Time-lapse changes in ambient noise interferometry and dispersion analysis at the Stanford DAS Array

Eileen Martin and Biondo Biondi

ABSTRACT

Ambient noise interferometry allows us to extract signals that mimick active source surveys without the cost and permitting requirements of a true active survey for near-surface imaging. In many environments, seismic velocities in the near surface may change seasonally, reacting to temperature and saturation, and even subsidence. We analyze time-lapse changes in virtual source response estimates extracted from ambient seismic noise recorded at the Stanford Distributed Acoustic Sensing Array (SDASA-1) between September 2016 and August 2017. Our analysis indicates that only one week of noise is enough to yield stable virtual source response estimates when compared with the estimate from the same full month of noise. The virtual source response estimates we extract throughout one year appear to show an improvement in signal-to-noise-ratio during months when the ground is more saturated. The Rayleigh wave dispersion images show velocities in the same range as active source geotechnical surveys (Thomas et al., 2013). Further, their Rayleigh wave dispersion images suggest changes in near surface velocity tied to those saturation changes. But these apparent velocity changes are also accompanied by power spectrum changes, so further investigation into the ambient noise field is needed before these velocity shifts can be interpereted with certainty.

INTRODUCTION

For the purpose of near-surface characterization, we are interested in processing ambient noise to avoid the cost, time, and permitting requirements involved in active surveys. Ambient noise interferometry has successfully been used with point sensors to create data mimicking active surveys at the scale of a city (Chang et al., 2016) and time-lapse surveys at the reservoir scale (de Ridder, 2014). In urban areas, the long-term use of traditional sensors, nodes or otherwise is logistically and economically prohibitive for many applications: permission and permits for installation must be obtained separately for each location, wired sensors can only span a short distance, node batteries must be replaced regularly (often monthly), few cost-effective sensors have wireless data communication capabilities, and failed or stolen sensors must be replaced.

Distributed acoustic sensing (DAS) installed in existing telecommunications conduits, similar to the Stanford Distributed Acoustic Sensing Array (SDASA-1), may provide a cost-effective alternative on the receiver side: the cost per sensor is below \$1/meter, higher density can be achieved through changes in the interrogator unit, a single power source and communication point is required for the entire system of potentially thousands of sensors, and sensors can be left in place indefinitely in a conduit as secure as existing utilities with little to no maintenance cost. However, DAS systems measure strain rates, a tensor quantity, distributed over subsets of fiber. This leads to some potentially challenging theoretical issues that we must account for in processing (Martin et al., 2017).

Passive Rayleigh wave interferometry has been applied recently to trenched-fiber DAS data for imaging at geotechnical scales (Zeng et al., 2017; Martin et al., 2016), and it appears to yield the proper velocities, but this yields no information between lines in the array. In hopes of increasing our ray coverage throughout the space between fiber lines, we have extended the calculation of virtual source response estimates to pairs of fiber channels throughout 2D arrays, and have specifically applied this theory at SDASA-1 (Martin and Biondi, 2017). In this report, we test whether the virtual source response estimates throughout the 2D array show significant changes throughout the first year of recording.

First, we establish that stacking roughly one week of cross-correlations is enough to ensure that virtual source response estimates throughout the array are stable relative to the signal extracted from one month of noise. Next, we look at one year of minimally pre-processed cross-correlations responding to two virtual source locations (one east-west oriented and one north-south oriented) on SDASA-1. Virtual source response estimates from epochs when the ground is more saturated show a clear increase in signal-to-noise ratio at longer distances. We calculate dispersion images for the Rayleigh wave portion of the virtual source response estimates (channels colinear to the virtual source channel), showing how much energy travels at each velocity for each frequency, and clarifying the apparent velocity changes throughout this one year period. Still, the changes in apparent velocities are accompanied by changes in the power spectrum of the background noise field, so further analysis of the ambient noise field is needed before the picked velocities can be interpreted.

CONVERGENCE ANALYSIS

To test how long we need to average cross-correlations to get stable virtual source response estimates, we study the similarity of short-term averages from varying window lengths to month-long averages throughout September 2016. Note that convergence of virtual source response estimates does not necessarily mean a more accurate signal, just a more repeatable one. As in (Seats et al., 2012), for any virtual source v receiver r pair, we use the normalized zero time-lag correlation between the long term average

virtual source response estimate, $l(v, r, \tau)$, where τ is time lag, and any shorter virtual source response estimate, $s(v, r, \tau; t, t + w)$ averaging cross-correlations for windows that start and end between time t and $t + w$. Thus, if the virtual source response estimates are compactly supported on $(-\tau, \tau)$, their correlation coefficient is:

$$R_C(v, r; t, t + w) = \frac{\int_{-\tau}^{\tau} l(v, r, \tau') \cdot s(v, r, \tau'; t, t + w) d\tau'}{\left(\int_{-\tau}^{\tau} l^2(v, r, \tau') d\tau'\right)^{1/2} \left(\int_{-\tau}^{\tau} s^2(v, r, \tau; t, t + w) d\tau'\right)^{1/2}} \quad (1)$$

Note that in practice R_C is an approximation with sums over discretized signal vectors instead of integrals of continuous functions. If $R_C(v, r; t, t + w) \approx 1$ for nearly all window start times, t, this suggests little is gained by continuing to average (v, r) cross-correlations averaged over a longer window, w.

As seen in Figures 1a to 1g, we calculated R_C to see how it varied throughout our 30 days of cross-correlations. We calculated s for all receivers responding to channel 75 as a virtual source for each contiguous subset of w hours throughout the 30 days. This was tested for $w = 6, 12, 24, 48, 96, 192$, and 384 hour sliding windows. Note that windows with $w > 1$ overlap, so when $w = 6$, there is a window from midnight to 6 am on the first day, a window from 1 am to 7 am on the first day, and so on. The horizontal stripes in R_C plots for 6 and 12 hour windows repeatedly show lower R_C values during the daytime, even in parts of the array farther from roads. These daily variations are barely picked up by the auto-correlation of channel 75, indicating that processing decisions based on convergence require a measure of convergence throughout the array, not just auto-correlations. Overall improvement could be quantified by an R_C matrix with higher entropy (more even convergence throughout the array and over time), higher average values, or higher minimum values. One possibility is not using any daytime noise, but this increases the recording time required. By the time one week of data has been integrated, we have stable cross-correlations.

CALCULATING TIME-LAPSE CHANGES AND RESULTS

Previously, we showed that coherence virtual source response estimates could be extracted throughout the array out of one week of data (Martin and Biondi, 2017). Using similar processing, we extended those analyses to calculate a monthly correlation using six days of data from each month from September 2016 to April 2017, including all 24 hours of the day. Even when averaging over multiple days of data, ambient noise pre-processing decisions can cause significant biases in the estimated Green's functions retrieved by (Fichtner, 2014). Thus, we did minimal pre-processing: data were divided into 5 minute windows with 50% overlap, were bandpassed from 0.5 -24 Hz, were thresholded to +/- 1, cross-correlated, then stacked for each hour. After saving each hour's average cross-correlations throughout the week, we normalized the cross-correlations by their L2 norms. Finally these virtual source response estimates were stacked over six days, yielding a virtual source response estimate.

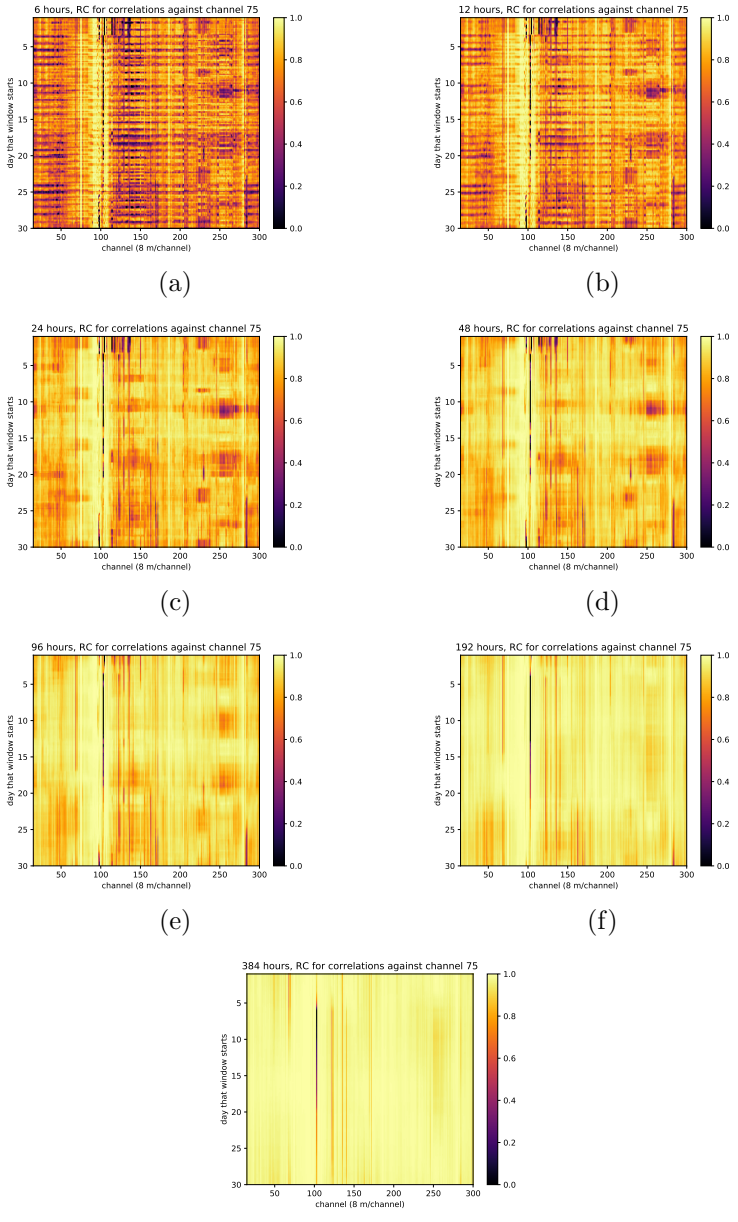

Figure 1: Calculations of R_c for cross-correlations between channel 75 and the rest of the array throughout September 2017 for (a) 6 hour windows, (b) 12 hour windows, (c) 24 hour windows, (d) 48 hour windows, (e) 96 hour windows, (f) 192 hour windows, (g) 384 hour windows show convergence on nearly all receiver channels after one week (indicated by R_c near 1). [**CR**]

Cross-correlation changes

Our convergence analysis indicates that just four days of data are be enough to yield stable virtual source response estimates when compared to estimates from the same full month of data. Still, to be on the safe side with respect to convergence, we show cross correlations for one week of data from each month in Figures 2a to 3l. Although these correlations may be relatively constant in the one month time scale, changes do emerge over a longer time scale, indicating either significant change in the noise field or velocity profile. Data collected in November through April yielded higher signal to noise ratios at longer distances than data in June, July, August, September and October. This can be seen comparing Figures 2h and 3h to Figures 2a and 3a, which are the virtual source response estimates to the same virtual source channels shown in September in Martin and Biondi (2017). We hypothesize that as the ground became more saturated with rain starting in the late fall, the coupling between the conduits and soil improved, which meant an improvement in sensitivity on both the virtual source and the receiver side. We are continuing to investigate time-lapse changes in this challenging environment with many transient noise sources. Even for static virtual source response estimates, the issues with pre-processing related biases are especially pertinent in environments with transient noise sources, so we are working towards scalable methods for automatic identification and filtering of these sources in urban environments (Huot et al., 2017).

Dispersion image changes

We calculate Rayleigh wave dispersion images from virtual source response estimates limited to channels along the same line as these, as seen in Figures 4a to 5l. These dispersion images were calculated via tau-p transforms followed by a Fourier transform along the tau axis. These dispersion images tell us how much energy is traveling at each velocity for a given frequency, summarizing the Rayleigh wave speeds in each of the cross-correlation plots. Energy in the 10-20 Hz range tends to travel at faster velocities in the summer months in response to channel 75, and this trend sort of continues in response to channel 35 (but some months this energy is not well-focused enough to differentiate). At lower frequencies, 3-10 Hz, energy travels faster in the wetter months in response to channel 35, but there is no clear trend in response to channel 75. Below 3 Hz and above 20 Hz there is no focused peak in these dispersion images.

Due to its apparently higher signal-to-noise ratio, we focus on results from April, seen for the whole month in Figure 6a to 6d. To better distinguish signal from noise, we do not normalize each frequency so that the value at the peak velocity in the dispersion image is 1. These show velocities and frequencies in the correct range for geotechnical surveys, suggesting this may be a reasonable tool for earthquake hazard analysis. At 5 Hz the main velocity near channel 61 is 700 m/s, and such a wave would be sensitive to features in the top 45 to 70 meters. Over at channel 141, the

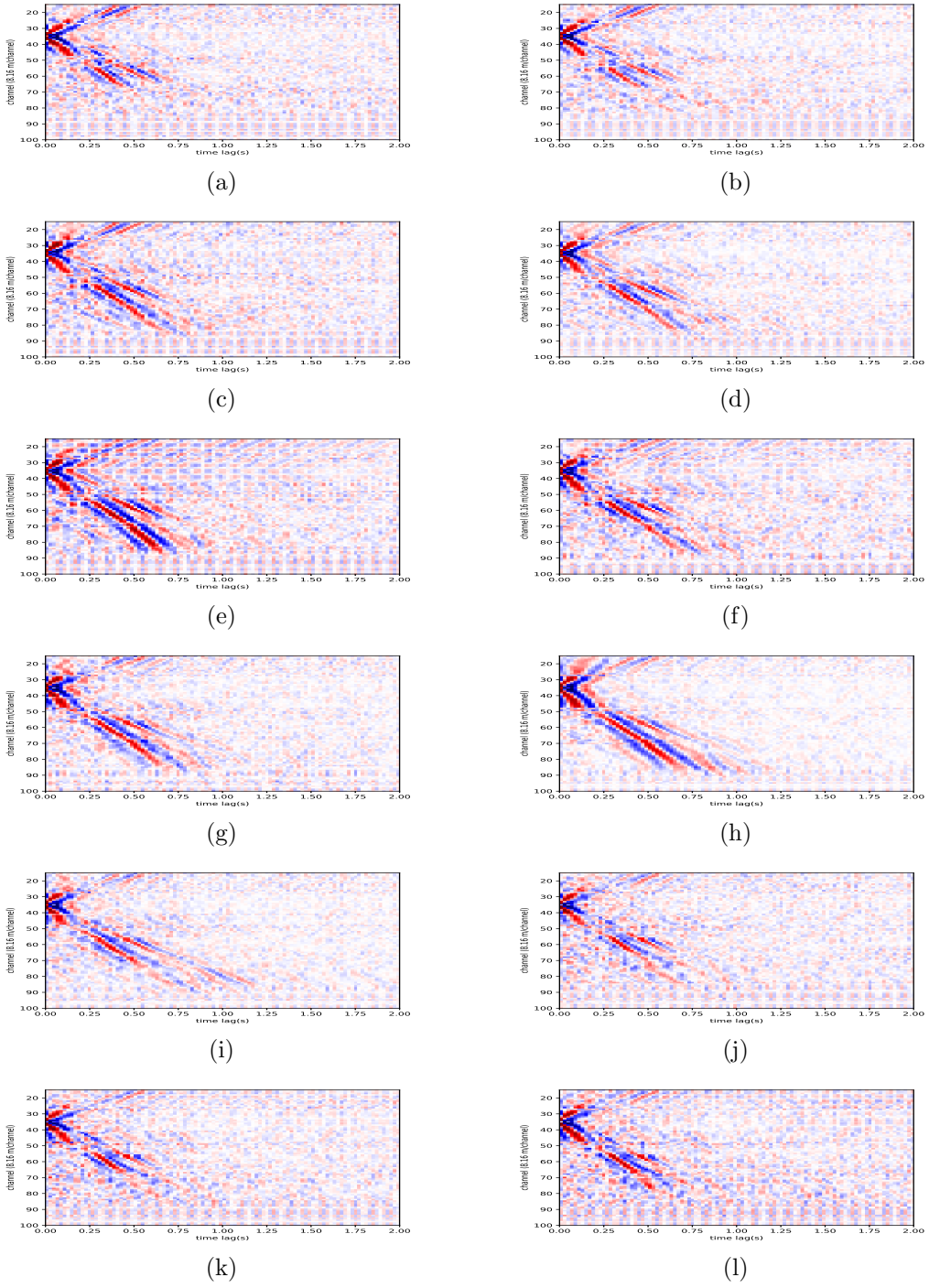

Figure 2: Cross-correlations around one corner of the array responding to a virtual source at channel 35 show an increase in signal to noise ratio at longer distances from September 2016 (top left) and October 2016 (top right) through July 2017 (bottom left) and August 2017 (bottom right). Each month's cross-correlations were calculated independently (not stacked as more months were recorded). [**CR**]

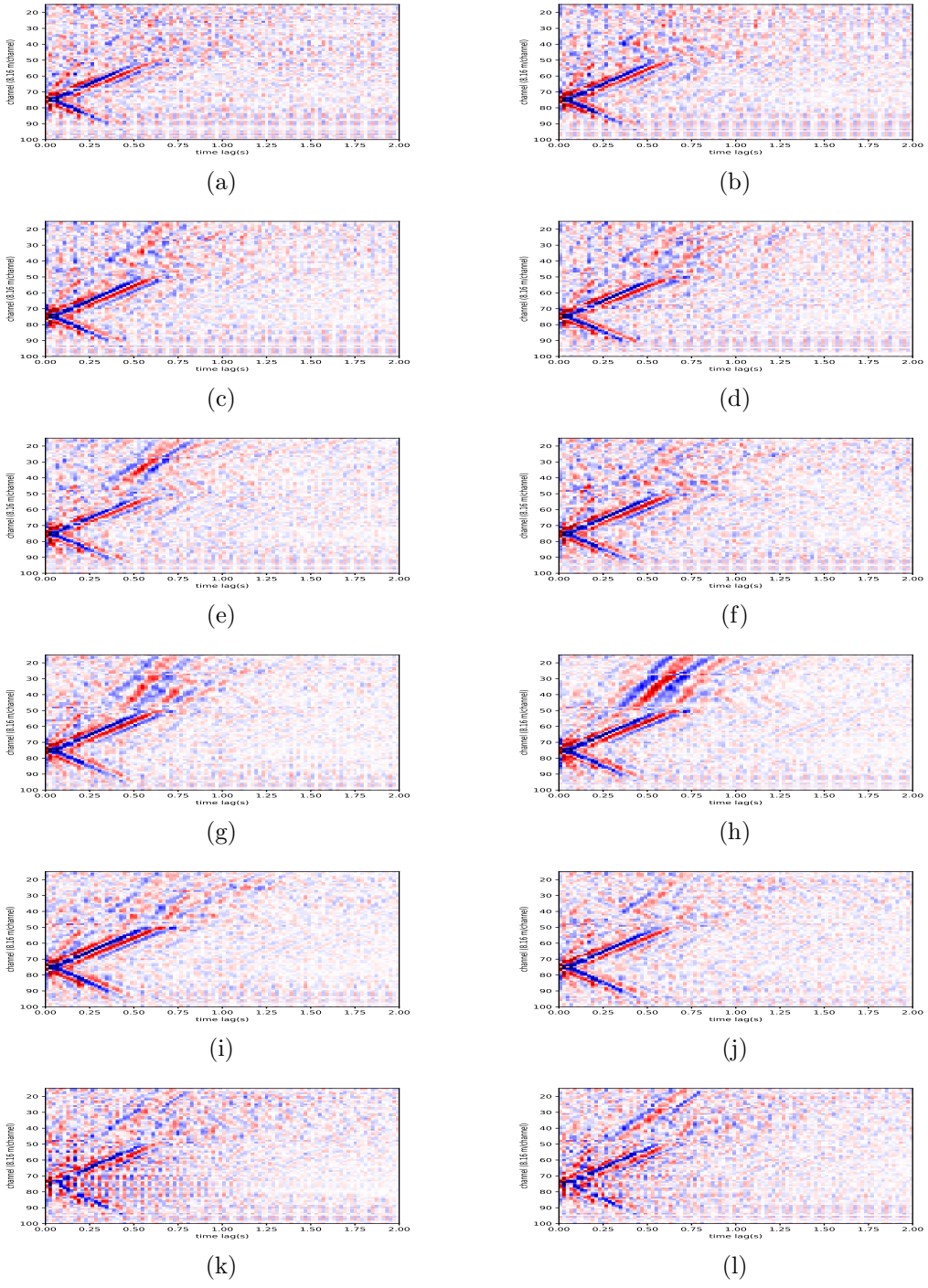

Figure 3: Cross-correlations around one corner of the array responding to a virtual source at channel 75 show an increase in signal to noise ratio at longer distances from September 2016 (top left) and October 2016 (top right) through July 2017 (bottom left) and August 2017 (bottom right). Each month's cross-correlations were calculated independently (not stacked as more months were recorded). [**CR**]

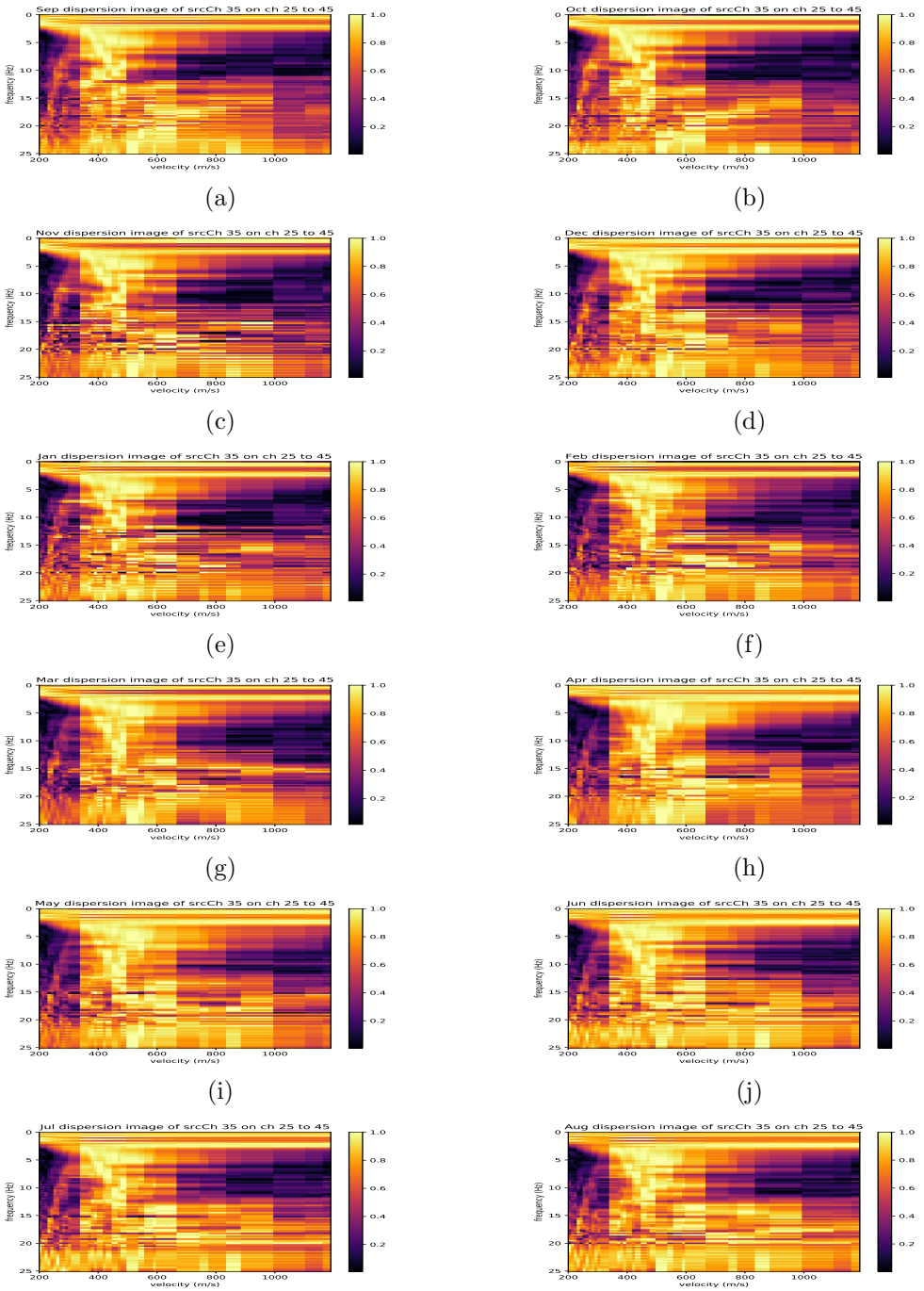

Figure 4: We calculated dispsersion images from the Rayleigh-wave parts of the cross-correlations in Figures 2a to 2l between September 2016 and August 2017. The Rayleigh waves are expected to primarily be extracted from channels 15 to 50. The peak velocity for each frequency is normalized to 1 (so energy being spread in the 20-25 Hz range actually indicates that we do not extract much coherent energy in that range). [**CR**]

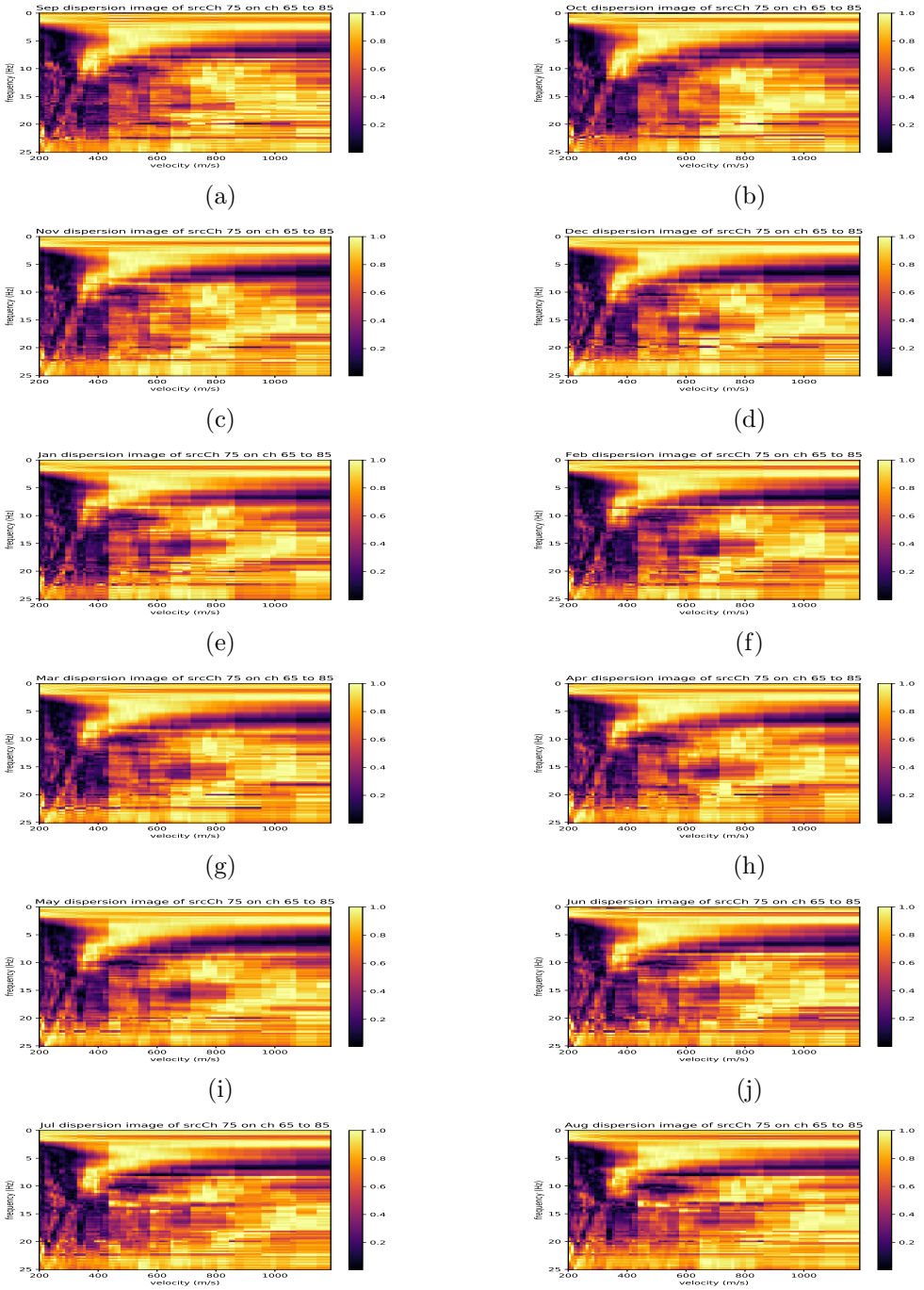

Figure 5: We calculated dispsersion images from the Rayleigh-wave parts of the cross-correlations in Figures 3a to 3l. The Rayleigh waves are expected to primarily be extracted from channels 50 to 100. The peak velocity for each frequency is normalized to 1 (so energy being spread in the 20-25 Hz range actually indicates that we do not extract much coherent energy in that range). [**CR**]

peak velocity for 5 Hz waves is closer to 400 m/s, which would be sensitive to features at a scale of 25 to 40 meters. At both locations, the peak at 5 Hz is continuous over a range of frequencies and is likely the fundamental mode. At 10 Hz, the main velocity near channel 61 is around 400 m/s, which would be sensitive to features at a scale of roughly 13 to 20 meters. Farther north, at channel 141, the peak velocity for 10 Hz is closer to 300 m/s, which would be sensitive to features in the top 10 to 15 meters. For earthquake hazard analysis, engineers must estimate surface wave and S-wave velocity in the top 30 meters of the subsurface. Thus, the frequencies at which we extract signals are in the right range for geotechnical studies, and the ability to cut the cost of these surveys could enable more widespread near surface stability studies.

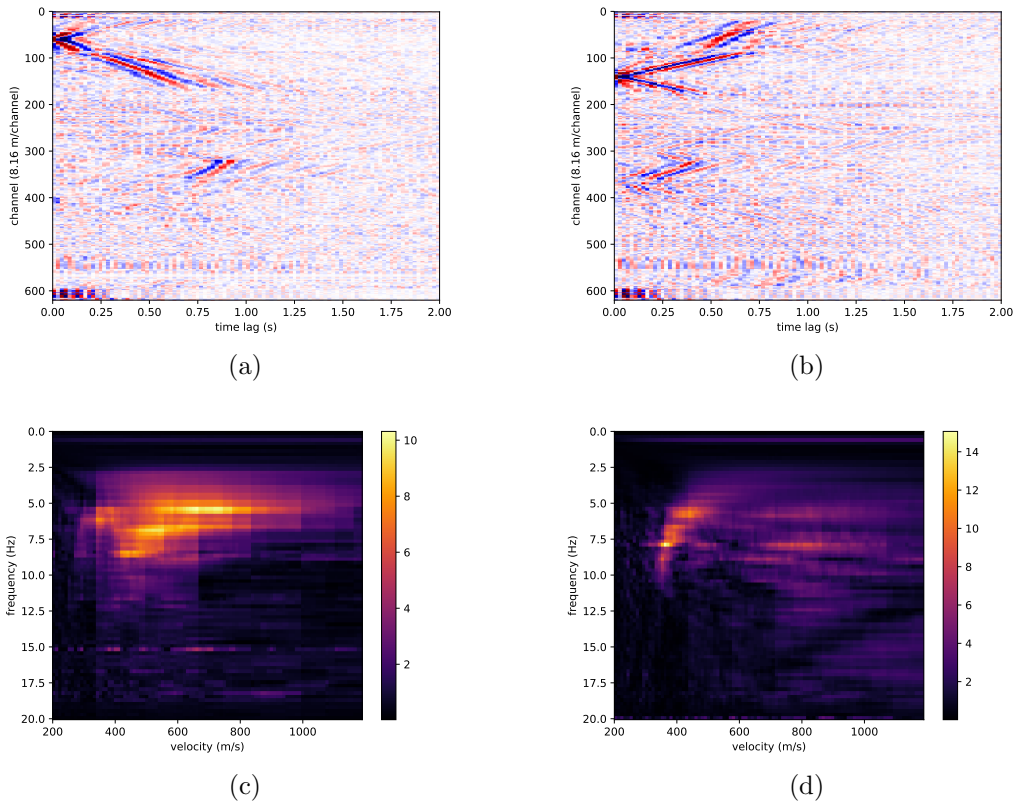

Figure 6: Throughout April we calculated virtual source response estimates throughout both loops of the entire array, and the dispersion images just for fiber in-line with the virtual source channel's orientation. In addition to the in-line channels yielding Rayleigh waves, we can also see channels parallel to the virtual sources yielding something made primarily of Love waves. Responding to virtual source channel 25 (left, marked as 69 when both loops are there), we see a reaction between 0.6 and 1.0 seconds on parallel channels 300 to 370. Responding to a virtual source at channel 75 (right, marked as 141 when both loops are there) we see a reaction between channels 320 and 400 between 0.2 and 0.5 seconds which is likely to be primarily converted wave energy due to the orthogonal orientation between the virtual source and receivers. [**CR**]

DISCUSSION

The findings here indicate the likely potential of DAS interferometry to detect near-surface changes, but further work must be done to ensure that spatially and temporally heterogeneous noise sources are not the cause of these changes. In this report we used very minimal preprocessing, but new tools developed within SEP need to be incorporated into ambient noise analysis to remove nearby vehicles (Huot et al., 2017) and earthquake recordings (Yuan et al., 2017). Furthermore, it has been documented that even power spectrum changes can cause false apparent velocity changes, so we need to test the robustness of our velocity changes to spectral whitening and cross-coherence (Daskalakis et al., 2016). Further, we intend to build on these results by performing surface wave inversion.

While the methods used have their limitations, these findings establish that changes in the virtual source response estimates in an urban area throughtout one year can be significant, whether by changes in the background noise field or by changes in the velocity profile due to saturation or other effects. We believe this is the first year-long study of ambient noise interferometry from a dense array in an urban area, and it was made possible primarily by our novel acquisition method: DAS in existing telecom conduits. This method of deploying many dense sensors is much easier to maintain in an urban area over long periods of time than geode or node systems. As we further probe the cause of these changes in interferometry, we will no doubt learn more about trends in anthropogenic noise on campus, and potentially even our urban hydrology system.

ACKNOWLEDGEMENTS

We thank sponsors of the Stanford Exploration Project for funding this work. E. Martin has also been supported in part by the DOE CSGF under grant number DE-FG02-97ER25308, and a Schlumberger Innovation Fellowship. We would like to thank OptaSense, particularly M. Karrenbach and S. Cole for the use of their ODH-3 interrogator unit and for many helpful suggestions. We would also like to thank Jonathan Ajo-Franklin (Lawrence Berkeley National Lab), Nate Lindsey (University of California Berkeley), George Papanicolaou, Jason Chang and Bob Clapp for their insights and assistance.

REFERENCES

Chang, J., S. de Ridder, and B. Biondi, 2016, High-frequency rayleigh-wave tomography using traffic noise from long beach, california: Geophysics, **81**, 843–854.

Daskalakis, E., C. Evangelidis, J. Garnier, N. Melis, G. Papanicolaou, and C. Tsoga, 2016: Geophysics Journal International, **205**, 1926–1936.

de Ridder, S., 2014, Passive seismic surface-wave interferometry for reservoir-scale imaging: PhD thesis, Stanford University.

Fichtner, A., 2014, Source and processing effects on noise correlations: Geophysics Journal International, **197**, 1527–1531.

Huot, F., Y. Ma, R. Cieplicki, E. Martin, and B. Biondi, 2017, Automatic noise exploration in urban areas: Expanded Abstracts of the 87th Ann. Internat. Mtg.

Martin, E. and B. Biondi, 2017, Ambient noise interferometry across two-dimensional das arrays: Expanded Abstracts of the 87th Ann. Internat. Mtg.

Martin, E., G. Fabien-Oullet, R. Clapp, and B. Biondi, 2017, Sensitivity analysis of distributed acoustic sensing arrays: SEP Report, **170**.

Martin, E., N. Lindsey, S. Dou, J. Ajo-Franklin, T. Daley, B. Freifeld, M. Robertson, C. Ulrich, A. Wagner, and K. Bjella, 2016, Interferometry of a roadside das array in fairbanks, ak: Expanded Abstracts of the 86th Ann. Internat. Mtg., 2725–2629.

Seats, K., J. Lawrence, and G. Prieto, 2012, Improved ambient noise correlation functions using welch's method: Geophys. Journal Internat., **188**, 512–523.

Thomas, P., I. Wong, J. Zachariasen, R. Darragh, and W. Silva, 2013, 2013 update to the site-specific seismic hazard analyses and development of seismic design ground motions: Stanford university, california: URS Corporation and Pacific Engineering & Analysis.

Yuan, S., E. Martin, J. Chang, S. Cole, and B. Biondi, 2017, Catalog of northern california earthquakes recorded by das: SEP Report, **170**.

Zeng, X., C. Thurber, H. Wang, D. Fratta, E. Matzel, and P. Team, 2017, High-resolution shallow structure revealed with amibient noise tomography on a dense array: Proceedings, 42nd Workshop on Geothermal Reservoir Engineering, SGP–TR–212.