

# Velocity-Based Earthquake Detection Using Downhole Distributed Acoustic Sensing—Examples from the San Andreas Fault Observatory at Depth

by Ariel Lellouch, Siyuan Yuan, William L. Ellsworth, and Biondo Biondi

**Abstract** Conventional seismographic networks sparsely sample the wavefields excited by earthquakes. Thus, standard event detection is conducted by analyzing separate stations and merging their results. Emerging distributed acoustic sensing recording technologies allow for unbiased spatial sampling of the wavefield and, as a result, array-based processing of the recorded signals. Using a cemented fiber in the San Andreas Fault Observatory at Depth main hole, 800 virtual receivers are sampled at a 1 m interval from the surface to 800 m depth. Recorded earthquakes are approximated as plane waves reaching the bottom of the array first. Following this assumption, the relative travel times of the recorded event depend on the local velocity at the array location and the angle of incidence at which the planar wavefront reaches it. Given the seismic velocity, a newly proposed detection algorithm amounts to a single-parameter scan of the incidence angle and measurement of data coherency along the different possible travel-time curves. Using the entire recording array, a much higher effective signal-to-noise ratio can be obtained when compared to individual channel processing. About 20 days of recorded seismic activity from the San Andreas Fault is analyzed. Using a downhole single array, the majority of cataloged events in the area are detected. In addition, a previously unknown event is unveiled. We estimate its magnitude at roughly  $-0.5$ .

## Introduction

Conventional seismographic networks are sparse. For example, the total number of permanent stations in California, one of the most seismically active regions in the world, is less than 1700 (July 2019, according to the Federation of Digital Seismograph Networks foundation). On average, this number amounts to a station every 15 km. Therefore, recorded wavefields are spatially discontinuous and aliased at frequencies of interest for studying local earthquakes. Event detection has to be performed on each station individually. It is usually conducted by detection of amplitude changes across the network or by associating automatically measured phase picks (Johnson *et al.*, 1994; Kao and Shan, 2004; Grigoli *et al.*, 2018). Decades of research have been invested in the development of reliable automatic detection and picking strategies.

Using distributed acoustic sensing (DAS) for seismology has been emerging during the last few years (Lindsey *et al.*, 2017; Jousset *et al.*, 2018; Wang *et al.*, 2018; Ajo-Franklin *et al.*, 2019). It offers a high-resolution temporal and spatial sampling of the seismic wavefield. Previously studied DAS arrays used for earthquake seismology have been mostly based on shallow ( $<10$  m) horizontal arrays, often using

telecommunication cables. For such cables, coupling with the earth is usually poor. In addition, because the arrays are installed at shallow depth, they suffer from velocity heterogeneity, strong anelastic dissipation, and local scattering (Wang *et al.*, 2018). In this study, we apply a wavefield-based method for event detection to a downhole vertical DAS array. Such deployments are common in the oil and gas industry (Mateeva *et al.*, 2014; Correa *et al.*, 2017; Jin *et al.*, 2019; Karrenbach *et al.*, 2019) for different purposes. Their deployment is much simpler than that of geophones, they are more resistant to temperature and pressure, and they can be installed on the outside of wells without prohibiting hydrocarbon production. In addition, once they are installed, they can be used for a virtually infinite amount of time, in contrast to conventional sensors, which are costly to maintain in place. This fact is especially beneficial for monitoring and time-lapse projects.

Thanks to the high-resolution sampling of the propagating wavefield using DAS, an array processing approach can be applied to earthquake detection. Trace picking is not required, and signal-to-noise ratio issues are overcome by concurrently processing a large number of receivers. The vast majority of picking algorithms do not utilize detailed prior

knowledge about the velocity structure of the subsurface (Withers *et al.*, 1998; Rost and Thomas, 2002; Zhang *et al.*, 2003). The structure becomes important only later during the location process (Thurber, 1992; Kissling *et al.*, 1994). In this study, we use the underlying velocity model as part of the detection procedure. It enforces a physical constraint on the recorded wavefield, and events not adhering to it are rejected. For example, wavefronts originating at, or close to, the surface can be filtered out. We illustrate this array-based approach using a simple acquisition scenario with an 800 m long vertical fiber installed in the San Andreas Fault Observatory at Depth (SAFOD). It is based on DAS measurements but can be applied to any dense vertical acquisition system. For such setups, useful and greatly simplifying assumptions of wavefront propagation can be made.

### DAS Recording

DAS measurements are inherently different from those obtained using conventional mechanical seismometers. They are based on backscattering of light traveling in an optical fiber. Because of fiber imperfections, some of the propagating photons are reflected to their source. When the fiber is stationary and under a fixed temperature, the photon Rayleigh backscattering is constant. Brillouin and Raman scattering can also occur nonetheless and is part of the noise. However, when the fiber is strained due to an incoming seismic wave, the light reflection pattern changes. Such variations can be measured by an interrogator unit connected to one end of the fiber. There are many technical details regarding the design of such systems (see, e.g., Grattan and Meggitt, 2000; Parker *et al.*, 2014; Dean *et al.*, 2017; Papp *et al.*, 2017; Martin *et al.*, 2018), but they are not the focus of this study.

DAS measures strain (or strain rate) and not particle displacement, velocity, or acceleration. In this study, the interrogator is measuring strain. In addition, the measurement of strain is conducted along the direction of the fiber. However, measuring strain along a single axis induces angle-dependent sensitivity. For  $P$  waves, the sensitivity is roughly  $\cos^2(\theta)$ , in which  $\theta$  is measured in relation to the fiber axis (Martin *et al.*, 2018). For a vertical well,  $P$  waves propagating along the array ( $\theta = 0^\circ$ ) will be perfectly captured. If they propagate perpendicularly to the array ( $\theta = 90^\circ$ ), they will not induce any strain along the fiber and thus will not be visible. For  $S$  waves, the behavior is proportional to  $\sin(2\theta)$ . Therefore, we expect a maximal  $S$ -wave signal at a propagation direction of  $\theta = 45^\circ$ . Finally, DAS suffers from a gauge length effect, arising from the optical design, which effectively acts as a moving window averaging strain. DAS does not yield a point measurement like conventional sensors, but rather a spatially averaged value, following the direction of the fiber. The size of the spatial window is commonly set to 10 m, with separate measurements conducted every 1 m for commercial interrogators such as the one used in this study. In practice, it means that recorded data are not independent, as would be with conventional point sensors. The gauge length effect varies for different wavelengths

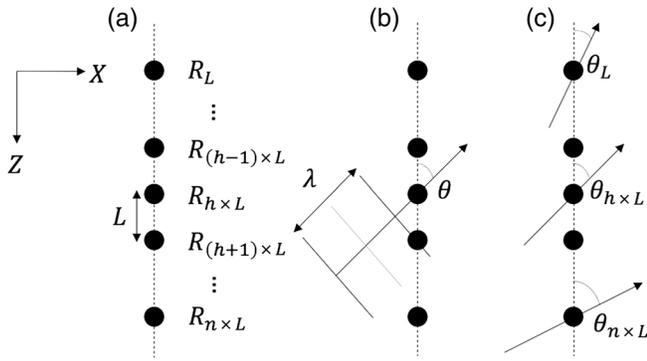
and apparent propagation velocities, measured along the array. Long wavelengths are less affected by it (Dean *et al.*, 2017). Events with high-apparent velocity will be significantly less affected than those with slower velocity. Therefore, vertically propagating earthquakes will be most affected by the gauge length, as they have the slowest possible apparent velocity.

A direct comparison between DAS and conventional seismic sensors is thus challenging. Several studies have compared DAS and conventional seismic sensors, either downhole or at the surface (Daley *et al.*, 2013, 2016; Correa *et al.*, 2017; Lindsey *et al.*, 2017; Olofsson and Martinez, 2017; Wang *et al.*, 2018; Spikes *et al.*, 2019). It is generally accepted that DAS offers lower trace-by-trace quality but compensates for it by a significantly higher number of channels, which allows for better signal processing and stronger array focusing power.

### Travel-Time Differences in a Vertical Array

We focus on predicting the relative arrival travel times of a given event at different DAS receiver locations along a vertical array. We refer to such a travel-time pattern as moveout. The moveout depends solely on the source and receiver locations and the subsurface model. The source focal mechanism, receiver axis, and type of measurement may influence the recorded signal's amplitude, but not propagation travel time (Lellouch, Yuan, *et al.*, 2019). Using a single vertical array and assuming a 1D model of the earth, there is a full azimuthal symmetry in terms of moveout. Only the source depth and horizontal distance from the well are, in theory, recoverable. For detection purposes, the estimation of these parameters is unnecessary. In addition, we assume the wavefront arrives on the bottom of the array first. This is reasonable as earthquakes usually occur much deeper than recording arrays. Moreover, the velocity structure of the earth will bend propagating wavefronts such that waves originating from a shallow earthquake located far from the array will reach the deepest receiver first.

In Figure 1a, we analyze the relative travel times of a plane-wave incident on a downhole array from below. The plane-wave approximation is acceptable if the size of the recording array is negligible compared to the radius of curvature of the wavefront. Because of the representative 3D distances of estimated earthquake hypocenters, this assumption is acceptable. Assuming a constant velocity  $V$ , the travel-time difference between any two adjacent receivers is, geometrically,  $dt = \frac{L \times \cos(\theta)}{V}$ , in which  $L$  is the distance between receivers,  $\theta$  is the angle of incidence, and  $V$  is the velocity. The wavefront will first reach the bottom receiver, as assumed, and arrivals at shallower receivers (depth  $Z$ ) will be delayed by  $T = \frac{(H-Z) \times \cos(\theta)}{V}$ , in which  $Z$  is the depth of the shallower receiver, and  $H$  is the depth of the bottom receiver. All depths are measured as positively increasing away from the surface. This result arises from a simple integration of delays between all adjacent receiver pairs. However, in real-life scenarios, velocity changes with depth. Therefore, we cannot use a constant velocity for the entire array.



**Figure 1.** (a) A 2D vertical acquisition geometry. Axes are displayed on the left. Receivers are idealized as point sensors, denoted by full circles. There are  $n$  receivers and their depths are denoted in subscripts. Distance between receivers is  $L$ . (b) An incident plane wave in a constant velocity ( $V$ ) medium. The propagation direction denoted by an arrow is at an angle of  $\theta$  relative to the array. The wavelength is denoted by  $\lambda$ . (c) Varying incidence angle along the array. The angle  $\theta$  is depth dependent and varies with receiver depth, denoted in subscript.

For a vertical fiber, we assume a 1D velocity model given by  $V(Z)$ . Only the velocity along the array is required. A simplistic yet useful approach is to assume the angle of incidence  $\theta$  remains constant throughout the array. In this case, the travel-time delay  $T$  of a receiver at depth  $Z$  is given by

$$T = \int_{z=H}^{z=Z} L \times \frac{\cos(\theta)}{V(z)} dz. \quad (1)$$

This assumption is reasonable if the velocity changes slowly with depth. However, as we will show later, at SAFOD the velocity increases rapidly with depth over the length of the fiber. Therefore, it is incorrect to assume a constant angle of incidence throughout the array, because the wavefronts will bend while propagating along it. This is illustrated in Figure 1c. The importance of this effect grows with the incidence angle, measured for reference at the bottom of the array. For a vertical propagation ( $\theta = 0^\circ$ ), it does not influence the moveout at all. For all other cases, the incidence angle  $\theta$  is a function of depth. However, for all incidence angles, the  $\cos(\theta)$  scaling assuming constant velocity is still locally true between two adjacent receivers if  $L$  is small enough. When using DAS records,  $L$  is usually 1 m or so. In other words, we safely assume that the angle of incidence does not change within the 1 m distance between adjacent receivers. As a result, equation (1) becomes

$$T = \int_{z=H}^{z=Z} L \times \frac{\cos(\theta[z])}{V(z)} dz. \quad (2)$$

The remaining issue is the computation of the depth dependence of the incidence angle  $\theta$ , given by  $\theta[z]$ . Naturally,  $\theta[z]$  depends on the initial incidence angle at the bottom of the array. Therefore, there is a 2D mapping to the local incidence angle, which is a function of both the initial incidence angle

and the depth of the receiver. We compute this mapping by assuming the ray approximation. This is a high-frequency approximation that is valid only when the scale of changes in the subsurface is much longer than the propagating wavelength. We show later that this condition is only partially fulfilled. Nonetheless, it is an approximation that is often used in seismology, and velocity structures in the area of the study are relatively smooth and do not exhibit sharp changes, which are often the cause of the ray approximation breaking down. By consecutively applying Snell's law between virtual layers of thickness  $L$ , we find that the angle of incidence as a function of depth  $\theta[z]$  is given by

$$\theta[z] = \sin^{-1} \left( \frac{V_z}{V_H} \times \sin(\theta[H]) \right), \quad (3)$$

in which  $V_H$  is the velocity at the bottom of the array and  $\theta[H]$  is the angle of incidence at that point. If we combine equations (2) and (3), we obtain full mapping between the angle of incidence at bottom of the array and travel-time differences, or moveout, along with the receiver depth

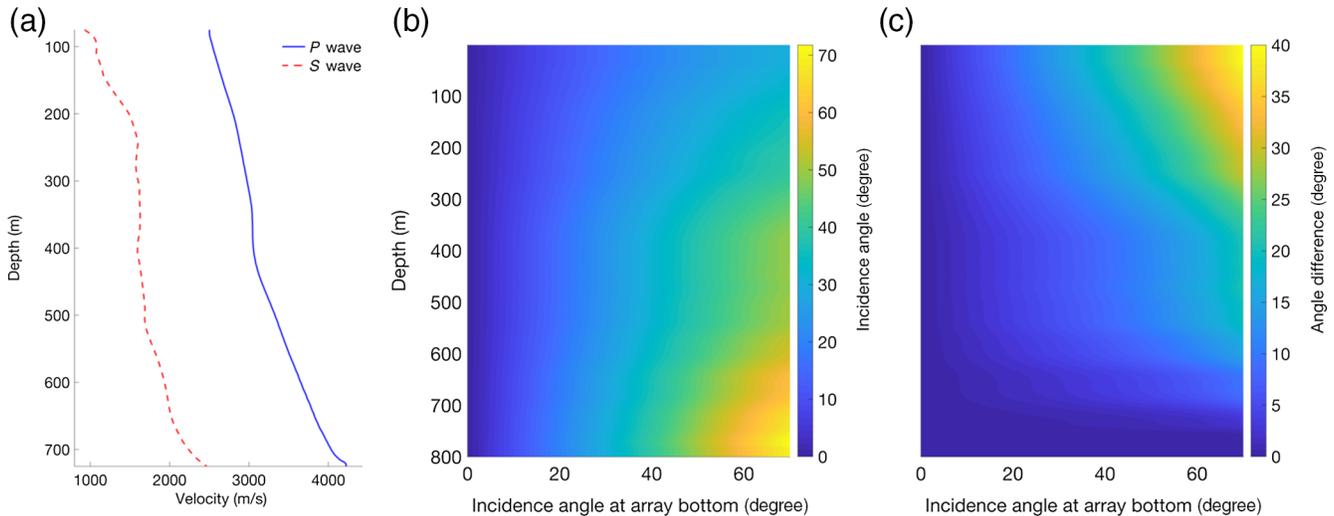
$$T = \int_{z=H}^{z=Z} L \times \frac{\cos \left( \sin^{-1} \left( \frac{V_z}{V_H} \times \sin(\theta[H]) \right) \right)}{V(z)} dz. \quad (4)$$

This mapping depends on the 1D velocity model of the subsurface, the depth of the receivers, and the angle of incidence at the bottom of the array. This mapping can naturally be computed for both  $P$ - and  $S$ -wave velocities.

### The SAFOD DAS Array and Its Surroundings

As part of the SAFOD experiment (Zoback *et al.*, 2011), two optical fibers were cemented in between casing strings for use as an interferometric strainmeter (Blum *et al.*, 2010). The fibers terminate at depths of 782 and 864 m. Because of failure of the loop at the end of the longer fiber, it could not be used for interferometry but provided a well-coupled fiber for DAS. During June–July 2017, OptaSense were contracted to deploy a model ODH 3.1 DAS interrogator connected to this fiber. The interrogator has to be deployed in a temperature-controlled room with a standard electrical power supply. The fiber can extend from the wellhead to such a location for distances of several kilometers, if needed. In this temporary setup, the fiber on the surface was not anchored to the ground and suffered from ambient common-mode vibrations in the recording hut. The interrogator has data storage capabilities of several to tens of terabytes (TB). Data are accumulated at roughly 1 TB/day, so added storage might be required.

For the fiber-equipped portion, the well is close to vertical and reaches a depth of 864 m. Our study is limited to depths of up to 800 m due to the fiber breaking, which induces strong optical noise at nearby channels. The recording system was operated for about a month, but due to technical issues, only about 22 discontinuous days of recorded data are



**Figure 2.** (a) Used velocity models (*P*-line, *S*-dashed line) used in this study, adapted from Lellouch, Yuan, *et al.* (2019). Note the strong depth variation. (b) 2D mapping of incidence angle at the bottom of the array and receiver depth to the local incidence angle, computed using the *P*-wave velocity model in (a). (c) Difference between incidence angle at the bottom and the local incidence angle. For higher initial incidence angles, the difference is more substantial. Near-vertical arrivals are practically unaffected. The color version of this figure is available only in the electronic edition.

useful. Data were recorded with a 1 m spatial sampling, 10 m gauge length at 2500 samples per second. We convert them to their strain-rate equivalent by applying a time derivative. This is a convenient choice as it eliminates the DC component in the strain records, which is the background backscattering of the fiber. In addition, the strain rate is the closest measure to particle velocity, which is what commonly used downhole geophones measure. To save computational effort, data have been downsampled to 250 Hz with an antialias filter prior to running the detection algorithm. All data shown in this article are otherwise unfiltered.

We previously used DAS recordings to develop the *P*- and *S*-wave velocities along the array (Lellouch, Yuan, *et al.*, 2019). Examples of earthquakes used for velocity analysis can also be seen in that study. Here, we use the *P* and *S* velocities estimated using the slant-stack approach, as in our opinion it yields the most stable results. At the surface and deepest levels of the array, we use the picking-based model because there is not enough aperture for the slant-stack approach. The velocity models are shown in Figure 2a. Using the *P*-wave velocity model, we compute the varying angle of incidence along the array using equation (3). The mapping is shown in Figure 2b, whereas the angle difference relative to the initial angle of incidence at the bottom of the array is shown in Figure 2c. The same process is repeated using the *S*-wave velocity profile, but not shown here because of the velocity structure, which is increasing with depth, the change in angles is the most noticeable for large initial incidence angles, as they are strongly bent toward the vertical direction. Near-vertical events are almost unaffected. As a side note, this is the reason that near-vertical events can be reliably used for velocity estimation (Lellouch, Yuan, *et al.*, 2019).

## Earthquake Detection

Given a 1D velocity model along the array, we can compute different trial moveouts as a function of receiver depth and initial incidence angle for a plane wave at the bottom of the array (equation 4). The detection algorithm scans through possible incidence angles, taken within a certain range and angular resolution, and finds the angle yielding the maximal match between the trial moveouts and recorded data. If the match is better than a certain threshold, a detection is declared. This procedure is conducted separately for *P*- and *S*-wave velocities, and results are eventually combined.

Data are processed in 60 s intervals. Before applying the algorithm, we automatically remove noisy traces. Quite surprisingly, the traces that are faulty change throughout the survey. The mechanism that controls this behavior is still opaque to us, and we thus treat each 60 s interval individually. We iterate that the fiber used for this experiment was installed 12 yr before the acquisition. In addition, it broke down at its loop, thus inducing unwanted optical reflection noise. There are many examples of DAS acquisitions in which such noise does not exist, and it is not a limitation of the technology. For an example, see Lellouch, Biondi, *et al.* (2019).

The median energy of noisy channels is at least an order of magnitude higher than for functioning channels, and they can thus be easily detected. Quantitatively, about 5% of the channels are deleted, depending on the specific record. To remove geometrical spreading, receiver coupling, and nonlinear effects (see Miah and Potter, 2017 for a detailed review) of the fiber affecting recorded amplitudes, we normalize each 60 s trace by its median energy. This is important as wave-form-based coherency measures are sensitive to amplitude

variations. After that normalization, possibly remaining noisy channels will see their detrimental effect strongly mitigated.

There are two approaches for measuring the match between trial moveouts and recorded data. The first is to pick arrivals in the data and measure their time differences from the expected moveout. However, picking continuous records is challenging, ambiguous, and time consuming. In addition, it operates on single recorded traces, thus increasing its sensitivity to noise (Lellouch and Reshef, 2019). Instead, we opt for a new method, which is waveform-based, fast, and fully automatic. Data are shifted according to different trial moveouts, computed using the velocity model and a range of incidence angles. We use a range of 0–70° sampled at a 1° interval. In other words, for each scanned incidence angle, each recorded trace is shifted in time by the computed time delay at its matching receiver location. This shift can be implemented at once for the whole record. We used 1-min-long files, sampled at 250 Hz after downsampling. After the time shift has been applied to the recorded data, a process which is repeated for each possible incidence angle, we measure the multichannel coherency of the moveout-corrected record. Ideally, if the correct shift has been applied, the first arrival phase should be aligned. This principle is often used in exploration seismology (Biondi, 2006) as a tool to measure velocity model correctness. Here, we assume the velocity is correct and use it instead as a tool to find the incidence angle of the event. Coherency along the different receivers can be measured by various methods. We opt for semblance (Neidell and Taner, 1971), which is one of the simplest and most widespread measures of multichannel coherency. More sophisticated coherency measurement methods, specifically tailored for DAS response, may be beneficially used, but they are not in the scope of this study.

We define recorded data after moveout correction as  $d_{r,t}$ , with  $r$  being the receiver number (out of  $N$ ) and  $t$  being the time sample. We compute the semblance within a temporal window of size  $W$ , which is chosen to roughly match the recorded wavelet. In this study, we use a 32 ms window. The windowed semblance is defined as

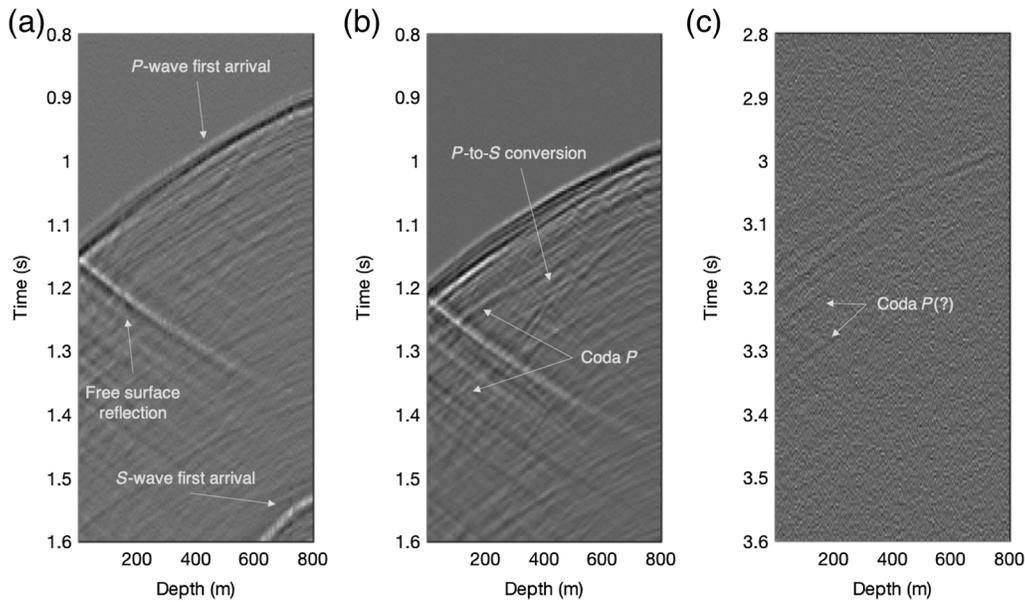
$$S(t) = \frac{\sum_{j=t-W/2}^{t+W/2} (\sum_{i=1}^N d_{i,j})^2}{N \times \sum_{j=t-W/2}^{t+W/2} \sum_{i=1}^N d_{i,j}^2}. \quad (5)$$

Semblance has the benefit of having a value in the interval [0,1]. It is 1 for perfectly coherent data and tends to 0 for uncorrelated zero-mean noise. As a result, it is convenient for thresholding applications. In addition, it measures an average property of the entire record, thus increasing its ability to handle low signal-to-noise ratio records. Because our array contains 800 virtual receiver points, the potential improvement in detection is substantial, and single faulty receivers are not as detrimental as when using trace-by-trace picking. The semblance is computed for every time sample. The result is a 2D matrix, which represents semblance as a function of time sample and trial incidence angle. At each time sample,

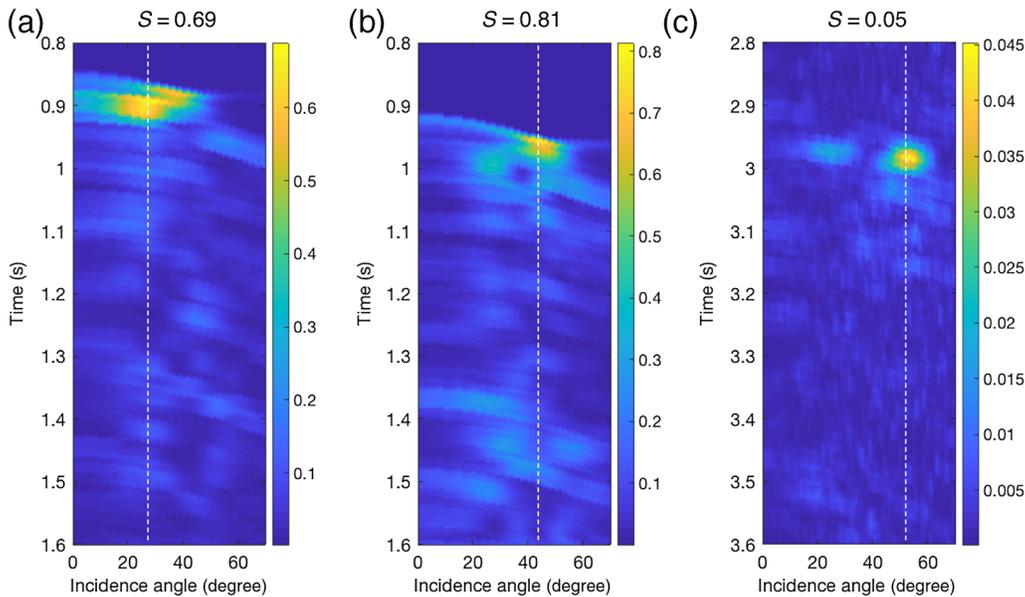
the incidence angle yielding the maximal semblance is chosen. If this value is above a certain fixed threshold, which we chose to be 0.018, a detection is declared. The threshold was set by running the detection on several minutes of ambient noise and choosing a threshold value two standard deviations above it. Semblance noise levels are very low thanks to the large number of used receivers. The whole process is separately conducted for both  $P$ - and  $S$ -wave velocities. Detection with either  $P$  or  $S$  is sufficient. However, if both detections are present within than 10 s of each other, we attribute them to the same source. This way, we are not spuriously introducing redundant events to our detection catalog.

We exemplify the procedure using three recorded earthquakes, chosen using the U.S. Geological Survey (USGS) catalog in the area. The records are shown zoomed in on the  $P$ -wave arrivals in Figure 3. Their magnitudes and distances from the wellhead are taken from the USGS catalog. The earthquakes in Figure 3a,b are clearly visible, whereas the one in Figure 3c is harder to see. The frequency content of recorded events is relatively high and reaches a central frequency of 60–70 Hz at the bottom of the array. Propagating wavefields are sampled with a high resolution, and the various phases are locally coherent and easily followed across the array. Therefore, we can safely use waveform-based methods.

We first demonstrate the angle scanning procedure for the  $P$  arrivals. In Figure 4, we show the semblance matrices obtained for the events shown in Figure 3. For all events, the semblance is maximal at or close to the time of the first arrival of the  $P$  phase. Semblance values are, as expected, significantly higher for the strong earthquakes (Fig. 4a,b). However, even a relatively weak signal can be coherently aligned after optimal moveout correction. Because semblance is applied over the entire array, it can successfully handle a low signal-to-noise ratio. The angle at which the maximal semblance is obtained varies between events. If we compute straight lines between the bottom of the array and the USGS estimated source location, the incidence angles are (a)  $-34.6^\circ$ , (b)  $-42.9^\circ$ , and (c)  $-89.7^\circ$ . The same trend can be seen in the semblance analysis. However, a direct comparison of angles cannot be made without accounting for the 3D velocity structure of the subsurface, at least up to the depth of the earthquakes (e.g., Zhang *et al.*, 2009). Nonetheless, for events (a) and (b), the retrieved angles are close to the geometrically computed angles. For the far event (c), which is practically horizontal using straight lines, we observe the largest difference between the estimated incidence angles. Because it occurs farther away from the source, the velocity field below the array, which we assume is increasing with depth, has a larger effect on the incidence angle. In addition, due to Snell's law, angle variations are most prominent when the straight line from source to receivers is close to horizontal, which is the case for this event. This is why the propagating wavefronts eventually reach the array at a very different angle than the one predicted by a straight line—roughly  $55^\circ$  instead of  $89.7^\circ$ .



**Figure 3.** Three recorded earthquakes. Properties from U.S. Geological Survey (USGS) catalog: (a)  $M = 1.12$ ,  $R = 5.2$  km, (b)  $M = 1.56$ ,  $R = 5.4$  km, and (c)  $M = 0.7$ ,  $R = 13.9$  km. Distances  $R$  are measured as the 3D distance from the wellhead. The earthquake in (c) has low signal-to-noise ratio. Arrows and labels mark notable events of interest.

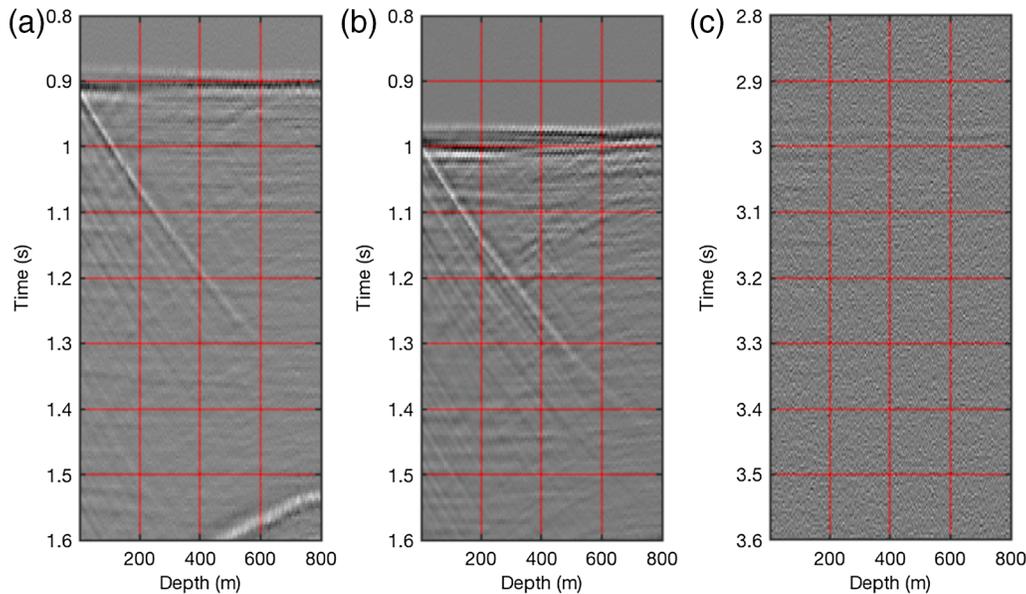


**Figure 4.** Moveout scans of earthquakes using  $P$ -wave velocity. The time scale is the same as for Figure 3. The color scale is the semblance and is different for each event. Semblance values are significantly higher for the (a,b) strong earthquakes than for the (c) weak one. The maximal semblance value for each record is written on top of the plots. The angle yielding the maximal overall semblance is denoted by a dashed line. The color version of this figure is available only in the electronic edition.

In Figure 5, we show the  $P$ -wave arrivals in the recorded events after optimal moveout correction. Such gathers are only used for the semblance computation and not outputted during the application of the detection algorithm, but we plot them here for illustration purposes. After moveout correction with the optimal incidence angles, first arrivals have little to no residual moveout and appear flat. Some coda events are also visible in Figure 5a,b and have a similar flat pattern, but

they undergo interference with free surface reflections and converted modes. The flatness of the first arrivals along with the reasonable incidence angles at which they are obtained indicates the reliability of the used velocity model.

Although an individual virtual receiver can be quite noisy, processing the entire array simultaneously greatly increases the effective signal-to-noise ratio of the recorded data. For example, the weak event in Figure 3c would be challenging



**Figure 5.** Records shown in Figure 3 after time shift with optimal moveout, estimated from semblance analysis. First arrival events have little to no residual moveout and appear as flat. In (a,b), some coda events are also flattened but are harder to distinguish due to interference with free surface reflections and converted modes. (c) The moveout-corrected event is much clearer than in the original record and centered around the 3 s mark. The color version of this figure is available only in the electronic edition.

to detect using first break picking. However, it is visible and largely above the detection threshold when analyzed along the entire array after moveout correction (Fig. 5c). For the strong events, we can also clearly see the surface reflection in the flattened data, as well as *P*-to-*S* converted waves that are likely from the basement transition located just below the array.

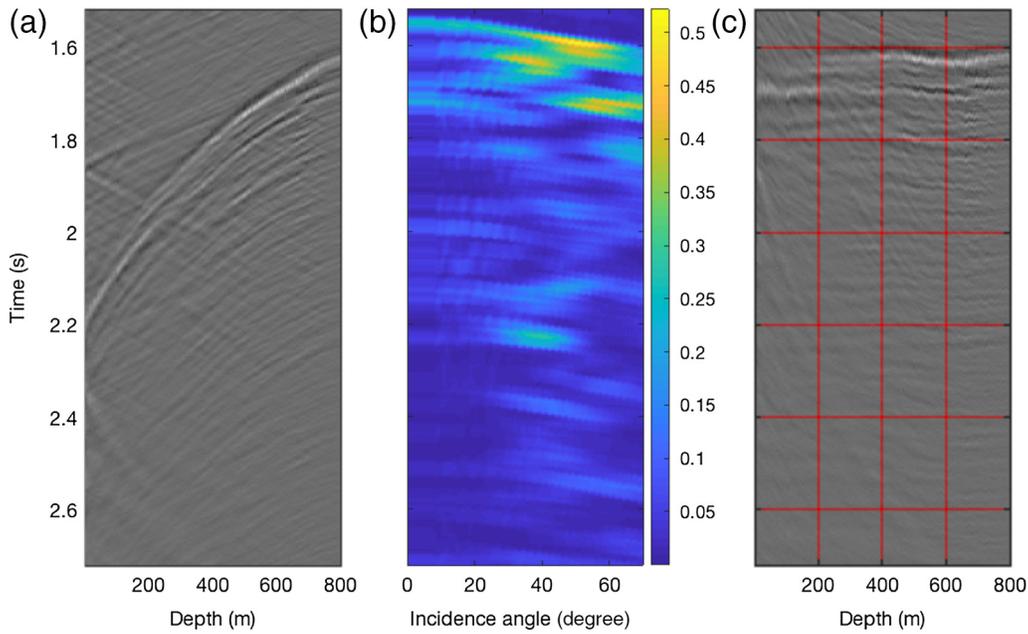
In Figure 6, we show semblance analysis using *S* waves for the  $M = 1.56$ ,  $R = 5.4$  km event, shown in Figures 3b–5b. The first arrival of the *S* phase is clear but interferes with many different phases—coda *P*, *SP* conversions, free-surface reflections, and so forth (Fig. 6a). The semblance scan (Fig. 6b) is not as straightforward as for the *P* waves. There are multiple events with different semblance values, appearing at varying incidence angles. However, data after optimal moveout correction (Fig. 6c) are only approximately flat. The principal assumption of a plane wave reaching the array remains correct, as indicated by the short residuals (<10 ms). Primary causes for such deviations from flatness are imperfections in the used velocity model and interference between the different seismic phases that alter the recorded signals. Such interference is clearly visible in Figure 6c at the shallowest (<200 m) receivers, where it severely degrades the flat part of the signal.

In general, the recorded signal is much more complex than for the previously shown *P*-wave first arrivals. This could be at least partially due to the DAS response of *S* waves in a vertical array. The fiber loses sensitivity for near-vertical propagation of *S* waves. Because of ray bending, this is often the case in near-surface receivers. *P* waves, on the contrary, are far from the nonsensitivity zone. The most extreme incidence angles we observe are about  $60^\circ$ , which would yield a loss of only 3.5 dB for the *P* waves. The semblance measure we use is sensitive to the recorded waveform. Therefore,

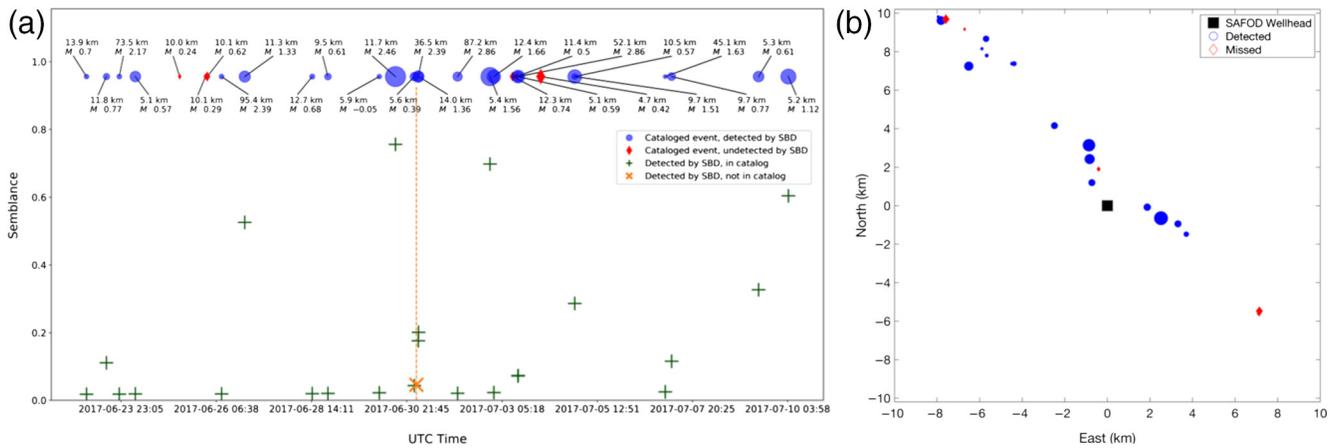
phase changes, induced by the DAS response, will deteriorate its value. The data also exhibit changing waveforms. Although semblance is a straightforward choice, it is possible that other coherency measures can outperform it.

## Detection Results

We apply the earthquake detection algorithm to about 22 days of usable recorded passive data from SAFOD. We plot events surpassing the semblance threshold using the *P*- or *S*-wave velocities after manual quality control. We had to override about 20 detections. Most of them were due to DAS-induced noise bursts that had a rectangular pattern wider in time than the computed moveouts. A few false alarms were not identifiable upon manual observation and are due to the statistical variation of the noise. Naturally, the false alarm and detection performances are dictated by the threshold choice, and a few false alarms over three weeks of recording are reasonable. To compare our detection results with the USGS catalog, we define any event within 15 km of the well-head, based on the USGS location, as detectable by the array. We compare all events within that range to our detection results, yielding either a detection or a miss. For events that we detect outside that range, we compare the estimated origin time with the USGS timing, taking a 1 min window to assure events are not missed. If a cataloged detection is present, we add the event as a correct detection. However, as stated, any cataloged event located farther away than 15 km will not count as a missed detection. If a detected event is not in the catalog, it is defined as a previously unknown earthquake. We summarize these results in Figure 7. Detections are plotted according to their semblance values, and the USGS catalog is plotted on top.



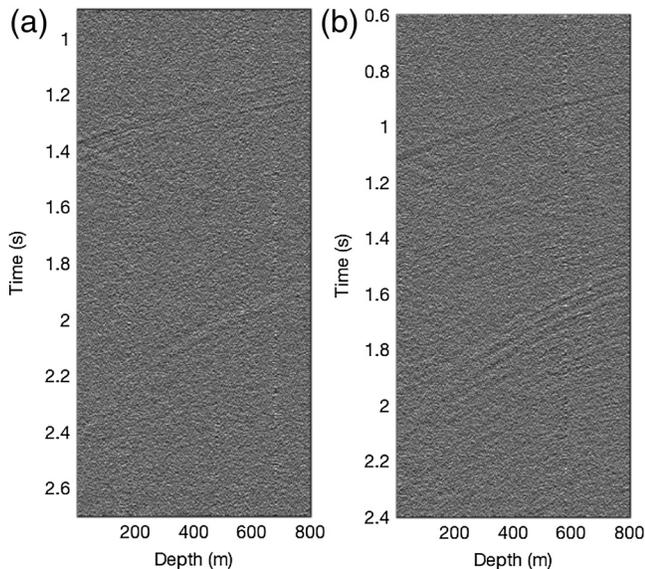
**Figure 6.** *S*-wave analysis for the  $M = 1.56$ ,  $R = 5.4$  km event, shown in Figures 3b–5b. (a) Recorded event zoomed on the *S*-wave first arrivals, (b) semblance scan, and (c) event after moveout-correction with optimal incidence angle. The analysis is noisier and contains many events. The event is only approximately flattened using the optimal incidence angle. The color version of this figure is available only in the electronic edition.



**Figure 7.** (a) Detection results. Semblance results above the detection threshold are plotted in green crosses. USGS catalog results are plotted as either circles (if they were detected by the semblance scan) or diamonds (if they were not), along with their cataloged magnitude and distance from the wellhead. The size of the markers is proportional to the detectability of the event. The uncataloged event detected by the semblance scan is plotted as an orange asterisk. SBD, semblance-based detection. (b) Map view of cataloged events within a 15 km radius from the wellhead (square). Detected (circles) and missed (diamonds) events are plotted, with their size proportional to the event detectability. The events are distributed along the fault. Missed events are mostly peripheral, except one weak event close to the borehole. SAFOD, San Andreas Fault Observatory at Depth. The color version of this figure is available only in the electronic edition.

The size of the markers is computed by  $\log(10^M/R^2)$ , in which  $M$  is the cataloged magnitude and  $R$  is the distance from the wellhead. This value represents an approximate detectability criterion (large, easy to detect, and small, hard to detect) following a simple energy estimation of recorded event. Only relatively hard-to-detect events are missed by the semblance scan, and all except one are peripheral. We detect 18 out of 24 of the events within a range of 15 km from the wellhead and five

cataloged earthquakes outside that range. All six missed events have a magnitude of less than 0.75. Five out of six of the undetected events are peripheral, showing that for source distances of less than 10 km the DAS detection is very reliable. The nearby event that was missed is discussed later, but it is visually detectable, and a slight improvement in signal-to-noise ratio would have probably allowed for its automatic detection. In addition, a previously uncataloged event is detected.



**Figure 8.** (a) Uncataloged earthquake detected by semblance scan. The  $P$  and  $S$  arrivals are visible, along with the free surface reflection of the  $P$  phase. (b) An  $M = 0.42$  event missed by semblance scan.  $P$  and  $S$  arrivals are again visible. For both events, traces have been scaled individually and amplified. Events have a similar moveout, but (b) is slightly below the threshold in the semblance analysis.

We display two recorded events of interest in Figure 8. In Figure 8a, we show the detected uncataloged event. It has clear  $P$ - and  $S$ -phase arrivals, and the free surface reflection of the  $P$  phase is also visible. Using the borehole station CCRB, we can analyze continuous seismograms at the time of this event. The uncataloged event is visible in the borehole seismogram. For reference, we compare DAS and borehole records to a cataloged event originating 70 min earlier. In both types of records, the two events have similar  $P$ – $S$  travel-time differences, and we thus assume they are located at similar distances ( $\sim 5.5$  km) from the array. We convert the CCRB borehole seismogram to the Wood–Anderson response. The uncataloged event appears to be of a magnitude of roughly  $M = -0.5$ , explaining why it is not cataloged. In Figure 8b, we show one of the missed events. This event is cataloged and estimated at  $M = 0.42$  with a distance of 4.7 km from the array. Although both  $P$  and  $S$  phases are clearly visible to the naked eye, signal-to-noise ratio is too low for the automatic detection threshold. Computed values are just below the 0.018 semblance threshold we used, and thus the event is missed. This illustrates the limitations of the DAS acquisition system in terms of the signal-to-noise ratio. However, the underlying assumption of plane-wave propagation is still valid, as the  $P$  phase of the event is reasonably flattened when an incidence angle of about  $30^\circ$  is used.

This study is the first application of downhole DAS to earthquake seismology. Therefore, it is too early to compare it with the performance of conventional surface seismic sensors. Recorded data are conceptually different. Although conventional sensors offer 3C measurements and a better

signal-to-noise ratio, DAS arrays can increase the number of channels by three orders of magnitude. It allows for a much deeper understanding of the different seismic phases in action. In addition, DAS technology is still new and keeps improving, whereas seismic sensors have a relatively fixed design. The fiber used for this study is not ideal, as it is broken and was installed more than a decade before the acquisition for non-DAS purposes. Newer interrogators can reduce noise and decrease the gauge length. Helical designs of fibers can offer a more omnidirectional response and, theoretically, 3C measurements (Ning and Sava, 2016). Better physical designs of the optical fiber properties may improve sensitivity, as well as looping the fiber to improve signal-to-noise ratio. Overall, we expect the recorded signal quality to improve with time. In addition, the suggested processing workflow can be further refined to better match DAS records' properties such as frequency response, gauge length, noise mechanisms, and more.

## Conclusions

The high-resolution sampling of the DAS allows for array-based techniques in seismology, which employ the local spatial coherency of recorded wavefields. It cannot be directly compared to a broadband station as it offers different data. We show how knowledge of the velocity model along a vertical DAS array can be used in designing a useful earthquake detection procedure with some simplifying assumptions. The method is waveform-based, automatic, simple to implement, and fast. It operates on the entire array and measures an average signal property, thus improving its ability to handle low signal-to-noise ratio records. As a by-product, it also estimates the angle of incidence at the base of the array. In addition, our understanding of the different seismic phases is significantly improved, thanks to the continuous spatial sampling. This procedure can be applied separately to  $P$  and  $S$  waves. We show its application on 22 days of recorded data in the vertical part of the SAFOD main hole, spanning DAS array depths of 0–800 m. The method recovers 18 out of 24 of cataloged events within a radius of 15 km from the array, detects five farther away cataloged events, and detects a previously unknown earthquake, which is estimated at  $M = -0.5$ . Despite signal-to-noise ratio issues causing some missed events, these encouraging results set the path for incorporating downhole DAS arrays into event detection workflows for induced seismicity monitoring as well as earthquake seismology, possibly in conjunction with horizontal DAS arrays as well as conventional broadband stations. Improvements in interrogator technology and fiber design are likely to improve signal-to-noise ratio and subsequent processing.

## Data and Resources

None of the authors have any conflict of interest. Records around cataloged or detected events can be accessed directly at <https://github.com/ariellelouch/DASDetection> (last accessed October 2019).

## Acknowledgments

This study is supported in part by the Israeli Ministry of Energy under the program for postdoctoral scholarships in leading universities.

## References

- Ajo-Franklin, J. B., S. Dou, N. J. Lindsey, I. Monga, C. Tracy, M. Robertson, V. R. Tribaldos, C. Ulrich, B. Freifeld, T. Daley, and X. Li (2019). Distributed acoustic sensing using dark fiber for near-surface characterization and broadband seismic event detection, *Sci. Rep.* **9**, 1328–1341.
- Biondi, B. (2006). *3D Seismic Imaging*, Society of Exploration Geophysicists, Tulsa, Oklahoma.
- Blum, J., H. Igel, and M. Zumberge (2010). Observations of Rayleigh-wave phase velocity and coseismic deformation using an optical fiber, interferometric vertical strainmeter at the SAFOD borehole, California, *Bull. Seismol. Soc. Am.* **100**, 1879–1891.
- Correa, J., A. Egorov, K. Tertyshnikov, A. Bona, R. Pevzner, T. Dean, B. Freifeld, and S. Marshall (2017). Analysis of signal to noise and directivity characteristics of das VSP at near and far offsets—A CO2CRC Otway Project data example, *The Leading Edge* **36**, 994a1–994a7.
- Daley, T. M., B. M. Freifeld, B. Ajo-Franklin, S. Dou, R. Pevzner, V. Shulakova, S. Kashikar, D. E. Miller, J. Goetz, J. Hennings, and S. Lueth (2013). Field testing of fiber-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring, *The Leading Edge* **32**, 699–706.
- Daley, T. M., D. E. Miller, K. Dodds, P. Cook, and B. M. Freifeld (2016). Field testing of modular borehole monitoring with simultaneous distributed acoustic sensing and geophone vertical seismic profiles at Citronelle, Alabama, *Geophys. Prospect.* **64**, 1318–1334.
- Dean, T., T. Cuny, and A. H. Hartog (2017). The effect of gauge length on axially incident P-waves measured using fibre optic distributed vibration sensing, *Geophys. Prospect.* **65**, 184–193.
- Grattan, L. S., and B. T. Meggitt (2000). *Optical Fiber Sensor Technology: Fundamentals*, Springer, Boston, Massachusetts.
- Grigoli, F., L. Scarabello, M. Böse, B. Weber, S. Wiemer, and J. F. Clinton (2018). Pick- and waveform-based techniques for real-time detection of induced seismicity, *Geophys. J. Int.* **213**, 868–884.
- Jin, G., K. Mendoza, B. Roy, and D. G. Buswell (2019). Machine learning-based fracture-hit detection algorithm using LFDAS signal, *The Leading Edge* **38**, 520–524.
- Johnson, C. E., A. G. Lindh, and B. Hirshorn (1994). Robust regional phase association, *U.S. Geol. Surv. Open-File Rept.* 94-621, doi: [10.3133/ofr94621](https://doi.org/10.3133/ofr94621).
- Jousset, P., T. Reinsch, T. Ryberg, H. Blanck, A. Clarke, R. Aghayev, G. P. Hersir, J. Hennings, M. Weber, and C. M. Krawczyk (2018). Dynamic strain determination using fibre-optic cables allows imaging of seismological and structural features, *Nature Comm.* **9**, 2509–2519.
- Kao, S., and S. Shan (2004). The source-scanning algorithm: Mapping the distribution of seismic sources in time and space, *Geophys. J. Int.* **157**, 589–594.
- Karrenbach, M., S. Cole, A. Ridge, K. Boone, D. Kahn, J. Rich, K. Silver, and D. Langton (2019). DAS microseismic, strain and temperature monitoring during hydraulic fracturing, *Geophysics* **84**, D11–D23.
- Kissling, E., W. L. Ellsworth, D. Eberhart-Phillips, and U. Kradolfer (1994). Initial reference models in local earthquake tomography, *J. Geophys. Res.* **99**, 19,635–19,646.
- Lellouch, A., and M. Reshef (2019). Velocity analysis and subsurface source location improvement using moveout-corrected gathers, *Geophysics* **84**, KS119–KS131.
- Lellouch, A., B. Biondi, S. Home, M. A. Meadows, and T. Nemeth (2019). DAS observation of guided waves in a shale reservoir generated by perforation shots, *SEG Technical Program Expanded Abstracts 2019*, 943–947.
- Lellouch, A., S. Yuan, Z. Spica, B. Biondi, and W. L. Ellsworth (2019). Seismic velocity estimation using passive downhole distributed acoustic sensing records—Examples from the San Andreas Fault Observatory at Depth, *J. Geophys. Res.* **124**, 6931–6948.
- Lindsey, N. J., E. R. Martin, D. S. Dreger, B. Freifeld, S. Cole, S. R. James, B. Biondi, and J. B. Ajo-Franklin (2017). Fiber-optic network observations of earthquake wavefields, *Geophys. Res. Lett.* **44**, 792–799.
- Martin, E. R., N. J. Lindsey, and B. Biondi (2018). Introduction to interferometry of fiber optic strain measurements, *EarthArXiv*, 1–33.
- Mateeva, A., J. Lopez, H. Potters, J. Mestayer, B. Cox, D. Kiyashchenko, P. Wills, S. Grandi, K. Hornman, B. Kuvshinov, et al. (2014). Distributed acoustic sensing for reservoir monitoring with vertical seismic profiling, *Geophys. Prospect.* **62**, 679–692.
- Miah, K., and D. K. Potter (2017). A review of hybrid fiber-optic distributed simultaneous vibration and temperature sensing technology and its geophysical applications, *Sensors* **17**, 2511.
- Neidell, N. S., and M. T. Taner (1971). Semblance and other coherency measures for multichannel data, *Geophysics* **36**, 482–497.
- Ning, I. L. C., and P. Sava (2016). Multicomponent distributed acoustic sensing, *SEG Technical Program Expanded Abstracts, 2016*, 5597–5602.
- Olofsson, B., and A. Martinez (2017). Validation of DAS data integrity against standard geophones—DAS field test at Aquistore site, *The Leading Edge* **36**, 981–986.
- Papp, B., D. Donno, J. E. Martin, and A. H. Hartog (2017). A study of the geophysical response of distributed fibre optic acoustic sensors through laboratory-scale experiments, *Geophys. Prospect.* **65**, 1186–1204.
- Parker, T., S. Shatalin, and M. Farhadiroushan (2014). Distributed acoustic sensing—A new tool for seismic applications, *First Break* **32**, 61–69.
- Rost, S., and C. Thomas (2002). Array seismology: Methods and applications, *Rev. Geophys.* **40**, 2-1–2-7.
- Spikes, K. T., N. Tisato, T. E. Hess, and J. W. Holt (2019). Comparison of geophone and surface-deployed distributed acoustic sensing seismic data, *Geophysics* **84**, A25–A29.
- Thurber, C. H. (1992). Hypocenter-velocity structure coupling in local earthquake tomography, *Phys. Earth Planet. In.* **75**, 55–62.
- Wang, H. F., X. Zeng, D. E. Miller, D. Fratta, K. L. Feigl, C. H. Thurber, and R. J. Mellors (2018). Ground motion response to an ML 4.3 earthquake using co-located distributed acoustic sensing and seismometer arrays, *Geophys. J. Int.* **213**, 2020–2036.
- Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bull. Seismol. Soc. Am.* **88**, 95–106.
- Zhang, H., C. Thurber, and P. Bedrosian (2009). Joint inversion for Vp, Vs, and Vp/Vs at SAFOD, Parkfield, California, *Geochem. Geophys. Geosys.* **10**, Q11002, doi: [10.1029/2009GC002709](https://doi.org/10.1029/2009GC002709).
- Zhang, H., C. Thurber, and C. Rowe (2003). Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings, *Bull. Seismol. Soc. Am.* **93**, 1904–1912.
- Zoback, M., S. Hickman, and W. L. Ellsworth (2011). Scientific drilling into the San Andreas fault zone—An overview of SAFOD's first five years, *Sci. Drill.* **22**, 14–28.

Department of Geophysics  
Stanford University  
397 Panama Mall  
Stanford, California 94035 U.S.A.  
ariellel@stanford.edu  
syuan@stanford.edu  
wellsworth@stanford.edu  
biondo@stanford.edu

Manuscript received 28 June 2019;  
Published Online 12 November 2019