Geophysical Journal International

Geophys. J. Int. (2024) **237,** 1609–1623
Advance Access publication 2024 March 27
GJI Applied and Marine Geophysics

https://doi.org/10.1093/gji/ggae122

DAS for 2-D MASW imaging: a case study on the benefits of flexible subarray processing

Michael B. S. Yust [1], Brady R. Cox,[2] Joseph P. Vantassel [3] and Peter G. Hubbard[4]

[1] *Department of Civil Architectural and Environmental Engineering, The University of Texas at Austin,* 301 *E Dean Keeton Street, Austin, TX* 78712*, USA. E-mail: yustm@utexas.edu*
[2] *Department of Civil and Environmental Engineering, Utah State University,* 4110 *Old Main Hill, Logan, UT* 84322*, USA.*
[3] *Charles Edward Via, Jr, Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University,* 750 *Drillfield Drive, Blacksburg, VA* 24061*, USA.*
[4] *FiberSense,* 66 *Franklin Street, Suite* 300*, Oakland, CA* 94607*, USA.*

Accepted 2024 March 23. Received 2024 March 18; in original form 2023 June 20

SUMMARY

Distributed acoustic sensing (DAS) is a relatively new technology for recording the propagation of seismic waves, with promising applications in both engineering and geophysics. DAS's ability to simultaneously collect high spatial resolution waveforms over long arrays suggests that it is well-suited for near-surface imaging applications such as 2-D multichannel analysis of surface waves (MASWs), which require, at a minimum, long, linear arrays of single-component receivers. The 2-D MASW method uses a large number of sensor subarrays deployed along a linear alignment to produce 1-D shear-wave velocity (V_S) profiles beneath each subarray. The 1-D V_S profiles are then combined to form a pseudo-2-D V_S image beneath the entire linear alignment that can be used for the purpose of identifying and characterizing lateral variations in subsurface layering. Traditionally, 2-D MASW is conducted using arrays consisting of either 24 or 48 geophones. While additional receivers could easily be incorporated into the testing configuration, it is rare for researchers and practitioners to have access to greater numbers of seismographs and geophones. When a limited number of geophones are available for deployment, there is a need to pre-determine the geophone spacing and subarray length prior to field data acquisition. Studies examining how the choice of subarray geometry impacts the resulting pseudo-2-D V_S cross-sections have been largely limited to synthetic data. In response, this study utilizes DAS data to examine the effects of using various subarray lengths by comparing pseudo-2-D V_S cross-sections derived from active-source waveforms collected at a well-characterized field site. DAS is particularly useful for 2-D MASW applications because the subarray geometry does not need to be determined prior to field data acquisition. We organize the DAS waveforms into multiple sets of overlapping MASW subarrays of differing lengths, ranging from 11 to 47 m, along the same alignment, allowing for direct comparison of the derived pseudo-2-D V_S results at the site. We show that the length of the individual MASW subarrays has a significant effect on the resulting V_S cross-sections, including the resolved location of a strong impedance contrasts at our study site, and evaluate the results relative to ground truth from invasive testing. Our results suggest that the choice of subarray length is important and should be carefully chosen to meet project-specific goals. Furthermore, analysts may consider using multiple subarray geometries during the data processing stage, as is made possible by DAS, to properly evaluate the uncertainty of 2-D MASW results. This study demonstrates the potential of using DAS to collect data for 2-D MASW in a manner that is efficient and flexible, and can be easily scaled up for use with very long arrays.

Key words: Tomography; Downhole methods; Surface waves and free oscillations; Waveform inversion.

1 INTRODUCTION

Surface wave methods are powerful tools for non-invasive seismic site characterization. One of the most popular testing methods is the multichannel analysis of surface waves (MASW), which is capable of producing a 1-D shear-wave velocity (V_S) profile of the subsurface (Park *et al.* 1999; Foti 2000). MASW is traditionally performed using waveforms recorded by a linear array of geophones that are used to sense surface waves generated by an active source (i.e. hammer, weight drop, or vibroseis shaker) located off one or both ends of the array. As the number of geophones is often fixed by the availability of 24-channel seismographs to either 24 or 48, and equipment costs are directly proportional to the number of channels used, the analyst must balance the finer vertical layer resolution provided by smaller receiver spacings and a shorter array with the greater characterization depth provided by larger receiver spacings and a longer array (Foti *et al.* 2018). Soon after the introduction of MASW, engineers began to explore how this new method could be used to characterize 2-D variations of subsurface V_S (Miller *et al.* 1999). Xia *et al.* (2000) proposed the use of MASW to construct pseudo-2-D V_S cross-sections by combining multiple 1-D V_S profiles resulting from multiple MASW subarrays along a common linear alignment. This approach came to be known as 2-D MASW and has since been used successfully on a variety of near-surface imaging projects.

Xia *et al.* (2000) were able to successfully identify multiple subsurface features using 2-D MASW, including a known steam tunnel at the University of Kansas, and a bedrock channel in Olathe, Kansas, which was later confirmed with drilling. Thitimakorn *et al.* (2005) utilized 2-D MASW to survey a 1950 m segment of Interstate 70 in St Louis, Missouri. They used 12, 4.5-Hz vertical geophones at a 3-m spacing, with all 12 geophones used in each subarray (i.e. 33-m long), to construct a pseudo-2-D V_S cross-section using a subarray interval of 12 m between 1-D V_S profiles. Based on this 1950-m-long cross-section, Thitimakorn *et al.* (2005) were able to identify depths to bedrock ranging from 6 to 13.4 m, which agreed well with 19 boreholes drilled along the testing alignment. Mohamed *et al.* (2013) performed 24 2-D MASW surveys at a site outside of Cairo, each with 13 subarrays consisting of 24 geophones with 1-m spacings (i.e. 23-m long) and a subarray interval of 4 m. The cross-sections they developed identified low-velocity regions which, when compared to borehole logs, were found to correspond to claystone layers experiencing swelling due to nearby water sources. Ismail *et al.* (2014) successfully mapped the depth to bedrock, including identification of near-surface faults, using 2-D MASW surveys of two alignments totalling 3.7 km. Their subarrays consisted of 48 geophones at 1.5-m spacings (i.e. 70.5-m long) and a subarray interval of 7.5 m. They found that the 2-D MASW approach was easier to use than producing 2-D traveltime images using colocated shear-wave reflection testing.

While not exhaustive by any means, the above-cited studies are indicative of successful applications of 2-D MASW. All of those studies, however, utilized a single, predetermined survey geometry (i.e. number and spacing of geophones, subarray length, subarray interval, shot location, etc.), which is generally the standard of practice for 2-D MASW. This is a potential cause for concern, as multiple studies of synthetic data (Park 2005; Mi *et al.* 2017; Arslan *et al.* 2021; Crocker *et al.* 2021) have found that MASW array geometry, particularly array length, has a significant impact on the vertical and horizontal resolution of the subsurface and the ability of 2-D MASW to accurately resolve layer boundaries and V_S anomalies in the subsurface. Therefore, despite its successful use in many projects, one potential improvement to 2-D MASW would be the ability to easily consider multiple subarray geometries during processing, without being constrained by choices made during data acquisition (e.g. fixing the subarray length) when using traditional geophone arrays with a limited number of receivers and seismographs.

This study aims to demonstrate how distributed acoustic sensing (DAS) can be used to collect field data for 2-D MASW, allowing for greater flexibility in data acquisition and processing than possible with most traditional geophone surveys. While DAS will not replace geophones as the most practical array option for many applications, particularly with standalone, nodal seismometers providing significant versatility for non-linear arrays (e.g. Lin *et al.* 2013), it can provide the opportunity to collect data in ways that were not previously practicable, particularly for applications requiring very long linear arrays of single-component receivers. DAS has the ability to simultaneously record seismic data along the entire length of the fibre-optic cable, without the need to pre-determine subarray lengths, allowing analysts to consider all possible combinations of sources and receivers. In this study, we demonstrate some of the benefits of DAS for 2-D MASW using a 200-m long fibre-optic cable installed at a well-characterized site. Specifically, we examine how changing the length of each MASW subarray, which is trivial for data collected using DAS, affects the resulting V_S cross-sections. While this is a rather short survey alignment for even traditional equipment, we hope the reader can appreciate that this application could be scaled-up to distances of 10 km or greater without the loss of spatial resolution by simply installing more fibre-optic cable in the ground. However, before the testing performed in this study can be fully discussed, it is important to first cover additional background information about how traditional 2-D MASW testing is performed, such that modifications to the traditional approach discussed later in the paper in regard to DAS data can be better understood.

2 TRADITIONAL 2-D MASW

Traditional 2-D MASW surveys typically use the roll-along method (Mayne 1962), with 24 or 48 geophones often mounted on a land streamer system at a fixed receiver spacing. The number of available geophones and 24-channel seismographs, when coupled with the choice of receiver spacing, pre-determines the length of the subarray used during data acquisition and processing. The land streamer is pulled behind a vehicle, allowing the geophone array to be moved along the survey alignment and incrementally stopped at a predetermined horizontal distance called the subarray interval. The subarray interval is typically set equal to some portion of the total subarray length (e.g. 1/4 or 1/3), such that there is significant spatial overlap between adjacent subarrays. The subarray interval also determines the horizontal distance between the 1-D V_S profiles that will ultimately be interpolated to obtain the pseudo-2-D V_S image.

Typically, a source such as a weight drop is attached to the towing vehicle to actively generate seismic energy. In the interest of rapid data acquisition, generally only a single shot location is used at a fixed offset distance for each MASW subsurvey. Even when data is collected using larger, mobile or stationary geophone arrays, the traces are often reorganized to mimic the roll-along method with a single shot location per subarray (Thitimakorn *et al.* 2005; Park & Miller 2005a,b). If the subsurface layering at a site tends to generate significant higher-mode surface wave energy (e.g. sites

with shallow and strong impedance contrasts), using only a single shot location may result in misinterpretation of the dispersion data (Yust 2018). While using multiple shot locations helps minimize the risk of mode misidentification, it is not easy to implement when using the roll-along method with traditional equipment.

All surface wave testing, including that involved with using individual subarrays for 2-D MASW, follows the same three main steps: acquisition of surface-wave data in the field, processing the collected records to extract experimental dispersion data, and inverting the experimental dispersion data to produce a 1-D subsurface V_S profile (Park *et al.* 1998; Park *et al.* 1999, 2007; Foti *et al.* 2014, 2018; Vantassel & Cox 2022). Due to the nature of the dispersion processing and inversion stages, which are inherently 1-D, the resulting V_S profile represents a spatial average of the conditions beneath the array and can most reasonably be assumed to represent the 1-D layering and material properties beneath the centre of the MASW array (Luo *et al*. 2009). 2-D MASW adds a fourth step: combining multiple 1-D V_S profiles to create a pseudo-2-D velocity cross-section along the testing alignment. The large amount of data required to produce these cross-sections (e.g. tens to hundreds 1-D V_S profiles obtained from sequential subarrays) is what necessitates using the roll-along method of data acquisition. When faced with a limited number of channels for data acquisition, evaluation of multiple subarray geometries is highly impractical, as it would require repeating the entire data acquisition step after changing, for example, the subarray length or subarray interval. This restriction to primarily using a single subarray geometry with a single shot location is the greatest limitation to the potential of 2-D MASW and could have unforeseen impacts on the resulting V_S cross-section.

Park (2005) examined the impact of subarray length and subarray interval using synthetically generated waveforms. Park (2005) found that the optimal array length is a balance between maximizing length, to improve dispersion data quality and maximize characterization depth, and minimizing length to reduce the amount of spatial averaging that occurs within each subarray. This spatial averaging caused smearing of the subsurface details, resulting in reduced lateral resolution for longer subarrays. Park (2005) also found that the subarray interval should be shorter than the subarray length, such that there is some overlap between successive subarrays.

Mi *et al.* (2017), Arslan *et al.* (2021), Crocker *et al.* (2021) all evaluated the ability of 2-D MASW to detect anomalous structures within the subsurface. Mi *et al.* (2017) analysed a combination of synthetic and field data sets and concluded that the velocity of an anomalous feature could not be accurately resolved if the feature was shorter than the subarray length. Arslan *et al.* (2021) and Crocker *et al.* (2021) utilized over 3000 synthetic data sets to examine the detection and resolution abilities of MASW depending on anomaly size and depth, finding that anomalies less than half the length of the subarray were unlikely to be detected. They also cautioned against blind application of 2-D MASW to detect anomalies due to the inherently 1-D nature of all surface wave methods and the effects of array geometry. Despite the important influence of subarray length on characterization depth, dispersion data quality, and lateral resolution, its effects on 2-D MASW results have not been extensively studied, particularly using experimental field data. Due to the impracticality of adjusting geophone subarray lengths in the field using traditional MASW equipment, the limited studies that have been performed are based on synthetic data sets. DAS technology has enabled us to address that limitation by evaluating the impact of subarray length using experimental surface wave data collected at a well-characterized field site.

3 DAS FOR 1-D AND 2-D MASW

DAS is a relatively new technology for recording seismic wave propagation, with promising applications in both engineering and geophysics (Daley *et al*. 2013, Lindsey *et al.* 2017; Spikes *et al.* 2019; Hubbard *et al.* 2021a). DAS uses backscatter interferometry (Hartog 2017) of light propagated through fibre-optic cables to collect data over very large scales (e.g. kilometres) while still maintaining very high spatial resolution (e.g. metres), which is much less practical with geophone arrays (Soga & Luo 2018). DAS measures the axial strain within a fibre-optic cable using an interrogator unit (IU) which produces the source light and measures Rayleigh backscattering (Nakazawa 1983).

In DAS, backscattered light originating from two locations along a sensing cable are compared to determine the relative change in length of the cable between them as a function of time. The baseline distance between these reflection points is the gauge length set by the IU. The change in optical phase between two backscatter source locations as a function of time is used to calculate the change in length between those points and, by extension, the strain in the cable (Giallorenzi *et al.* 1982; Hartog 2017; Hubbard *et al.* 2022). The fibre-optic cable acts as a linear array of sensors (channels) and can be either laid across the ground surface (Spikes *et al.* 2019) or buried to improve coupling with the soil (Galan-Comas 2015; Vantassel *et al.* 2022). Importantly, each DAS channel is a distributed measurement over the gauge length, unlike the point measurements provided by individual geophones.

The relatively high upfront cost of purchasing an IU and the fact that DAS is presently limited to horizontal, in-line sensing, means that it may not currently be practical for some projects, such as those with limited scales or requirements for multi-component sensing. Nevertheless, many previous studies have demonstrated that DAS can be successfully used as an alternative to traditional geophone arrays to collect high-quality data for various geophysical and engineering applications, including MASW. Galan-Comas (2015), Cole *et al.* (2018), Lancelle *et al.* (2021), Vantassel *et al.* (2022), Song *et al.* (2020) all utilized DAS for active-source 1-D and 2-D MASW surveys covering various lateral extents. Lancelle *et al.* (2021) collected MASW data from a DAS array in Garner Valley, CA, as well as geophone and accelerometer arrays collocated with various sections of the DAS array. Lancelle *et al.* (2021) found that the dispersion data from all three receiver types agreed well overall, with excellent agreement above 6 Hz. Lancelle *et al.* (2021) inverted the DAS data and found that the resulting V_S profile agreed well with an inverted profile from the accelerometer data and a borehole log from the site. Galan-Comas (2015) performed a rigorous MASW survey using collocated 70-m-long geophone and DAS arrays with multiple shot locations off each end of the arrays. The dispersion data extracted from both arrays agreed from 6 to 23 Hz, where the dispersion data resolved from the DAS records ended, with the only significant difference being that the geophone data extended up to 56 Hz.

Vantassel *et al.* (2022) successfully extracted multimodal Rayleigh-wave dispersion data from colocated DAS and geophone arrays, with all three extracted modes and the resulting inverted V_S profiles being in good agreement. Vantassel *et al.* (2022) demonstrated that converting the units of the DAS data from strain to particle velocity, like those of the geophones, was not necessary when using frequency-dependent normalization in the MASW processing, simplifying the analysis workflow for DAS data. Additionally, Vantassel *et al.* (2022) identified gauge length and channel spacing as two critical parameters for consideration when using DAS

to collect MASW data. The importance of channel spacing is not surprising, given it is an analog to geophone spacing, which has been shown to be the primary factor that controls the minimum resolvable wavelength, requiring at least two sensors per wavelength (Park *et al.* 1999; Foti *et al.* 2018). Less-intuitively, Vantassel *et al.* (2022) found that gauge length strongly limits the minimum resolvable wavelength, as DAS channels are unable to accurately measure signals with wavelengths near and below the gauge length due to their distributed nature. As a result, the minimum resolvable wavelength of DAS is the longer of either two channel spacings or the gauge length, which is consistent with the limitation of extracting dispersion data at higher frequencies, as observed by Galan-Comas (2015), if using an IU with a long gauge length. Given that Vantassel *et al.* (2022) collected their DAS data using a 1-m channel separation and a 2-m gauge length, it is not surprising that they obtained great agreement with the geophone dispersion data, which was collected using a 2-m receiver spacing.

Cole *et al.* (2018) and Song *et al.* (2020) utilized DAS to perform partial and full 2-D MASW surveys, respectively, following the traditional roll-along approach. Cole *et al.* (2018) utilized an existing 5-km fibre near a rail line in the UK to collect DAS data along an approximately 500-m long alignment. They used 50 shot locations spaced roughly 10 m apart for both DAS (1-m channel spacing and unspecified gauge length) and a 24-channel geophone (2-m spacing) array on a land streamer. Cole *et al.* (2018) found that they were able to record coherent wave energy out to significantly higher offsets with the DAS array (125 m) than the geophone array (50 m), which was limited by the maximum number of geophones the study team was able to deploy practically. They also found that they were able to extract dispersion data from both arrays but did not proceed to the inversion stage. Unfortunately, they do not state what portion(s) of the DAS array were used for dispersion processing, so the utility use of this study when considering the effectiveness of DAS for 2-D MASW is limited. Song *et al.* (2020) collected DAS data for 2-D MASW using a section of cable approximately 500-m long with 26 active shot locations. For each shot location, Song *et al.* (2020) used the 30 DAS channels (2-m channel spacing and 10-m gauge length) with a minimum offset of 40 m. The shot locations were placed at 20-m intervals along the cable, resulting in overlaps of approximately 40 m (66 per cent) between adjacent subarrays. Song *et al.* (2020) observed distortion of the extracted dispersion data for wavelengths shorter than about 11 m, which is not surprising given their gauge length of 10 m. Song *et al.* (2020) then inverted the dispersion data to produce a pseudo-2-D V_S cross-section below the array.

Notably, neither the Cole *et al.* (2018) or Song *et al.* (2020) groups fully utilized what we believe to be one of the key strengths of performing 2-D MASW with DAS; its flexibility as related to the choice of subarray processing geometry. Rather, both groups mimicked the traditional roll-along approach of selecting a single subarray geometry relative to a single shot location. This is unnecessary with DAS, as the entire length of the fibre-optic cable can be recorded simultaneously. The sensor spacing and gauge length are also constant along the entire length of the fibre, removing the trade-off between subarray length and sensor spacing imposed by traditional seismic equipment. Additionally, any number of the available shot locations can be considered for any given subarray, further increasing flexibility and reducing the potential for misinterpretation (Yust 2018). Therefore, the exact location and length of the individual MASW subarrays, including the subarray interval, can be varied during the processing stage after field testing has been performed. By taking advantage of this flexibility, we are able to demonstrate how DAS can be used to collect a wealth of 2-D MASW data quickly and efficiently by enabling evaluation of the effects of subarray geometry, which would be highly impractical with traditional equipment.

4 DAS TESTING SETUP

The DAS data for this study were collected at the NHERI@UTexas (Stokoe *et al.* 2020) Hornsby Bend test site in Austin, TX in October 2021 with the help of NHERI@UTexas personnel. The site is located approximately 1 km north and east of the lower Colorado River with surficial soils dated to the late Pleistocene and early Holocene (Blum & Valastro 1994). A quarry located approximately 1 km northeast of the site has reported the presence of a shale layer at a depth of about 12 m. An overview of the testing layout is shown in Fig. 1. The endpoints of a 200-m-long linear array were first established using a total station. Second, a trenching machine was used to excavate a 10–15 cm trench along the length of the array. A fibre-optic cable manufactured by NanZee was installed in a 200-m long, 10- to 15-cm deep trench. Installing the fibre-optic cable in a trench was important to allow for adequate coupling of the cable to the surrounding soil, as well as to protect it from environmental effects. While this did increase the effort of installation somewhat, the shallow depth and narrow width required meant it was not a significant challenge or expense. This is the same experimental installation that was used in the studies by Vantassel *et al.* (2022) and Hubbard *et al.* (2022). The cable from NanZee Sensing Technologies (NZS-DSS-C02) is a single-mode, tightly-buffered cable which is specifically designed for strain sensing applications and has been shown to work well in a variety of civil engineering projects (Zhang *et al.* 2019; Hubbard *et al.* 2021a, b). The tightly buffered optical fibre is surrounded by braided steel reinforcement to protect the fibre and increase the strength of the cable. A textured polyethylene sleeve encases the reinforced fibre and aids in the transfer of soil strains to the fibre.

The data analysed in this study were collected using an OptaSense ODH4 + IU owned by NHERI@UTexas. The data were recorded continuously at a ping rate of 20 kHz and then decimated down to 1 kHz prior to dispersion processing. The ODH4 + was configured with a gauge length and channel separation of 2.04 and 1.02 m, respectively, the minimum possible with the unit. This results in a 50 per cent overlap with each adjacent channel. Unlike geophones, where the location of each channel is known, the exact location of each DAS channel must be established in the field using a tap test. A tap test (i.e. lightly hitting the ground at known locations along the cable) was performed to determine which channels along the NanZee cable were associated with specific locations on the ground surface. This allowed the absolute spatial locations of the DAS channels to be determined within a maximum error of half the gauge length (1.02 m). Tap test results from each end of the array were used to determine that the DAS array includes a total of 196 usable channels, as shown by the blue line in Fig. 1.

The source used in this study was the Thumper vibroseis mobile shaker truck from the NHERI@UTexas experimental facility (Stokoe *et al.* 2020). Thumper is a moderate force shaker designed for testing in urban areas and has a maximum force output of about 27 kN. Thumper was configured to shake vertically, producing a chirp signal with the frequency of shaking sweeping linearly from 5 to 200 Hz over 12 s and a 0.5 s cosine taper on each end. Shaking was performed at 32 shot locations along the alignment of the array, beginning at −24 m and continuing every 8–224 m. These

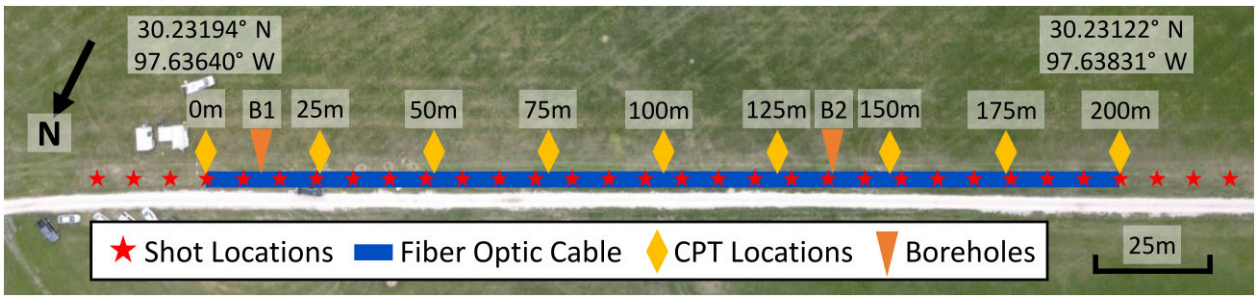

Figure 1. Aerial view of the Hornsby Bend test site showing the locations of CPT tests and boreholes as well as the DAS fibre-optic cable and the vibroseis shot locations.

shot locations are shown as red stars in Fig. 1. Three sweeps were performed with Thumper at each shot location, resulting in a total of 96 shot records. A GPS time break and the measured dynamic ground force output from Thumper were recorded for each sweep. The time break was used to extract the individual shot records from the continuous and GPS-synchronized DAS recording, which totalled about 1.5-hr long. The individual source and shot records were then set to a common time scale for ease of processing.

5 INVASIVE SITE CHARACTERIZATION DATA

In support of this and other studies being performed at the Hornsby Bend site, the NHERI@UTexas team performed invasive subsurface characterization tests, including cone penetration testing (CPT) and drilling geotechnical borings. Nine CPT soundings were performed adjacent to the fibre-optic cable at 25-m intervals from 0 to 200 m, as shown in Fig. 1. The method developed by Robertson (2009) was used to determine the normalized soil behaviour type index value (I_c) for all the CPT data collected. These results were then used to develop a soil-type cross-section of the subsurface conditions along the cable alignment, which is shown in Fig. 2. The soil behaviour type at each CPT sounding is shown as a colour-coded column at the location of each sounding. Cohesive soil types have been coloured blue and green, while granular materials are shown with tan, brown, and grey. Additionally, the measured tip resistance (q_c) profile from each sounding is shown with a red line and shaded area to the right of each column at a scale of 20 MPa (scale bar shown on the plot). When identifying sections of different soil behaviour types for each column, zones with thicknesses less than 10 cm (approximately 6 samples) were ignored for simplicity. An interpretation of the potential layer boundaries beneath the site was made based on the soil behaviour types, as shown in Fig. 2.

The CPT soil behaviour type results indicate the presence of three main layers above the depth of CPT refusal (i.e. the depth where the cone encountered too much resistance to be pushed deeper), which is shown as a heavy black line in Fig. 2. The three layers consist of a shallow (roughly 0–4 m depth) granular layer of sand and sand mix, an intermediate (roughly 4–7 m depth) cohesive layer of clay and silt mix, and a deeper (roughly 7–10 m depth) granular layer of sand and sand mix. The depth of refusal along the cable ranged from 7.96 to 10.56 m, with an average depth of 9.15 m. The depth of refusal generally trended deeper as the distance along the array increased to the southwest, with the exception of the sounding at 75 m and the final sounding at 200 m.

Two geotechnical borings were drilled at 12.5 m (B1) and 137.5 m (B2) along the alignment of the DAS array (see Fig. 2). These borings were drilled after the 2-D MASW analyses had been completed as a means to determine how deep the shale layer was across the site. B1 was drilled to 24.4 m and B2 was drilled to 15.2 m. Disturbed sampling was performed in both boreholes every 1.5 m and the collected material was classified according to the Unified Soil Classification System (ASTM 2017). The lithographic logs for both boreholes are shown in Fig. 2. Except for the thin high-plasticity clay (CH) layer at the very top of log B2, both boreholes show the same sandy–silty–clay (CL-ML) layer for the top 8.5–10 m of depth. This layer is noted on the boring log reports as alternatively being classified as silty–clayey–sand (SC-SM), which is consistent with the silt and sand mixes characterized by the CPT soil behaviour types for much of these depths. The bottom depth of the CL-ML layer (8.5 m for B1 and 10 m for B2) also agrees well with the depths to CPT refusal. Below this depth, the lithology in B1 transitions to dense clayey-sand (SC), while B2 encounters dense clayey-gravel with sand (GC) and well-graded gravel with sand (GW) layers. Finally, after the dense granular layers, both boreholes encounter a shale layer at depths of 13.4 m for B1 and 14 m for B2, relatively consistent with the report from a nearby quarry.

Downhole seismic testing was performed in B1 by the NHERI@UTexas team to a depth of 24 m using a receiver interval of 1 m. An aluminium shear plank offset 1-m away from the borehole was used to generate vertically propagating shear waves in both forward and reverse directions, which were recorded with a Geostuff BHG-3 3-component borehole geophone. The NHERI@UTexas team used the slope method on the corrected vertical traveltimes (Stolte & Cox 2019) to identify four velocity layers: a 3.5-m-thick layer at the surface with $V_S = 140\ \mathrm{m\,s^{-1}}$, a second layer with $V_S = 248\ \mathrm{m\,s^{-1}}$ down to 6.5 m, a third layer with $V_S = 359\ \mathrm{m\,s^{-1}}$ down to a depth of 13.5 m, and a final layer with $V_S = 445\ \mathrm{m\,s^{-1}}$ starting at 13.5 m and continuing to the bottom of the borehole. The boundary below the third layer agrees very well with the top of the shale layer identified in both boring logs. These downhole V_S measurements are used later in the paper as a means to evaluate the 2-D MASW V_S profiles.

6 2-D MASW PROCESSING

Individual MASW subarrays were generated from the DAS data by extracting groups of contiguous channels of various lengths from the full DAS array of 196 channels. The three sets of MASW subarrays used in this study to examine the effects of subarray length on the resulting 2-D V_S cross-sections consisted of subarrays with 12, 24 and 48 channels. Each set of subarrays started on the first channel and used the same subarray interval (i.e. the number of channels between the starting points of two adjacent subarrays) of 4 channels (approximately 4 m). For instance, the first 12-channel subarray

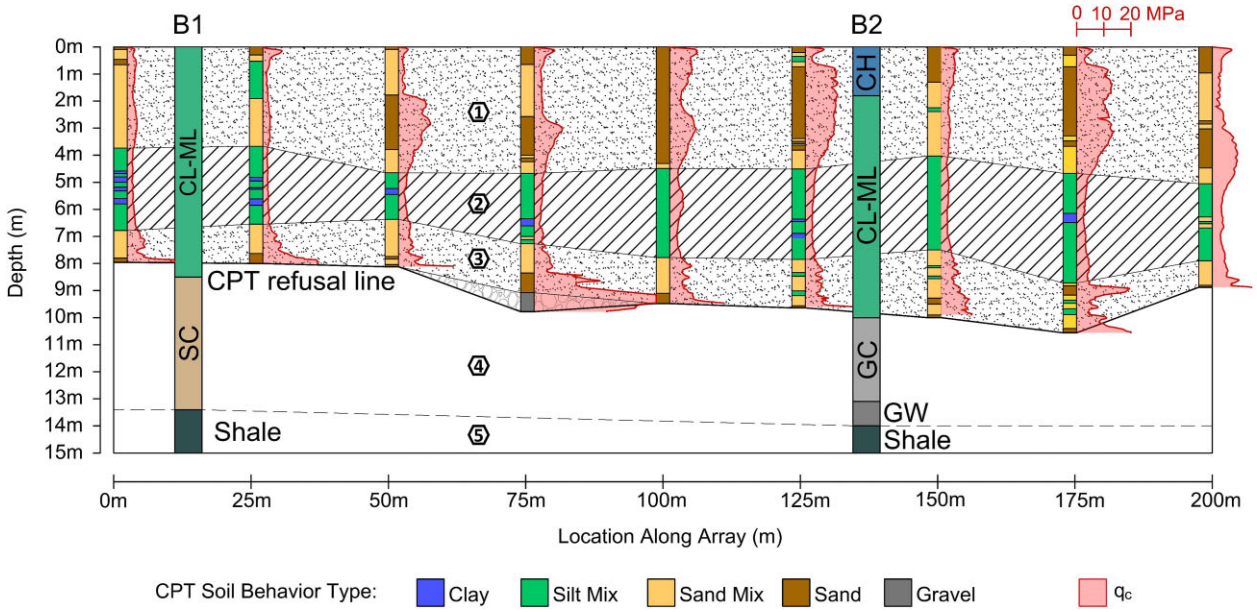

Figure 2. A cross-section of subsurface soil conditions beneath the DAS array determined from invasive testing at the site. The CPT soil behaviour type is indicated at each CPT sounding location, and the corresponding tip resistance (q_c) profiles are shown to the right of each sounding. A preliminary interpretation of the soil layering has been performed based on soil behaviour type as follows: 1- surface granular layer, 2- cohesive layer, 3- intermediate granular layer, 4- dense granular layer and 5- shale layer. The lithographic logs with classifications of disturbed samples for both boreholes are also shown (CL-ML: sandy silty clay, SC: clayey sand, CH: fat clay, GC: clayey gravel with sand and GW: well-graded gravel with sand).

Table 1. Characteristics of the three sets of MASW subarrays analysed in this study.

Number of channels	Number of arrays	Overlap (%)
12	47	67
24	44	83
48	38	92

included channels 1–12, the second subarray included channels 5–16, etc. The resulting quantity of subarrays in the three sets, and the percentage of array overlap between them, are summarized in Table 1, with a total of 129 individual MASW subarrays analysed in this study.

The waveforms from each individual subarray were processed and inverted separately to produce a single 1-D V_S profile located at the midpoint of that subarray (Luo *et al*. 2009). The flexibility and efficiency of the DAS acquisition method allowed for more extensive and rigorous processing and inversion procedures than would have been practicable for most traditional pseudo-2-D MASW surveys. Six different shot locations were considered to generate the dispersion data for each subarray, consisting of the three closest shot locations located further than 4 m off either end. These locations were selected so that a variety of shot offset distances would be used, while also ensuring that the shot location was not too close to the first channel in the subarray as a means to mitigate near-field effects. Shot offsets for any given subarray ranged from approximately 4–24 m. An example of one of the 24-channel subarrays is shown in Fig. 3(a), along with the locations of the six shots used for that subarray. This subarray will henceforth be referred to as the example subarray. The example subarray was the second 24-channel MASW subarray along the DAS line and ranged from 4.08 to 27.54 m. The raw DAS records from each shot location were correlated to the ground force output by the vibroseis shaker truck, as recommended by Xia *et al.* (2000). The three correlated records from each shot location were then stacked in the time domain to improve the signal to noise ratio of the records prior to dispersion processing. The stacked, correlated records from the shot location at −8 m are shown in Fig. 3(b).

The MASW processing for each subarray followed the workflow documented by Vantassel & Cox (2022) and utilized the *swprocess* Python package developed by Vantassel (2021). The frequency-domain beamformer method with a cylindrical-wave steering vector and square-root-distance weighting (Zywicki & Rix 2005) was used to extract Rayleigh wave dispersion for each stacked record for the source location and subarray combination. The resulting dispersion image for the example records is shown in Fig. 3(c). This image shows the relative energy for different phase velocity–frequency pairs. The phase velocity value with the highest energy at each frequency was then selected automatically and saved for calculating dispersion statistics across all six shot locations. For this particular example, a clear fundamental mode trend was only resolved between approximately 10–30 Hz. Several higher-mode trends are observable at frequencies greater than 30 Hz. The picked dispersion data from all six shot locations was combined to form a single data set that could be used to calculate dispersion statistics (refer to Fig. 4). Prior to calculating dispersion statistics, the combined set of dispersion points was trimmed to remove obvious outliers and points that did not appear to correspond to slowest apparent modal trend, which was assumed to represent the fundamental Rayleigh mode (R0). No pre-determined frequency or velocity bounds were used to trim the dispersion data at this stage, and any points that could reasonably be considered part of the coherent R0 trend were retained and included within the combined dispersion data set used to quantify dispersion uncertainty. Additional information on the methods used to calculate dispersion data statistics may be found in Vantassel & Cox (2022). The data were then resampled in terms of frequency and

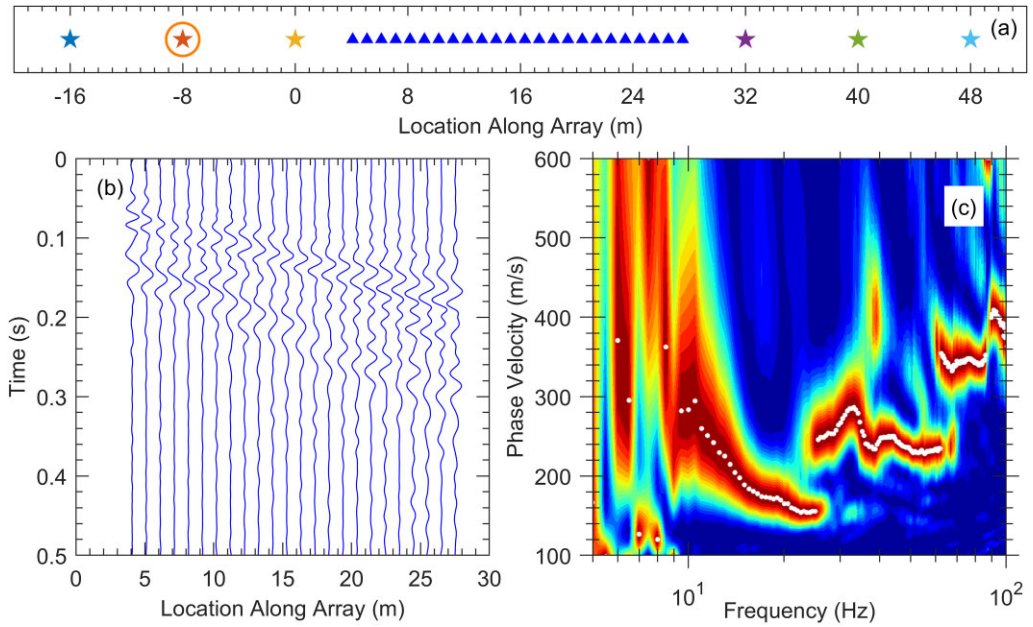

Figure 3. Example 24-channel MASW subarray and the data extracted from it. (a) The location of DAS channels (blue triangles) and the 6 shot locations (stars) analysed for the subarray. (b) Waterfall plot of the correlated and stacked vibroseis truck waveforms recorded with the subarray for the highlighted shot location at −8 m. (c) Rayleigh dispersion image extracted from the waveforms for the highlighted shot location. The peak energy for each frequency value is marked with a white circle.

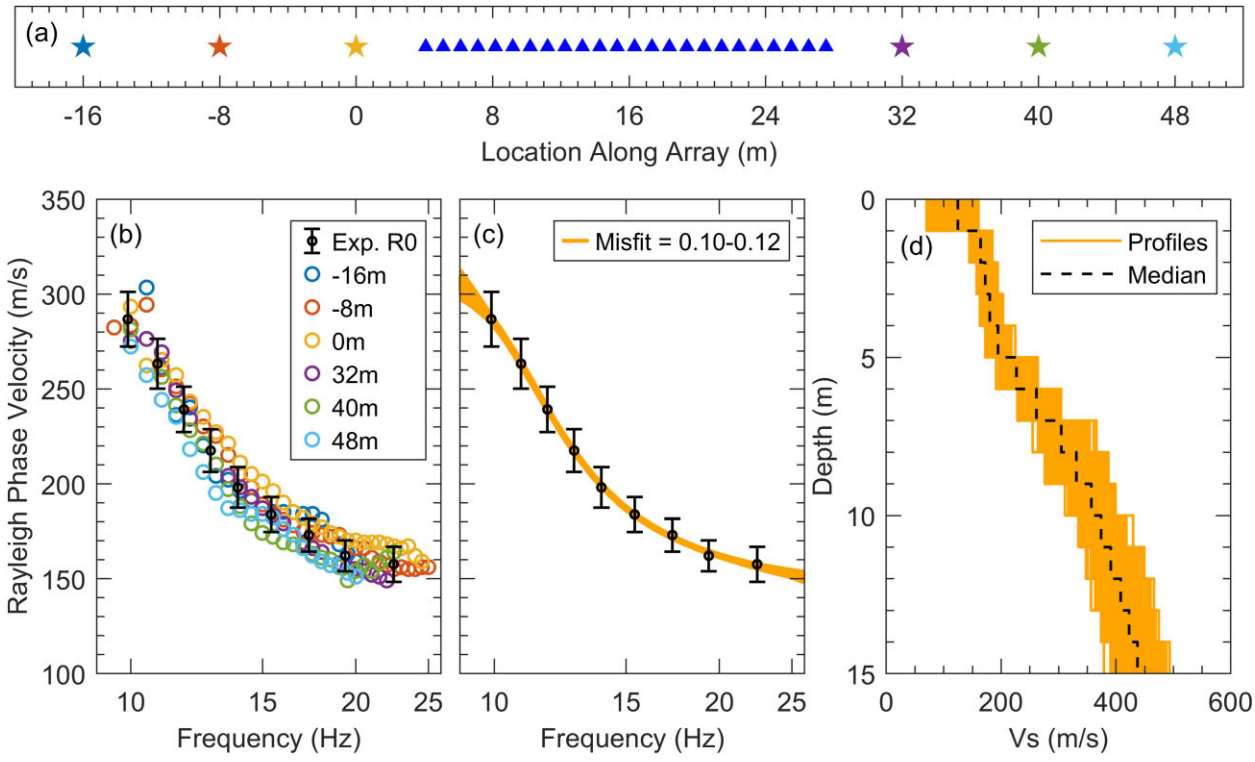

Figure 4. Example 24-channel MASW subarray and the dispersion processing and inversion results from it. (a) The location of DAS channels (blue triangles) and the 6 shot locations (stars) analysed for the example subarray. (b) Picked Rayleigh dispersion data from all six shot locations after trimming and calculating statistics. (c) The fundamental Rayleigh mode (R0) experimental dispersion data and inversion-derived theoretical dispersion curves for the 300 lowest misfit trial models. The ranges of misfit values for the 300 best models from each parametrization are shown in brackets in the legend. (d) The 300 lowest misfit inversion-derived V_S profiles.

the mean and standard deviation phase velocity of the points within each frequency bin were calculated, as shown with black error bars in Fig. 4(b).

The experimental dispersion data for each subarray was inverted using the Dinver module of the open-source Geopsy software (Wathelet *et al.* 2020), which uses the Neighborhood global search

algorithm (Sambridge 1999). The model parametrization consisted of 14 layers with a uniform thickness of 1 m each, overlying a half-space. This parametrization was chosen because it is similar to those often used in 2-D MASW analyses, where a large number of relatively-thin, fixed-boundary layers are used during inversion (e.g. Seshunarayana & Sundararajan 2004; Park & Miller 2005a; Song *et al.* 2020). As the extracted dispersion data does not indicate the presence of any significant velocity reversals, the velocities of each layer were restricted to increasing with depth. The inversions followed the SWinvert workflow and used the Dinver tuning parameters recommended by Vantassel & Cox (2021). Each inversion considered a total of 60 000 trial models and was performed three times to account for the variations that may result from the random starting seed of the search algorithm. Thus, a total of 180 000 trial models were searched for each individual subarray. The 1000 'best' (i.e. lowest misfit) trial models from each inversion were saved, from which the 300 with the lowest overall misfit were then selected to represent uncertainty in the subsurface V_S. The 300 best V_S profiles resulting from the inversion of the example subarray are shown in Fig. 4(d), with the corresponding theoretical dispersion curves shown relative to the experimental dispersion data in Fig. 4(c). Finally, the median V_S profile for each subarray was calculated, as shown in Fig. 4(d).

7 RESULTS

7.1. Dispersion results

While the final results of the 2-D MASW surveys are pseudo-2-D V_S cross-sections, the experimental dispersion data from each subarray is an intermediate result of the processing step that can still provide useful information beyond its use in the subsequent inversions. The experimental dispersion data from the three sets of subarrays provide an opportunity to examine the effects of subarray length on the MASW processing results. All 129 statistical dispersion data sets are shown in Fig. 5 in terms of both frequency and wavelength.

The most visually obvious difference between the dispersion data resulting from subarrays of different lengths is the bandwidth of their resolved wavelengths (i.e. the range of wavelengths actually recorded by the arrays). While there is some variation in these values among subarrays of the same length, averages for each set can be used to examine overall trends between them. The average minimum wavelength ($\lambda_{min,avg}$) resolved by each subarray set was very consistent, with values of 6.6, 6.5 and 6.4 m for the 12-, 24- and 48-channel subarrays respectively. This is consistent with expectations based on findings in the existing literature, as the channel (sensor) spacing and gauge length, which generally control the minimum resolvable wavelength (Park *et al.* 1999; Foti *et al.* 2018; Vantassel *et al.* 2022), are constant in all three sets of subarrays. This ability to maintain the same channel spacing and gauge length across subarrays of any length is one of the key strengths of utilizing DAS for 2-D MASW, as it allows for data to be collected over long arrays without sacrificing the ability to resolve short wavelengths by increasing receiver spacing. While this can enable more efficient recording for very long arrays, there will still be some applications, particularly for projects of smaller lateral extent, where geophone arrays may be the best option.

The differences in dispersion data are most pronounced in terms of the average maximum wavelength ($\lambda_{max,avg}$), with values of 20, 28 and 39 m for the 12-, 24- and 48-channel subarrays, respectively. The theoretical maximum resolvable wavelength ($\lambda_{a,max}$) which can be recorded by an MASW array is a function of both the shot location and the array length (L) (Yoon & Rix 2009; Vantassel & Cox 2021). If the shot location is too close to an array of a given length, near-field effects can cause phase velocities at longer wavelengths to be incorrectly measured (typically underestimated), which is one of the reasons why multiple shot locations were used in this study to mitigate these effects. The exact effect of array length on maximum resolvable wavelength is disputed, with Foti *et al.* (2018) arguing that $\lambda_{a,max} \leq L$, while Park (2005) suggests a less restrictive value of $\lambda_{a,max} \leq 3\ L$. Here, the ratio λ_{max}/L ranged from 0.8 to 1.8, within the bounds prescribed by previous studies. The maximum resolved wavelength is especially important, as it controls the depth to which the subsurface can be accurately characterized. For example, it is common to assume the approximately-maximum, well-resolved depth of the inverted V_S profile is equal to one-half to one-third of $\lambda_{a,max}$ (Foti *et al.* 2018). This is the cause of the trade-off between increased characterization depth made possible by using longer arrays, and the desire to minimize spatial averaging by using shorter arrays.

While the Hornsby Bend site considered in this study does not appear to have a large amount of lateral variability based on the CPT results presented in Fig. 2, the effects of spatial averaging can nonetheless still be observed in the dispersion data by examining the phase velocity values from different subarray lengths. We examine this variability quantitatively by computing the standard error of the mean Rayleigh phase velocities over the shared wavelength band for the 12-, 24- and 48-channel subarray lengths as 8.2, 7.1 and 6.1 m s^{-1}, respectively. This is the variability of the mean Rayleigh-wave velocity extracted for each subarray at each shared wavelength. As wavelength is generally considered to be reasonable proxy for depth in dispersion data, this allows for an examination of the variability of the velocities recorded at common depths within each set of subarrays. The 12-channel subarrays produced the most variable dispersion data (i.e. highest standard error in the mean), while the 48-channel subarrays produced the least variable data (i.e. lowest standard error in the mean), indicative of more spatial averaging (i.e. smoothing of the dispersion data) for the longer subarrays than their shorter counterparts. This is consistent with the behaviour described by Park (2005) and confirms that the trends previously observed only in synthetic data are applicable to the field.

7.2. Pseudo-2-D V_S cross-sections

Once all of the individual subarray inversion analyses were completed, the resulting median V_S profiles (see Fig. 4d) were used to produce pseudo-2-D V_S cross-sections of the subsurface beneath the 200-m long DAS array. Each V_S profile was first assigned a location along the DAS array equal to the midpoint of the corresponding MASW subarray (Luo *et al.* 2009). The V_S profiles were then discretized at 0.1-m intervals, first in terms of depth (0–15 m) and then laterally along the array (0–200 m) using linear interpolation. Finally, a Gaussian filter ($\sigma = 1$ m) was applied to the entire grid of points to smooth out some of the abrupt and unrealistic discontinuities between profiles while still maintaining as much of the behaviour identified in each individual profile as possible.

The result is that each pseudo-2-D V_S cross-section consists of a 2001 by 151 grid. Because each set of subarrays used a different number of channels, the lateral extent of the 2-D cross-sections also varies. For example, the cross-section derived from the 24-channel subarrays has a total lateral extent of 175.44 m, ranging from 11.73 to 187.17 m (i.e. the mid-points of the first and last subarrays).

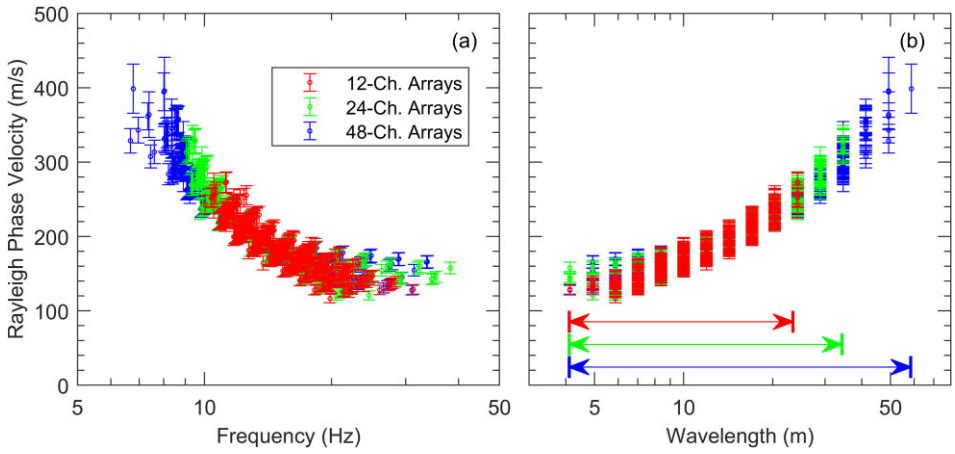

Figure 5. The fundamental-mode Rayleigh (R0) experimental dispersion data extracted for all 12-, 24- and 48-channel MASW subarrays shown in terms of: (a) frequency and (b) wavelength. The bars at the base of (b) show the range of wavelengths resolved for each set of subarrays.

The pseudo-2-D V_S cross-sections produced with the 12-, 24- and 48-channels subarrays are shown in Figs 6(a)–(c), respectively. The variation in lateral extent between the cross-sections is clearly visible, with extents of 187.68, 175.44 and 150.96 m for the 12-, 24- and 48-channel subarrays, respectively. While this reduction in lateral extent for longer subarrays may be insignificant for projects covering significant distances where the loss is a negligible portion of the total characterized length, it is important to consider this limitation when selecting subarray length for projects with tight lateral constraints, where this loss of characterized length may be more consequential.

The cross-sections shown in Fig. 6 demonstrate how the chosen subarray length can affect how V_S varies both with depth and laterally along the array. In general, the top 10 m of the subsurface appears to be quite similar between cross-sections determined using different subarray lengths, with V_S values starting at approximately 100 m s^{-1} at the ground surface and increasing to approximately 200 m s^{-1} at a depth of approximately 4 m. The V_S values in all cross-sections increase from approximately 200–300 m s^{-1} between depths of approximately 4–10 m. The spatial variability of this transition can be quantified by examining the mean and standard deviation values of the depth at which individual V_S profiles within each cross-section exceed 250 m s^{-1}. Within the 12-channel subarray cross-section, V_S exceeds 250 m s^{-1} at a mean depth of 6.74 m with a standard deviation of 0.68 m. For the 24-channel cross-section the mean depth of exceedance and standard deviation are 6.93 and 0.53 m, respectively, and for the 48-channel cross-section the mean depth of exceedance and standard deviation are 6.17 and 0.48 m, respectively. This reduction in lateral variability as subarray length increases is consistent with the dispersion behaviour observed above and the behaviour described by Park (2005).

Below 10 m, the V_S values for each cross-section vary considerably. While all three cross-sections have similar maximum velocities, generally ranging from 400 to 600 m s^{-1} along the array, the average depth at which higher velocities are reached varies between approximately 10–13 m for the different cross-sections. For example, the depths where $V_S > \sim 400$ m s^{-1} from the 12-channel subarrays are generally the shallowest, and these depths consistently increase with each subsequent increase in subarray length (i.e. number of channels). Specifically, the mean depths for m s^{-1} exceeding 400 m s^{-1} are 11.9, 12.8 and 13.8 m for the 12-, 24- and 48-channel cross-sections, respectively. Due to the proximity of the deep impedance contrast in the 48-channel cross-section to the half-space at the base of the model parametrization, the inversions of the 48-channel subarray results were repeated with a similar 1-m layering down to a depth of 20 m to investigate the possibility of the deep contrast being an inversion artifact. Those inversions also indicated a significant impedance contrast at approximately 13–14 m depth, suggesting that it is a real feature within the subsurface and not an artefact of the inversion parametrization.

When examining lateral variation of V_S, values in the 12-channel cross-section vary more significantly along the array, with transitions happening gradually at some locations and much more suddenly at others. In contrast, the lateral V_S transitions seen in the 48-channel cross-section are much smoother, with only a few isolated profiles varying significantly from those around them. With that in mind, it is not surprising that the 24-channel cross-section shows behaviour between the other two, with less variation than the 12-channel one and more than the 48-channel one. This behaviour is consistent with the lateral smearing described by both Park (2005) and Mi *et al.* (2017). As the length of each subarray is increased, more soil is being averaged within the MASW processing scheme. This results in a smearing effect, where $V_S$variations of relatively small lateral extent can no longer be resolved.

While these results show that the choice of subarray length does have a significant impact on real 2-D MASW results, they raise a new issue of needing to determine which set of subarrays produce a cross-section that most reasonably characterizes the actual subsurface conditions at the site. To address this, the results from the invasive testing performed at the site can be used as an additional point of comparison.

7.3. Comparisons to invasive testing

The most comprehensive invasive information about the subsurface conditions at the site is provided by the nine CPT soundings that were performed along the study array. However, the information recorded by the CPT cannot be directly compared to the 2-D MASW V_S results. Therefore, the CPT data were converted to V_S estimates

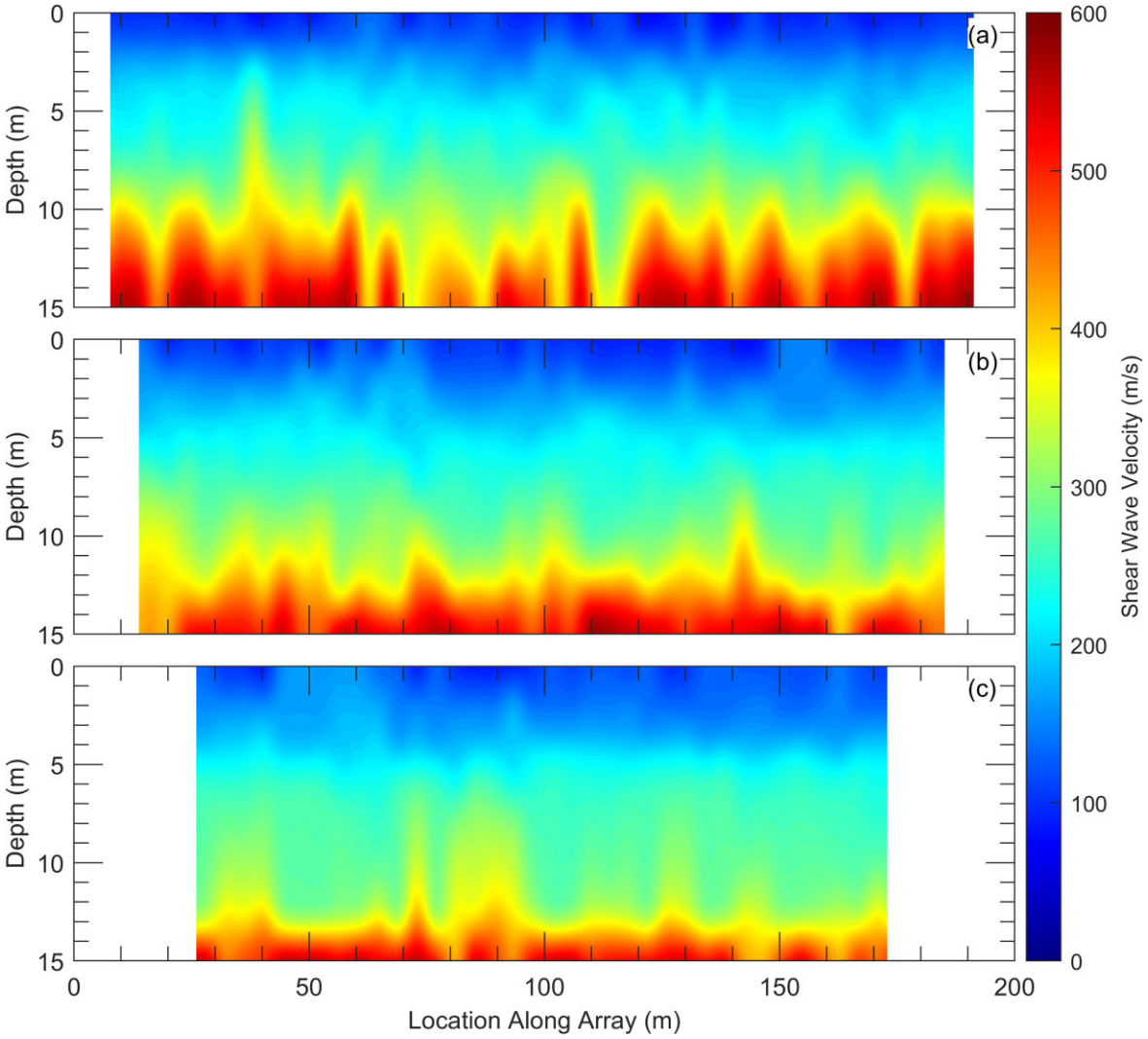

Figure 6. Pseudo-2-D V_S cross-sections inverted using a 14-layer parameterization and dispersion data extracted from: (a) the 47, 12-channel MASW subarrays, (b) the 44, 24-channel MASW subarrays, and (c) the 38, 48-channel MASW subarrays.

using three CPT-to-V_S correlations for each CPT sounding. The three CPT-to-V_S correlations developed by Mayne (2007), Andrus *et al.* (2007) and Robertson (2009) are shown as eqs (1), (2) and (3), respectively, and were recommended by Wair *et al.* (2012) for their 'All Soils' method.

$$V_S = 118.8 \log (f_s) + 18.5 \tag{1}$$

$$V_S = 2.62\, q_t^{0.395} I_c^{0.912} D^{0.124} SF \tag{2}$$

$$V_S = \left[10^{(0.55 I_c + 1.68)} (q_t - \sigma_v) / p_a\right]^{0.5}. \tag{3}$$

These relationships estimate V_S as a function of a variety of factors, including sleeve friction (f_s), corrected tip resistance (q_t), soil behaviour type index (I_c), depth (D), total vertical stress (σ_v) and atmospheric pressure (p_a). The total vertical stress used in the Robertson (2009) relationship was calculated using the depth and the unit weight relationship developed by Robertson & Cabal (2010). Wair *et al.* (2012), modified each of the equations to use consistent units, with f_s, q_t and σ_v in kilopascals, D in metres and $p_a = 100$ kPa. The Mayne (2007) and Robertson (2009) relationships were developed for Quaternary soils in general, while the Andrus *et al.* (2007) relationship uses an age scaling factor (SF) to distinguish between Holocene (SF = 0.92) and Pleistocene (SF = 1.12) soils. Because the age of the soils at the Hornsby Bend test site varies between Holocene and Pleistocene, an average value of SF = 1 was used, as recommended by Wair *et al.* (2012). The resulting V_S profiles from each of these relationships were averaged together to produce a single, mean V_S profile for each CPT sounding. The mean V_S profiles for each CPT sounding within the characterized extent of the array (i.e. soundings from 25 to 175 m) are presented in Fig. 7. Also shown in each panel for comparison are the 1-D V_S profiles from each 2-D MASW subarray centered most closely to the respective CPT soundings.

The 1-D V_S profiles presented in Fig. 7 demonstrate that, in general, there is good agreement between the velocities produced from the CPT correlations and those produced by the individual 2-D MASW subarrays. A notable exception is that the correlated profiles from the CPT soundings at 125 and 175 m indicate higher velocities over the top 4 m, consistent with the higher tip resistances encounter there (refer to Fig. 1). The most likely explanation for this discrepancy is that those two CPT soundings encountered localized pockets of stiffer material. It is important to note that while the CPT

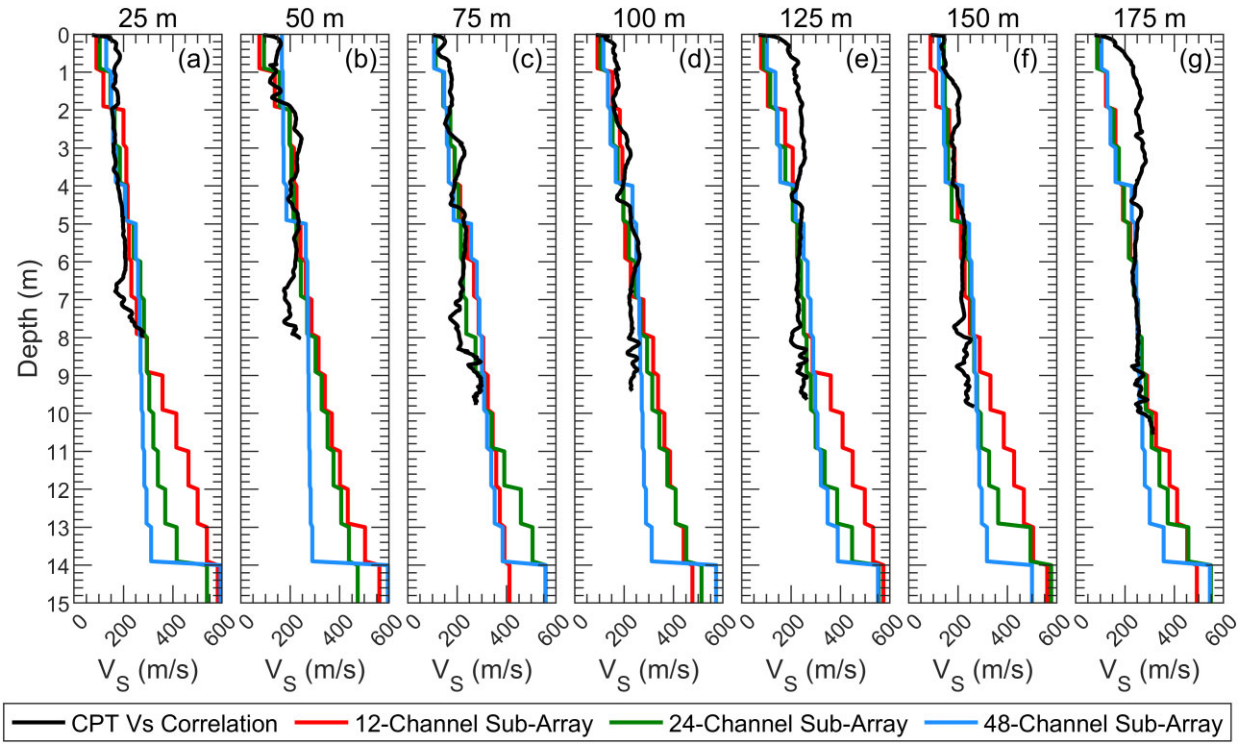

Figure 7. Correlated V_S profiles for the CPT tests performed at (a) 25 m, (b) 50 m, (c) 75 m, (d) 100 m, (e) 125 m, (f) 150 m and (g) 175 m along the DAS array at the Hornsby Bend test site. Also shown in each panel for comparison are the 1-D V_S profiles from each 2-D MASW subarray centred most closely to the respective CPT soundings.

sounding is a discrete measurement of the soil properties at a single location, each MASW subsurvey is a distributed measurement that results in the spatial averaging of the material under the entire length of the subarray.

At shallower depths, the 1-D V_S profiles from the 2-D MASW subarrays of different lengths are generally quite consistent with each other. However, they begin to diverge from one another at depths ranging from approximately 8–10 m, with the 12-channel profiles generally indicating the highest velocities at shallower depths and the 48-channel profiles indicating the lowest velocities at greater depths. This behaviour is consistent with that observed for the overall 2-D MASW V_S cross-sections in Fig. 6. Unfortunately, while the CPT 1-D V_S correlations do provide some valuable information about the validity of the subsurface conditions determined from the 2-D MASW V_S cross-sections, they cannot be used to assess the quality of the 2-D MASW results over the full characterized depth.

While not as comprehensive as the CPT soundings, the B1 and B2 boreholes drilled along the array, and particularly the downhole (DH) testing performed in borehole B1, provide an addition and deeper reference for the subsurface conditions. The lithology logs from boreholes B1 and B2, and the DH 1-D V_S profile acquired from borehole B1, are plotted in Fig. 8 along with the 2-D MASW V_S cross-sections determined from the 12-, 24- and 48-channel subarrays. The 1-D V_S profile from DH testing is plotted as a vertical colour-bar immediately to the left of the borehole B1 lithology log and using the same colour scale as the 2-D MASW V_S cross-sections. The subsurface layering determined from the CPT soundings and SPT boring logs (refer to Fig. 2) is also superimposed on the 2-D MASW V_S cross-sections.

It is clear that the V_S values and layering from the DH testing agree best with the results from the 48-channel subarrays, particularly in regards to the location of the shale layer (layer 5). While the V_S values for the shale layer are generally higher in the pseudo-2-D V_S cross-section than in the downhole results, the depth to the top of the shale layer has been resolved quite accurately with the 48-channel subarrays. The same cannot be said for the pseudo-2-D V_S cross-section obtained from the 12-channel or 24-channel subarrays. At most locations, the V_S values below the depth of CPT refusal (i.e. the boundary between layers 3 and 4) have been overestimated in the 12- and 24-channel V_S cross-sections relative to what was obtained from DH testing. While this appears to be an obvious error in overestimating the V_S in layer 4, the location of the strong impedance contrast indicated in the 12-channel V_S cross-section seems to agree quite well with the depth to CPT refusal, which roughly corresponds with the depth to the top of the dense granular layers that overly the shale. Thus, if one were simply trying to identify the depth to the top of the dense granular layers/CPT refusal, using 12-channel subarrays would work quite well. However, the V_S values for the granular layers would be significantly overestimated. Conversely, if one were trying to identify the depth to the top of the shale layer, using 48-channel subarrays would work best, and the V_S values of the shale would only be slightly overestimated.

The stratigraphy of the top layer, which is based on the CPT soil behaviour types, appears to correspond well with a V_S of approximately 200 m s^{-1}. This layer seems to be consistently resolved by all three subarray lengths. The stratigraphy of the second and third layers is not distinct in terms of changes in V_S within any of the 2-D MASW cross-sections.

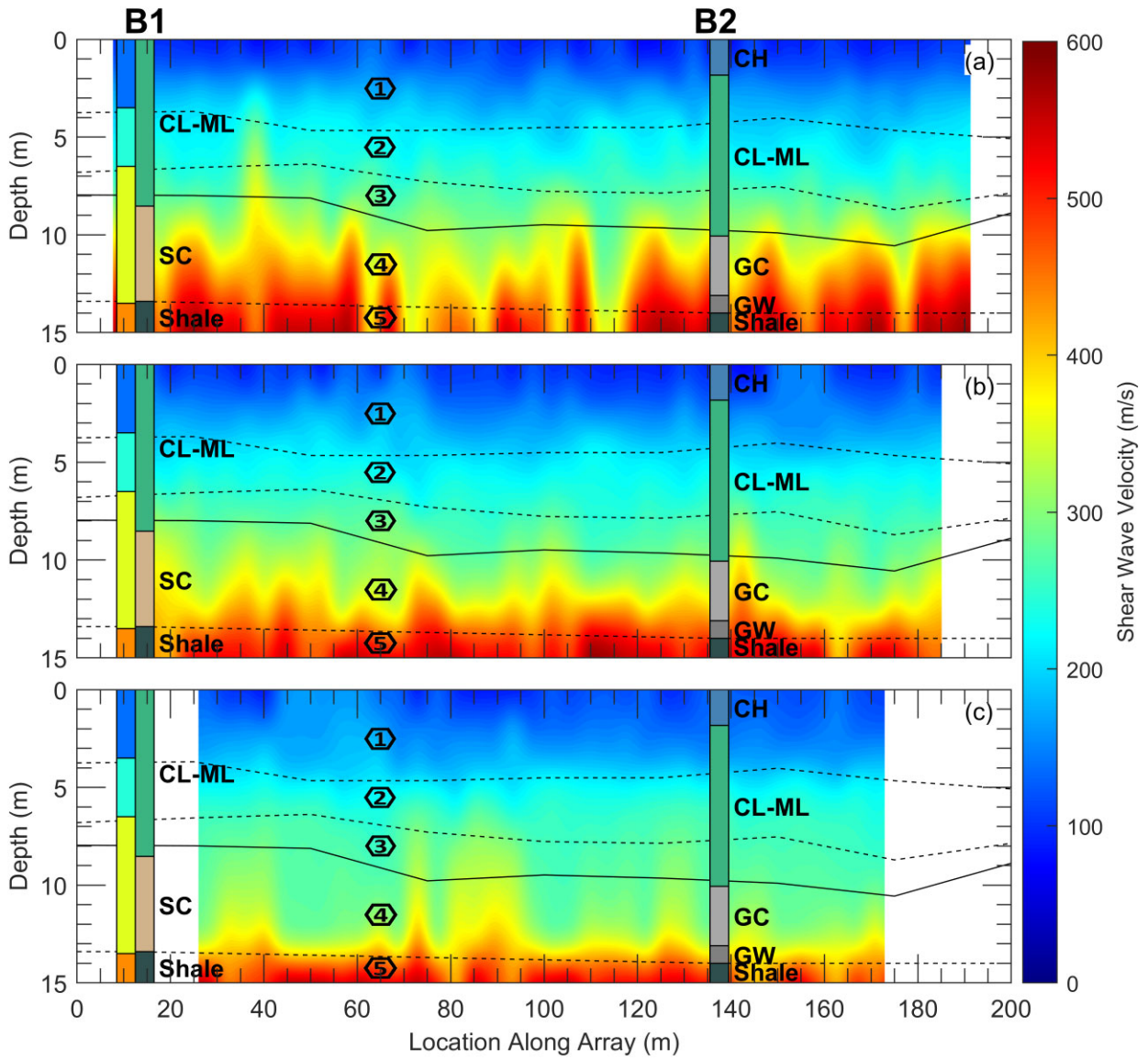

Figure 8. Pseudo-2-D V_S cross-sections from the (a) 47, 12-channel, (b) 44, 24-channel and (c) 38, 48-channel subarrays inverted using a 14-layer parametrization and the lithographic logs for boreholes B1 and B2 shown at their respective locations along the array. The four velocity layers from downhole testing in borehole B1 are shown next to the lithographic log for that borehole. The depths of refusal for 9 CPT soundings along the array are shown with a solid black line and other stratigraphy derived from the boring logs and CPT results are shown with dashed black lines on all three plots.

8 DISCUSSION

These results clearly indicate that changing the subarray length can result in significant changes to the resulting V_S cross-sections, particularly at greater depths. Thus, calibration to a limited number of invasive field-testing results is likely needed to determine the best 2-D MASW processing and inversion scheme for a given application. For example, if one were looking to estimate the depth to CPT refusal across this site without needing accurate V_S results, using a criteria of $V_S > 400$ m s^{-1} with a 12-channel array would likely be an effective way to do this (refer to Fig. 8a), even if the resulting V_S cross-section should not be considered a true representation of subsurface conditions. Alternatively, if one's goal was to identify the depth of the shale layer, applying the same criteria ($V_S > 400$ m s^{-1}) to the 48-channel subarray results shown in Fig. 7(c) would yield the best results. Conversely stated, if locating the top of shale were the desired testing outcome and 12-channel subarrays were used, the results would have been misinterpreted significantly. However, one would not know this until several different processing schemes with variable subarray length had been investigated relative to a limited number of invasive CPT soundings and/or boreholes.

It is important to keep in mind the goals of a specific project when deciding what subarray geometry to use for 2-D MASW. While this is possible with traditional roll-along geophone surveys conducted with a limited number of acquisition channels, it is inconvenient to have to alter the physical setup in order to perform multiple surveys at the same site. Additionally, if there is no *a priori* information about the site layering and velocity, multiple subarray geometries will need to be considered for robust characterization, which is highly impractical with traditional equipment. Furthermore, using both shorter and longer subarrays could be highly beneficial at sites with significant lateral variability at both shallow and greater depths. In this situation, using shorter subarrays would be very beneficial

for resolving the shallow variability without excessive spatial averaging/smoothing, while longer subarray would be more useful for accurately resolving deeper variability. In both cases, the ability of DAS to allow the analyst to select multiple subarray geometries during the processing stage rather than needing to predetermine the survey geometry before data acquisition is a positive benefit for pseudo-2-D MASW imaging.

Collecting DAS data for pseudo-2-D MASW imaging is also highly efficient once the fibre-optic cable has been deployed. To cover the full 200-m extent of the alignment, we used 47, 12-channel subarrays, 44, 24-channel subarrays and 38, 48-channel subarrays, with six shot locations per subarray (three off either end). Only 96 individual active-source shots (three shots at each of 32 locations) were needed to complete this survey, as the same shot can be used for multiple subarrays. Attempting to complete this same survey using geophone subarrays on a land streamer would require 774 shots, with 264 shots required to perform just the 24-channel tests. Furthermore, many 2-D MASW surveys cover alignments significantly longer than 200 m.

While DAS has many positive benefits for collecting and processing 2-D MASW data, there are definitely situations where installing the fibre-optic cable to enable DAS measurements could be overly time-consuming and costly, such as when testing in paved or rocky areas. DAS is particularly well suited to sites where testing spans long distances and/or may be repeated multiple times, for which the cable can be left in the ground. However, it may not be cost-effective for sites with limited spatial extents. Hence, the choice to use DAS for 2-D subsurface imaging projects may be practically based on a cost-benefit analysis, while considering project requirements and site conditions. Overall, as much data as possible should be collected in the field to allow for flexible subarray processing, whether through the use of larger and/or more geophone arrays or DAS. By collecting the data required for flexible subarray processing, analysts are able to properly consider subarray geometry as a potential source of epistemic uncertainty and, when possible, calibrate the analysis based on project-specific goals and supporting information from a limited number of invasive boreholes or CPT soundings.

9 CONCLUSIONS

This study examined the effects of 2-D MASW subarray length using DAS data collected at a single well-characterized site. Specifically, three sets of subarrays of varying lengths were used to develop pseudo-2-D V_S cross-sections along the same 200-m-long DAS line: (i) 12-channel subarrays approximately 11-m long, (ii) 24-channel subarrays approximately 23-m long and (iii) 48-channel subarrays approximately 47-m long. Rigorous dispersion processing and inversion techniques not used in previous, documented applications of 2-D MASW were used to characterize the subsurface as comprehensively as possible, while accounting for multiple sources of uncertainty. The V_S cross-sections resulting from each set of subarrays were visibly different from one another, illustrating how sensitive the results of 2-D MASW can be to the choice of subarray length. However, the results from different cross-sections were shown to better correlate with different key features of the subsurface, as determined from comparisons with invasive site characterization data collected along the array alignment, including correlated V_S values from CPT soundings, borehole lithography logs and downhole seismic testing results.

For this site, the longer, 48-channel subarrays agreed best with the invasive test results, indicating a deeper impedance contrast consistent with the depth of a shale layer identified in the boreholes with only a slight overestimation of the shale velocity. The shorter, 12- and 24-channel subarrays still performed reasonably well, indicating a shallower impedance contrast, which could be used to locate the top of a dense granular layer, but with significant overestimation of V_S for compared to the downhole testing. These results further support the conclusion that *a priori* information, such as invasive testing results, should be used to calibrate the 2-D MASW analysis procedures to specifically address project-specific goals whenever possible. Further, in the absence of *a priori* information that conclusively constrains subsurface layering, analysts may need to consider multiple 2-D MASW subarray configurations to comprehensively characterize the subsurface conditions and properly evaluate the uncertainty of the 2-D MASW results, which has not been considered by previous studies on 2-D MASW using field data.

This study also demonstrates the abilities of using DAS data to perform pseudo-2-D MASW imaging and highlights certain benefits of DAS in terms of enabling flexible subarray processing. The ability of DAS to collect high spatial resolution records simultaneously along significant lengths of fibre-optic cable for each shot allows for data acquisition at rates much faster and more efficiently than traditional roll-along geophone surveys (at least once the cable has been installed), while allowing the subarray length, subarray interval, channel separation and processed shot locations to be selected or changed after acquisition during the processing stage. This is not easy done with traditional roll-along equipment, as the subarray length, subarray interval, and shot location are typically predetermined for the acquisition stage and cannot easily be changed during processing. Overall, DAS is a promising new technology for recording surface-wave data that is very well suited to 2-D MASW testing. DAS's efficiency and flexibility will allow for future studies that would not be practicable with traditional equipment.

ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation (NSF) Graduate Research Fellowship under Grant No. DGE-2137420 and by NSF grants CMMI-2037900, CMMI-1520808 and CMMI-1931162. However, any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF. Special thanks to Dr Kevin Anderson at Austin Water—Center for Environmental Research for the access to the Hornsby Bend Biosolids Management Plant test site. Special thanks to Dr Kenichi Soga for the contribution of the NanZee cable used in this study.

AUTHOR CONTRIBUTIONS

Michael B. S. Yust (Conceptualization [equal], Data curation [lead], Formal analysis [lead], Investigation [equal], Methodology [equal], Software [lead], Visualization [lead], Writing – original draft [lead], Writing – review and editing [equal]), Brady R. Cox (Conceptualization [equal], Funding acquisition [lead], Investigation [equal], Methodology [equal], Project administration [lead], Resources [lead], Supervision [lead], Writing – review and editing [equal]), Joseph P. Vantassel (Data curation [supporting], Investigation [supporting], Methodology [supporting], Software [supporting], Writing – review and editing [supporting]), and Peter G. Hubbard (Data curation [supporting], Investigation [supporting],

Methodology [supporting], Resources [equal], Writing – review and editing [supporting])

DATA AVAILABILITY

The data underlying this paper are available in the NHERI Design-Safe Data Depot as part of the project titled Characterization of the NHERI@UTexas Hornsby Bend Test Site (Yust *et al.* 2023, doi:10.17603/ds2-6ap5-sk09). This includes the individual DAS and source records for each shot as well as detailed documentation of the array geometry and shot locations.

REFERENCES

Andrus, R.D., Mohanan, N.P., Piratheepan, P., Ellis, B.S. & Holzer, T.L., 2007. Predicting shear-wave velocity from cone penetration resistance, in *Proceedings of the 4th International Conference on Earthquake Geotechnical Engineering,* 25–28 June, Thessaloniki, Greece, Paper No. 1454.

Arslan, U., Crocker, J.A., Vantassel, J.P. & Cox, B.R., 2021. Ability of the multichannel analysis of surface waves method to resolve subsurface anomalies, in *IFCEE 2021,* pp. 360–371.

ASTM Committee D-18 on Soil and Rock, 2017. *Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System) 1,* ASTM international.

Blum, M.D. & Valastro Jr, S., 1994. Late quaternary sedimentation, lower Colorado River, Gulf Coastal Plain of Texas, *Geol. Soc. Am. Bull.,* **106**(8), 1002–1016.

Cole, S., Karrenbach, M., Gunn, D. & Dashwood, B., 2018. MASW analysis of DAS Fibre-Optic data for active and passive seismic sources, in *Proceedings of the 80th EAGE Conference and Exhibition 2018,* Vol. **2018,** pp. 1–5, EAGE Publications BV.

Crocker, A.J., Vantassel, J.P., Arslan, U. & Cox, B.R., 2021. Limitations of the multichannel analysis of surface waves (MASW) method for subsurface anomaly detection, in *Proceedings of the 6th International Conference on Geotechnical and Geophysical Site Characterization,* 26–29 September, Budapest, Hungary.

Daley, T.M., Freifeld, B.M., Ajo-Franklin, J., Dou, S., Pevzner, R., Shulakova, V., Kashikar, S., Miller, D.E., Goetz, J., Henninges, J. & Lueth, S. 2013 Field testing of fiber-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring *The Leading Edge* **32**(6): 699–706.

Foti, S. *et al.* 2018. Guidelines for the good practice of surface wave analysis: a product of the InterPACIFIC project, *Bull. Earthq. Eng.,* **16,** 2367–2420.

Foti, S., 2000. Multistation methods for geotechnical characterization using surface waves, *PhD dissertation,* Polytechnic University of Turin, Turin, Italy.

Foti, S., Lai, C.G., Rix, G.J. & Strobbia, C., 2014. *Surface Wave Methods for Near-Surface Site Characterization,* CRC Press.

Galan-Comas, G., 2015. Multichannel analysis of surface waves using distributed fiber optic sensors, *Master's thesis,* Mississippi State University, Starkville, MS.

Giallorenzi, T.G., Bucaro, J.A., Dandridge, A., Sigel, G.H., Cole, J.H., Rashleigh, S.C. & Priest, R.G., 1982. Optical fiber sensor technology, *IEEE Trans. Microwave Theory Tech.,* **30**(4), pp. 472–511.

Hartog, A.H., 2017. *An Introduction to Distributed Optical Fibre Sensors,* CRC Press.

Hubbard, P., Luo, L., Yeskoo, A., Soga, K., Araica, K.M., Cicala, G. & McLeod, M. 2021b. Design of a distributed strain monitoring system for HDPE pipelines crossing an earthquake fault, in *Proceedings of the 20th Plastic Pipes Conference,* 6–8 September 2021, Amsterdam, Netherlands.

Hubbard, P.G. *et al.* 2021a. Dynamic structural health monitoring of a model wind turbine tower using distributed acoustic sensing (DAS), *J. Civil Struct. Health Monitor.,* **11**(3), pp. 833–849.

Hubbard, P.G., Vantassel, J.P., Cox, B.R., Rector, J.W., Yust, M.B. & Soga, K., 2022. Quantifying the surface strain field induced by active sources with distributed acoustic sensing: theory and practice, *Sensors,* **22**(12), doi:10.3390/s22124589.

Ismail, A., Denny, F.B. & Metwaly, M., 2014. Comparing continuous profiles from MASW and shear-wave reflection seismic methods, *J. appl. Geophys.,* **105,** 67–77.

Lancelle, C.E., Baldwin, J.A., Lord, N., Fratta, D., Chalari, A. & Wang, H.F., 2021. Using distributed acoustic sensing (DAS) for multichannel analysis of surface waves (MASW), *Distrib. Acoust. Sens. Geophys.,* 213–228.

Lin, F.C., Li, D., Clayton, R.W. & Hollis, D., 2013. High-resolution 3D shallow crustal structure in Long Beach, California: application of ambient noise tomography on a dense seismic array, *Geophysics,* **78**(4), Q45–Q56.

Lindsey, N.J., Martin, E.R., Dreger, D.S., Freifeld, B., Cole, S., James, S.R., Biondi, B.L. & Ajo-Franklin, J.B., 2017. Fiber-optic network observations of earthquake wavefields, *Geophys. Res. Lett.,* **44**(23), 11–792.

Luo, Y., Xia, J., Liu, J., Xu, Y. & Liu, Q., 2009. Research on the middle-of-receiver-spread assumption of the MASW method, *Soil Dyn. Earthq. Eng.,* **29**(1), 71–79.

Mayne, P.W., 2007. In-situ test calibrations for evaluating soil parameters, *Character. Eng. Propert. Nat. Soils,* **3,** 1601–1652. doi: 10.1201/NOE0415426916.ch2.

Mayne, W.H., 1962. Horizontal data stacking techniques, *Geophysics,* **27,** 927–938.

Mi, B., Xia, J., Shen, C., Wang, L., Hu, Y. & Cheng, F., 2017. Horizontal resolution of multichannel analysis of surface waves, *Geophysics,* **82**(3), EN51–EN66.

Miller, R.D., Xia, J., Park, C.B., Ivanov, J. & Williams, E., 1999. Using MASW to map bedrock in Olathe, Kansas, in *SEG Technical Program Expanded Abstracts 1999,* pp. 433–436. Society of Exploration Geophysicists.

Mohamed, A.M., Abu El Ata, A.S.A., Abdel Azim, F. & Taha, M.A., 2013. Site-specific shear wave velocity investigation for geotechnical engineering applications using seismic refraction and 2-D multi-channel analysis of surface waves, *NRIAG J. Astron. Geophys.,* **2**(1), 88–101.

Nakazawa, M., 1983. Rayleigh backscattering theory for single-mode optical fibers, *JOSA,* **73**(9), 1175–1180.

Park, C.B. & Miller, R.D., 2005a. Seismic characterization of wind turbine sites near Lawton, Oklahoma, by the MASW method. Open File Report, Kansas Geological Survey, University of Kansas.

Park, C.B. & Miller, R.D., 2005b. Seismic characterization of wind turbine sites in Kansas by the MASW method, Kansas geological survey open file report 2005-23, Report to Barr Engineering Company, Minneapolis.

Park, C.B., 2005. MASW horizontal resolution in 2-D shear-velocity (Vs) mapping, Open-File Report, Kansas Geologic Survey, Lawrence, 36pp.

Park, C.B., Miller, R.D. & Xia, J., 1998. Imaging dispersion curves of surface waves on multi-channel record, in *SEG Technical Program Expanded Abstracts 1998,* pp. 1377–1380, Society of Exploration Geophysicists.

Park, C.B., Miller, R.D. & Xia, J., 1999. Multichannel analysis of surface waves, *Geophysics,* **64**(3), 800–808.

Park, C.B., Miller, R.D., Xia, J. & Ivanov, J., 2007. Multichannel analysis of surface waves (MASW)—active and passive methods, *Leading Edge,* **26**(1), 60–64.

Robertson, P.K. & Cabal, K.L., 2010. Estimating soil unit weight from CPT, in *Proceedings of the 2nd International Symposium on Cone Penetration Testing,* Huntington Beach, CA, USA, May 2010.

Robertson, P.K., 2009. Interpretation of cone penetration tests—a unified approach, *Can. Geotech. J.,* **46**(11), 1337–1355.

Sambridge, M., 1999. Geophysical inversion with a neighbourhood algorithm-I. Searching a parameter space, *Geophys. J. Int.,* **138,** 479–494.

Seshunarayana, T. & Sundararajan, N., 2004. Multichannel analysis of surface waves (MASW) for mapping shallow subsurface layers—a case study, Jabalpur, India, in *Proceedings of the 5th International Conference on Petroleum Geophysics,* Hyderabad, India, pp. 642–646.

Soga, K. & Luo, L., 2018. Distributed fiber optics sensors for civil engineering infrastructure sensing, *J. Struct. Integ. Mainten.,* **3**(1), 1–21.

Song, Z., Zeng, X., Xu, S., Hu, J., Sun, T. & Wang, B., 2020. Distributed acoustic sensing for imaging shallow structure I: active source survey, *Chinese J. Geophys.,* **63**(2), 532–540.

Spikes, K.T., Tisato, N., Hess, T.E. & Holt, J.W., 2019. Comparison of geophone and surface-deployed distributed acoustic sensing seismic data, *Geophysics,* **84**(2), A25–A29.

Stokoe, K.H., Cox, B.R., Clayton, P.M. & Menq, F., 2020. NHERI@ UTexas experimental facility with large-scale mobile shakers for field studies, *Front. Built Environ.,* **6,** doi:10.3389/fbuil.2020.575973.

Stolte, A.C. & Cox, B.R. 2019. Towards consideration of epistemic uncertainty in shear wave velocity measurements obtained via seismic cone penetration testing (SCPT), *Can. Geotech. J.,* **57**(1), 48–60.

Thitimakorn, T., Anderson, N.L., Stephenson, R. & Liu, W., 2005. 2-D shear-wave velocity profile along test segment of Interstate I-70, St. Louis, Missouri, in *Site Characterization and Modeling,* pp. 1–9.

Vantassel, J.P. & Cox, B.R., 2021. SWinvert: a workflow for performing rigorous 1-D surface wave inversions, *Geophys. J. Int.,* **224**(2), 1141–1156.

Vantassel, J.P. & Cox, B.R., 2022. SWprocess: a workflow for developing robust estimates of surface wave dispersion uncertainty, *J. Seismol.,* **26**(4), 731–756.

Vantassel, J.P. 2021. jpvantassel/swprocess: latest (Concept), Zenodo. doi: 10.5281/zenodo.4584128.

Vantassel, J.P., Cox, B.R., Hubbard, P.G. & Yust, M. 2022. Extracting high-resolution, multi-mode surface wave dispersion data from distributed acoustic sensing measurements using the multichannel analysis of surface waves, *J. appl. Geophys.,* **205,** doi:10.1016/j.jappgeo.2022.104776.

Wair, B.R., DeJong, J.T. & Shantz, T., 2012. *Guidelines for Estimation of Shear Wave Velocity Profiles,* Pacific Earthquake Engineering Research Center.

Wathelet, M., Chatelain, J.L., Cornou, C., Giulio, G.D., Guillier, B., Ohrnberger, M. & Savvaidis, A., 2020. Geopsy: a user-friendly open-source tool set for ambient vibration processing, *Seismol. Res. Lett.,* **91**(3), 1878–1889.

Xia, J., Miller, R.D., Park, C.B. & Ivanov, J., 2000 Construction of 2-D vertical shear-wave velocity field by the multichannel analysis of surface wave technique, in *13th EEGS Symposium on the Application of Geophysics to Engineering and Environmental Problems,* cp–200, European Association of Geoscientists & Engineers.

Yoon, S. & Rix, G.J., 2009. Near-field effects on array-based surface wave methods with active sources, *J. Geotech. Geoenviron. Eng.,* **135**(3), 399–406.

Yust, M., Cox, B., Menq, F., Hubbard, P. & Vantassel, J.. 2023 Active-source, near-surface, surface-wave measurements using distributed acoustic sensing (DAS), cone penetration testing (CPT), and downhole (DH) testing, in *Characterization of the NHERI@UTexas Hornsby Bend Test Site*. DesignSafe-CI.

Yust, M.B.S. 2018. Dynamic site characterization of TexNet ground motion stations, *Master's thesis,* The University of Texas at Austin, Austin, TX. http://hdl.handle.net/2152/68144.

Zhang, C.C., Shi, B. & Soga, K., March. Distr 2019 Distributed fiber optic sensing of land deformation: methods and case studies, in *Geo-Congress 2019: Engineering Geology, Site Characterization, and Geophysics,* pp. 188–196, American Society of Civil Engineers, Reston, VA.

Zywicki, D.J. & Rix, G.J., 2005. Mitigation of near-field effects for seismic surface wave velocity estimation with cylindrical beamformers, *J. Geotech. Geoenviron. Eng.,* **131**(8), 970–977.