AGU Advances

Authors' Response to Peer Review Comments on

Rapid Response to the 2019 Ridgecrest Earthquake with Distributed Acoustic Sensing

Zefeng Li, Zhichao Shen, Yan Yang, Ethan Williams, Xin Wang, and Zhongwen Zhan

Seismological Laboratory, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA

Authors' Response to Peer Review Comments on Original Version of Manuscript (2021AV000395)

Dear *AGU Advances* Editor:

We thank you and the referees for the useful comments on our manuscript (#2018GL077870) entitled "*Rapid Response to the 2019 Ridgecrest Earthquake with Distributed Acoustic Sensing*". We are submitting a revised manuscript that addressed all the points raised by you and the referees. In this revision, we update the noise data set with more difficult cases for EEW (the noise signal closely resemble the earthquakes). We obtained essentially the same results but the numbers changed slightly. Below are the detailed responses to the comments. We marked the original comments as **bold**, our reply as regular fonts, and new text inserted in the paper as *italic*.

Reviewer #1 Evaluations:

This manuscript describes the data recorded with distributed acoustic sensing (DAS) for the 2019 Ridgecrest aftershock sequence. Using DAS technology to detect earthquakes is an exciting advancement, and it's even more impressive to see the technology deployed rapidly to capture an aftershock sequence. I think this justifies

eventual publication in AGU Advances, and the current manuscript and figures are mostly clear. However, in my opinion, the manuscript could be strengthened with modifications to more clearly emphasize the unique strengths of DAS data. For example, the authors could more strongly emphasize the potential benefits of dense spatial sampling. This could be very valuable for structural studies in particular? Other applications?

Thanks for the suggestions. We revised the Introduction and Discussion sections to emphasize more the unique strengths of DAS. We enrich the discussion of the benefits of dense spatial sampling in terms of structural studies and fault detection, which are also important tasks in rapid response monitoring. We add the following in the Introduction and Discussion:

Recent applications have demonstrated that DAS can record high-fidelity wavefields of local and regional earthquakes (Lindsey et al., 2017; Wang et al., 2018). The ultra-dense recordings provide unprecedented resolution for earthquake monitoring (e.g. Li & Zhan, 2018) and subsurface imaging (Lindsey et al., 2019). However, rapid response after a major earthquake demands not only high resolution but also fast installation; the potential of DAS as a rapid response system has not been fully explored before.

DAS has several outstanding advantages as a rapid response system (Figure 5). First, DAS provides large-N monitoring capability (e.g. Catchings et al., 2020) with meter-scale spacing and unaliased recording of the higher frequency wavefields. The dense spatial sampling offers high-resolution earthquake detectability, as demonstrated in this study. Also, recent results have shown that ambient noise correlation on DAS channels can be used to obtain high-resolution subsurface structures (Cheng et al., 2021; Dou et al., 2017). Yan et al. (2021) applied ambient noise tomography to the Ridgecrest array and observed strong lateral variations of the velocity structures that correlate well with shaking intensity. The dense spatial sampling is also valuable for mapping previously unknown faults (Jousset et al., 2018; Lindsey et al., 2019).

As a related point, it's worth pointing out that, in the case of Ridgecrest, the event detections achieved here could very likely have been achieved using just the permanent seismic network. Typically, many more new events can be detected than located, so the numbers of events in the Ross et al. (2019) and Shelly (2020) catalog would have been significantly higher for detection only. However, I think the authors could make a very convincing argument that DAS data in some cases could allow a similar catalog to be constructed even in areas lacking the high-quality permanent seismic network that's present around the Ridgecrest sequence.

We thank the reviewer for the pointing this out. We agree with the reviewer that there would be more new detections without locations in the Ross et al. (2019) and Shelly (2020) catalog, and that the DAS could allow high-quality catalogs in areas lacking high-quality permanent seismic networks. In order to estimate the performance of DAS in the less instrumented areas, we perform template matching tests on DAS with only large-magnitude events (to mimic the detectability of a sparse network). We find that, with 147 templates with M>3, the Ridgecrest DAS array produces 49,936 events; with 482 templates with M>2.5, it produces xxx events. This demonstrates that, in the areas lacking a high-quality permanent seismic network, DAS can allow a similar catalog to be constructed and thus play a relatively more important role in aftershock monitoring.

It is noteworthy that Ridgecrest is a well instrumented areas within the SCSN (Hauksson et al., 2020), with station spacing of 10-20 km. Our catalog constructed from the 10-km DAS array is substantially better than the standard catalog and is comparable to the Ross et al. (2019) and Shelly (2020) template matching catalogs. To evaluate the performance of a similar DAS array in less instrumented areas, we test how the final catalog changes with only the bigger template events. With 147 M>3 templates, the Ridgecrest DAS array produces 49,936 detections (a 340 times increase); with 482 M>2.5 templates, it produces 61,155 detections (a 127 times increase). This demonstrates that, in areas without a high-quality permanent seismic network, DAS can be of even greater value and thus play a comparatively more important role in seismic monitoring.

Furthermore, if DAS is to be used in rapid aftershock monitoring, the processing side needs to be addressed, but I didn't see much discussion of this in the current manuscript. For example, can DAS data be integrated into real-time processing of the Southern California Seismic Network? Or is some other scheme required? Some discussion of this issue seems required, unless the authors wish to de-emphasize "rapid aftershock monitoring" in their manuscript. Of course, the data can still be valuable for studies of earthquake physics even if processing in only applied after some delay.

We thank the reviewer for this important question. In principle, DAS is integratable into real-time processing of the Southern California Seismic Network, because all the data from the channels distributed over tens of kilometers along cable are collected and processed at the terminal system. The real-time data telemetry and the feasibility of integration into permanent network is a unique advantage for DAS, as emphasized in the manuscript. In addition, DAS appears to be the only large-N technology with inexpensive real-time telemetry. The major technical challenge lies in the processing of the large data volume, which however could be approached by efficient parallel computing and streamlining of the data processing procedure. We add the following discussion:

Moreover, DAS allows real-time data telemetry because all the data from the channels distributed over tens of kilometers along cable are processed and collected at the terminal DAS system. Therefore, DAS data can be uploaded to cloud computer centers in near real-time and integrated into the existing seismic networks that are also hosted on cloud (e.g., SCSN). In practice, the major challenge lies in processing large volume of data, which calls for leveraging the power of parallel computing and efficient down-sampling algorithms. Despite this, DAS appears to be the only large-N technology with plausible real-time telemetry options, critical for closely tracking ongoing aftershock sequences.

Minor comments:

Details of the template-matching procedure seem to be missing. For example, how were the waveforms windowed for the templates? How exactly is the SNR criterion for templates applied? How long were these windows? Over what time period was the detection threshold (10 times median absolute deviation above the median noise level) established? Daily? Hourly? Were duplicates removed (i.e. different templates detecting the same event in the data)?

Thanks for the comments. We add the details of the template-matching procedure in main text:

Here, we use 22,465 aftershocks from the SCSN catalog during the same period of time as template candidates. The signal-to-noise (SNR) ratio is calculated on each channel with a noise window 5.5-0.5 seconds before S arrivals and a signal window 0-5 seconds after the S arrivals. The final template library consists of 9,318 events with SNR>5 dB on at least 200 DAS channels. The waveform windows for cross-correlation are 1 second before and 5 seconds after the S arrivals. Cross-correlation of the templates with the continuous data leads to about 37 million hourly traces averaged over the array. We use a peak threshold of 10 times the median absolute deviation above the median noise level on hourly cross-correlations and remove duplicate detections from different templates. Finally 133,453 events are detected, which are 6 times more than those in the SCSN catalog (Figure 2).

The authors should probably define any acronyms used in the "highlights" bullet points (DAS, SCSN)

Fixed.

What is the reason for bad "channels" in DAS data (i.e. Figure 1c, vertical stripes)?

Figure 1c is primarily for showcasing the data quality. In template matching detection, a SNR thresholding (SNR> 5 dB, Line 115) is applied to each channel and these bad channels would be abandoned.

Units are missing in table S1, for channel spacing (m) and Sampling rate (Hz).

Added.

Figure 5: Is "quiescence" supposed to indicate noise level? I have mixed feelings about this figure overall - on the one hand, it's a cool graphical comparison that might be useful to non-specialists. On the other hand, the fact that each "axis" is subjective may limit its utility.

Yes, quiescence is for noise level. We change it to “noise control” to avoid confusion. We agree that the comparison is qualitive because it is related to specific applications. But in most application scenarios the qualitative comparison is generally useful and Figure 5 provides a straightforward coordinate for the strengths and weakness of different instruments. We mark three performance levels *(poor, good, and excellent*) in the figure and caption to show comparison is intended to be qualitative and approximate.

Lines 126-127: For magnitude computation, does this study also use the empirical calibration between amplitude ratio and local magnitude (ML) difference (the factor of 0.831) as Shelly (2020), or does it use something else? If this magnitude approximates ML, it might be helpful to clarify magnitude type on relevant figures.

Yes, we use the empirical factor of 0.831 between amplitude ratio and local magnitude (ML) difference. We mark the magnitude type ML in the text and Figures 2 and 3.

The calibration procedure follows Shelly (2020): we first calculate the amplitude ratio of body waves between the templates and the newly detected events, and then convert it to local magnitude (ML) difference empirically using $ML_2 = ML_1 + 0.831\, log_{10} \frac{A_2}{A_1}$. *The results show that the magnitudes of most newly detected events are ML -1~ML 1 (Figure 2).*

Lines 146-147: "likely suggesting a higher rate of aftershock production than the

average rate along the main fault". Can you make any quantitative comparison to support this? Some measure of seismicity per fault area, perhaps?

Thanks for the comment. As shown in Figure 4c, The peaks corresponding the cross faults shown 2500-5000 events in a 1 km bin, which is much higher than the median along the main fault (~1300). The accurate number of events on individual faults is nonetheless unavailable, because the complex intertwined fault system make it difficult to partition the seismicity into their causative faults.

Reviewer #3 Evaluations:

This paper describes the first rapid response to an earthquake using a das array. Overall, the paper is very concise, well written, and reads easily. The paper's content is of interest, although the merit of the method has already been published in Li & Zhan 2018. Indeed, Li & Zhan 2018 showed that template matching with DAS could dramatically increase the number of earthquake detections compared to a traditional seismic network. Therefore, the main novelty of this paper is that Li & Zhan 2018 pipeline is applied to one major California earthquake dataset (Ridgecrest) and those new waveforms "reveal abundant aftershocks on multiple crosscutting faults near the epicenters of the mainshock." I am not a Ridgecrest earthquake specialist, but I think such a finding was highlighted in previous studies, although with a lower density of aftershocks.

We agree with the reviewer on that this study and Li & Zhan (2018) employ a similar earthquake detection method (template matching). Both studies demonstrate that the DAS's ultra-dense spatial sampling can significantly improve earthquake detectability. However, there are several novel contributions of this work that make it fundamentally different from previous studies:

1. This paper focuses on the demonstration of DAS's potential for rapid response to large earthquakes. The Ridgecrest DAS is particularly special in that takes advantage of existing telecommunication networks to build a response system to a major earthquake.

This study is intended to use the Ridgecrest as an example to shed light on the DAS's prospect for rapid response in other places around the US and the world. With this motivation, we show the unique strengths of DAS as rapid response systems in terms of ultra-dense sampling, deployment promptness and real-time telemetry, and its great generalization capability in other places. In comparison, Li & Zhan (2018) focuses on a local array (1.5 km in aperture) and investigate the changes of induced seismicity due to geothermal operations.

2. The Ridgecrest DAS data set is unique compared to the PoroTomo and other DAS data sets in that it recorded extremely intense aftershock sequence that offers broad magnitude, distance and azimuth ranges. We take advantage of this strength and go beyond detection to investigate the impact factors of DAS detectability spatially (distance, azimuth, variations with channels) and temporally (periodic cultural noise). We are also able to compare a catalog constructed from a 10-km DAS array to one of the highest quality networks in the world and estimate the performance in other areas with less instrumental coverages. These findings have not been reported in previous DAS studies.

We revise throughout the manuscript accordingly to further highlight these novelties.

The authors claim they "demonstrate the monitoring capability of DAS," but I think this statement, which is repeated several times in the text, is misleading. Since about 2016, many studies have proven das' capability to provide high-fidelity earthquake waveforms but under certain conditions. For example, the coupling of the cable with the ground matters a lot. The angle between the wavefield and the cable matters too, and in fact, it can dramatically deteriorate the response of the DAS array. DAS is generally sensitive to higher frequencies than traditional sensors. Therefore, the distance between the array and the earthquakes is important. I feel the authors had an excellent opportunity to discuss all these points in detail, but unfortunately, they did not. I am not saying that the paper is not of interest but that they did not adequately demonstrate the monitoring capabilities of DAS.

We have quantitatively evaluated several impact factors, i.e. source-receiver distance, incident angles, and periodic cultural noise, on the earthquake detectability on the Ridgecrest array (see the sections Impact Factors and Discussion). The major findings are below:

1. The Ridgecrest DAS array (10 km long) has optimal detection capability (DAS/SCSN event ratio>5) within about 20 km from the array center. Outside the 20 km radius, the performance starts to drop significantly.
2. The events from north (90°incident angle) and from southeast (-10-40°incident angle) differ in SNR of 5-10 dB, suggesting strong influence of incidence angle.
3. The diurnal variations of traffic noise significantly modulated the DAS earthquake detection capability, resulting in a detectability change at a factor of 3.

In the revision, to investigate the variations of channel-dependent detectability, we calculate the SNRs of the template events on individual channels. We find that the overall variations of SNRs show a significant channel-dependent pattern. The number of detectable events (defined as SNR > 5 dB) varies between 4000 to 6000, a change of ~20% relative to the mean, which could be associated with the ground coupling effect. Please refer to the reply to Comment 3 for details.

Some ideas that could help to understand the DAS monitoring capabilities better
- Why are the other temporal das arrays not used in this study? What are the TM results for one or two of these other arrays? How do the earthquake detection method capabilities change with respect to the distance of the main chock epicenter? Comparing these array capabilities would highlight a minimum distance required to use DAS for aftershock studies.

Thanks for the suggestion. We choose to perform earthquake detection on the Ridgecrest array because it is closest to the aftershock zone of the M 7.1 earthquake, which is of most interest. Our results on the Ridgecrest array show that a 20-km radius is the optimal detection range for a 10-km fiber. We believe this would be a useful reference for further aftershock studies with DAS. Following this reference, the other arrays are more than 40 km away and the aftershock zone is out of their optimal detection range. For example, on the Ridgecrest array, there are 9,318 out of 22,465 events as qualified templates; the qualified templates on the Olancha and Goldstone arrays would be much less, making template matching ineffective. In addition, performing template matching on the other three arrays unlikely impacts the main conclusions but requires several folds of increase in workload. Although we agree with the reviewer that performing template matching on other arrays could provide additional information on the effect of distance, we would like to save the template matching of other arrays as future work.

- Which section of the cable detected most of the earthquakes? Is it the section closest to the epicenter, in the middle, or is it random? Does it depend on the coupling of the cable with the ground, or is it a function of the distance?

Thanks for the questions. To investigate the variations of channel-dependent detectability, we calculate the SNRs of the template events on individual channels. In Figure S4, the SNRs of 9300 cataloged events on each channel show the variations of detectability within the array. Despite the loops (low detected event counts and low SNRs), the overall variations of SNRs show a significant channel-dependent pattern. From this plot, we observe no significantly higher SNRs on the eastern section of the fiber which is generally closer to the earthquakes. This could be in part because the fiber is relatively short and the 3-sec window to calculate SNR consists of surface and scatter waves that decays less rapidly.

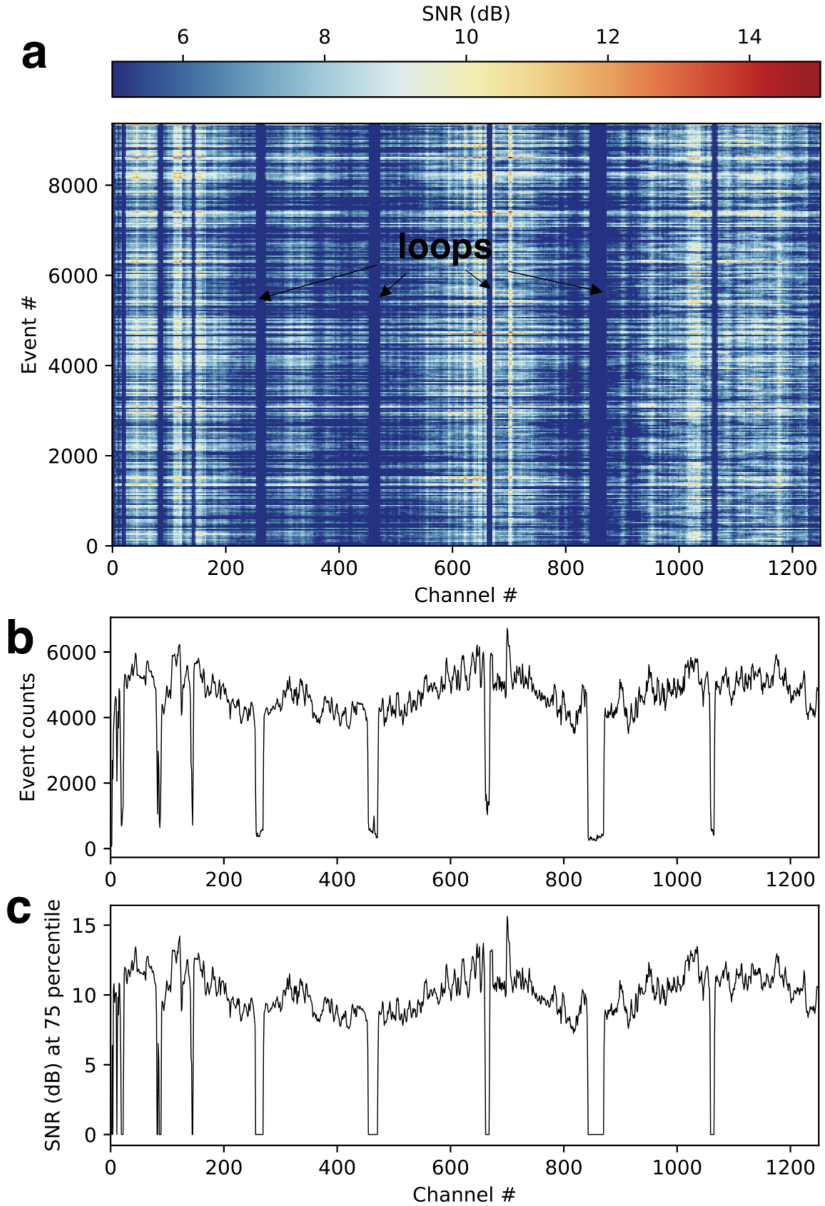

Figure S4. Detectability variations within the Ridgecrest DAS. a SNRs of 9370 events across the array. **b** The number of recordings with SNR>5 dB (defined as detectable) on individual channels. **c** SNRs at 75 percentile of all the events on individual channels.

We add the follow discussion in the *Impact Factors* section.

To evaluate the variations of detectability within the array, we calculate the SNRs of all template events on each channel (Figure S4). In addition to fiber service loops with low ground coupling, the overall variations show a significant channel dependence. The number of detectable events (defined as SNR > 5 dB) varies between 4000 to 6000, a *change of ~20% relative to the mean. A similar range of variation is also shown in the SNR at 75 percentile of all events. Furthermore, we observe no significantly higher SNRs on the eastern section of the fiber which is generally closer to the earthquakes, likely because the fiber is relatively short and the SNR windows consist of less rapidly decaying scatter and surface waves.*

- Why have fewer earthquakes been detected in the northern region of the fault (Fig 3)? Is it because the cable is orthogonal to the earthquake wavefields coming from the north while parallel to earthquakes from the east? This might be a sensitivity issue, and it looks that when the angle between the cable and the earthquake is >45*, the detection capability of the das sharply drops. Does the Olancha south array show similar findings while being parallel to the earthquake wavefield?

Yes, this is an effect of directional sensitivity. We compare the SNRs of the aftershocks from the north (back azimuth from -10° to 20°) and from the southeast (back azimuth from 80° to 130°) in similar ranges of magnitudes (M 2 - 3.5) and epicentral distance (20-35 km). Figure 3b shows that the SNRs of P and S waves from the southeast are systematically higher than those from the north. As the reviewer pointed out, this is consistent with theoretical prediction that both P and S waves perpendicular to DAS arrays have minimal amplitudes. Such comparison cannot be done in the Olancha south array, because most of the aftershocks are from the along-strike direction and data from the array-perpendicular direction are lacking.

- Kind of unrelated comment: How does your TM work? Do you cross-correlate all the templates with the 1200 channels of the das? If yes, why is it important to consider 1200 channels and not.. 120? the 120 with the highest SNR, for example?

We only select the channels with SNR>5 dB and one event must have at least 200 channels with SNR> 5 dB to be used as a template. Overall, the total channels used in cross-correlation vary from 200 to 1250 for different templates. As shown by Li and Zhan (2018), using the channels at this number can achieve similar performance to that of all the channels, because the spatial sampling is so dense that the waveforms are similar on adjacent channels. We add the template matching details in the manuscript:

Here, we use 22,465 aftershocks from the SCSN catalog during the same period of time as template candidates. The signal-to-noise (SNR) ratio is calculated on each channel with a noise window 5.5-0.5 seconds before S arrivals and a signal window 0-5 seconds after the S arrivals. The final template library consists of 9,318 events with SNR>5 dB on at least 200 DAS channels. The waveform windows for cross-correlation are 1 second before and 5 seconds after the S arrivals. Cross-correlation of the templates with the continuous data leads to about 37 million hourly traces averaged over the array. We use a peak threshold of 10 times the median absolute deviation above the median noise level on hourly cross-correlations and remove duplicate detections from different templates. Finally 133,453 events are detected, which are 6 times more than those in the SCSN catalog (Figure 2).

Minor comments attached in the pdf:

How did you get access to these fibers? Sharing a bit of your experience would be much appreciated by future das users.

The fibers are owned by Digital 395 and JPL, as acknowledged in the end of the paper. We gained access to the fiber at Day 3~5 after the M 7.1 event. The process was actually complicated and also different from the two owners. For JPL, as a research institute well connected to Caltech, the main challenge was quick physical access through the military base. For Digital 395, which serves the communities near the earthquakes, it was through some lucky connections. While the stories are interesting, we do not think it should be a part of the current paper, because the experience might not be transferable because the fibers in the States are owned by a broad range of agencie, such as governments, telecommunication companies, and universities. The approach to gain access from different parties may vary a lot. However, we agree that gaining access to the fibers is an important step in rapid response deployment. We expect that with more rapid response cases like Ridgecrest in the future, negotiation with different parties for fiber access might become easier. For example, after the Long Beach array, it seems more common to deploy nodal arrays in urban area to study subsurface structures and potential seismic hazards.

Line 220: Not completely true.. I think some Swedish companies developped a 100km das. There is also the option to have das-specific repeaters. APsensing has a 70-km das. To be less specific I would say something like "beyond several tens of km"

Now it is fixed.