Fault Zone Imaging with Distributed Acoustic Sensing: Surface-to-Surface Wave Scattering

Yan Yang[1], Zhongwen Zhan[1], Zhichao Shen[1], and James Atterholt[1]

[1]California Institute of Technology

November 18, 2022

Abstract

Fault zone complexities contain important information about earthquake physics. High-resolution fault zone imaging requires high-quality data from dense arrays and new seismic imaging techniques that can utilize large portions of recorded waveforms. Recently, the emerging Distributed Acoustic Sensing (DAS) technique has enabled near-surface imaging by utilizing existing telecommunication infrastructure and anthropogenic noise sources. With dense sensors at several meters' spacing, the unaliased wavefield can provide unprecedented details for fault zones. In this work, we use a DAS array converted from a 10-km underground fiber-optic cable across Ridgecrest City, California. We report clear spurious arrivals and coda waves in ambient noise cross-correlations caused by surface-to-surface wave scattering. We use these scattering-related waves to locate and characterize potential faults. The mapped fault locations are generally consistent with those in the USGS Quaternary Fault database of the United States but are more precise. We also use waveform modeling to infer that a 35-m wide, 90-m deep fault with 30% velocity reduction can best fit the observed scattered coda waves for one of the identified fault zones. These findings demonstrate the potential of DAS for passive imaging of fine-scale faults in an urban environment.

Fault Zone Imaging with Distributed Acoustic Sensing: Surface-to-Surface Wave Scattering

Yan Yang, Zhongwen Zhan, Zhichao Shen, James Atterholt

Seismological Laboratory, California Institute of Technology, Pasadena, CA

Corresponding author: Yan Yang (yanyang@caltech.edu)

Key Points:

- Ambient noise interferometry with distributed acoustic sensing captures scattered surface waves from fault zones.
- The fault locations mapped with spurious arrivals are generally consistent with previous models but with higher resolution.
- We constrain the fault zone geometry and velocity reduction with the amplitudes of the scattered surface waves.

Abstract

Fault zone complexities contain important information about earthquake physics. High-resolution fault zone imaging requires high-quality data from dense arrays and new seismic imaging techniques that can utilize large portions of recorded waveforms. Recently, the emerging Distributed Acoustic Sensing (DAS) technique has enabled near-surface imaging by utilizing existing telecommunication infrastructure and anthropogenic noise sources. With dense sensors at several meters' spacing, the unaliased wavefield can provide unprecedented details for fault zones. In this work, we use a DAS array converted from a 10-km underground fiber-optic cable across Ridgecrest City, California. We report clear spurious arrivals and coda waves in ambient noise cross-correlations caused by surface-to-surface wave scattering. We use these scattering-related waves to locate and characterize potential faults. The mapped fault locations are generally consistent with those in the USGS Quaternary Fault database of the United States but are more precise. We also use waveform modeling to infer that a 35-m wide, 90-m deep fault with 30% velocity reduction can best fit the observed scattered coda waves for one of the identified fault zones. These findings demonstrate the potential of DAS for passive imaging of fine-scale faults in an urban environment.

Plain Language Summary

Fault zones are complex networks of fractures that can host earthquakes. The fractured rock surrounding the faults in the top hundreds of meters can amplify earthquake shaking intensity. Therefore, locating and characterizing faults is important for evaluating seismic hazards, especially in urban settings. But it is challenging to identify small hidden faults in the absence of surface evidence or cataloged seismicity. High resolution, high frequency seismic experiments may provide a solution. Distributed acoustic sensing (DAS) is an emerging technique that can turn existing fiber-optic cables into cost-effective seismic networks with meter-scale spacing. In this work, we show how we image the fault zones at shallow depth using seismic noise generated by traffic along a DAS cable in Ridgecrest City, CA. The results can detect and distinguish faults at sub-kilometer scales. We also show we can use DAS data to characterize fault zone properties. These results demonstrate the potential of DAS in fine-scale fault imaging without needing earthquakes.

1 Introduction

Faults are characterized as damaged material that accommodate localized deformation of rocks (Ben-Zion, 2008). The deformation of fault zone rocks is associated with earthquake generation and rupture process (Perrin et al., 2016; Thakur et al., 2020). The fault material with reduced seismic velocity and altered rheological properties can also amplify ground shaking and influence the migration of hydrocarbons and fluids (Caine et al., 1996; Spudich & Olsen, 2001). Thus, mapping the location and properties of faults is critical for understanding earthquake process and assessing seismic hazard. One common method of mapping faults is the observation of exhumed faults in the field (e.g., Collettini et al., 2009; Faulkner et al., 2003; Mitchell & Faulkner, 2009), which utilizes slices through the fault outcrops. Fault zone drilling projects can extend the examination of fault structure to greater depths and be used to monitor long-term changes in physical properties (e.g., Hickman et al., 2004; Hung et al., 2009). These methods provide precise measurements at single points of observation but require considerable labor and resources. Seismological methods can help develop a more complete picture of subsurface fault characteristics. Earthquake locations and focal mechanisms shed light on fault locations and structural complexities (Ross et al., 2017; Wang & Zhan, 2020). Seismic tomography can produce images of seismic velocity and attenuation near a fault zone (e.g., Allam et al., 2014; Liu et al., 2021; Y. Wang et al., 2019). Fault zone trapped waves recorded by the sensors within the fault zones can be used to model fault zone geometries and properties in detail (e.g., Lewis et al., 2007; Li et al., 2004; Li & Malin, 2008).

The methods above give detailed information on large faults that are visible at the surface or faults with abundant seismicity. Small, buried faults that are not readily visible in the terrain and have little cataloged seismicity may be difficult to discern yet can contribute to the hidden hazards in urban settings. With the deployment of dense arrays, improved spatial coherence at high frequencies allows noise-based tomography to capture finer details of the subsurface (AlTheyab et al., 2016; Castellanos & Clayton, 2021). Distributed acoustic sensing (DAS) is an emerging technique that can repurpose pre-existing telecommunication fiber-optic cables into cost-effective, long-term dense seismic arrays in urban areas (Lindsey & Martin, 2021; Zhan, 2019). DAS measures strain or strain rate along the fiber by applying optical interferometry to laser light back-scattered from the fiber's intrinsic inhomogeneities. With an effective channel spacing of a few meters, DAS can record unaliased high-frequency wavefields and capture the waves that attenuate too rapidly to be detected by conventional networks. In practice, DAS-recorded ambient noise wavefields have been used successfully for near-surface imaging and fault zone identification (e.g., Cheng et al., 2021; Yang et al., 2021).

In this study we use a DAS array rapidly deployed after the 2019 Ridgecrest M7.1 earthquake (Li et al., 2021). The Ridgecrest earthquake ruptured the Little Lake and the Airport Lake fault zones, and produced numerous aftershocks (Ross et al., 2019). The Little Lake fault zone (LLFZ) is part of the Eastern California shear zone, which is composed of a network of dextral, normal, and dextral-oblique faults (Amos et al., 2013). The DAS array at Ridgecrest City was converted from a underground dark fiber in the city of Ridgecrest, which crossed the southern end of the LLFZ (Figure 1). The three mapped fault traces across the DAS array, unlike the northern part of the LLFZ, are not well constrained by the current USGS fault maps and are only inferred with large uncertainty (Figure 1). While the primary goal of this DAS array was to study the aftershocks (Li et al., 2021), the unprecedented spatial resolution also offers an opportunity to improve our knowledge of the fault locations and properties.

In this work, we first report evident spurious arrivals and coda waves in noise cross-correlations related to surface wave scattering. We then use waveform modeling to confirm that the cause of the scattering waves can be faults. With the travel times of the spurious arrivals, we map the fault locations and compare them with current fault maps in this region. With the amplitudes of the coda waves, we constrain the geometry and property of one of the identified faults.

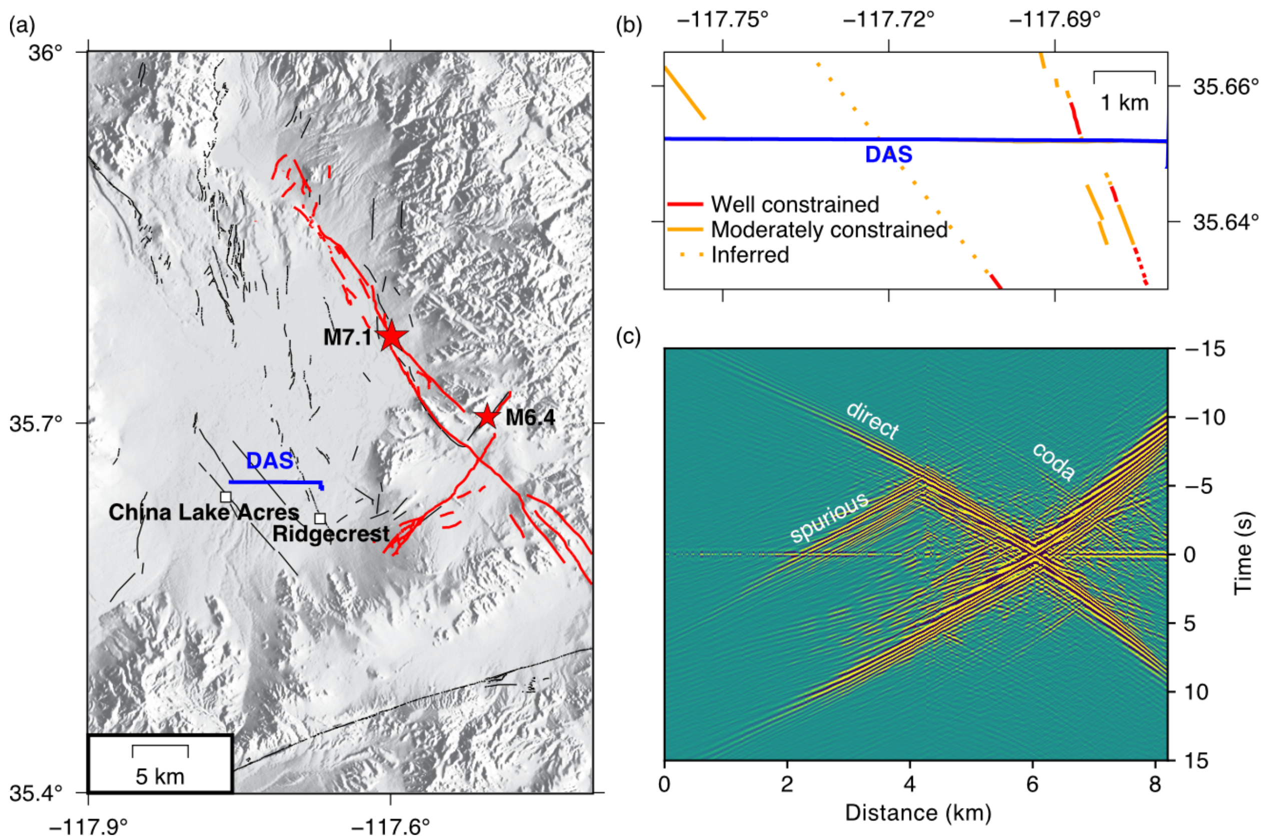

Figure 1 Study region and noise cross-correlation example. (a) Map view of the Ridgecrest DAS array. The ground trace of the Ridgecrest DAS array is shown in blue. The M7.1 mainshock and the M6.4 foreshock are marked with the red stars. The fault zones are marked with black lines (Jennings, 1975). The surface rupture of 2019 Ridgecrest M7.1 earthquake is marked in red lines (Brandenberg et al., 2019); (b) A zoomed-in view of the DAS array and the Little Lake fault zone across the array; (c) Example wavefield of ambient noise cross-correlation. The channel at 6-km distance is used as a virtual source. In addition to the direct Rayleigh waves, we observe scattered surface coda waves and spurious arrivals appeared as precursors.

2 Surface wave scattering

2.1 Observation in noise cross-correlations

The DAS array was converted from a 10-km telecommunications cable across Ridgecrest, California, and has 1250 channels with 8-meter spacing. In this work we focus on the 8-km segment along W Inyokern Road which is approximately a linear array in the east-west direction.

We use three-month continuous data for ambient noise cross-correlation. We follow the conventional workflow that has been developed over the last two decades for imaging earth interior with larger scale and longer period data (Bensen et al., 2007). Preprocessing includes removing mean and linear trends, applying a band pass filter, temporal normalization, and spectral whitening. Then the cross-correlation functions are computed in frequency domain, transformed back to time domain, and stacked over time.

An example wavefield of cross-correlations using a channel at 6 km distance as the virtual source to all the channels is shown in Figure 1c. In addition to the direct surface waves, we can also observe secondary signals exhibiting either precursory energy (arriving at correlation times earlier than the direct wave) or coda energy (arriving at correlation times later than the direct wave). The precursory and coda energy always emerges from several fixed locations when we move the virtual source along the linear array.

2.2 Interpretation with synthetics

Secondary signals have been observed in noise cross-correlations and attributed to a persistent active source or passive scattering from heterogeneities (Ma et al., 2013; Retailleau & Beroza, 2021; Zeng & Ni, 2010; Zhan et al., 2010). The cause of the secondary arrivals in our case of a linear array can be simplified as a 1D scenario for the following reasons: 1) The dominant contribution to the empirical Green's function comes from the constructive interference of waves generated by the stationary points along the receiver line (Snieder, 2004); 2) The primary noise source is the traffic noise with weekly periodicity (Yang et al., 2021). The colinear geometry of the DAS array and highway means that the vast majority of vehicle-generated surface waves are along the DAS array; 3) The directional sensitivity of DAS emphasizes longitudinal Rayleigh waves along the station pairs more than conventional seismometers (Martin et al., 2018).

The origins of the direct and secondary phases are illustrated in Figure 2. For the direct waves, the arrival times are the surface-wave travel times from one receiver to the other (Figure 2a, e). For the coda waves, the later arrival times are caused by the cross-correlation between waves traveling from the noise source to one receiver and waves traveling from the noise source to the other receiver but reflected by a passive scatterer. The coda waves' arrival times are the summation of the travel times from the scatterer to both receivers (Figure 2b, e). Both direct and coda waves are part of the true Green's functions and their travel times are symmetrical on the positive and negative lag times. It is more appropriate to refer to precursory energy as 'spurious arrivals', as it is not part of the true Green's functions between the receivers. For the 1D scenario here, the spurious arrivals appear when there exists a persistent noise source or a passive scatterer between the receivers (Figure 2c, d; Ma et al., 2013). The earlier spurious arrival times in the cross-correlations are the difference between the travel times of the waves from the active source/scatterer to the two receivers (Figure 2c, d, e), and are not symmetrical between the positive and negative sides. Note that the intersection of the scattering waves (including spurious arrivals and coda waves) and the direct waves is the location of the active source/passive scatterer. The direct and scattering waves arrive at the same time because the virtual receiver is overlapping with the active source/passive scatterer. Both active sources and passive scatterers can generate spurious arrivals, whereas coda waves can be ascribed only to passive scatterers. Given the clearly observed coda waves in our noise cross-correlations, we believe that scattering from passive scatterers must be the primary cause, if not the only one. Additionally, the aftershocks recorded by the DAS array also display clear body-to-surface converted waves, further confirming the presence of passive scatterers (the companion paper Atterholt et al., in review).

We find the location of the scatterers generally coincide with the fault traces across the array, for example, the faults in the middle and the east in Figure 1b are close to the interception of direct and scattered waves in Figure 1c. To verify that the presence of a fault can result in the observed scattering-related phases, we simulate noise cross-correlations using a fully elastic GPU-based two-dimensional finite difference code (Li et al., 2014). Our background velocity model is based on a recent tomography study along this DAS array (Yang et al., 2021) and superimposed by a 20-m wide, 40-m deep, rectangular fault with 40% velocity reduction at the distance of 4 km. We place two in-plane noise sources 40 km away from each end of the array. Receivers have the same layout as the DAS array. The simulated wavefield is accurate up to 10 Hz with the grid spacing of 4 m and the time increment of 0.0008 s. We then cross-correlate the synthetic seismogram recorded at the receiver at 1.6 km with the synthetic seismograms from all the other receivers. Both spurious arrivals and coda waves are visible in the synthetic noise cross-correlation (Figure 2e), confirming that the observed scattering waves can be caused by faults.

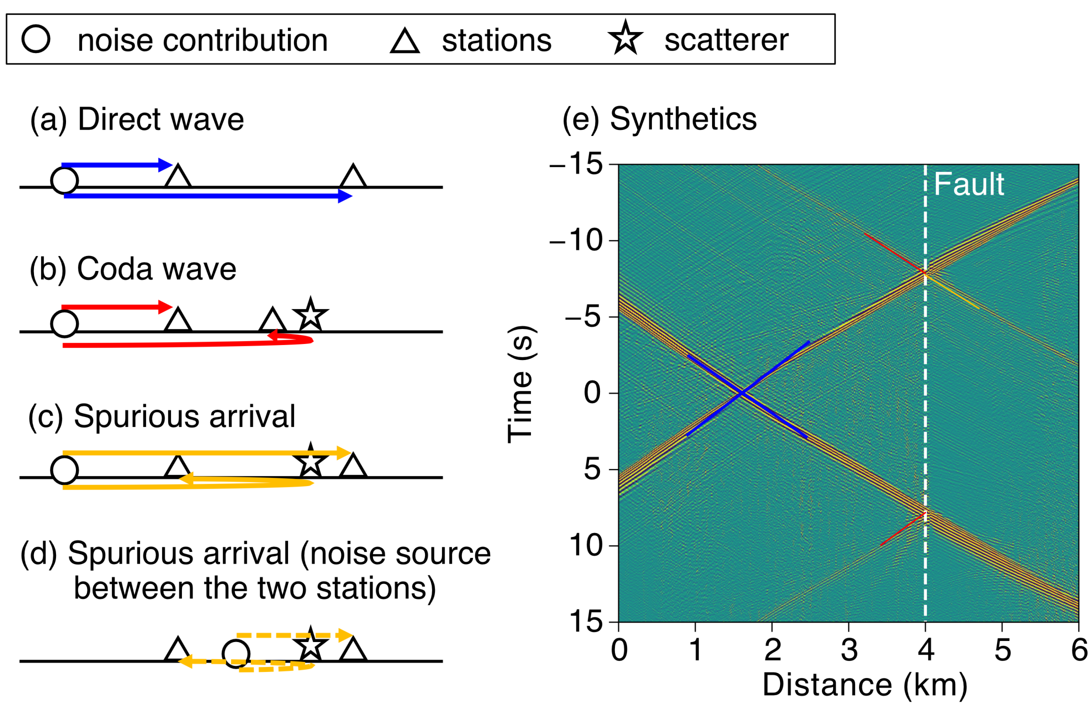

Figure 2 Explanation for the cause of the observed scattering waves. (a)-(d) Schematic cartoon showing the generation of the direct waves, coda, and spurious arrivals appeared as precursors in the cross-correlation. (e) Synthetic noise cross-correlation using waveform modeling. The noise source is put 40 km away from the array and the fault is at 4 km distance. The virtual source is at 1.6 km distance.

3 Locate the faults with the spurious arrivals

3.1 Group velocity inversion for travel-time prediction

Previous regional studies of passive noise scatterers focus on longer periods and usually assume a homogeneous background velocity model to locate the scatterers (Ma et al., 2013; Zeng & Ni, 2010). The lateral variation of the shallow subsurface structure in our case, on the other hand, could have a substantial effect on the mapping resolution. Yang et al., (2021) showed that the shear velocity in the top 30 meters along the Ridgecrest DAS profile has a lateral variation up to ~30% over only 8-km distance. This is illustrated well by the bending in the arrival times of the direct wave group as shown in Figure 1c. Therefore, we invert for the group velocity model along the profile. For each channel pair, we apply frequency-time analysis on the envelop of the cross-correlations and get the group velocity dispersion in the period [0.1, 1] s (or the frequency band [1, 10] Hz) averaged over the distance between the channel pair. The approximately one thousand channels provide half million channel pairs for a dense coverage of the profile. We invert for the group velocity dispersion at the 8-m spacing grids along the profile using linear inversion with second-order Tikhonov regularization. The group velocity model shows a slow section in the east end of the profile (Figure 3a), which is consistent with the microbasin imaged in the shear wave velocity model using phase velocity (Yang et al., 2021).

Given the group velocity model and assuming all surface waves' ray paths are in-plane, we can predict the frequency-dependent (1-10 Hz) arrival times of direct, spurious, and coda waves for any trial scatterer location. For a channel as virtual source at distance x_{src}, and a receiver channel at distance x_{rec}, the arrival times of the direct waves at frequency f will be

$$t_{\mathrm{direct}}(f) = \pm \int_{x_{\mathrm{src}}}^{x_{\mathrm{rec}}} \frac{1}{v(x,f)}\,dx, \tag{1}$$

where $v(x,f)$ denotes the group velocity at the distance x and the frequency f, respectively. As described in Section 2, if a fault located at distance x_{scat} can scatter the seismic waves from the ambient noise, we will observe spurious arrivals or coda waves. If the fault is between the source and receiver channels, there will be spurious arrivals arriving at

$$t_{\mathrm{spurious}}(f) = \left|\int_{x_{\mathrm{scat}}}^{x_{\mathrm{rec}}} \frac{1}{v(x,f)}\,dx\right| - \left|\int_{x_{\mathrm{src}}}^{x_{\mathrm{scat}}} \frac{1}{v(x,f)}\,dx\right|, \tag{2}$$

If the fault is located on the same side as the source and receiver channels, there will be coda waves arriving at

$$t_{\mathrm{coda}}(f) = \pm\left(\left|\int_{x_{\mathrm{scat}}}^{x_{\mathrm{rec}}} \frac{1}{v(x,f)}\,dx\right| + \left|\int_{x_{\mathrm{src}}}^{x_{\mathrm{scat}}} \frac{1}{v(x,f)}\,dx\right|\right). \tag{3}$$

In this section we will only use the spurious arrivals for fault localization as they are typically stronger than the coda waves and hence more suitable for stacking. An example of predicted travel times is shown in Figure 3b. We calculate the arrival times for direct waves and spurious arrivals at 4 Hz, assuming a fault at 4 km. We can see the spurious arrival times are well predicted, as is the bending feature of the direct waves at the 7-8 km distance.

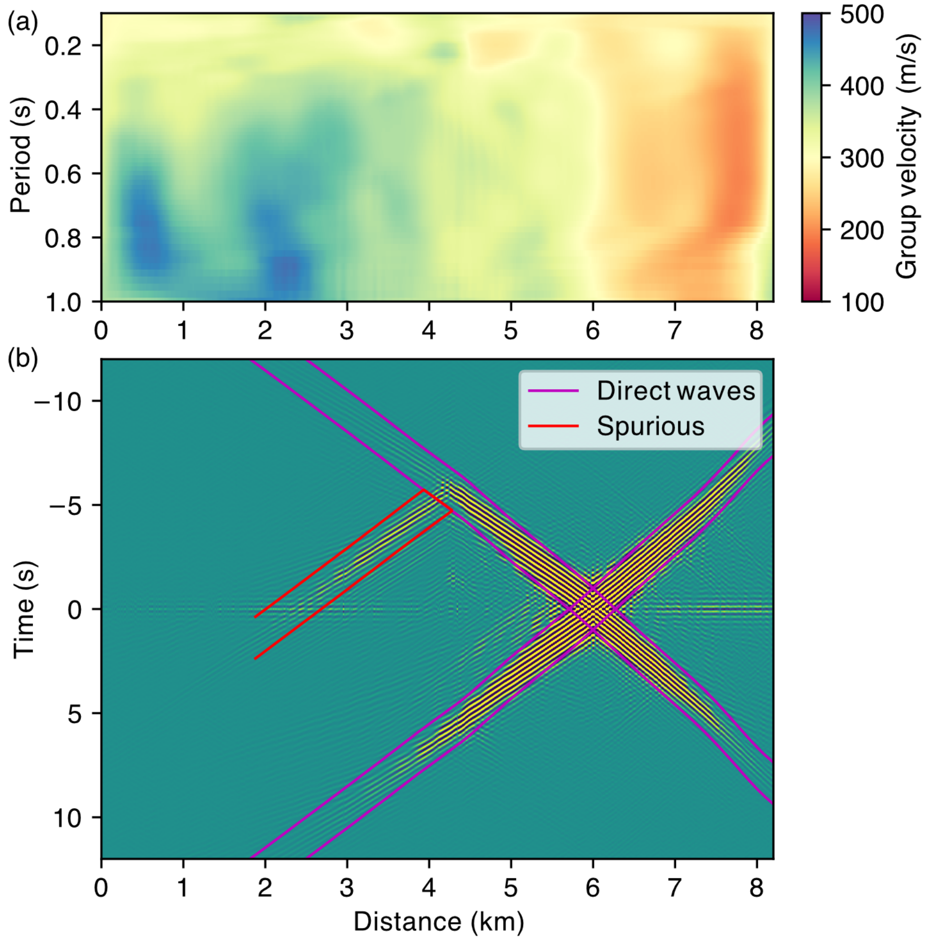

Figure 3. Group velocity model and an example of predicted travel times. (a) Group velocity dispersion along the DAS array in the period of [0.1, 1] s inverted from direct surface wave arrival times; (b) Cross-correlation with the virtual source at 6 km, filtered in a narrow frequency band around 4 Hz. The purple and red lines mark the 2-sec time windows around arrival times of direct waves and spurious arrivals, respectively. The arrival times are calculated by the group velocity model in (a) assuming a scatterer at 4.3 km.

3.2 Fault mapping results

We perform a grid search for the scatterer with an 8-m grid spacing. For each trial scatterer location, we calculate the arrival times of the spurious arrivals using equation (2). We stack the envelope amplitudes of the cross-correlation over a four-period time window centered on the predicted arrival times and get the maximum stacked amplitude. The stacking is done for narrow frequency bands between 1 Hz and 10 Hz, using frequency-dependent group velocities. All channels can be considered as virtual sources while only the receivers within 1 km distance from the assumed scatterer are used for stacking. We take the median of the maximum stacked amplitude from all virtual sources and create a 'scattering amplitude' profile as shown in Figure

4a. We detect multiple stripes with high scattering amplitudes in the grid search result, for example, at 1 km, 4.3 km, and 7.3 km. To be more quantitative, we find the local maxima of the scattering amplitudes as indicative of the presence of fault scatterers. We calculate the peak prominence (how much a peak deviates from the surrounding baseline of the signal) for the scattering amplitudes at each frequency. If the peak prominence exceeds a certain threshold, we consider the peak to be a fault candidate.

From the scattering amplitude profile, we can identify several scattering peaks marked with ‘A’, ‘B’, ‘C’, ‘D’, and the most obvious one throughout all frequencies marked with ‘X’ (Figure 4b). Notable is the closeness of the discovered faults A-D to the USGS-mapped Quaternary faults a-d (Figure 4c, Jennings, 1975). In particular, the two closely spaced fault branches ‘c’, and ‘d’ in the east that are classified as ‘well constrained’ are closely located with the two peaks ‘C’ and ‘D’ (Figure 4b) in our data, with different frequency dependences. The fault in the west (‘a’ in Figure 4c) classified as ‘moderately constrained’ seems associated with the peak marked with ‘A’ in Figure 4b. For the middle zone where the location is inferred rather than directly observed as stated in the USGS database, we identified two scattering peaks (‘B’ and ‘X’ in Figure 4b), one at closer location with fault ‘b’ (Figure 4c) and the other one about 1 km to the east. In the companion paper using earthquake body-to-surface wave scattering (Atterholt et al., in review), the located fault here is also offset to the east, consistent with the more obvious scattering peak ‘X’ in our mapping. Based on the observation and comparison, we believe that the scatterers are indeed related with faults even though their precise positions deviate when there is a lack of constraint in the USGS database.

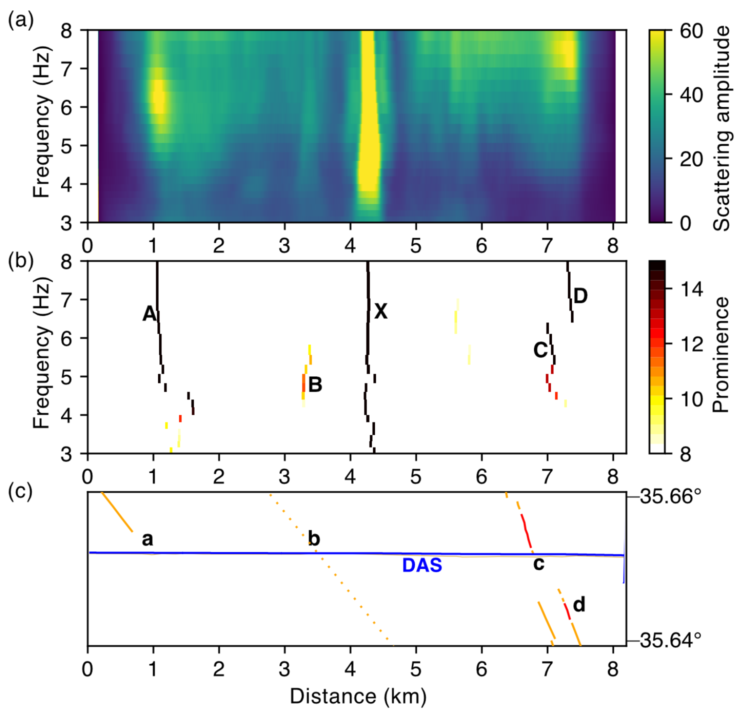

Figure 4 Fault mapping results using spurious arrivals. (a) Grid search results for the scatterer location using the stacked amplitudes along the predicted spurious arrival times; (b) Peak

prominence of the scattering amplitudes in (a), which is calculated individually for each frequency; (c) The DAS array with the USGS mapped fault traces. The legend is the same as that seen in Figure 1b. The fault traces are closely aligned with some of the detected scatterers in (b).

4 Resolving fault zone property with coda waves

With fault locations being accurately mapped, we aim to further investigate the fault zone properties. However, the strength of the stacked spurious arrival amplitudes in Section 3 does not necessarily represent fault zone properties. As shown in Figure 2d, spurious arrivals can be caused not only by far-field noise sources within stationary zones, but also by noise sources between receiver pairs. In the case of the Ridgecrest DAS array, which is located alongside a highway with traffic as the dominant source of noise, the variation of amplitude among the scatterers might be due to noise source attributes rather than the scatterer strength. Therefore, the spurious arrivals' amplitudes are affected largely by their noise sources and are difficult to quantify because they don't share the same noise source as the direct waves (Figure 2d; see section 5.1 for more detailed discussion). In contrast, the coda waves are part of the true Green's function between the two sensors and share the same contributions from noise sources within the stationary zones as the direct waves. In this section, we develop a framework to use the coda waves in noise interferometry to resolve fault zone characteristics.

4.1 Reflection/transmission coefficient ratio

Given a virtual source, the direct wave amplitude in the cross-correlation of the channel on the opposite side of the fault from the source channel can be written as

$$A_{\mathrm{direct}}(f) = A_{\mathrm{CCproc}}(f) A_{\mathrm{src}}(f) A_{\mathrm{path}}(x, f) T(f), \tag{4}$$

where f is the frequency, A_{CCproc} is the amplitude response due to cross-correlation processing, A_{src} is the source effect on the amplitude, A_{path} is the path attenuation effect, x is the location of the receiver channel, T is the transmission coefficient related to the fault properties. Similarly, the coda wave amplitude in the cross-correlation of the channel on the same side as the fault from the source channel can be expressed as

$$A_{\mathrm{coda}}(f) = A_{\mathrm{CCproc}}(f) A_{\mathrm{src}}(f) A'_{\mathrm{path}}(x, f) R(f), \tag{5}$$

where R is the reflection coefficient related to the fault properties. Although it has long been debated whether the absolute amplitude in cross-correlations is usable, taking the amplitude ratio can cancel out the A_{CCproc} term caused by the common processing in the cross-correlation calculation. In addition, if we carefully select two receiver channels that are symmetrical and close enough to the located fault, the path-related attenuation term A_{path} and A'_{path} should be almost identical. The ray paths of the direct and coda waves are shown in Figure 5a. Now, if we divide coda wave amplitudes by direct wave amplitudes recorded on two symmetrical channels, we have

$$\frac{A_{\text{coda}}(f)}{A_{\text{direct}}(f)} = \frac{R(f)}{T(f)}. \tag{6}$$

The concept is that $\frac{A_{\text{coda}}(f)}{A_{\text{direct}}(f)}$ represents the fault properties and should be independent of source or receiver location.

4.2 R/T dispersion measurements and modeling results

Given the locations of virtual source, receivers, and faults, we can predict the travel times of direct and coda waves using equations (1) and (3). We cut a window with a frequency-dependent length around the predicted travel times and measure the peak envelop amplitude. Then the reflection/transmission coefficient R/T is determined with equation (6). As shown in Figure 5, we select a virtual source and filter the cross-correlations in narrow frequency bands. For each pair of channels with the same distance to the fault, we can get the associated R/T ratio. We avoid the channels closest to the fault because coda waves overlap with direct waves. When we shift the channel pair further away from the fault, the measured R/T remains steady (Figure 5c, e). We can also shift virtual sources and repeat the process. The measurements confirm our statement in Section 4.1 that R/T is independent of source location and receiver-to-fault distance. We can see a distinct increase of R/T from 0.12 at 2.5 Hz to 0.16 at 4.5 Hz, indicating clear frequency dependency (Figure 5b-e). Using all available virtual sources and symmetrical channel pairs within 1.2 km from the fault, we can construct the R/T dispersion curve with uncertainty (Figure 6d). The dispersion curve is between 1.5 and 6 Hz because coda waves are difficult to observe outside of this frequency range.

To better understand what the observed R/T dispersion means for fault properties, we simulate the R/T dispersion curves for different fault models using waveform modeling. Many fault parameters, such as fault zone width, depth extent, dipping angle, velocity, attenuation, and country-rock velocities, can influence seismic observations (Lewis & Ben-Zion, 2010; Li et al., 2004; Thurber, 2003). With only the R/T dispersion curve, there will certainly be trade-offs among the many model parameters. In this work, we simplify a fault zone as a rectangular shape with three parameters: fault zone width w, depth extent h, and shear velocity reduction Δv(Figure 5a).

We use a high-resolution shear velocity model along the DAS array as background velocity and embed the rectangular fault in the mapped locations (Yang et al., 2021). The P-wave and density models are calculated with empirical relations in the crust (Brocher, 2005). We perform a rough grid search for the three parameters. For each set of the parameters, we use the fully elastic two-dimensional finite difference code with a grid spacing of 4 m and a time increment of 0.0008 s to ensure accurate simulations up to 10 Hz (Li et al., 2014). Since the coda waves in cross-correlations correspond to the fault-reflected waves in the true Green's function, we directly put the source at the virtual source location without calculating cross-correlations to expedite the grid search process. For the simulated wavefield, we apply the same procedure that we apply to the data to track the travel times of direct and reflected waves and then calculate the R/T dispersion using the peak envelop amplitudes. We use grid search to find the set of parameters that can minimize the $\ell 2$ norm of the misfit between observed and synthetic R/T dispersion. Our grid search results show that the data is best fitted by a 35-m wide, 90-m deep fault with 30% reduction in shear velocity (Figure 6). When we set each of the three parameters to the value of the best-fitted model and examine the two-dimensional grid search results, we find that the fault width and velocity reduction are both well resolved whereas the depth extent is the least resolved (Figure 6a~c).

The resolved fault zone parameters hold important information for fault dynamics. We refer to the characterized low-velocity zone as the fault damage zone. According to field studies on outcrops over different regions, damage zone width can vary from tens of meters to kilometers and is thought to have a scaling law with fault displacement. Even though different regressed scaling relations including linear, logarithm and power laws can span over three orders of magnitude, the damage zone width generally have a positive correlation with fault displacement (Choi et al., 2016; Faulkner et al., 2011; Fossen & Hesthammer, 2000). Our resolved 35-meter wide damage zone could imply a medium-size fault with a fault displacement-damage zone width ratio close to 1 (Torabi & Berg, 2011). On the other hand, our estimated 30% shear wave reduction of the fault damage zone is surprisingly comparable to that of those major faults (20%~60%) studied by fault zone trapped waves (e.g., Lewis & Ben-Zion, 2010; Li et al., 2004). The velocity reduction together with damage zone width and depth can guide numerical modeling of earthquake dynamic ruptures and even long-term earthquake behaviors such as the earthquake cycle duration and potential maximum magnitudes (Huang et al., 2014; Thakur et al., 2020; Weng et al., 2016)

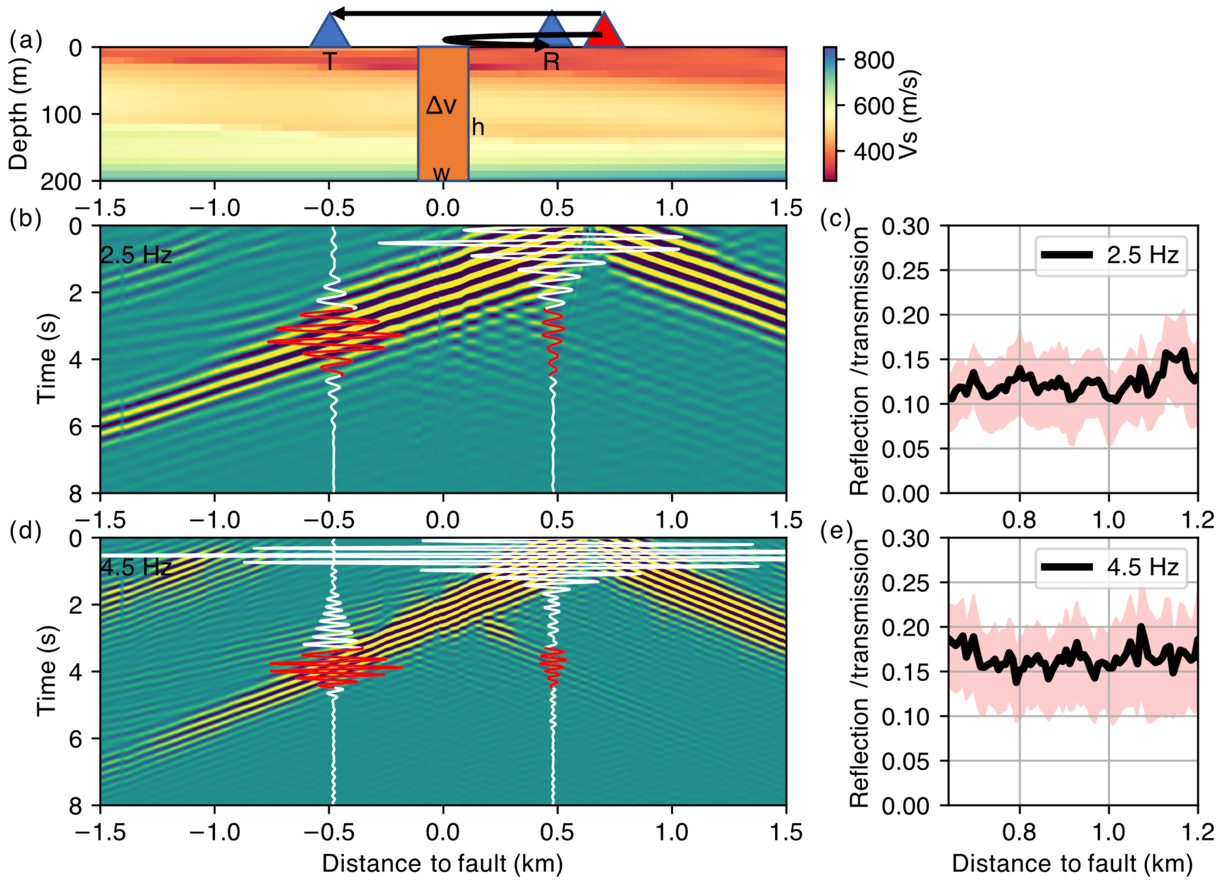

Figure 5 The illustration of reflection/transmission coefficient ratio and the observed frequency dependency. (a) A two-dimensional background shear velocity model with a simplified

rectangular fault in the center. The red triangle represents the channel as virtual source. The blue triangles represent two symmetrical receiver channels regarding the fault. R: reflected wave amplitude, which can be measured by the coda wave amplitude; T: transmitted wave amplitude, which can be measured by the direct wave amplitude; Δv: velocity reduction; w: fault width; h: fault depth; (b) 2.5 Hz cross-correlation record section, the waveforms of the two symmetrical channels are plotted in white lines, with the red portion of the waveform used to measure R and T. (c) R/T measurements at symmetric channel pairs at different distances from the fault. The uncertainty is determined by using 100 different virtual sources; (d) (e) are similar to (b) (c) respectively but for the frequency of 4.5 Hz.

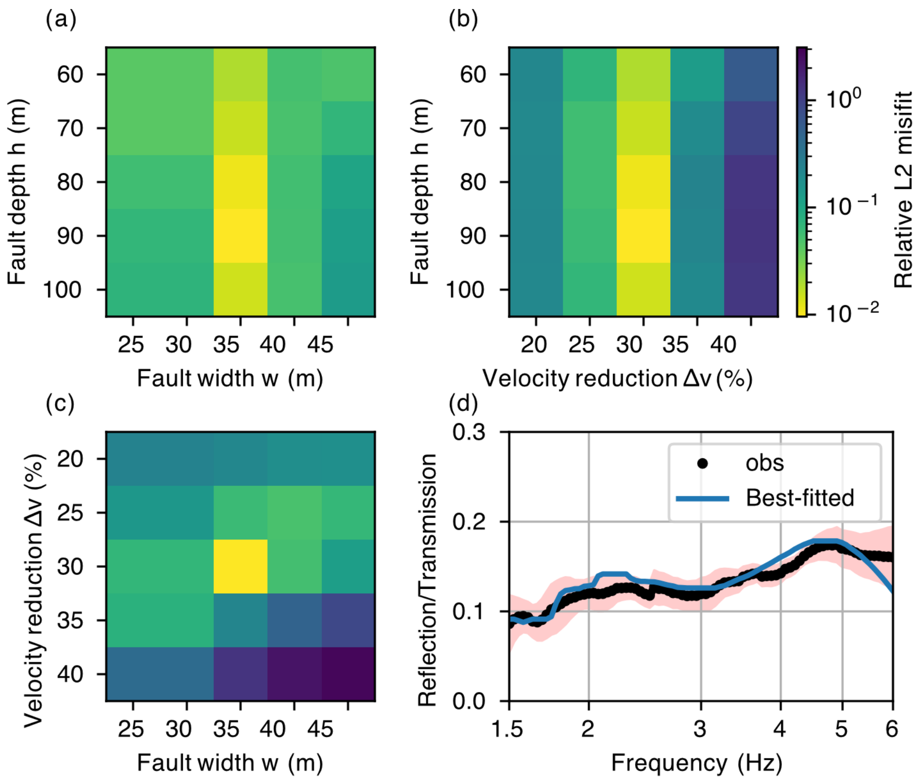

Figure 6 Grid search results of fitting the observed R/T data using waveform modeling. (a)(b)(c) are two-dimensional slices showing the misfit variation with fixed velocity reduction, fault width, and fault depth, respectively. The parameters are fixed at the value of the best-fitted model. (d) The R/T dispersion curve measured by the observed data (black) and the synthetic data using the best-fitted model (blue). The data uncertainty in the red shaded area is calculated by the two times the standard deviation of the measurements from all available virtual sources and symmetrical channel pairs.

5 Discussion

5.1 Understanding the amplitude of spurious arrivals

For a passive scatterer, both spurious arrivals and coda waves are generated by the scattered seismic waves, which is expected to have less coherence and thus weaker amplitudes in the cross-correlations compared to the direct waves. In our observation, all the coda waves have less than 20% amplitude of the direct waves. Some spurious arrivals are stronger than coda waves but remain weaker than direct waves, e.g., at 5.3 km and 7 km (Figure 3b). Some spurious arrivals have exceptionally high amplitudes that are comparable to, if not higher than, the amplitudes of direct wave, e.g., at 4.3 km (Figure 1c, Figure 3b). It was also observed in a recent analysis of the Wasatch fault in Salt Lake City that spurious arrivals arising exactly at the fault have amplitudes comparable to direct waves (Gkogkas et al., 2021).

Here we show that the high amplitudes of spurious arrivals do not necessarily indicate a particularly strong fault or the presence of active source at the located fault. Instead, the cause could be near-field noise sources. As shown in Figure 2d, noise sources between the cross-correlated channel pairs can contribute to spurious arrivals but not to direct waves. This is most certainly the case in our instance because the primary noise source is traffic everywhere along the cable. The less attenuation of the seismic energy from near-field sources may add to the high coherence and subsequent strong spurious arrivals in the cross-correlations. We perform a synthetic test using the two-dimensional finite difference simulation. For this conceptual test, we use a one-dimensional velocity model averaged from the tomography model along this DAS array and add a rectangular fault. The fault parameters are the same as the one used in Section 2.2. We put 20 far-field sources 40 km away from each end of the array and 15 near-field sources evenly distributed from 2 km to 5 km distance (Figure 7a). The synthetic seismogram is then cross correlated between the receiver at 6.2 km and all other receivers. The simulated cross-correlation wavefield confirms that the within-array noise sources can produce spurious arrivals stronger than direct waves, even though no noise source is placed right at the fault (Figure 7b). This explains why we must use the weaker coda waves to characterize the fault zone structures, rather than the spurious arrivals.

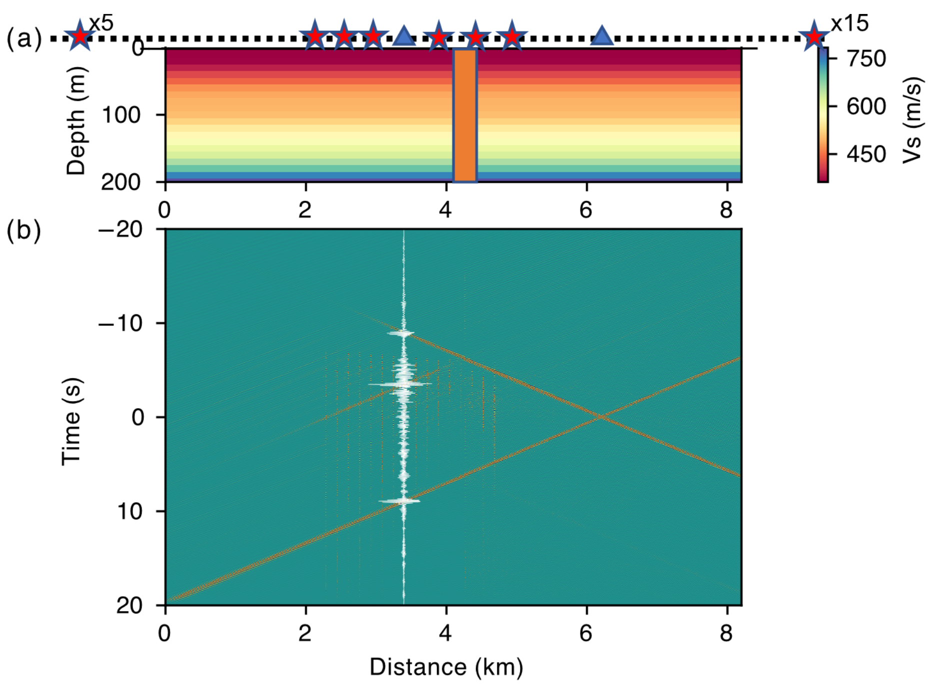

Figure 7 Synthetic noise cross-correlation using waveform modeling. (a) The velocity model used in simulation. Red stars denote the noise sources. In addition to the far-field noise sources in the stationary points, we put several noise sources inside the array. Blue triangles show the

two channels, between which the cross-correlation is plotted as the white waveform in (b). The orange rectangle represents the fault. (b) The cross-correlation wavefield. The blue triangle on the right in (a) is the virtual source.

5.2 Implications for fault imaging at shallow depth

Shallow structures in the top hundreds of meters in general have low seismic velocities, high attenuation, high Vp/Vs ratios, and heterogeneities across very small distances that are challenging to study (e.g., Liu et al., 2015; Qin et al., 2020). Noise interferometry with high-resolution, high-frequency seismic experiments can help enhance our visions on the shallow structure and associated seismic hazards (Castellanos & Clayton, 2021; Yang et al., 2021). Shallow fault complexities such as the splayed features and localized fault-related shallow sources, can further contribute to seismic hazards (e.g., Gradon et al., 2021; Huang & Liu, 2017). Our study using high-frequency surface wave scattering in DAS noise interferometry can capture the faulting structure at the top 100 m and discern different faults at sub-kilometer scales (Figure 4). We also show the accuracy of the mapped fault locations by comparing to the USGS Quaternary fault map and the results from earthquake body-to-surface wave scattering. We then use coda wave amplitudes to give the best-fitting model of the identified shallow fault. The resulted fault geometry is close to that characterized by earthquake body-to-surface wave scattering, which is a good verification of this method (the companion paper by Atterholt et al., in review). The two methods using scattering from different types of waves have complementary sensitivity kernels. For example, body-to-surface wave scattering can discern fault depth of burial by using high-frequency waves while the reflection/transmission coefficient ratio in this work is particularly sensitive to fault zone velocity reduction. Although we do not know whether these located shallow faults are branches that are connected at depth, the mapped shallow locations indicate possible paths that the earthquake rupture can propagate to the surface.

DAS is particularly useful for studying logistically difficult regions, including marine, volcanic, and glacial regions (Lindsey et al., 2019; Nishimura et al., 2021; Walter et al., 2020). This method of passive imaging with DAS can be beneficial for fault detection and imaging for the cases that surface evidence or seismicity catalog is not accessible.

Conclusion

In this work we apply noise interferometry on a 10-km DAS array with 8 meters' spacing in Ridgecrest, California. The dense nature of DAS allows for the recovery of unprecedented wavefield details. We report clear surface wave scattering, including spurious arrivals and scattered coda waves, in noise cross-correlation functions. We use waveform modeling to show that the observed scattered waves can be caused by faults with velocity reduction. We use travel times of the spurious arrivals to map the fault locations. We locate several strong fault scatterers that are generally consistent with the USGS fault map but with refined locations. We further use amplitudes of the coda waves to characterize the geometry and velocity reduction of the mapped faults. We identify a 35-m wide, 90-m deep fault with 30% velocity reduction for one of the identified fault zones. Our results suggest a viable application of DAS for refining prior fault maps or imaging hidden faults at top 100 meters at high lateral resolution in urban areas.

Acknowledgments

This study is funded by the National Science Foundation (NSF) through the Faculty Early Career Development (CAREER) award number 1848166, the Braun Trust, and the United States Geological Survey (USGS) Earthquake Hazards Program (EHP) award number G22AP00067. JA is supported by Graduate Research Fellowships Program (GRFP) number DGE-1745301. We would also like to thank the California Broadband Cooperative for fiber access in the Digital 395 project.

Open Research

Data to reproduce the main results are publicly available (http://doi.org/10.22002/D1.20035). Fault zone data in Figure 1 are downloaded from https://www.usgs.gov/natural-hazards/earthquake-hazards/faults (accessed August 1, 2019).

References

Allam, A. A., Ben-Zion, Y., Kurzon, I., & Vernon, F. (2014). Seismic velocity structure in the Hot Springs and Trifurcation areas of the San Jacinto fault zone, California, from double-difference tomography. *Geophysical Journal International*, *198*(2), 978–999. https://doi.org/10.1093/gji/ggu176

AlTheyab, A., Lin, F. C., & Schuster, G. T. (2016). Imaging near-surface heterogeneities by natural migration of backscattered surface waves. *Geophysical Journal International*, *204*(2), 1332–1341. https://doi.org/10.1093/gji/ggv511

Amos, C. B., Brownlee, S. J., Rood, D. H., Burch Fisher, G., Bürgmann, R., Renne, P. R., & Jayko, A. S. (2013). Chronology of tectonic, geomorphic, and volcanic interactions and the tempo of fault slip near Little Lake, California. *Bulletin of the Geological Society of America*, *125*(7–8), 1187–1202. https://doi.org/10.1130/B30803.1

Atterholt, J. W., Zhan, Z., & Yang, Y. (Companion paper in review). Fault Zone Imaging with Distributed Acoustic Sensing: Body-to-Surface Wave Scattering

Ben-Zion, Y. (2008). Collective Behavior of Earthquakes and Faults. *Reviews of Geophysics*, *46*, 1–70. https://doi.org/10.1029/2008RG000260.1.INTRODUCTION

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., et al. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, *169*(3), 1239–1260. https://doi.org/10.1111/j.1365-246X.2007.03374.x

Brandenberg, S. J., Wang, P., Nweke, C. C., Hudson, K., Mazzoni, S., Bozorgnia, Y., et al. (2019). Preliminary Report on Engineering and Geological Effects of the July 2019 Ridgecrest Earthquake Sequence, Report of the NSF-Sponsored Geotechnical Extreme Event

Reconnaissance Association, (July), 1–77. https://doi.org/10.18118/G6H66K

Brocher, T. M. (2005). Empirical Relations between Elastic Wavespeeds and Density in the Earth's Crust. *Bulletin of the Seismological Society of America*, *95*(6), 2081–2092. https://doi.org/10.1785/0120050077

Caine, J. S., Evans, J. P., & Forster, C. B. (1996). Fault zone architecture and permeability structure. *Geology*, *24*(11), 1025–1028. https://doi.org/10.1130/0091-7613(1996)024<1025:FZAAPS>2.3.CO;2

Castellanos, J. C., & Clayton, R. W. (2021). The Fine-Scale Structure of Long Beach, California, and Its Impact on Ground Motion Acceleration. *Journal of Geophysical Research: Solid Earth*, *126*(12). https://doi.org/10.1029/2021jb022462

Cheng, F., Chi, B., Lindsey, N. J., Dawe, T. C., & Ajo-Franklin, J. B. (2021). Utilizing distributed acoustic sensing and ocean bottom fiber optic cables for submarine structural characterization. *Scientific Reports*, *11*(1), 1–14. https://doi.org/10.1038/s41598-021-84845-y

Choi, J. H., Edwards, P., Ko, K., & Kim, Y. S. (2016). Definition and classification of fault damage zones: A review and a new methodological approach. *Earth-Science Reviews*, *152*, 70–87. https://doi.org/10.1016/j.earscirev.2015.11.006

Collettini, C., Niemeijer, A., Viti, C., & Marone, C. (2009). Fault zone fabric and fault weakness. *Nature*, *462*(7275), 907–910. https://doi.org/10.1038/nature08585

Faulkner, D. R., Lewis, A. C., & Rutter, E. H. (2003). On the internal structure and mechanics of large strike-slip fault zones: Field observations of the Carboneras fault in southeastern Spain. *Tectonophysics*, *367*(3–4), 235–251. https://doi.org/10.1016/S0040-1951(03)00134-3

Faulkner, D. R., Mitchell, T. M., Jensen, E., & Cembrano, J. (2011). Scaling of fault damage zones with displacement and the implications for fault growth processes. *Journal of Geophysical Research: Solid Earth*, *116*(5), 1–11. https://doi.org/10.1029/2010JB007788

Fossen, H., & Hesthammer, J. (2000). <JSG 2000 Fossen Hesthammer.pdf>, *22*, 851–863.

Gkogkas, K., Lin, F. C., Allam, A. A., & Wang, Y. (2021). Shallow damage zone structure of the wasatch fault in Salt Lake City from ambient-noise double beamforming with a temporary linear array. *Seismological Research Letters*, *92*(4), 2453–2463. https://doi.org/10.1785/0220200404

Gradon, C., Roux, P., Moreau, L., Lecointre, A., & Ben Zion, Y. (2021). Characterization with dense array data of seismic sources in the shallow part of the San Jacinto fault zone. *Geophysical Journal International*, *224*(2), 1133–1140. https://doi.org/10.1093/gji/ggaa411

Hickman, S., Zoback, M., & Ellsworth, W. (2004). Introduction to special section: Preparing for the San Andreas Fault Observatory at Depth. *Geophysical Research Letters*, *31*(12), 10–13. https://doi.org/10.1029/2004GL020688

Huang, L., & Liu, C. Y. (2017). Three Types of Flower Structures in a Divergent-Wrench Fault Zone. *Journal of Geophysical Research: Solid Earth*, *122*(12), 10,478-10,497. https://doi.org/10.1002/2017JB014675

Huang, Y., Ampuero, J., & Helmberger, D. V. (2014). Earthquake ruptures modulated by waves in damaged fault zones. *Journal of Geophysical Research: Solid Earth*, *119*(4), 3133–3154. https://doi.org/10.1002/2013JB010724

Hung, J. H., Ma, K. F., Wang, C. Y., Ito, H., Lin, W., & Yeh, E. C. (2009). Subsurface structure, physical properties, fault-zone characteristics and stress state in scientific drill holes of Taiwan Chelungpu Fault Drilling Project. *Tectonophysics*, *466*(3–4), 307–321. https://doi.org/10.1016/j.tecto.2007.11.014

Jennings, C. W. (1975). *Fault map of California with locations of volcanoes, thermal springs, and thermal wells: California Division of Mines and Geology California Geologic Data Map 1, scale 1:750,000*.

Lewis, M. A., & Ben-Zion, Y. (2010). Diversity of fault zone damage and trapping structures in the Parkfield section of the San Andreas Fault from comprehensive analysis of near fault seismograms. *Geophysical Journal International*, *183*(3), 1579–1595. https://doi.org/10.1111/j.1365-246X.2010.04816.x

Lewis, M. A., Ben-Zion, Y., & McGuire, J. J. (2007). Imaging the deep structure of the San Andreas Fault south of Hollister with joint analysis of fault zone head and direct P arrivals. *Geophysical Journal International*, *169*(3), 1028–1042. https://doi.org/10.1111/j.1365-246X.2006.03319.x

Lewis, Michael A., & Ben-Zion, Y. (2010). Diversity of fault zone damage and trapping structures in the Parkfield section of the San Andreas Fault from comprehensive analysis of near fault seismograms. *Geophysical Journal International*, *183*(3), 1579–1595. https://doi.org/10.1111/j.1365-246X.2010.04816.x

Li, D., Helmberger, D., Clayton, R. W., & Sun, D. (2014). Global synthetic seismograms using a 2-D finite-difference method. *Geophysical Journal International*, *197*(2), 1166–1183. https://doi.org/10.1093/gji/ggu050

Li, Y., & Malin, P. E. (2008). San Andreas Fault damage at SAFOD viewed with fault-guided waves. *Geophysical Research Letters*, *35*(8), 1–6. https://doi.org/10.1029/2007GL032924

Li, Y., Vidale, J. E., & Cochran, E. S. (2004). Low-velocity damaged structure of the San Andreas Fault at Parkfield from fault zone trapped waves. *Geophysical Research Letters*, *31*(12), 1–5. https://doi.org/10.1029/2003GL019044

Li, Z., Shen, Z., Yang, Y., Williams, E., Wang, X., & Zhan, Z. (2021). Rapid Response to the 2019

Ridgecrest Earthquake With Distributed Acoustic Sensing. *AGU Advances*, *2*(2). https://doi.org/10.1029/2021av000395

Lindsey, N. J., & Martin, E. R. (2021). Fiber-Optic Seismology. *Annual Review of Earth and Planetary Sciences*, *49*(1), 309–336. https://doi.org/10.1146/annurev-earth-072420-065213

Lindsey, N. J., Craig Dawe, T., & Ajo-Franklin, J. B. (2019). Illuminating seafloor faults and ocean dynamics with dark fiber distributed acoustic sensing. *Science*, *366*(6469), 1103–1107. https://doi.org/10.1126/science.aay5881

Liu, X., Ben-Zion, Y., & Zigone, D. (2015). Extracting seismic attenuation coefficients from cross-correlations of ambient noise at linear triplets of stations. *Geophysical Journal International*, *203*(2), 1149–1163. https://doi.org/10.1093/gji/ggv357

Liu, X., Beroza, G. C., Yang, L., & Ellsworth, W. L. (2021). Ambient noise Love wave attenuation tomography for the LASSIE array across the Los Angeles basin. *Science Advances*, *7*(22), 2–8. https://doi.org/10.1126/sciadv.abe1030

Ma, Y., Clayton, R. W., Tsai, V. C., & Zhan, Z. (2013). Locating a scatterer in the active volcanic area of Southern Peru from ambient noise cross-correlation. *Geophysical Journal International*, *192*(3), 1332–1341. https://doi.org/10.1093/gji/ggs103

Martin, E. R., Lindsey, N. J., Ajo-Franklin, J. B., & Biondi, B. (2018). Introduction to Interferometry of Fiber Optic Strain Measurements. *EarthArXiv*, *2*, 1–33.

Mitchell, T. M., & Faulkner, D. R. (2009). The nature and origin of off-fault damage surrounding strike-slip fault zones with a wide range of displacements: A field study from the Atacama fault system, northern Chile. *Journal of Structural Geology*, *31*(8), 802–816. https://doi.org/10.1016/j.jsg.2009.05.002

Nishimura, T., Emoto, K., Nakahara, H., Miura, S., Yamamoto, M., Sugimura, S., et al. (2021). Source location of volcanic earthquakes and subsurface characterization using fiber-optic cable and distributed acoustic sensing system. *Scientific Reports*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-85621-8

Perrin, C., Manighetti, I., Ampuero, J.-P., Cappa, F., & Gaudemer, Y. (2016). Location of largest earthquake slip and fast rupture controlled by along-strike change in fault structural maturity due to fault growth. *Journal of Geophysical Research: Solid Earth*, *121*(5), 3666–3685. https://doi.org/10.1002/2015JB012671

Qin, L., Ben-Zion, Y., Bonilla, L. F., & Steidl, J. H. (2020). Imaging and Monitoring Temporal Changes of Shallow Seismic Velocities at the Garner Valley Near Anza, California, Following the M7.2 2010 El Mayor-Cucapah Earthquake. *Journal of Geophysical Research: Solid Earth*, *125*(1), 1–17. https://doi.org/10.1029/2019JB018070

Qiu, H., Allam, A. A., Lin, F. C., & Ben-Zion, Y. (2020). Analysis of Fault Zone Resonance Modes

Recorded by a Dense Seismic Array Across the San Jacinto Fault Zone at Blackburn Saddle. *Journal of Geophysical Research: Solid Earth*, *125*(10), 1–21. https://doi.org/10.1029/2020JB019756

Retailleau, L., & Beroza, G. C. (2021). Towards structural imaging using seismic ambient field correlation artefacts. *Geophysical Journal International*, *225*(2), 1453–1465. https://doi.org/10.1093/gji/ggab038

Ross, Z. E., Hauksson, E., & Ben-Zion, Y. (2017). Abundant off-fault seismicity and orthogonal structures in the San Jacinto fault zone. *Science Advances*, *3*(3). https://doi.org/10.1126/sciadv.1601946

Ross, Z. E., Idini, B., Jia, Z., Stephenson, O. L., Zhong, M., Wang, X., et al. (2019). 2019 Ridgecrest earthquake sequence. *Science*, *3665*(October), 346–351.

Snieder, R. (2004). Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*. https://doi.org/10.1103/PhysRevE.69.046610

Spudich, P., & Olsen, K. B. (2001). Fault zone amplified waves as a possible seismic hazard along the Calaveras Fault in central California. *Geophysical Research Letters*, *28*(13), 2533–2536. https://doi.org/10.1029/2000GL011902

Thakur, P., Huang, Y., & Kaneko, Y. (2020). Effects of Low-Velocity Fault Damage Zones on Long-Term Earthquake Behaviors on Mature Strike-Slip Faults. *Journal of Geophysical Research: Solid Earth*, *125*(8), 1–20. https://doi.org/10.1029/2020JB019587

Thurber, C. (2003). Earthquake locations and three-dimensional fault zone structure along the creeping section of the San Andreas fault near Parkfield, CA: Preparing for SAFOD. *Geophysical Research Letters*, *30*(3), 10–13. https://doi.org/10.1029/2002gl016004

Torabi, A., & Berg, S. S. (2011). Scaling of fault attributes: A review. *Marine and Petroleum Geology*, *28*(8), 1444–1460. https://doi.org/10.1016/j.marpetgeo.2011.04.003

Walter, F., Gräff, D., Lindner, F., Paitz, P., Köpfli, M., Chmiel, M., & Fichtner, A. (2020). Distributed acoustic sensing of microseismic sources and wave propagation in glaciated terrain. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15824-6

Wang, X., & Zhan, Z. (2020). Seismotectonics and Fault Geometries of the 2019 Ridgecrest Sequence: Insight From Aftershock Moment Tensor Catalog Using 3-D Green's Functions. *Journal of Geophysical Research: Solid Earth*, *125*(5), 1–15. https://doi.org/10.1029/2020JB019577

Wang, Y., Allam, A., & Lin, F. C. (2019). Imaging the Fault Damage Zone of the San Jacinto Fault Near Anza With Ambient Noise Tomography Using a Dense Nodal Array. *Geophysical Research Letters*, *46*(22), 12938–12948. https://doi.org/10.1029/2019GL084835

Weng, H., Yang, H., Zhang, Z., & Chen, X. (2016). Earthquake rupture extents and coseismic slips promoted by damaged fault zones. *Journal of Geophysical Research: Solid Earth*, *121*(6), 4446–4457. https://doi.org/10.1002/2015JB012713

Yang, Y., Atterholt, J. W., Shen, Z., Muir, J. B., Williams, E. F., & Zhan, Z. (2021). Sub-kilometer correlation between near-surface structure and ground motion measured with distributed acoustic sensing. *Geophysical Research Letters*. https://doi.org/10.1029/2021gl096503

Zeng, X., & Ni, S. (2010). A persistent localized microseismic source near the Kyushu Island, Japan. *Geophysical Research Letters*, *37*(24), 8–13. https://doi.org/10.1029/2010GL045774

Zhan, Z. (2019). Distributed acoustic sensing turns fiber-optic cables into sensitive seismic antennas. *Seismological Research Letters*, *91*(1), 1–15. https://doi.org/10.1785/0220190112

Zhan, Z., Ni, S., Helmberger, D. V., & Clayton, R. W. (2010). Retrieval of Moho-reflected shear wave arrivals from ambient seismic noise. *Geophysical Journal International*, *182*(1), 408–420. https://doi.org/10.1111/j.1365-246X.2010.04625.x

U.S. Geological Survey and New Mexico Bureau of Mines and Mineral Resources, Quaternary fault and fold database for the United States, accessed August 1, 2019, at: https://www.usgs.gov/natural-hazards/earthquake-hazards/faults