

Fiber Seismic Tomography for Geothermal Exploration

Ettore Biondi*, Jiaxuan Li, Valey Kamalov, and Zhongwen Zhan

1200 E. California Blvd., MC 252-21 South Mudd Building, Pasadena, CA 91125

ebiondi@caltech.edu

Keywords: Seismic tomography, distributed acoustic sensing, Long Valley caldera

ABSTRACT

High-resolution delineation of subsurface anomalies associated with geothermal reservoirs is a costly and significant geophysical and geological effort. We present a novel seismic tomographic approach that employs machine-learning traveltimes picks obtained on distributed acoustic sensing (DAS) data recorded on existing telecommunication fibers located within the Long Valley caldera, California. Our imaging methodology can resolve the seismic features related to the shallow hydrothermal system present within the caldera.

1. INTRODUCTION

The identification of the subsurface structures where high-temperature fluids are present within geothermal systems is critical to developing an effective exploitation strategy of the heat sources for energy production (Huenges and Ledru, 2011; Dobson, 2016). Magnetotellurics methods can characterize shallow subsurface structures but lack the resolution to characterize the connectivity between heat reservoirs or deep hydrothermal systems (Aizawa et al., 2009). On the other hand, seismic imaging methods can resolve the mesoscale velocity structures characterizing a geothermal system (Paulatto et al., 2022). However, the success of seismic imaging and monitoring efforts is contingent upon the spatial density of the deployed recording devices.

In recent years, DAS deployed on telecom optical cables has shown the potential of recording seismic data with an unprecedented spatio-temporal resolution that could not be achieved with conventional instruments (Ajo-Franklin et al., 2019). A modern DAS instrument can transform up to 100 km of a telecom fiber into a dense array of channels with meter-level spacing, which is opening new opportunities for monitoring natural and induced earthquakes and characterizing subsurface structures by transforming the ever-expanding telecommunication infrastructure into sensing antennas (Zhan, 2020).

Besides the deployment of DAS on telecom fibers, such technology can be employed on cables placed within boreholes, which permits the recording of deformation signals close to induced seismicity (Jin and Roy, 2017). Such recordings enable the imaging of the natural or stimulated fracture network necessary to estimate rock permeability and monitor induced seismicity due to enhanced geothermal system (EGS) activities. Compared to conventional borehole instrumentation, DAS employs fiber cables that can withstand high-temperature environments (Ma et al., 2022), making this technology suitable for recording seismic signals during EGS operations. We developed a subsurface imaging workflow with km spatial resolution enabling the identification of high-temperature formations. Our methodology makes use of thousands of seismic channels in known locations obtained by converting existing optical fiber into sensitive acoustic sensors to achieve ultimate spatial resolution and AI-based algorithms to convert optical signals into geophysical parameters. We first describe the mathematical formulation of our tomographic workflow and then showcase a synthetic example where a low-velocity anomaly can be resolved by inverting traveltimes from subsurface events recorded with surface instruments. We then describe a field-data application of our approach on DAS data recorded using pre-existing telecommunication cables located in the Long Valley caldera, California. Our method can resolve the low-velocity features associated with the shallow hydrothermal systems, which are currently used for geothermal energy production. This field example demonstrates the potential for exploiting the existing fiber network for geothermal prospect exploration, which in turn would significantly decrease the cost of geophysical analysis.

2. ADJOINT-STATE EIKONAL TOMOGRAPHY

Within our tomography workflow, we consider the Eikonal equation to accurately model event traveltimes. For a given event at location \mathbf{x}_s , its traveltime field τ at any location \mathbf{x} within the considered event is constrained by the following partial-differential equation:

$$\nabla \cdot \nabla \tau(\mathbf{x}) = \frac{1}{v^2(\mathbf{x})} \quad \text{s.t. } \tau(\mathbf{x}_s) = 0, \quad (1)$$

where v represents the velocity field. Similarly, as well-known tomographic packages (Zhang and Thurber, 2006), we perform our methodology using a double-difference (DD) approach (Waldhauser and Ellsworth, 2000). Compared to these methodologies, however, our inversion scheme is matrix-free and based on non-linear optimization iterative schemes. We compute the necessary objective function gradients using the adjoint-state method applied to the Eikonal equation (Plessix, 2006; Li et al., 2013; Tong, 2021). By following a matrix-free approach during the optimization process, we can invert millions to billions of traveltimes picks obtained using the thousands of channels composing any commonly deployed DAS arrays. Within a DD framework, the size of the considered least-squares matrices would be on the order of billions of elements, which on modern computational architectures would be impossible to invert within a reasonable time. For this reason, we implement our workflow using an operator-based optimization approach (Biondi et al., 2021). To alleviate the location and velocity structure trade-off, we perform the traveltime inversion in an alternate-direction fashion (Tong, 2021):

we first relocate the events for fixed P- and S-wave velocity structures, then we fix the earthquake locations and invert the P- and S-wave models. These two steps are alternated until convergence is reached based on locations and velocity model changes. For the relocation step, we minimize the following objective function:

$$\phi(\mathbf{x}_s) = \frac{1}{2} \left\| \begin{bmatrix} \lambda_A \mathbf{I} \\ \lambda_{DD} \mathbf{D} \end{bmatrix} [\mathbf{f}_v(\mathbf{x}_s) + \boldsymbol{\tau}_0 - \boldsymbol{\tau}_{obs}] \right\|_2^2 + \frac{\epsilon}{2} \|\mathbf{x}_s - \mathbf{x}_{s,0}\|_2^2, \quad (2)$$

where λ_A and λ_{DD} are the relative weights of the absolute and double-difference traveltimes (Zhang and Thurber, 2006), respectively, $\boldsymbol{\tau}_{obs}$ are the observed P- and S-wave traveltimes, \mathbf{f}_v is the solution to equation 1 for a fixed velocity structure (White et al., 2020), $\boldsymbol{\tau}_0$ is event origin times, and \mathbf{D} is the double-difference operator (Waldhauser and Ellsworth, 2000). The regularization term based on the initial event locations $\mathbf{x}_{s,0}$ and weight by the scalar ϵ is necessary to avoid inversion instabilities when $\lambda_A=0$ (Waldhauser and Ellsworth, 2000). For each event, the optimal origin time is found at the beginning and at the end of the relocation process with the following equation:

$$\tau_0^j = \frac{1}{N_r} \sum_{k=1}^{N_r} [\tau_{obs}^{j,k} - f_v^{j,k}(\mathbf{x}_s)], \quad (3)$$

where optimal origin time for the j -th event is simply the average of the traveltimes residuals over the picked N_r stations.

The velocity models are obtained by minimizing the following cost function:

$$\psi(v) = \frac{1}{2} \left\| \begin{bmatrix} \lambda_A \mathbf{I} \\ \lambda_{DD} \mathbf{D} \end{bmatrix} [\mathbf{f}_{x_s}(v) + \boldsymbol{\tau}_0 - \boldsymbol{\tau}_{obs}] \right\|_2^2, \quad (4)$$

where \mathbf{f}_{x_s} is the Eikonal operator for fixed event locations \mathbf{x}_s . Smoothness in the model is imposed by applying a Gaussian filter to the computed gradient during the optimization process. We employ the Broyden–Fletcher–Goldfarb–Shanno (Liu and Nocedal, 1989) iterative scheme to minimize both objective functions and the iteration process is stopped when the considered model parameters are not changing significantly.

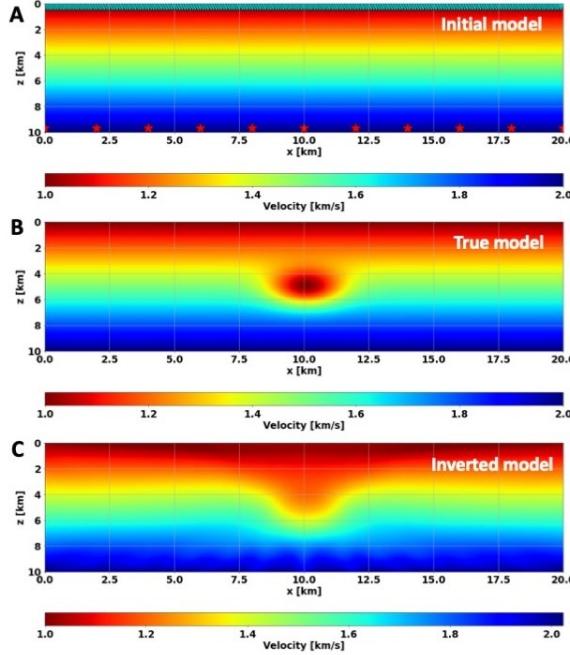


Figure 1: Synthetic tomography test. (A) Initial velocity model where the sources and receivers considered are plotted as red stars and cyan triangles, respectively. (B) True subsurface model assumed for the synthetic test. (C) Velocity model retrieved by our tomographic Eikonal matrix-free algorithm.

2.1 Synthetic example

We test our tomographic workflow on a simple 2D synthetic test. We assume to obtain traveltimes from 11 events located at the bottom of the domain and recorded by a surface array with a channel distance of 50 meters (Figure 1a). The true velocity model contains a Gaussian low-velocity anomaly located in the center of the domain (Figure 1b), while our initial guess is a laterally invariant velocity field (Figure 1a). The smoothness of the model is achieved by applying a Gaussian filter of 2 km during the inversion process. The iterative process converges after 26 iterations and achieves a traveltime matching of 0.1 ms on average starting from a maximum traveltime

error of approximately 6.6 s, highlighting the numerical accuracy of our inversion methodology. From the inverted model (Figure 1c), we can observe that our method can successfully image the missing anomaly despite underestimating the velocity decrease within its core.

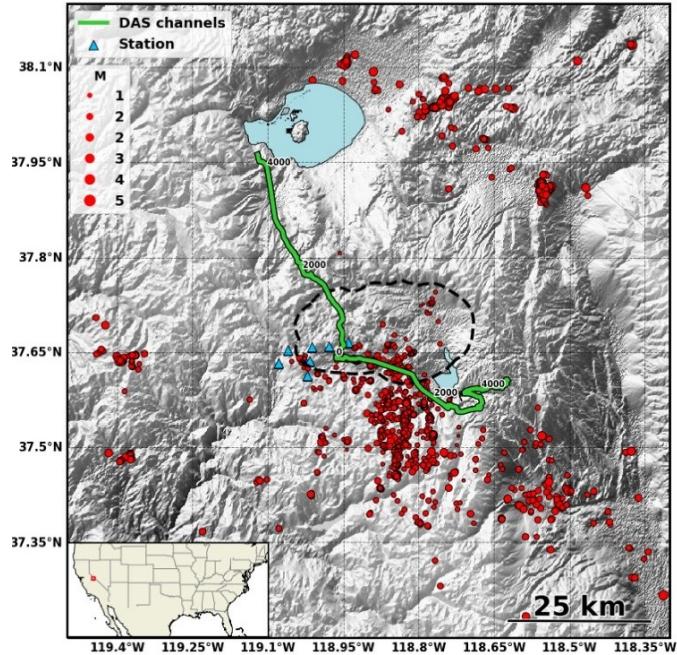


Figure 2: Area considered for the field experiment. Conventional stations are depicted by the blue triangles while the green line indicates the locations of the DAS channels. The earthquakes are indicated by the red dots whose sizes are proportional to their magnitude. The inset map shows the location of the study area.

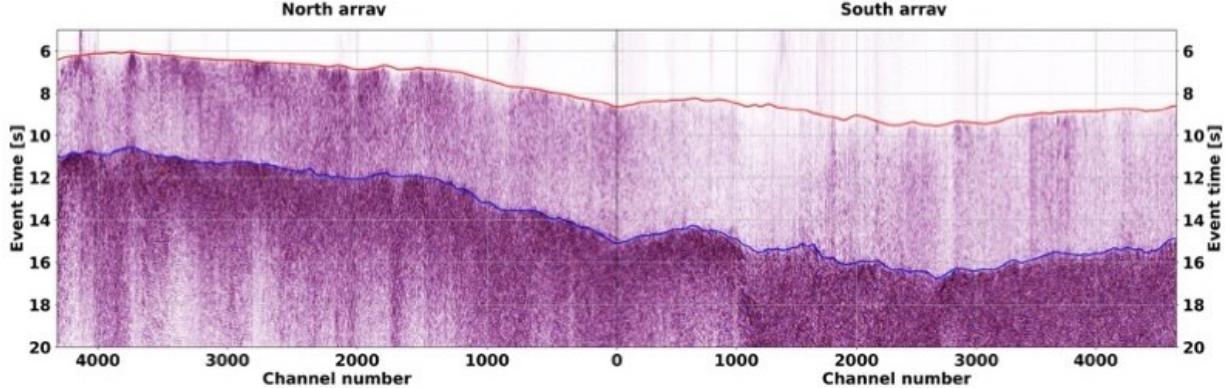


Figure 3: Representative example of travelttime picking. Strain recorded by the DAS arrays induced by a local event. The red and blue curves depict the P- and S-wave PhaseNet-DAS-picked traveltimes.

2.2 Field example: Long Valley caldera

We apply our tomographic workflow on DAS data recorded on two telecommunication fiber cables located in the Long Valley caldera (Figure 2), located in the Eastern Sierra, which is one of the largest calderas in the North American continent. This geological feature was formed approximately 767 ka by a single eruptive event that released 650 km³ of rhyolitic material into the atmosphere (Bailey et al., 1976; Hildreth and Wilson, 2007). The hydrothermal system of this caldera has been producing geothermal energy since 1984 with a current capacity of 45 MW (Sorey et al., 1995). Our two DAS arrays are composed of more than 9000 channels covering an approximately 100-km north-south transect across the caldera, whose locations are obtained using a precise tap-test methodology (Biondi et al., 2023a). Over a span of 12 months (between November 2020 and November 2021), we recorded more than 6000 local and regional events as part of the catalog maintained by the Northern California Earthquake Data Center (NCEDC) Waldhauser and Schaff (2008). From these events, we select 843 earthquakes whose signal-noise ratio (SNR) on the DAS data is on average approximately 40 dB. We estimate the SNR of each event by computing the noise and signal energy using 2-second and 0.8-second windows before and after the initially predicted P-wave travelttime, respectively.

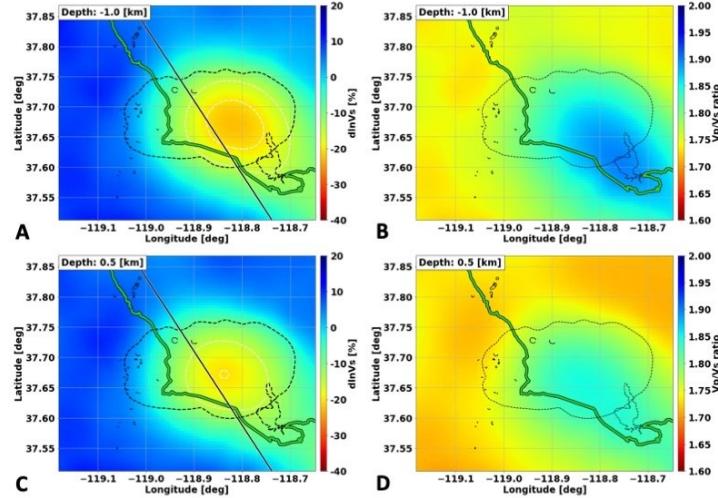


Figure 4: Initial Long Valley caldera properties. Long Valley subsurface properties extracted from the initial 3D model at two different depth levels with respect to the mean sea level. (A and C) S-wave wave speed perturbation with respect to a reference 1D velocity model. (B and D) Vp/Vs ratio for the same depth slices.

Accurate arrival times play a crucial role in any efficient and successful tomographic process. To obtain such traveltimes from DAS data, we utilized a machine-learning (ML) algorithm to precisely capture more than 12 million arrival times for P- and S-waves called PhaseNet-DAS (Zhu et al., 2023). In comparison to tables generated from conventional seismic arrays, even those with high density, DAS data yields arrival time tables with a substantially larger number of picks—by 2 to 3 orders of magnitude—posing a computational challenge for existing tomographic methodologies. Figure 3 displays an example of traveltime picking for a local event recorded on our DAS array. Using a cross-correlation methodology and assuming Gaussian distributed traveltime picking errors (Li et al., 2023), we estimate picking errors of our ML approach to have standard deviations of 49.5 ms and 99 ms for P- and S-wave, respectively.

To initialize our tomographic workflow for this field dataset, we employ velocity models based on full waveform inversion of surface waves between 6 and 20 s Lee et al. (2014). The panels in Figure 4 depict the VS anomalies (relative to a 1D velocity model) and the VP/VS ratio for two depth slices. In these panels, a single anomaly centered within the caldera is visible, with no discernible small-scale features.

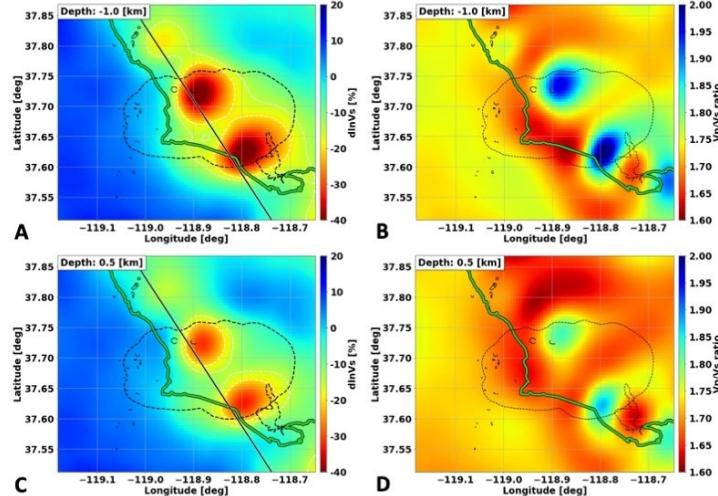


Figure 5: Inverted Long Valley caldera properties from fiber tomography. Long Valley subsurface properties extracted from our fiber-tomography results for the same depth levels as in Figure 4. (A and C) S-wave wave speed perturbation with respect to a reference 1D velocity model. (B and D) Vp/Vs ratio for the same depth slices.

After applying our tomographic workflow (Figure 5), the shallow structures are resolved, particularly well for shallow depths. In particular, two kilometric-scale low velocities anomalies with a high VP/VS ratio (greater than 1.8) are imaged within the bounds of the caldera. These features are likely associated with the shallow hydrothermal system currently utilized for geothermal energy production. Their locations align with previously identified resistivity anomalies from magneto-telluric methods and the presence of surface hydrothermal springs Pakiser et al. (1960); Peacock et al. (2015). A more detailed discussion of the system-scale images and their scientific

implications is provided in Biondi et al. (2023b). This field-data example demonstrates the efficacy of our approach to resolving the hydrothermal system commonly present within volcanic calderas.

3. CONCLUSIONS

We introduce an innovative matrix-free and computationally efficient adjoint traveltime tomography workflow. The mathematical formulation is outlined, and a synthetic example showcases the method's capability to delineate subsurface velocity anomalies using seismicity, whether natural or induced, recorded by surface seismic instrumentation. To illustrate its potential in geothermal geophysical exploration, we apply this methodology to seismic data obtained through DAS units deployed on telecommunication fibers within the Long Valley caldera. Our approach yields high-resolution images of the shallow hydrothermal system, offering valuable insights for guiding geothermal energy exploitation. Leveraging both natural seismicity and existing telecommunication infrastructure, this method has the potential to significantly reduce geophysical exploration costs in geothermal energy projects.

4. ACKNOWLEDGMENTS

We would like to thank OptaSense for the support provided for the Long Valley experiment. In particular, the authors thank Martin Karrenbach, Victor Yartsev, and Vlad Bogdanov. We also thank the California Broadband Cooperative for providing access to the Digital 395 telecommunication fibers. This work is supported by National Science Foundation (NSF) Faculty Early Career Development Program (CAREER) Award Number 1848166, the Resnick Institute of Sustainability, and the Gordon and Betty Moore Foundation.

REFERENCES

Aizawa, K., Y. Ogawa, and T. Ishido, 2009, Groundwater flow and hydrothermal systems within volcanic edifices: Delineation by electric self-potential and magnetotellurics: *Journal of Geophysical Research: Solid Earth*, 114.

Ajo-Franklin, J. B., S. Dou, N. J. Lindsey, I. Monga, C. Tracy, M. Robertson, V. Rodriguez Tribaldos, C. Ulrich, B. Freifeld, T. Daley, et al., 2019, Distributed acoustic sensing using dark fiber for near-surface characterization and broadband seismic event detection: *Scientific reports*, 9, 1328.

Bailey, R. A., G. B. Dalrymple, and M. A. Lanphere, 1976, Volcanism, structure, and geochronology of Long Valley caldera, Mono county, California: *Journal of Geophysical Research*, 81, 725–744.

Biondi, E., G. Barnier, R. G. Clapp, F. Picetti, and S. Farris, 2021, An object-oriented optimization framework for large-scale inverse problems: *Computers & Geosciences*, 154, 104790.

Biondi, E., X. Wang, E. F. Williams, and Z. Zhan, 2023a, Geolocalization of large-scale DAS channels using a GPS-tracked moving vehicle: *Seismological Society of America*, 94, 318–330.

Biondi, E., W. Zhu, J. Li, E. F. Williams, and Z. Zhan, 2023b, An upper-crust lid over the Long Valley magma chamber: *Science Advances*, 9, eadi9878.

Dobson, P. F., 2016, A review of exploration methods for discovering hidden geothermal systems: *GRC Transactions*, 40, 695–706.

Hildreth, W., and C. J. Wilson, 2007, Compositional zoning of the Bishop Tuff: *Journal of Petrology*, 48, 951–999.

Huenges, E., and P. Ledru, 2011, Geothermal energy systems: exploration, development, and utilization: John Wiley & Sons.

Jin, G., and B. Roy, 2017, Hydraulic-fracture geometry characterization using low-frequency DAS signal: *The Leading Edge*, 36, 975–980.

Lee, E.-J., P. Chen, T. H. Jordan, P. B. Maechling, M. A. Denolle, and G. C. Beroza, 2014, Full-3-d tomography for crustal structure in southern California based on the scattering-integral and the adjoint-wavefield methods: *Journal of Geophysical Research: Solid Earth*, 119, 6421–6451.

Li, J., W. Zhu, E. Biondi, and Z. Zhan, 2023, Earthquake focal mechanisms with distributed acoustic sensing: *Nature Communications*, 14, 4181.

Li, S., A. Vladimirs, and S. Fomel, 2013, First-break traveltime tomography with the double-square-root eikonal equation: *Geophysics*, 78, U89–U101.

Liu, D. C., and J. Nocedal, 1989, On the limited memory BFGS method for large scale optimization: *Mathematical programming*, 45, 503–528.

Ma, S., Y. Xu, Y. Pang, X. Zhao, Y. Li, Z. Qin, Z. Liu, P. Lu, and X. Bao, 2022, Optical fiber sensors for high-temperature monitoring: a review: *Sensors*, 22, 5722.

Pakiser, L. C., F. Press, and M. F. Kane, 1960, Geophysical investigation of Mono Basin, California: *Geological Society of America Bulletin*, 71, 415–448.

Paulatto, M., E. Hooft, K. Chrapkiewicz, B. Heath, D. Toomey, and J. Morgan, 2022, Advances in seismic imaging of magma and crystal mush: *EarthArXiv*.

Peacock, J. R., M. T. Mangan, D. McPhee, and D. A. Ponce, 2015, Imaging the magmatic system of Mono Basin, California, with magnetotellurics in three dimensions: *Journal of Geophysical Research: Solid Earth*, 120, 7273–7289.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: *Geophysical Journal International*, 167, 495–503.

Sorey, M., C. Farrar, G. Marshall, and J. Howie, 1995, Effects of geothermal development on deformation in the Long Valley Caldera, eastern California, 1985–1994: *Journal of Geophysical Research: Solid Earth*, 100, 12475–12486.

Tong, P., 2021, Adjoint-state traveltimes tomography: Eikonal equation-based methods and application to the Anza Area in Southern California: *Journal of Geophysical Research: Solid Earth*, 126, e2021JB021818.

Waldhauser, F., and W. L. Ellsworth, 2000, A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California: *Bulletin of the seismological society of America*, 90, 1353–1368.

Waldhauser, F., and D. Schaff, 2008, Large-scale cross correlation based relocation of two decades of northern California seismicity: *J. Geophys. Res.*, 113, B08311.

White, M. C., H. Fang, N. Nakata, and Y. Ben-Zion, 2020, Pykonal: a Python package for solving the eikonal equation in spherical and Cartesian coordinates using the fast marching method: *Seismological Research Letters*, 91, 2378–2389.

Zhan, Z., 2020, Distributed acoustic sensing turns fiber-optic cables into sensitive seismic antennas: *Seismological Research Letters*, 91, 1–15.

Zhang, H., and C. Thurber, 2006, Development and applications of double-difference seismic tomography: *Pure and Applied Geophysics*, 163, 373–403.

Zhu, W., E. Biondi, and J. e. a. Li, 2023, Seismic arrival-time picking on distributed acoustic sensing data using semi-supervised learning: *Nature Communications*, 14, 8192.