The stacking without time-shift suppresses signals from sources at the surface or very shallow depth. Figs 3(a) and (b) provide examples of stacked waveforms within the nine subarrays for two representative 18 s time windows and the entire array. Black dashed boxes bracket time intervals where stacked coherent signals are seen in all nine subarrays, while dashed grey boxes mark intervals with large-amplitude signals in four or less subarrays. The
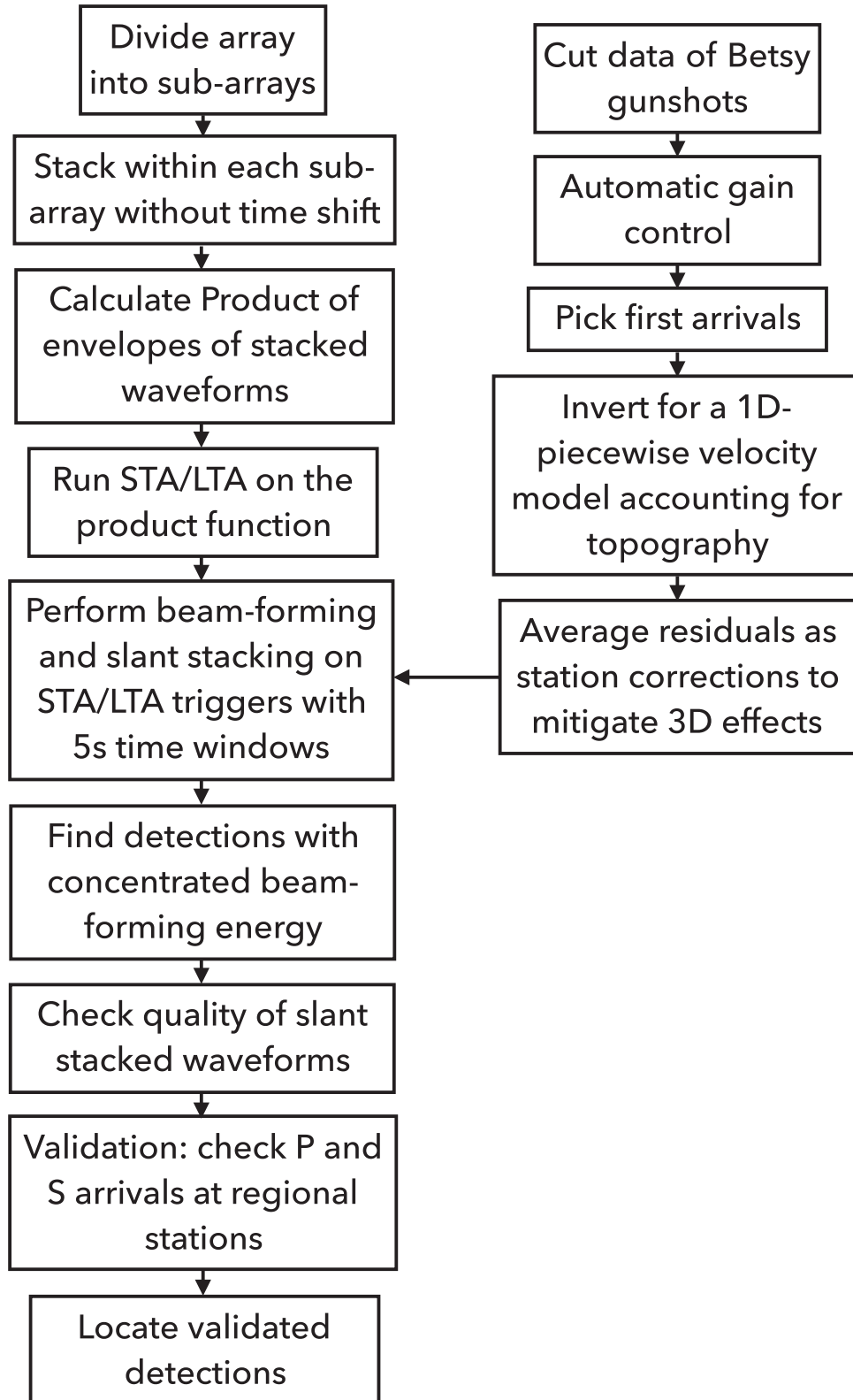
```
┌─────────────────┐              ┌─────────────────┐
│   Divide array  │              │ Cut data of Betsy│
│  into sub-arrays│              │    gunshots     │
└─────────────────┘              └─────────────────┘
         │                                │
         ▼                                ▼
┌─────────────────┐              ┌─────────────────┐
│ Stack within    │              │ Automatic gain  │
│ each sub-array  │              │    control      │
│ without time    │              └─────────────────┘
│ shift           │                       │
└─────────────────┘                       ▼
         │                       ┌─────────────────┐
         ▼                       │ Pick first      │
┌─────────────────┐              │    arrivals     │
│ Calculate       │              └─────────────────┘
│ Product of      │                       │
│ envelopes of    │                       ▼
│ stacked         │              ┌─────────────────┐
│ waveforms       │              │ Invert for a 1D-│
└─────────────────┘              │ piecewise       │
         │                       │ velocity model  │
         ▼                       │ accounting for  │
┌─────────────────┐              │ topography      │
│ Run STA/LTA on  │              └─────────────────┘
│ the product     │                       │
│ function        │                       ▼
└─────────────────┘              ┌─────────────────┐
         │                       │ Average residuals│
         ▼                       │ as station      │
┌─────────────────┐◄─────────────│ corrections to  │
│ Perform beam-   │              │ mitigate 3D     │
│ forming and     │              │ effects         │
│ slant stacking  │              └─────────────────┘
│ on STA/LTA      │
│ triggers with   │
│ 5s time windows │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Find detections │
│ with concentrated│
│ beam-forming    │
│ energy          │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Check quality of│
│ slant stacked   │
│ waveforms       │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Validation: check│
│ P and S arrivals│
│ at regional     │
│ stations        │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Locate validated│
│ detections      │
└─────────────────┘
```

Figure 2. A flow diagram summarizing the main steps in the detection method.

10th traces in Figs 3(a) and (b) labeled 'all' show that some of these latter signals are also present in a direct stack of data across the entire array.

Next we calculate the envelopes of the stacked waveforms within each of the nine subarrays using the square root of $v(t)^2 + H[v(t)]^2$, where $v(t)$ is the original signal and $H[v(t)]$ is its Hilbert transform. The resulting non-negative envelope functions are normalized for each subarray to the range 0 to 1 during the entire day. Multiplying the nine envelopes functions with each other produces a signal referred to as Product function in the range 0 to 1 that suppresses
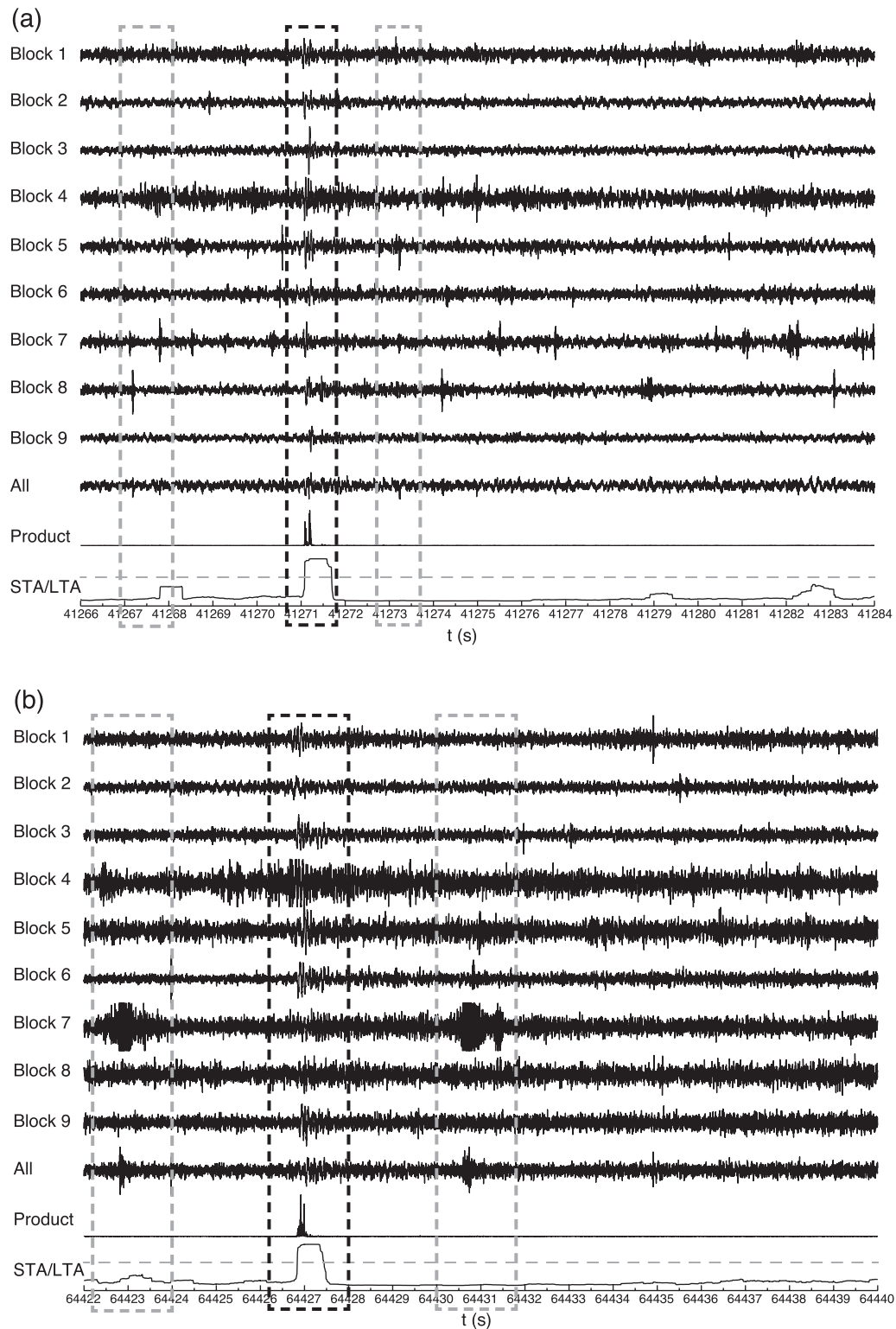
Figure 3. (a) Example associated with a time window on the test day indicated on the horizontal axes. The first nine time-series displays plain stacked waveforms within each subarray, while the 10th trace is the stacked waveform across the entire array. The 11th trace is the Product of the envelopes of the top nine traces. The bottom trace displays the result of running an STA/LTA on the Product function with a threshold for detection indicated by the horizontal grey dashed line. The vertical dashed black box highlights the detection, while the vertical dashed grey box shows strong signals in only some subarrays. (b) Same as (a) for another example time window. See the text for additional details.

significantly large-amplitude signals present only in the stacks of some of the nine subarrays. As seen in the 11th traces of Fig. 3, the Product functions have clear signals for the time windows associated with the black dashed boxes and have very low amplitude in all other time intervals. The bottom traces in Fig. 3 display results of running a standard detector on the Product function consisting of the ratio (e.g. Allen 1978) of a short-term moving average (STA) to a long-term moving average (LTA). Since we target events that are very small and relatively deep, we use 1 s for the short time window and 10 s for the long one. The STA/LTA signals have significant peaks only in the time intervals associated with the black dashed boxes.

Fig. 4(a) displays the Product function for the entire test day using a log scale. There are clear overall daily variations of the signals, with lower level at night time (from 22:00 to 06:00 local time), and sharp spikes that indicate possible detections. Figs 4(b) and (c) present zoomed-in views of one night and one day 10-min time windows. Fig. 4(d) gives the frequency–size statistics of the amplitudes of the Product function in Fig. 4(a). The results span 25 orders of magnitudes and resemble overall typical frequency–size statistics of earthquakes, with approximate power-law distribution over about 13 orders of magnitudes. To decide on potential detections, we run an STA/LTA detector on the Product function with the time intervals mentioned above. The results are shown in Fig. 5 along with a threshold (horizontal dashed red line) selected arbitrarily at a level of 5 times the median of the calculated STA/LTA values. This threshold choice leads to 723 potential detections (triggers) of small events tested further below using a beamforming approach.

The beamforming is performed together with slant stacking where, for each of the 723 potential detections, 5 s waveforms recorded by the entire array are shifted, aligned and stacked for different possible incident directions. The relatively short duration of the waveforms is appropriate for small target events, and each waveform starts 1 s before and ends 4 s after a trigger. The analysis employs a detailed 1-D piecewise P-wave velocity model for the shallow crust consisting of a surface layer of variable width over a half-space. The velocity model is derived from data associated with the Betsy gunshots and presented in Appendix A. The beamforming assumes a plane P wave with incident angle i and backazimuth baz at the top of the bottom medium (Fig. A1b and A3a), and uses second root slant stacking to suppress noise and enhance the coherent signals (Rost & Thomas 2002). Specifically, we calculate

$$u'(t; i, baz) = \frac{1}{N} \sum_{j=1}^{N} |v_j(t - t_j(i, baz))|^{1/2} \cdot \text{sgn}(v_j) \quad (1)$$

$$u(t; i, baz) = |u'(t; i, baz)|^2 \cdot \text{sgn}(u') \quad (2)$$

where N is the number of the array stations and $v_j(t)$ represents the seismogram at station j in a 5 s time window with 1 s before the trigger and 4 s after. The S-wave energy is suppressed in the stacking, since we use the local P-wave velocity model and residuals at each station (Appendix A) to align possible P arrivals. The total energy recorded by the array for each 5 s time window associated with each trigger is calculated by integrating the square of the signal obtained from the slant-stack procedure (eq. 2)

$$E(i, baz) = \int_{t_k - t_1}^{t_k + t_2} u^2(t; i, baz) dt \quad (3)$$

where t_k is the STA/LTA trigger time, $t_1 = 1$ s and $t_2 = 4$ s. The calculations are repeated in a grid search over the incident angle and backazimuth in the lower half-space with a $2°$ spacing. The

corresponding energy values calculated with eq. (3) are normalized using

$$e(i, baz) = \frac{E(i, baz)}{\max(E(i, baz))}. \quad (4)$$

Figs 6(a) and (c) present beamforming diagrams for the two STA/LTA triggers in Fig. 3, where each point represents the integrated power for given incidence angle and backazimuth, and the white crosses denote the strongest energy. The associated second-root-slant-stacked waveforms have considerably less noise than the corresponding linear slant stacks (Figs 6b and d). To evaluate the quality of a possible detection, we integrate the beamforming results for each trigger for directions with high energy on a lower hemisphere with unit radius using

$$ratio = \frac{1}{2\pi} \iint_{e \geq 0.8} dbaz \, di \sin(i) \, e(i, baz). \quad (5)$$

The integration is normalized by 2π so the ratios span the range 0 to 1. Low ratios associated with concentrated power distribution are taken to represent high-quality detections. Fig. 7(a) presents the histogram of the calculated ratios for all 723 STA/LTA triggers. Defining, for example, high-quality results to have ratios less than 0.1 leads to 220 detections. These detections include all 18 events during the test day found with standard techniques using the the Southern California Seismic Network (SCSN), ANZA network and a dense local PASSCAL deployment (YN) in and around the SJFZ (Vernon & Ben-Zion 2010). Using a less strict criterion associated with 0.25 leads to 339 detections. Figs 7(b) and (c) show, respectively, the azimuth distribution and incident directions of the detections associated with ratios ≤ 0.1.

As a validation step of the detection analysis, we check the waveforms at stations of the regional southern California and ANZA networks within ± 40 s from the predicted arrival time of each of the 220 high-quality detections. The predicted arrival times are calculated using the velocity model of Fig. A3 and the waveforms at regional stations are bandpass filtered from 2 to 15 Hz. Figs 8(a) and (b) display clear seismic signals at regional stations (red triangles at inset maps and name on left) close in time to predicted arrivals associated with the two STA/LTA triggers shown in Fig. 3 (dashed vertical red line). Of the 220 high-quality detections, 103 are confirmed as seismic events by finding corresponding phases at four or more seismic stations. The number of confirmed events is 5 times larger than the events in the ANZA and SCSN catalogues during the test day; many more of the 723 STA/LTA triggers in Fig. 5 may also be genuine seismic events.

To locate approximately the validated events, we manually pick the P or S arrivals and adopt the method of Geiger (1912), associated with ray tracing in a 1-D model obtained by averaging the 3-D velocity model of Lee *et al*. (2014) for the region shown in the top left inset of Fig. 1. To improve the convergence and stability of the algorithm, we grid search the focal depth with a 0.5 km step and derive tentative epicentre and origin time for each focal depth. The initial epicentre is set to the horizontal location of the station with the earliest P or S arrival and the initial origin time is set to the earliest P or S arrival. The location generating the minimum L2 norm misfit in the grid search is taken to be the source position. Fig. 9 presents the obtained locations of the validated events. The location uncertainties are estimated to be on the average $\sim$4 km horizontally and $\sim$4 km vertically considering the misfit curves and reasonable velocity variations from the adopted average 1-D model. Fig. 9 also shows estimated magnitudes of the validated detections, based on relative amplitudes to small reference events with available
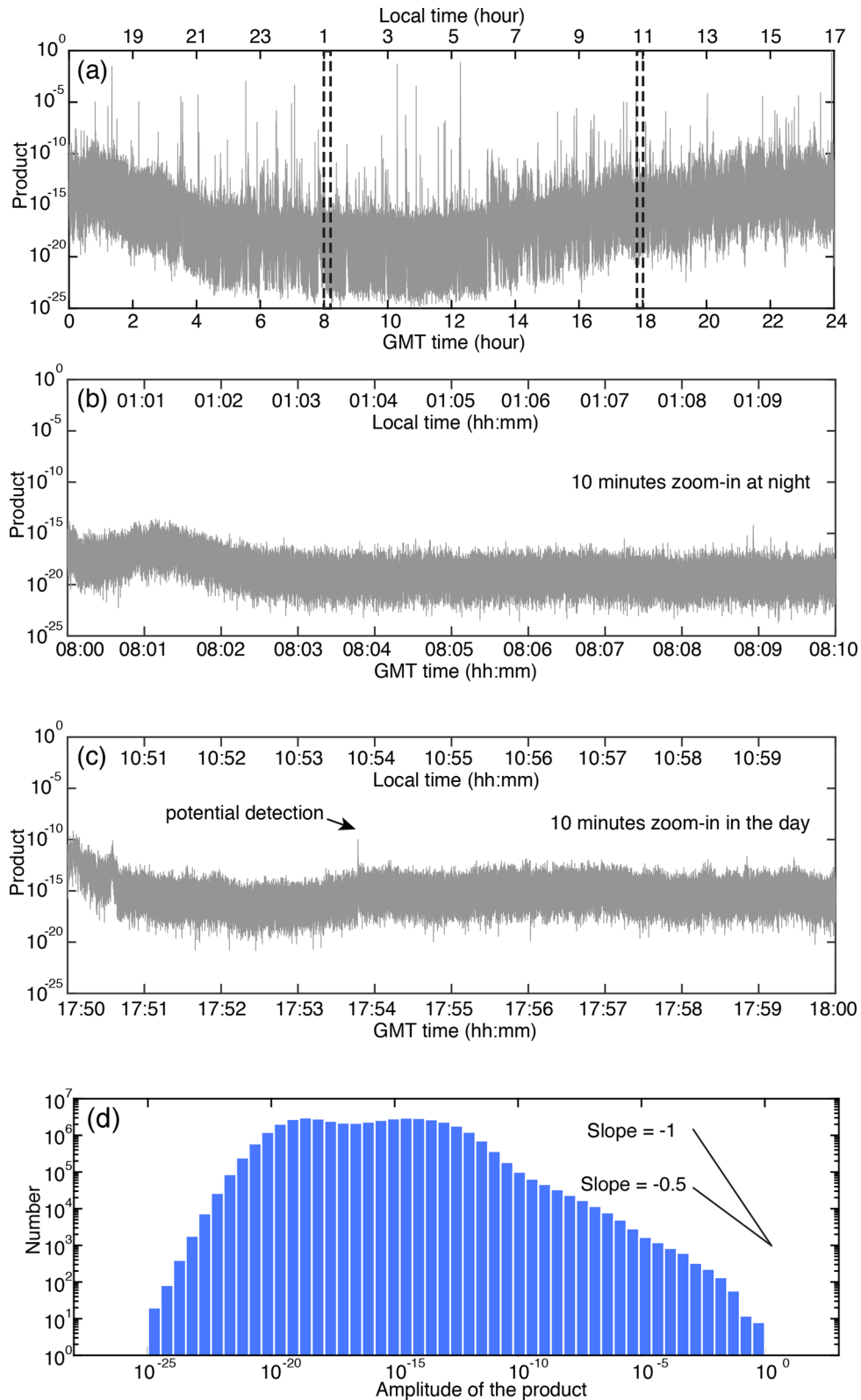
Figure 4. (a) The Product function in the range 0–1 for the test day with amplitudes in a log scale. The dashed black boxes are night and day 10 min intervals, shown with zoomed-in view in (b) and (c), respectively. (d) Histogram of Product function amplitudes. Note that the horizontal scale covers 25 orders of amplitudes.
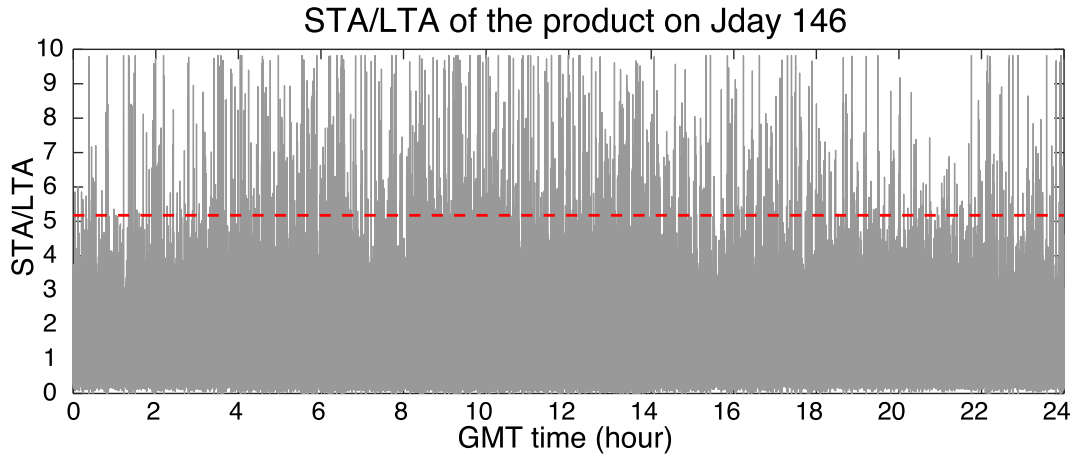
Figure 5. Results of running an STA/LTA on the Product function. The red dashed line denotes five times the daily median used as a threshold to accept STA/LTA triggers.
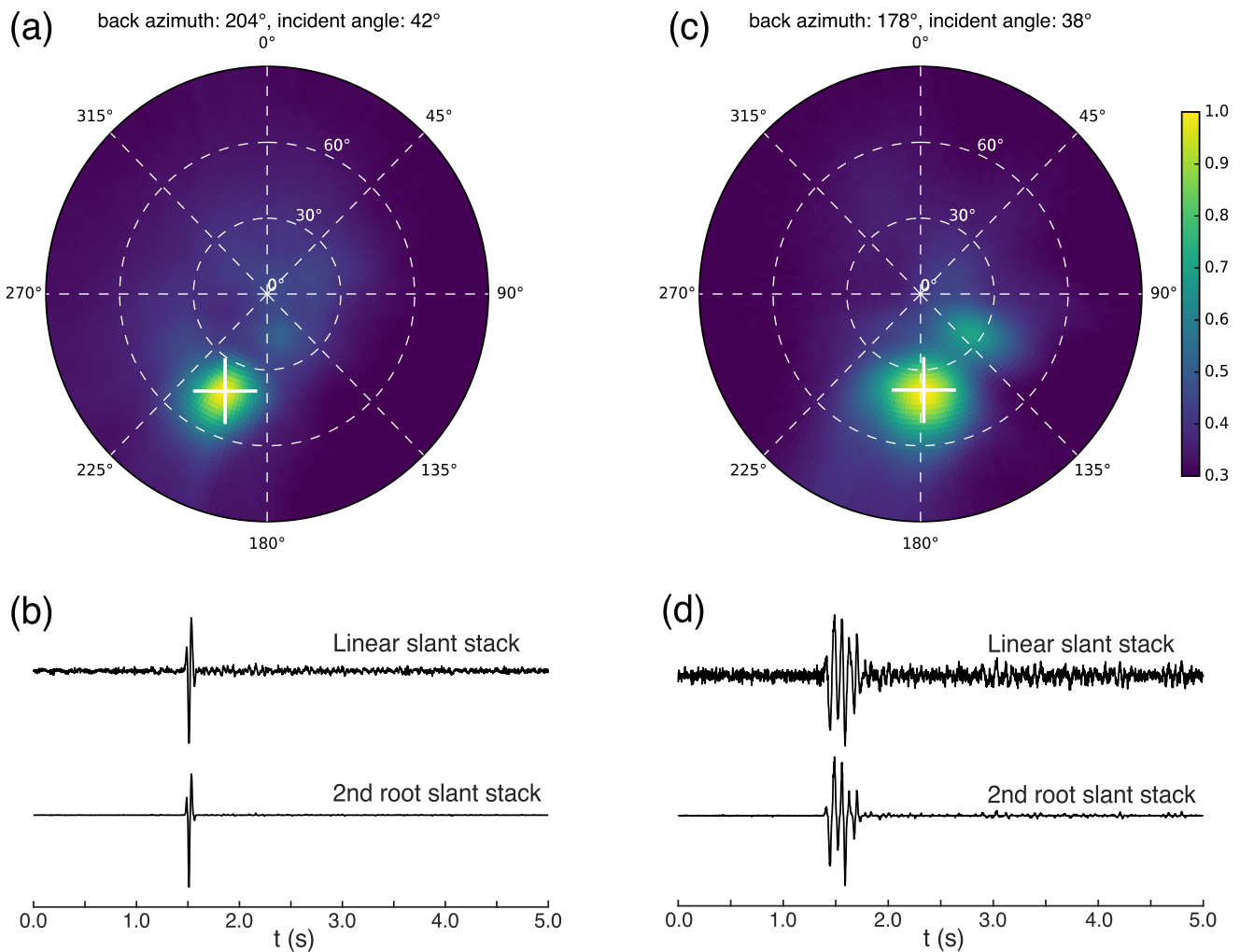
Figure 6. (a) A beamforming diagram showing the energy (eqs 3 and 4) of slant-stacked waveform within 5 s around the STA/LTA trigger of Fig. 3(a). The radii on the polar plot denote incident angles and the numbers on the outer circle indicate backazimuth angles. The white cross marks the maximum energy. (b) Slant-stack waveforms corresponding to the incident angle and backazimuth of the white cross. (c) and (d) Same as (a) and (b) for the STA/LTA trigger of Fig. 3(b). Note that the incident angle is defined on the top of bottom medium (Fig. A1b). The 1-D model used for beamforming is displayed in Fig. A3(b).
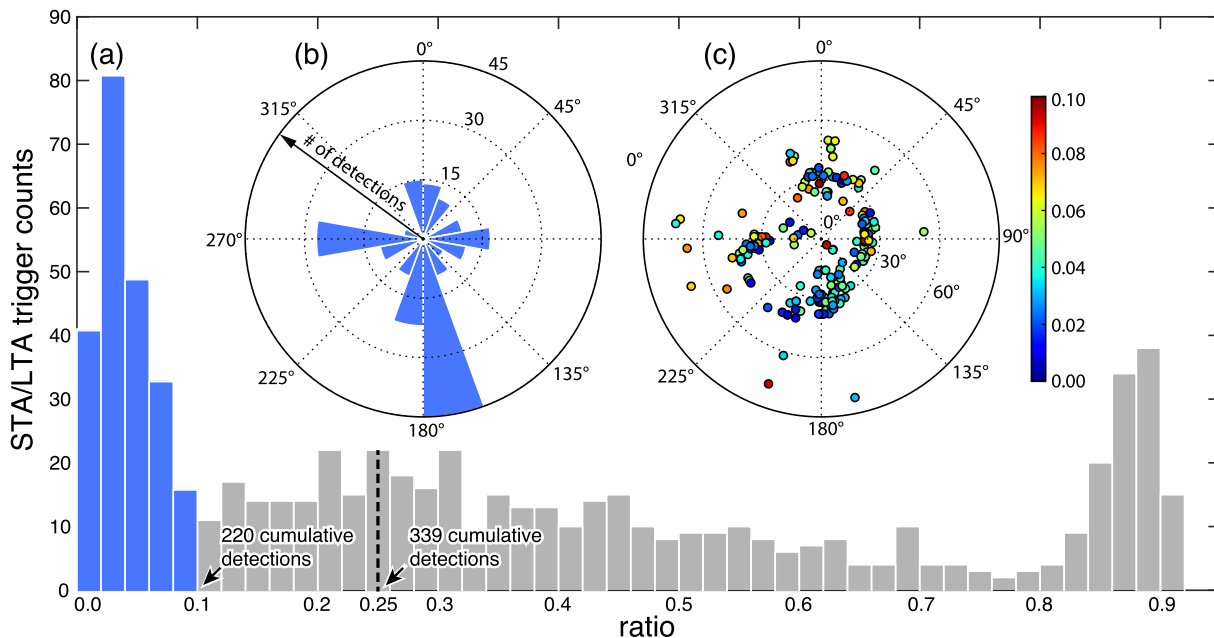
Figure 7. (a) A histogram of all 723 STA/LTA triggers for the test day. The ratio value on the horizontal axis indicates the quality of the detection in terms of the degree of localization in the beamforming diagram (Fig. 6). Blue bars denote high quality accepted STA/LTA triggers. (b) A rose diagram of the backazimuths of the 220 high-quality triggers. (c) The incident angle and backazimuth of high-quality triggers with ratio values denoted by colour.

catalogue magnitudes and focal mechanisms (Appendix B). Most detected events are very small with magnitudes around or below zero, but the results include also events with magnitudes up to 2.5 that are not detected (for some reason) by the standard techniques producing the regional catalogue.

3 DISCUSSION AND CONCLUSIONS

We developed a multistep procedure for detecting small earthquakes with continuous waveforms recorded by a dense seismic array, and illustrated the technique with data recorded by 1108 sensors in a tight configuration around the Clark branch of the SJFZ (Fig. 1). The different analysis steps (Fig. 2) are designed to suppress local sources affecting only subarrays (likely associated with surface and/or instrumental noise), and then detecting, validating and locating sources affecting the entire array. The implementation involves stacking data of nine subarrays without migration, multiplying the envelopes of the stacked subarray data, and performing the remaining steps on the Product function. Sources detected on the Product function have coherent energy across the entire array so are likely to be associated with events below the surface. The validation and location steps of the method demonstrate that at least a subset of these detections (103 during the test day) is indeed associated with small earthquakes in the region.

The choice of using nine subarrays in this study stems from the observed significant amplitude decay of Betsy gunshots over distance of 100 m (Fig. S1). Using fewer subarrays would lead to reduced suppression of surface sources, many more STA/LTA triggers of local sources and considerably more computation in the following beamforming step. On the other hand, using more subarrays would result in fewer stations in each subarray and reduced suppression of random noise. Appendix C presents a derivation of an optimal number of subarrays for detecting weak signals arriving with large incident angle (corresponding to deep events with general locations not directly below the array). The derivation assumes

stacking in subarrays without time-shifts and subsequent analysis of the Product function as in our procedure. The choice of 3×3 subarrays made in this study helps to maximize the array sensitivity to weak signals with large incident angles and facilitates the detection of more events. We recall that the direct stacking of data in each subarray in the early analysis procedure is motivated by our focus on deep events and helps to suppress surface and shallow sources. Efforts to detect shallow small events should incorporate migration in each subarray before stacking to enhance the Product function.

The detection procedure includes beamforming and slant staking using a 1-D piecewise local *P*-wave velocity model for the subsurface material (Fig. 2). The delay times for each pair of incident angle and backazimuth are calculated only once using eq. (A23) and then used for slant staking and beamforming for all triggers, making the process computationally efficient. In our case, computing the beamforming on a CPU node with 2 GB memory requires ∼3 min for one trigger and under 40 hr for all triggers. The model consists of a layer with a constant velocity gradient over a half space, and is derived in Appendix A from arrivals of Betsy gunshots and analytical expressions with four parameters. Although the model is simple, it accounts for topographic variations across the array and it explains the data well (Figs A3 and A4). Residuals from model predictions provide site corrections at each station that are incorporated in the beamforming and slant staking. The velocity variations across the array are consistent with features observed in previous imaging studies based on the array data (Ben-Zion *et al.* 2015; Hillers *et al.* 2016; Roux *et al.* 2016). The rapid velocity variations with depth in the very shallow crust (*P* velocities increasing from ∼550 to ∼1250 m s^{-1} in the top ∼100 m and then jumping to ∼2900 m s^{-1}) have important implications for properties of the subsurface material (e.g. Sleep 2011; Kurzon *et al.* 2014) and can be useful for future imaging and detection studies in the area.

The analysis procedure reduces the data redundancy of the dense array and zooms in progressively on small earthquakes (Figs 3–7) by calculating STL/LTA on the Product function representing
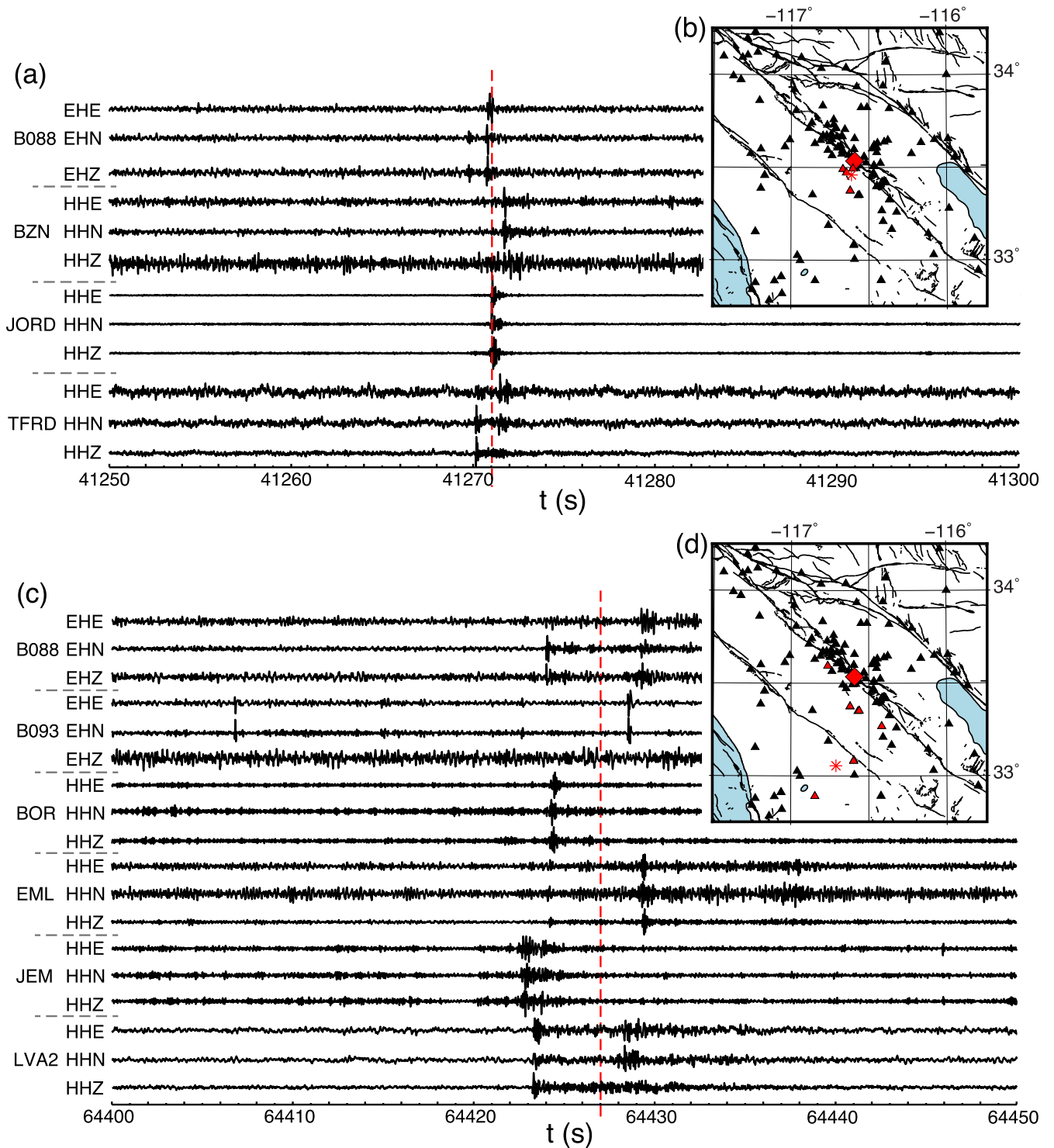
Figure 8. (a) Seismic records at several stations of the southern California and Anza networks around STA/LTA trigger time (red dashed line) in Fig. 3(a). The waveforms are self-normalized and bandpassed from 2 to 15 Hz. Station names and components are marked on the left. The inset in (b) shows with red triangles stations with clear phases. (c) and (d) Same as (a) and (b) for waveforms in a time window around STA/LTA trigger time in Fig. 3(b).

coherent signals across the array and performing beamforming on time windows around the STL/LTA triggers. The method generates 723 triggers during the test day, of which 220 are shown to be associated with localized energy sources below the surface (Fig. 7). Of these, 103 are validated by showing that they produce earthquake signals in four or more regional stations. Figs 3, 6 and 8 illustrate

with two examples STL/LTA triggers how detections are validated as genuine seismic events. The remaining 117 localized energy sources are also likely to be associated with seismic events, too small to be recorded by ≥4 standard seismic stations. The same likely holds for many of the ∼500 additional triggers in the test day. The slant stacking performed in this study is based on the local
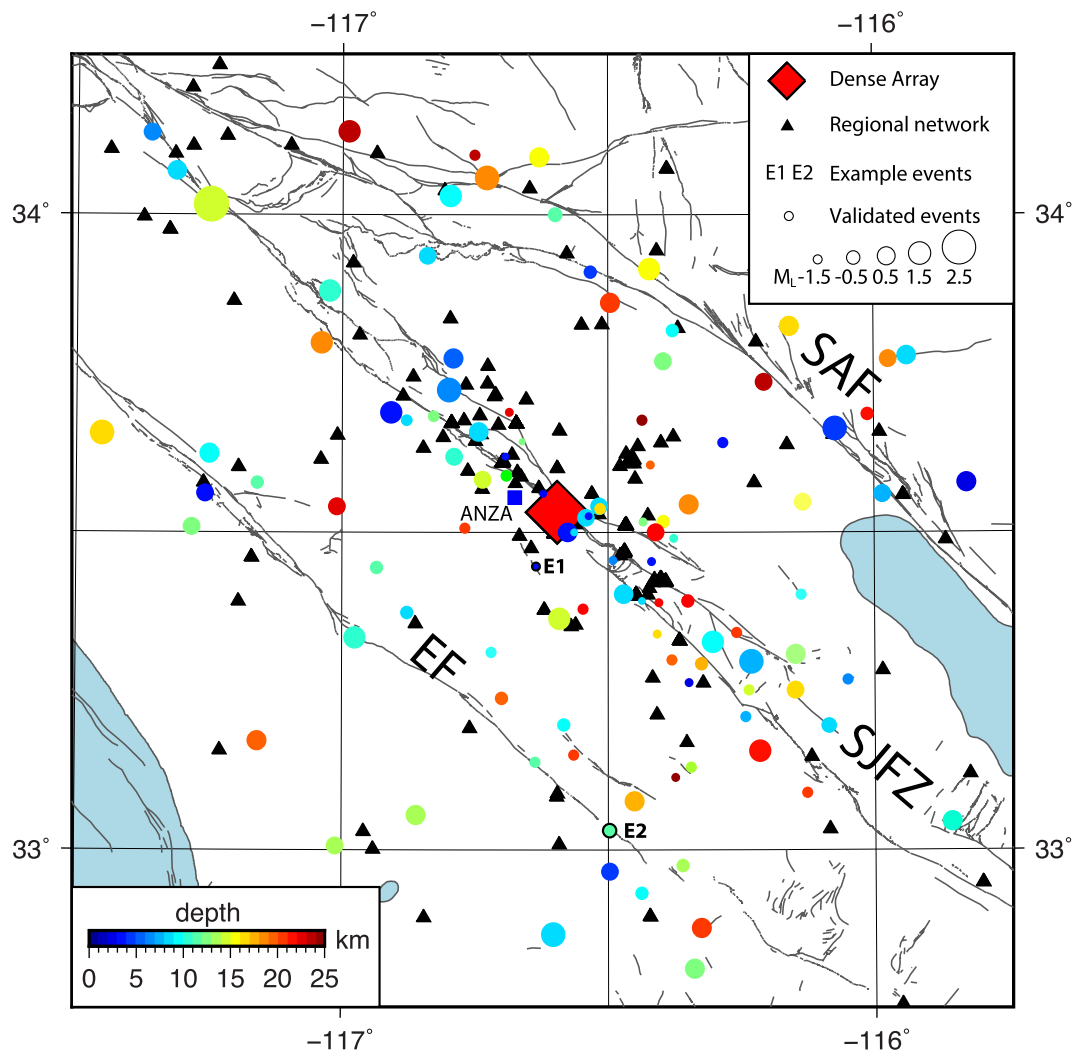
Figure 9. Approximate locations (colour) and magnitudes (circle size) of the 103 validated events. The bold circles mark the locations of the two example events E1 and E2 in Figs 3, 6 and 8. See the text for additional details.

P-wave velocity model and reduces signals associated with *S* waves. Performing in parallel beamforming and slant stacking associated with an *S*-wave velocity model can lead to additional detections and validations. Using the detected/validated events as templates will likely increase the number of detections by a factor of 20 or more (e.g. Shelly *et al.* 2016; Ross *et al.* 2017).

The approximate locations of the validated events (Fig. 9) are calculated using an average 1-D velocity model based on the 3-D tomographic results of Lee *et al.* (2014) for southern California. The event depths range from ∼3 to ∼25 km. The backazimuths of the events are generally consistent with the incident directions in Figs 7(b) and (c) with some discrepancies. For instance, the locations of the example events in Figs 6(a) and (c) are marked in Fig. 9. The backazimuth of the first event is within 2° to that estimated by the beamforming (Fig. 6a), while for the second event the discrepancy is 15° (Fig. 6c). The discrepancies reflect the significant variations of seismic velocities around the SJFZ (e.g. Allam & Ben-Zion 2012; Zigone *et al.* 2015). The uncertainties of event locations are relatively large because of the employed average 1-D velocity model and the fact that many events have only a limited number of picks. The magnitudes of the validated events are estimated by calculating relative amplitudes to reference small catalogue events near the

hypocentres of the detected events, with attention to likely radiation pattern effects on amplitudes (Appendix B).

The methodology and results of this paper can be improved in several ways. The locations can be refined by using a 3-D velocity model and attempting to obtain additional high-quality picks. The study can be extended by analysing more data and using both *P* and *S* waves in the detection stage. Another useful future research direction would be an effort to detect very shallow events. Focusing on shallow events can benefit from using a shorter STA interval (e.g. 0.1 s or less) and incorporating migration in the initial stacking of the subarrays. These additional improvements notwithstanding, the study demonstrates that dense arrays of the type used in this work can form 'supersites' that allow the creation of many new validated templates (>100 per day in the study area). Combining this additional set of small events not included in regular catalogues with standard matched-filter technique can increase the detection of small events further. We note that the magnitudes and source–receiver distances of the validated detections significantly outperform the theoretical limits to detection of Kwiatek & Ben-Zion (2016) based on simulations of wave propagation from various sources and assumed SNR > 1 at individual sensors, because the stacking enhances considerably the SNR of the array. A future

network configuration consisting of several spatially dense super-sites can increase significantly the detectibly of small events. Data from such a network will allow tracking active seismicity structures with much greater detail and increasing considerably the resolution of tomographic and other earthquake-based studies.

REFERENCES

Allam, A.A. & Ben-Zion, Y., 2012. Seismic velocity structures in the southern California plate-boundary environment from double-difference tomography, *Geophys. J. Int.,* **190,** 1181–1196.

Allen, R.V., 1978. Automatic earthquake recognition and timing from single traces, *Bull. seism. Soc. Am.,* **68,** 1521–1532.

Ben-Zion, Y. *et al.* 2015. Basic data features and results from a spatially dense seismic array on the San Jacinto fault zone, *Geophys. J. Int.,* **202,** 370–380.

Corciulo, M., Roux, P., Campillo, M. & Dubucq, D., 2012. Instantaneous phase variation for seismic velocity monitoring from ambient noise at the exploration scale, *Geophysics,* **77,** Q37–Q44.

Cros, E., Roux, P., Vandemeulebrouck, J. & Kedar, S., 2011. Locating hydrothermal acoustic sources at Old Faithful Geyser using Matched Field Processing, *Geophys. J. Int.,* **187,** 385–393.

Fink, M., 2006. Time-reversal acoustics in complex environments, *Geophysics,* **71,** SI151–SI164.

Geiger, L., 1912. Probability method for the determination of earthquake epicenters from the arrival time only, *Bull. St. Louis Univ,* **8**(1), 56–71.

Gibbons, S.J. & Ringdal, F., 2006. The detection of low magnitude seismic events using array-based waveform correlation, *Geophys. J. Int.,* **165,** 149–166.

Hillers, G., Roux, P., Campillo, M. & Ben-Zion, Y., 2016. Focal spot imaging based on zero lag cross-correlation amplitude fields: application to dense array data at the San Jacinto fault zone, *J. geophys. Res.,* **121,** 8048–8067.

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, *Geophys. Res. Lett.,* **42,** 6314–6323.

Kurzon, I., Vernon, F.L., Ben-Zion, Y., Atkinson & G., 2014. Ground motion prediction equations in the San Jacinto Fault Zone—significant effects of rupture directivity and fault zone amplification, *Pure appl. Geophys.,* **171,** 3045–3081.

Kwiatek, G. & Ben-Zion, Y., 2016. Theoretical limits on detection and analysis of small earthquakes, *J. geophys. Res.,* **121,** 5898–5916.

Larmat, C., Tromp, J., Liu, Q. & Montagner, J.P., 2008. Time reversal location of glacial earthquakes, *J. geophys. Res.,* **113,** doi:10.1029/2008JB005607

Lee, E.J., Chen, P., Jordan, T.H., Maechling, P.B., Denolle, M.A. & Beroza, G.C., 2014. Full-3-D tomography for crustal structure in southern California based on the scattering-integral and the adjoint-wavefield methods, *J. geophys. Res.,* **119,** 6421–6451.

Meng, X.F., Peng, Z.G. & Hardebeck, J.L., 2013. Seismicity around Parkfield correlates with static shear stress changes following the 2003 M(w)6.5 San Simeon earthquake, *J. geophys. Res.,* **118,** 3576–3591.

Nakata, N. & Beroza, G.C., 2015. Stochastic characterization of mesoscale seismic velocity heterogeneity in Long Beach, California, *Geophys. J. Int.,* **203,** 2049–2054.

Ross, Z.E., Hauksson, E. & Ben-Zion, Y., 2017. Abundant off-fault seismicity and orthogonal structures in the San Jacinto fault zone, *Sci. Adv.,* **3,** e1601946, doi:10.1126/sciadv.1601946

Rost, S. & Thomas, C., 2002. Array seismology: methods and applications, *Rev. Geophys.,* **40**(3), 1008, doi:10.1029/2000RG000100.

Roux, P., Moreau, L., Lecointre, A., Hillers, G., Campillo, M., Ben-Zion, Y., Zigone, D. & Vernon, F., 2016. A methodological approach towards high-resolution surface wave imaging of the San Jacinto Fault Zone using ambient-noise recordings at a spatially dense array, *Geophys. J. Int.,* **206,** 980–992.

Shelly, D.R., Ellsworth, W.L. & Hill, D.P., 2016. Fluid-faulting evolution in high definition: connecting fault structure and frequency-magnitude variations during the 2014 Long Valley Caldera, California, earthquake swarm, *J. geophys. Res.,* **121,** 1776–1795.

Slawinski, R.A. & Slawinski, M.A., 1999. On raytracing in constant velocity-gradient media: calculus approach, *Can. J. Explor. Geophys,* **35,** 24–27.

Sleep, N.H., 2011. Seismically damaged regolith as self-organized fragile geological feature, *Geochem. Geophys. Geosyst.,* **12,** doi:10.1029/2011GC003837.

Turek, G. & Kuperman, W., 1997. Applications of matched-field processing to structural vibration problems, *J. acoust. Soc. Am.,* **101,** 1430–1440.

Vernon, F. & Ben-Zion, Y., 2010. *San Jacinto Fault Zone Experiment*. International Federation of Digital Seismograph Networks, Other/Seismic Network.

Yang, W., Hauksson, E. & Shearer, P. M. 2012, Computing a large refined catalog of focal mechanisms for southern California (1981-2010): temporal stability of the style of faulting, *Bull. seism. Soc. Am.,* **102,** 1179–1194.

Zigone, D., Ben-Zion, Y., Campillo, M. & Roux, P., 2015. Seismic tomography of the Southern California plate boundary region from noise-based Rayleigh and Love waves, *Pure appl. Geophys.,* **172,** 1007–1032.

APPENDIX A: A 1-D PIECEWISE *P*-WAVE VELOCITY MODEL FOR THE VERY SHALLOW CRUST

A1 Ray tracing in a two-media model

The *P*-wave structure beneath the dense deployment is simplified as a surface layer over a half-space with a constant velocity gradient b in the top layer and a constant velocity w_0 in the bottom medium (Fig. A1). Downgoing waves in the top layer curve back to the surface and the constant velocity in the bottom solid is sufficient to account for the incident plane wave assumption of the beamforming approach. The remaining two parameters are the velocity v_0 at the bottom of the surface layer and the depth z_0 of the interface. The velocity v_{st} at a station is a function of z, so the topography (grey lines in Fig. A1a) is considered in the ray tracing.

A1.1 Downgoing wave in a solid with constant velocity gradient

We consider a model in which the *P*-wave velocity v increases linearly with depth z (Fig. A1a). By referring to the epicentre distance and traveltime of upgoing wave derived by Slawinski and Slawinski (1999), the epicentre distance X of downgoing wave in the top layer is

$$X = \frac{1}{bp} \left(\sqrt{1 - (pv_{src})^2} + \sqrt{1 - (pv_{st})^2} \right), \qquad (A1)$$

where p is the constant ray parameter, b is the constant velocity gradient in the top layer and v_{src} and v_{st} are the velocities at source and station positions. Solving eq. (A1) for p and ignoring non-physical solutions gives

$$p = \frac{2bX}{\sqrt{\left[(bX)^2 + v_{src}{}^2 + v_{st}{}^2 \right]^2 - (2v_{src}v_{st})^2}}. \qquad (A2)$$
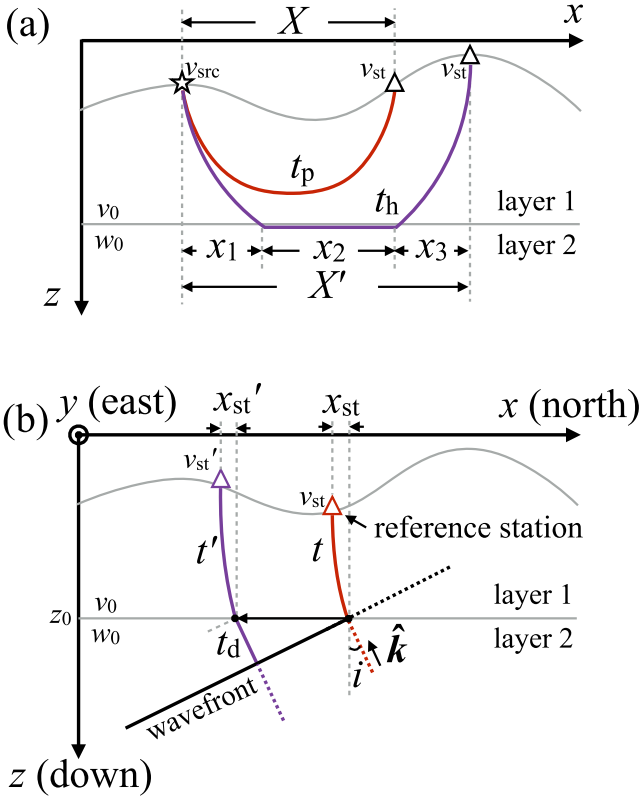
Figure A1 . A framework for deriving from Betsy gunshot data a shallow velocity model using a surface layer with topography denoted by grey lines over a half-space. (a) A diagram for ray tracing of a regular body wave (red line) and arrival associated with a head wave at the bottom of the layer (purple line). The star and triangles denote the source and two example stations, respectively. (b) A diagram for arrivals at two stations from a plane wave (bold black line) at the bottom layer. The red and purple curves denote rays at reference and target stations, respectively. See the text for additional explanations.

Thus, the traveltime for downgoing wave is

$$t_p = \frac{1}{b} \left| \ln \left[\frac{1 - \sqrt{1 - (pv_{\mathrm{src}})^2}}{pv_{\mathrm{src}}} \right] \right| + \frac{1}{b} \left| \ln \left[\frac{1 - \sqrt{1 - (pv_{\mathrm{st}})^2}}{pv_{\mathrm{st}}} \right] \right|. \tag{A3}$$

A1.2 Head wave

As the epicentre distance increases, the ray can refract along the interface between the media before turning up and reaching the surface (Fig. A1a). In this case, the first arrival is a head wave with ray parameter $p = w_0^{-1}$, where w_0 is the constant velocity in the bottom solid. For the downgoing segment of the head wave, the horizontal distance (x_1 in Fig. A1a) is

$$x_1 = \frac{1}{bp} \left(\sqrt{1 - (pv_{\mathrm{src}})^2} - \sqrt{1 - (pv_0)^2} \right), \tag{A4}$$

where v_0 is the velocity at the bottom of the top layer. The corresponding traveltime is

$$t_1 = \frac{1}{b} \left| \ln \left[\frac{v_0}{v_{\mathrm{src}}} \frac{1 - \sqrt{1 - (pv_{\mathrm{src}})^2}}{1 - \sqrt{1 - (pv_0)^2}} \right] \right|. \tag{A5}$$

For the upgoing segment, the horizontal distance (x_3 in Fig. A1a) and corresponding traveltime are

$$x_3 = \frac{1}{bp} \left(\sqrt{1 - (pv_{\mathrm{st}})^2} - \sqrt{1 - (pv_0)^2} \right) \tag{A6}$$

and

$$t_3 = \frac{1}{b} \left| \ln \left[\frac{v_0}{v_{\mathrm{st}}} \frac{1 - \sqrt{1 - (pv_{\mathrm{st}})^2}}{1 - \sqrt{1 - (pv_0)^2}} \right] \right|. \tag{A7}$$

For the ray segment at the interface between the media and epicentre distance X, the horizontal distance (x_2 in Fig. A1a) and corresponding traveltime are

$$x_2 = X - x_1 - x_3 \tag{A8}$$

and

$$t_2 = px_2. \tag{A9}$$

If $x_2 > 0$, the ray is a head wave; otherwise, the ray just propagates in the top layer. The total traveltime for the head wave is

$$t_h = t_1 + t_2 + t_3. \tag{A10}$$

A1.3 Traveltime from an incident plane wave at the bottom of the surface layer

To apply beamforming approach to data, we need traveltimes for each station from an incident plane wave. In the Cartesian coordinate system defined in Fig. A1(b), the location of a reference station is

$$r_{\mathrm{st}} = [\, x \ y \ z \,] \tag{A11}$$

while the location of a target station is

$$r_{\mathrm{st}}' = [\, x' \ y' \ z' \,]. \tag{A12}$$

For a plane wave with backazimuth *baz* and incident angle i, the ray parameter is $p = w_0^{-1} \sin i$ and the incident direction vector is

$$\hat{k} = [\, -\cos baz \, \sin i \ \ -\sin baz \, \sin i \ \ -\cos i \,]. \tag{A13}$$

The horizontal offsets between the intersections on the interface and the surface stations (x_{st} and x_{st}' in Fig. A1b) are

$$x_{\mathrm{st}} = \frac{1}{bp} \left(\sqrt{1 - (pv_{\mathrm{st}})^2} - \sqrt{1 - (pv_0)^2} \right) \tag{A14}$$

and

$$x_{\mathrm{st}}' = \frac{1}{bp} \left(\sqrt{1 - (pv_{\mathrm{st}}')^2} - \sqrt{1 - (pv_0)^2} \right). \tag{A15}$$

The corresponding traveltimes in the top layer are

$$t = \frac{1}{b} \left| \ln \left[\frac{v_0}{v_{\mathrm{st}}} \frac{1 - \sqrt{1 - (pv_{\mathrm{st}})^2}}{1 - \sqrt{1 - (pv_0)^2}} \right] \right| \tag{A16}$$

and

$$t' = \frac{1}{b} \left| \ln \left[\frac{v_0}{v_{\mathrm{st}}'} \frac{1 - \sqrt{1 - (pv_{\mathrm{st}}')^2}}{1 - \sqrt{1 - (pv_0)^2}} \right] \right|. \tag{A17}$$

For nearly vertical incident plane wave at the bottom layer, the ray parameter $p \to 0$. The limits of eqs (A16) and (A17) for this case are

$$\lim_{p \to 0} t = \frac{1}{b} \left| \ln \left(\frac{v_{\mathrm{st}}}{v_0} \right) \right| \tag{A18}$$
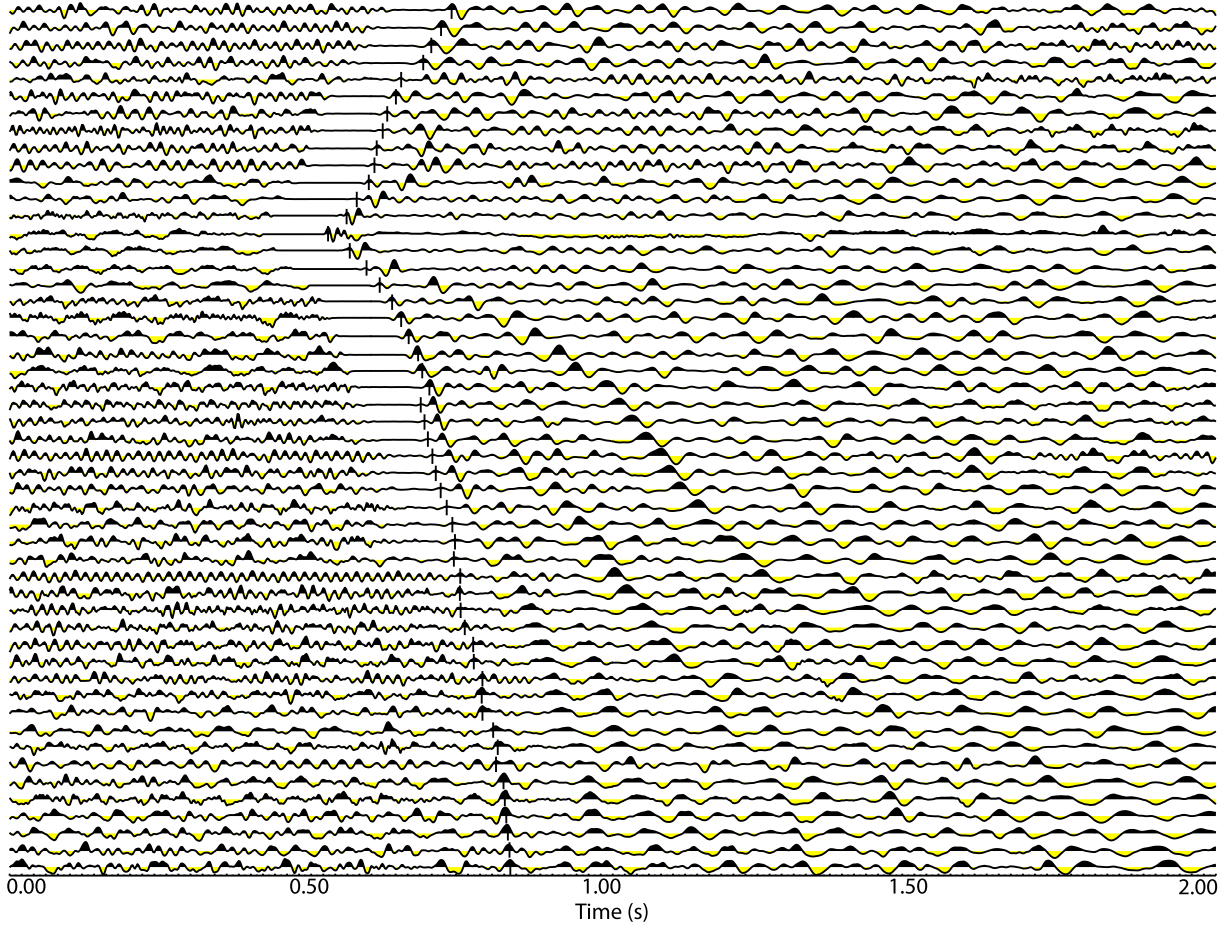
Figure A2. Waveforms with automatic gain control, generated by Betsy gunshot no. 1 (Fig. S1) and recorded at sensors along row 7 of the array. First arrivals are marked by short vertical line segments.

and

$$\lim_{p \to 0} t\prime = \frac{1}{b} \left| \ln \left(\frac{v'_{st}}{v_0} \right) \right|. \tag{A19}$$

The intersections for both rays at the interface are

$$r_{itsc} = \left[x + x_{st} \cos baz \ y + x_{st} \sin baz + z_0 \right] \tag{A20}$$

and

$$r'_{itsc} = \left[x' + x'_{st} \cos baz \ y' + x'_{st} \sin baz + z_0 \right]. \tag{A21}$$

For the target station, the traveltime t_d for the ray segment in the bottom layer is

$$t_d = \frac{\left(r'_{itsc} - r_{itsc} \right) \cdot \hat{k}}{w_0}. \tag{A22}$$

Thus, the delay time between arrivals at the target and reference stations is

$$\Delta t = t' + t_d - t. \tag{A23}$$

A2 Inversion of data for a 1-D model

For a given model $m_0 = [b \ v_0 \ w_0 \ z_0]^T$, the traveltime of the first P arrivals $T = T(m_0)$ at all stations are predicted by eqs (A3) and (A10). Starting with a reasonable initial model m_0 and observed

arrivals T_0, the final model m is inverted by improving data fit iteratively using

$$m = m_0 + \left(G^T G \right)^{-1} G^T (T_0 - T(m_0)) \tag{A24}$$

$$G = \left. \frac{\partial T}{\partial m} \right|_{m=m_0}. \tag{A25}$$

The partial derivatives can be easily calculated numerically. The mean residuals at individual stations are used as station correction to mitigate in the beamforming effects not accounted for in the inverted 1-D piecewise model.

A3 Imaging the structure below the array

We handpick first P arrivals generated by the 33 Betsy gunshots (yellow circles in Fig. 1) at all 1108 sensors. To make more accurate pick, we use an automatic gain controlled (AGC) algorithm to enhance the first arrivals of each gunshot. Fig. A2 displays the AGC waveforms and the corresponding arrival picks. In total, 22 406 high-quality picks are collected for 28 Betsy gunshots (three of the other gunshots are at repetitive location and two are bad). The inversion is highly overdetermined because the model has only four parameters. The final P velocity model and data fits are displayed in Fig. A3. The obtained model can be converted to a local S velocity model assuming a distribution of Poisson ratios. The velocity gradient in the top layer (5.1 s^{-1}) is high and the P-wave velocity jumps
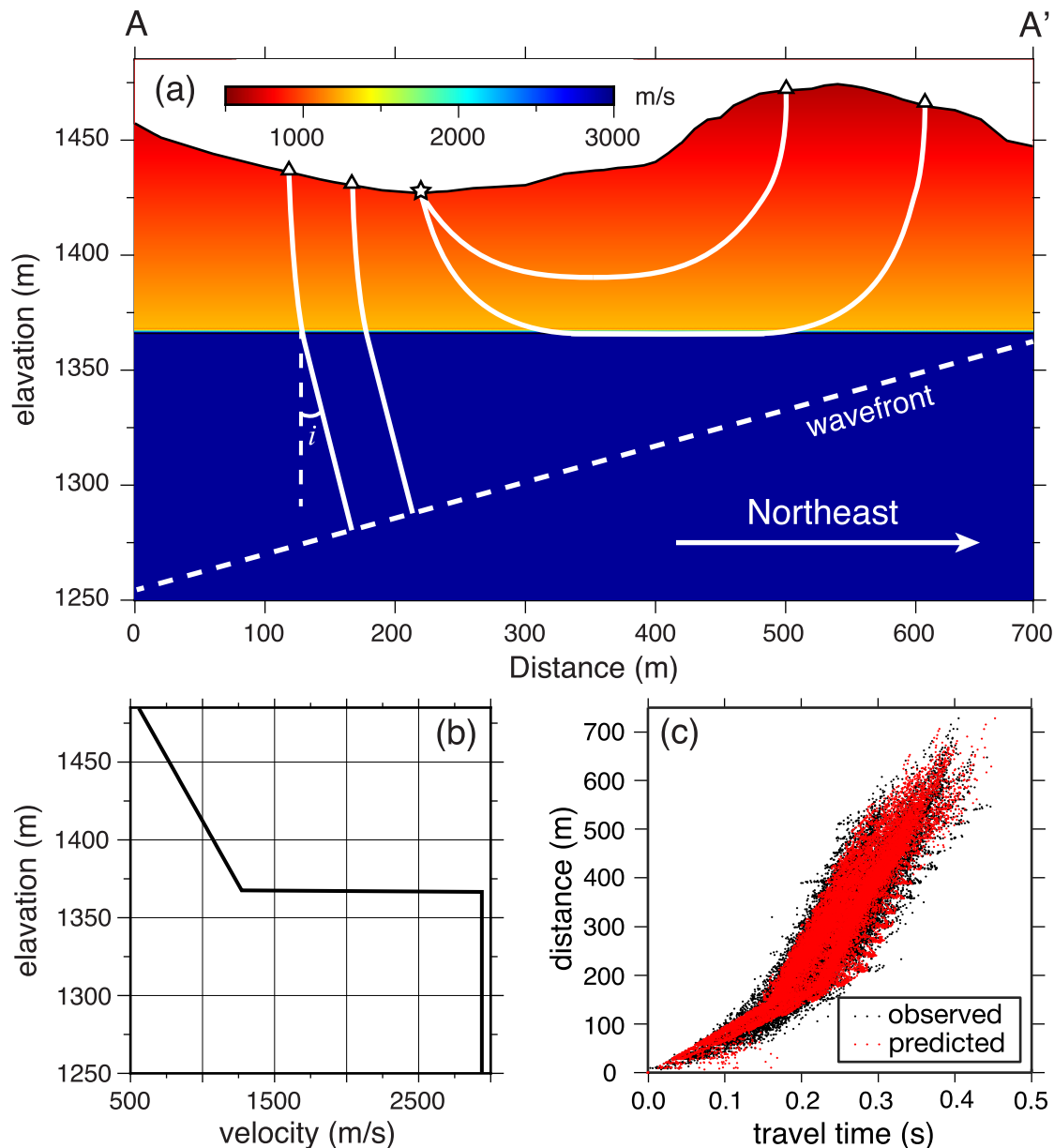
Figure A3. (a) A derived *P* velocity model along cross-section AA′ of Fig. 1. The star and triangles represent a Betsy gunshot and geophones. (b) A depth varying *P* velocities, starting from surface points with different elevations as in (a), inverted from the Betsy gunshots data. (c) Observed and predicted traveltimes of all Betsy gunshots data.

from 1120 to 2689 m s^{-1} across the interface at elevation of 1373 m, which is about 110 m below the local topographic high in the study area (Fig. A3b). The 1-D velocity model generally predicts well the traveltimes (Fig. A3c). The data scattering can mostly be predicted by the topography.

Fig. A4(a) shows the average residuals of traveltime at each station. Positive values represent arrivals later than predicted (slow local anomalies) and negative values represent the opposite situation. There are two major low-velocity anomalies. The one around the left fault trace is associated with a local sedimentary basin, while the one on the NE side of the right trace is associated with fault damage zone including a trapping structure (Ben-Zion *et al.* 2015). Since the topography is considered in the model, 3-D wave propagation effects contribute most of the remaining residuals. The mean residuals of Fig. A4(a) are used as station corrections in the slant stacking done in the paper. Fig. A4(b) dis-

plays the azimuthal distribution of residuals. Assuming *P*-wave azimuthal anisotropy, we fit the residuals with station correction removed (Fig. A4c) to cosine function res $= A \cdot \cos(2az + \phi)$. Although station corrections are removed, we still find weak but possible *P*-wave azimuthal anisotropy with fast axis parallel to the strike of SJFZ. This is consistent with noise-based imaging results of Hillers *et al.* (2016) based on the dense deployment data.

APPENDIX B: MAGNITUDE ESTIMATION

The magnitude estimation is motivated by the empirical Green's function (eGf) approach to account for propagation and site effects. We first find reference events within a 5 km radius sphere of the
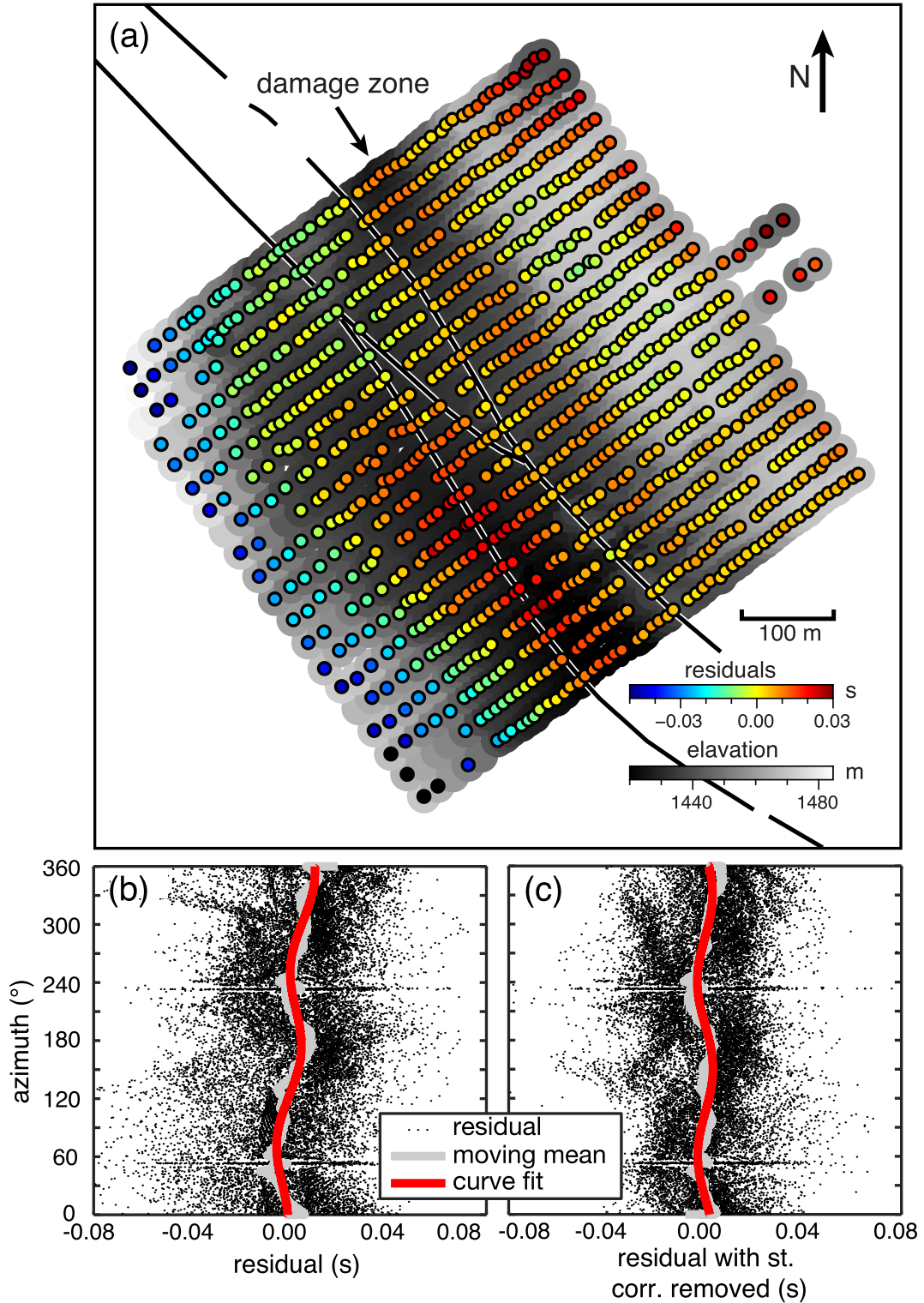
Figure A4. (a) Average residuals between observed and predicted first arrivals from Betsy gunshots data at sensors of the dense array. The background grey colours represent topography. (b) Residuals versus azimuth with red line showing the best fit of res $= A \cdot \cos(az + \varphi)$ with az denoting azimuth. (c) Residuals with station correction removed.

hypocentre of the target event using the focal mechanism catalogue of Yang *et al.* (2012) extended to later events. The magnitudes of reference events are smaller than 3 so the finite-source effects are negligible. We then measure the vector sum of the max velocity amplitude A^i at station i for the target event and A_0^i for the reference event using the seismic velocity time series (HH channels) filtered at 2–15 Hz to increase the signal-to-noise ratio (SNR). The magnitude M_l is estimated to be

$$M_l = \text{median}\left(\log\left(\frac{A^i/R^i}{A_0^i/R_0}\right)\right) + M_{l0}, \tag{B1}$$
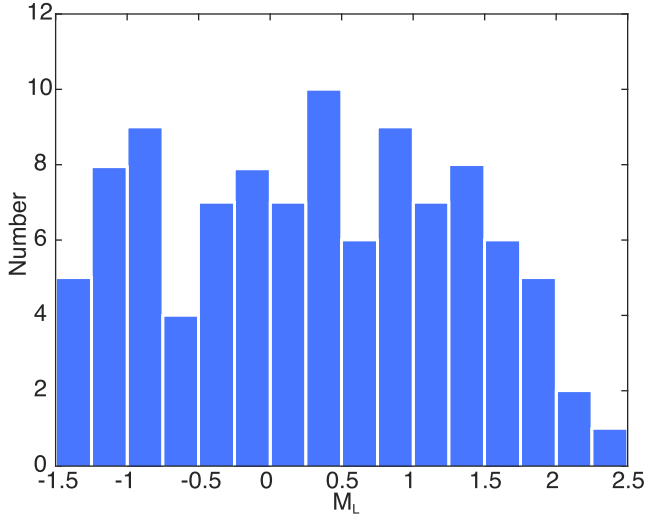
Figure B1. A histogram of magnitudes for the 103 validated events.

where M_{l0} is the magnitude of the reference (eGf) event and R_0, R^i are the radiation patterns of P or S phases of the reference and target events at station i, respectively. We use median instead of mean in eq. (B1) to reduce possible effects of outliers. Since it is not possible to determine the focal mechanism of the target event with the limited available measurements, we estimate the mechanism from a probability distribution calculated as the joint distribution of the focal mechanisms and their uncertainties of all the reference events. Specifically, we draw 1000 possible mechanisms of the target event from the distribution, calculate for each the radiation pattern R^i at station i and estimate the corresponding M_l for each reference event. The final magnitude of the target and its uncertainty are set to be the mean and standard deviation of the 1000 estimates of M_l. The estimated magnitudes of the validated detected events for Julian day 146 span the range -1.5 to 2.5 (Fig. B1).

APPENDIX C: SUBARRAY CONFIGURATION

Here, we analyse the optimal selection of subarrays with the procedure used on this work (Fig. 2), with the goal of increasing the sensitivity to weak coherent signals with large incident angle from sources below the surface. For simplicity, we assume that N stations are evenly installed in a dense deployment with square shape on a

flat surface. The signal from the source is approximated by a plane wave with incident angle i, backazimuth baz and duration longer than the maximum moveout across the array. The average SNR of waveforms at the array stations is r_0 and the entire array is divided into $m \times m$ subarrays. Assuming random noise, the SNR of the stacked waveform for each subarray is

$$r_{\text{sub}} = c(i, baz)\frac{\sqrt{N}}{m}r_0, \tag{C1}$$

where $c(i, baz)$ is a stacking coefficient related to the aperture and density of the array, signal duration and incident direction. We note that $m/\sqrt{N} < c(i, baz) \leq 1$ for stacking without time-shifts and $c(i, baz) = 1$ when $i = 0°$ and the waveforms are perfectly aligned. In general, $c(i, baz)$ decreases with increasing i or array aperture, and $c(i, baz)$ increases with increasing array density or signal duration. By multiplying $m \times m$ envelops of the stacked waveform, the SNR of the Product function is

$$r_P(m) \leq \left[c(i, baz)\frac{\sqrt{N}}{m}r_0 \right]^{m^2}. \tag{C2}$$

Using $I(m)$ to represent the right-hand side of eq. (C2), and $a > 1$ to represent $c(i, baz)\sqrt{N}r_0$, the maximum of $I(m)$ is found from the derivative

$$\frac{\partial I(m)}{\partial m} = m\left[\frac{a}{m}\right]^{m^2} \left[2\ln\left(\frac{a}{m}\right) - 1\right]. \tag{C3}$$

Setting eq. (C3) to zero, the optimal choice of m is $a/\sqrt{e}$. For our data set with $N = 1108$, r_0 is slightly larger than 1 with $m = 3$ (9 subarrays) when $c(i, baz) = 0.15$. Therefore, the choice $m = 3$ adopted in this paper helps maximizing the sensitivity of the array to weak signals with larger incident angles.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. Peak ground velocity (PGV) generated by a Betsy gunshot normalized across the array and colour coded using a log scale. The background grey colours represent topography.

Please note: Oxford University Press are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.