Geophysical Journal International

Geophys. J. Int. (2023) **232,** 1668–1683
Advance Access publication 2022 October 26
GJI Seismology

https://doi.org/10.1093/gji/ggac420

Understanding surface wave modal content for high-resolution imaging of submarine sediments with distributed acoustic sensing

Loïc Viens [1], Mathieu Perton [2], Zack J. Spica,[1] Kiwamu Nishida [3], Tomoaki Yamada [3] and Masanao Shinohara [3]

[1] *Department of E arth and Environmental Sciences, University of Michigan, Ann Arbor, MI* 48109, *USA. E-mail: lviens@umich.edu*
[2] *Instituto de Ingeniería, Universidad Nacional Autónoma de México, Coyoacán,* 04510 *Ciudad de México, CDMX, Mexico*
[3] *Earthquake Research Institute, The University of Tokyo, Bunkyo City, Tokyo* 113-0032, *Japan*

Accepted 2022 October 24. Received 2022 October 10; in original form 2022 April 20

SUMMARY

Ocean bottom distributed acoustic sensing (OBDAS) is emerging as a new measurement method providing dense, high-fidelity and broad-band seismic observations from fibre-optic cables deployed offshore. In this study, we focus on 35.7 km of a linear telecommunication cable located offshore the Sanriku region, Japan, and apply seismic interferometry to obtain a high-resolution 2-D shear wave velocity (V_S) model below the cable. We first show that the processing steps applied to 13 d of continuous data prior to computing cross-correlation functions (CCFs) impact the modal content of surface waves. Continuous data pre-processed with 1-bit normalization allow us to retrieve dispersion images with high Scholte-wave energy between 0.5 and 5 Hz, whereas spatial aliasing dominates dispersion images above 3 Hz for non-1-bit CCFs. Moreover, the number of receiver channels considered to compute dispersion images also greatly affects the resolution of extracted surface-wave modes. To better understand the remarkably rich modal nature of OBDAS data (i.e. up to 30 higher modes in some regions), we simulate Scholte-wave dispersion curves for stepwise linear V_S gradient media. For soft marine sediments, simulations confirm that a large number of modes can be generated in gradient media. Based on pre-processing and theoretical considerations, we extract surface wave dispersion curves from 1-bit CCFs spanning over 400 channels (i.e. ∼2 km) along the array and invert them to image the subsurface. The 2-D velocity profile generally exhibits slow shear wave velocities near the ocean floor that gradually increase with depth. Lateral variations are also observed. Flat bathymetry regions, where sediments tend to accumulate, reveal a larger number of Scholte-wave modes and lower shallow velocity layers than regions with steeper bathymetry. We also compare and discuss the velocity model with that from a previous study and finally discuss the combined effect of bathymetry and shallow V_S layers on earthquake wavefields. Our results provide new constraints on the shallow submarine structure in the area and further demonstrate the potential of OBDAS for high-resolution offshore geophysical prospecting.

Key words: Japan; Tomography; Scholte waves; Surface wave.

1 INTRODUCTION

More than 70 per cent of the Earth's surface is covered by oceans. Yet, the offshore Earth's structure remains greatly underexplored due to the high-pressure conditions and high costs involved in deploying seismic instruments on the ocean floor. Offshore cabled seismic arrays, such as DONET and S-net in Japan (Aoi *et al*. 2020), have recently been installed in highly seismically active regions and transmit their data in real-time via fibre-optic cables. While such networks can greatly contribute to earthquake early warning systems by detecting earthquakes seconds earlier than traditional onshore networks, their sensor's density is too low to provide high-resolutions images of the shallow subsurface. Nevertheless, detailed velocity models can provide constraints on local wave amplification effects and thus improve earthquake early warning systems by reducing the risk of overestimating the magnitude of earthquakes during real-time analyses (Nakamura *et al*. 2014). In this context, distributed acoustic sensing (DAS) appears as an enticing alternative to study the physical properties of Earth near the coast with an unprecedented spatial resolution by taking advantage of existing cabled networks.

DAS is a rapidly evolving technology in geophysics that turns standard optical fibres into seismic arrays measuring the Earth's vibrations over tens of kilometres with a spatial density of the order of a metre. The DAS technology uses an optoelectrical interrogator to probe fibre-optic cables with repeated laser pulses. Some of the light is reflected to the interrogator (i.e. backscattered Rayleigh light) as each laser pulse travels down the fibre. External forcing, such as seismic waves, generate phase shifts of the backscattered Rayleigh light, which are measured by the interrogator. Depending on the manufacturing design, the phase shifts are converted to longitudinal strain or strain-rate time-series along the axis of the cable over a sliding spatial distance (i.e. the gauge length). For an extensive review of the DAS technology, we refer the reader to Hartog (2017).

Millions of kilometres of fibre-optic cables have been deployed around the world over the past decades to support our modern telecommunication network. Many of these fibres are located offshore and could therefore compensate for the scarcity of seismic stations deployed on the ocean floor. Ocean bottom DAS (OBDAS) has recently emerged as a promising method to detect and monitor a multitude of physical marine phenomena. For example, OBDAS has been used to monitor the spatial evolution of near-coast microseisms (Guerin *et al*. 2022; Xiao *et al*. 2022; Lindsey *et al*. 2019), provide high-fidelity records of regional and teleseismic earthquake wavefields (Shinohara *et al*. 2019; Lior *et al*. 2021; Spica *et al*. 2022), detect T-phases and other acoustic waves (Rivet *et al*. 2021; Ugalde *et al*. 2021; Spica *et al*. 2022) and track ocean surface gravity waves and deep-ocean water mixing processes (Ide *et al*. 2021; Williams *et al*. 2022).

OBDAS measurements also offer new possibilities for imaging marine sediments at a spatial resolution and an extent previously unattainable with traditional passive seismic surveys (Spica *et al*. 2020; Williams *et al*. 2021; Cheng *et al*. 2021; Fukushima *et al*. 2022; Lior *et al*. 2022; Spica *et al*. 2022; Tonegawa *et al*. 2022). Moreover, Karrenbach *et al*. (2020) and Matsumoto *et al*. (2021) demonstrated that OBDAS can record active source shots with a fidelity similar to that of ocean-bottom seismometers. The success of these early studies can be attributed to the fact that travel-time based analysis, using earthquakes or the ambient seismic field (ASF), should theoretically yield the same results for both DAS and geophone-equivalent data (Zeng *et al*. 2017; Wang *et al*. 2018; Nayak *et al*. 2021). Therefore, traditional imaging techniques based on surface wave travel times are readily applicable to the OBDAS data sets; although they need to be adapted to the exceptional spatial sampling of OBDAS data.

Seismic interferometry is a well-established method to extract the seismic wave propagation between two sensors. Under certain conditions, cross-correlating ASF time-series recorded at two seismometers yields the elastodynamic response of the Earth between these stations (e.g. Shapiro & Campillo 2004). The surface wave component of the cross-correlations functions (CCFs) is generally better retrieved as the ASF sources mostly generate surface waves and seismometers are located on the Earth's surface (e.g. Spica *et al*. 2017). The dispersive properties of surface waves from CCFs have been extensively used to image the crust and the uppermost mantle worldwide (Sabra *et al*. 2005; Yao *et al*. 2006; Lin *et al*. 2007; Nishida *et al*. 2008; Stehly *et al*. 2009; Spica *et al*. 2016; Castellanos *et al*. 2018). During the last decades, dense offshore arrays have been deployed for seismic exploration purposes and ASF tomography has been applied to obtain high-resolution images of the submarine shallow structure (Bussat & Kugler 2011; de Ridder & Dellinger 2011; Mordret *et al*. 2013, 2014). While very successful, these studies primarily focused on the fundamental and first-higher modes of surface waves. Nevertheless, retrieving higher-order surface wave modes in marine environments could greatly enhance the resolution of these models and better constrain the deeper structure (e.g. Aki & Richards 2002; Socco & Strobbia 2004; Perton *et al*. 2019).

In this study, we retrieve multimode Scholte waves from the ASF recorded along a fibre-optic cable located offshore the Sanriku coast, Japan, using seismic interferometry. We first discuss the effect of data pre-processing on the retrieval of accurate dispersion images and perform numerical simulations to better understand the nature of the large number of surface wave modes (over 30 modes in some sections of the cable). Based on data processing and theoretical considerations, we then perform a multimode inversion to constrain the shallow shear wave velocity (V_S) structure along the cable. Finally, we investigate the stability of the results, compare the inverted 2-D model with that from a previous study, and discuss the impact of the subsurface on earthquake wavefields.

2 DATA

2.1 The Sanriku cable

A 120-km-long submarine cable offshore the Sanriku coast was installed in 1996 to record real-time data from an ocean-bottom observatory composed of three ocean-bottom accelerometers and two pressure gauges (Fig. 1; Kanazawa & Hasegawa 1997; Shinohara *et al*. 2021, 2022). An AP Sensing N5200A interrogator unit with a 70-km sensing range (Cedilnik *et al*. 2019) was used to record strain data with a 500 Hz sampling rate from 18 November to 1 December 2019. The gauge length was set to 40 m and the spatial sampling to 5.1 m. Following the cable installation report, the first 47.7 km of the cable from the landing station (i.e. the first ∼9350 channels) are buried under 60–70 cm of sediments below the ocean floor. The channel positions were recently precisely located using active sources (Takano *et al*. 2021) and we focus on the data recorded between channel numbers 1500 and 8500, corresponding to 7.65 and 43.35 km from the coast, respectively. The analysed section of the cable is relatively straight, perpendicular to the coast, and buried, which should guarantee a good coupling with the surrounding sediments. More details about the cable setup and measurement quality can be found in Shinohara *et al*. (2022).

3 METHODS

3.1 Cross-correlation functions and dispersion curves

For each channel, the 13-d data set is windowed into 20-min time-series, detrended, demeaned, bandpass filtered between 0.01 and 5 Hz (a four-pole two-pass Butterworth bandpass filter is used for all filtering operations), and downsampled to 10 Hz. We also process the data with and without 1-bit normalization (Bensen *et al*. 2007) to investigate the impact of this non-linear operation on the retrieval surface wave higher modes. One-bit normalization sets the positive and negative amplitudes of continuous signals to 1 and –1, respectively, and is used to the reduce the impact of transient signals (e.g. earthquakes) on the computed CCFs. We set up virtual sources every 10 channels (i.e. every 51 m) between channels 1500 and 8000 and compute CCFs with the following 500 channels from each virtual source (i.e. channels 2000–2500 for virtual source 2000). For each station pair, the CCFs are computed for each 20-min window

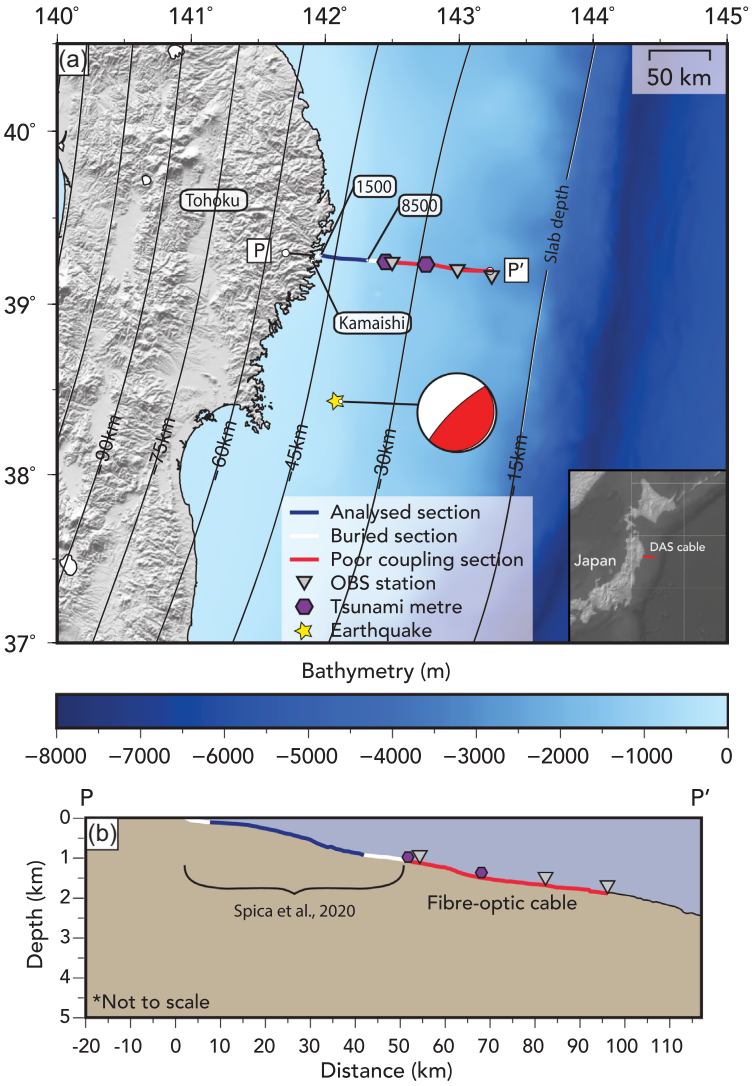

Figure 1. (a) Map showing the location of the Sanriku fibre-optic cable offshore the city of Kamaishi in Japan. The fibre sustains three ocean-bottom accelerometers (inverted grey triangles) and two pressure gauges (purple hexagons). The blue section of the cable highlights the buried section analysed in this study between channels 1500 and 8500. The white/blue and red sections of the cable depict areas where the fibre is buried and where it lays on the ocean floor by gravity, respectively. The subvertical solid lines show the subducting slab iso-depths (Hayes *et al*. 2018). The location of the M_w 3.7 earthquake discussed in Fig. 10 is highlighted by a star and its moment tensor. The inset map shows the Japanese Islands and the location of the cable in red. (b) Bathymetry profile along P-P' shown in (a). The location of the tomography performed by Spica *et al*. (2020) is also indicated.

in the frequency domain as

$$\mathrm{CCF}_{v-r}(t) = \mathfrak{F}^{-1}\left(\frac{\hat{s}_r \hat{s}_v^*}{\{|\hat{s}_v|\}\{|\hat{s}_r|\}}\right), \quad (1)$$

where $\hat{s}_v$ and $\hat{s}_r$ are the Fourier transform of 20-min strain records at the virtual source (s_v) and the receiver channel (s_r), respectively. The $*$ symbol represents the complex conjugate. Spectral whitening is applied and is represented by the denominator term of eq. (1) (i.e. $\{|\hat{s}_v|\}\{|\hat{s}_r|\}$), where $\{\cdot\}$ represents a smoothing of the absolute amplitude spectrum ($|\cdot|$) using a running-mean average algorithm over 30 discrete frequency samples (Bensen *et al*. 2007). The inverse Fourier transform ($\mathfrak{F}^{-1}$) is finally applied to retrieve the 20-min CCFs in the time domain.

Over the 13 d of continuous data, the 20-min CCFs are stable and exhibit a high signal-to-noise ratio (Fig. S1). For each channel pair, we stack 934 20-min CCFs calculated over the 13-d data set using a Phase-Weighed Stack method (Schimmel & Paulssen 1997) with a power of 2 and a smoothing of 0.5 s. The stacked CCFs are finally bandpass filtered between 0.25 and 5 Hz to remove the contribution of ocean surface gravity waves, which dominate the CCFs at frequencies below 0.25 Hz (Spica *et al*. 2020). In Fig. 2(a), we show an example of the CCFs calculated after applying 1-bit normalization to the continuous data recorded by channel 4000 (virtual source) and the following 400 channels. Clear propagating seismic waves can be observed in both the anticausal (negative) and causal (positive) parts of the CCFs. In this study, we only analyse the causal part of the CCFs (e.g. oceanward propagation), which has a stronger signal-to-noise ratio.

Martin *et al*. (2021) theoretically demonstrated that CCFs computed between strain data measured along the axis of the cable are mostly sensitive to surface waves travelling longitudinally to the fibre (e.g. Rayleigh and Scholte waves). Offshore Sanriku, we retrieve Scholte waves in the 0.25–5 Hz frequency band that are probably generated by wave–wave interactions of ocean waves travelling in

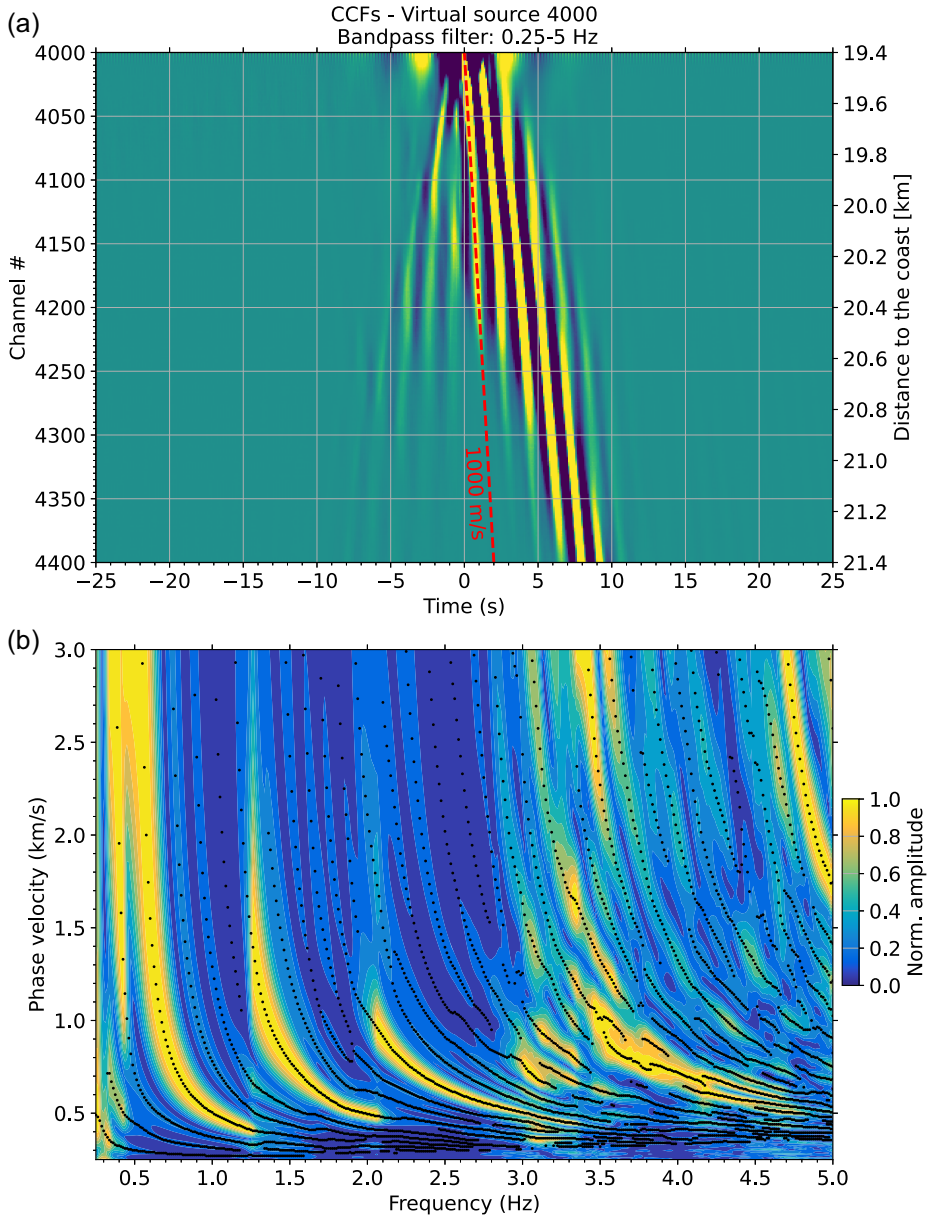

Figure 2. (a) Stacked CCFs computed after applying 1-bit normalization to the continuous data recorded by all channels between virtual source 4000 and channel 4400. The dashed red line highlights the 1000 m s^{-1} phase velocity moveout. All the waveforms are bandpass filtered between 0.25 and 5 Hz. (b) Dispersion image obtained from the causal part of the CCFs shown in (a). The black dots show the selected phase velocity dispersion points used to perform the inversion.

opposite directions (Longuet-Higgins 1950). As ocean waves tend to travel perpendicularly to the coast, ocean wave–wave interactions should generate Scholte waves that travel along the axis of the cable (i.e. perpendicularly to the coast). Therefore, the sensitivity of DAS to surface waves propagating along the axis of the cable combined with the potential generation mechanism of the Scholte waves indicates that the computed CCFs should capture the surface waves that propagate along the axis of the cable. In addition, the 20-min CCFs computed over the 13 d of data are stable (Fig. S1) and the stack of 20-min CCFs over 3 d yields similar waveforms as that obtained by stacking the 20-min CCFs over the entire data set (Fig. S2). This stability hints that potential non-physical arrivals due to the rapid changes of location of noise sources (Xiao *et al*. 2022) should be minimized. Therefore, we hypothesize that the computed Scholte waves capture the true surface wave propagation, but future work outside the scope of this study should be performed to confirm this hypothesis by fully characterizing the directionality of noise sources. We further discuss the implication of noise sources not aligned with the cable on inverted velocity models in Section 4.2.

We use a frequency-domain slant-stack method to compute dispersion images from the stacked CCFs (Park *et al*. 1998). A Fourier transform is first applied to the causal part of the time domain CCFs to retrieve the corresponding frequency-offset domain representation. Then, a slant-stack algorithm is used to retrieve the phase dispersion spectrum. In Fig. 2(b), we show the dispersion image computed from the 1-bit CCFs displayed in Fig. 2(a). Multiple

dispersive features (i.e. decreasing phase velocity with increasing frequency features in Fig. 2b) can be observed for this section of the cable between 0.5 and 5 Hz.

Most ambient noise tomography studies primarily focus on the surface wave fundamental mode, which is assumed to be the most energetic and to exhibit a global maximum at each frequency. In this study, dispersion images are exceptionally rich in high-energy higher modes. To take advantage of all possible modes and to avoid incorrect mode assignments, we follow the method proposed by Spica *et al*. (2018) and select the local maximum energy at all frequencies rather than the global maximum energy (Fig. 2b). To confirm that the selected dispersion points are stable and therefore physical, we show that dispersion points obtained from a 3-d stack of the 20-min CCFs yield similar results as those obtained from the 20-min CCF stack over the entire data set (Fig. S2c). Nevertheless, the phase velocity of some selected points strictly increases with increasing frequencies and are likely to be artefacts. In Sections 3.2.1 and 3.3, we further discuss the nature of such artefacts and present a way to refine and improve the picking process of dispersion points.

3.2 Data processing considerations

3.2.1 On the effect of 1-bit normalization

Pre-processing continuous DAS data with 1-bit normalization prior to computing CCFs is critical for retrieving dispersion curves (DCs) at high frequencies. In Fig. 3, we show the causal part of the CCFs computed with and without applying 1-bit normalization to the continuous data recorded by all channels between virtual source 4000 and channel 4400, as well as their respective dispersion images. In the time domain, CCFs computed with 1-bit normalization exhibit clearer propagating seismic waves than those computed without it (Figs 3a and b). Their respective dispersion images and extracted dispersion points are very similar for frequencies below 2 Hz, but clear differences can be observed at higher frequencies (Figs 3c and d). The dispersion image obtained from non-1-bit CCFs exhibits high-energy straight lines that increase with increasing frequency (Fig. 3d). These lines are caused by spatial aliasing and appear at high frequency where the incoherent noise energy is stronger than the surface wave energy. In this study, spatial aliasing obeys the relationship

$$c_i = \frac{\Delta x \cdot n_s}{0.5 + i} f, \tag{2}$$

where c_i, Δx, n_s and f are the phase velocity of the ith aliasing line, channel spacing (i.e. 5.1 m), number of channels used in the slant-stack analysis (i.e. array aperture) and frequency, respectively. i is a positive integer that defines the slope of spatial aliasing lines. Eq. (2) is very similar to that from Dai *et al*. (2018, their eq. 7), except that the slope of spatial aliasing lines depends on the array aperture used in the slant-stack analysis (Fig. S3). Nevertheless, applying 1-bit normalization to continuous DAS data generally allows us to retrieve surface wave higher modes with significant energy at high frequencies and reduces the effect of spatial aliasing.

3.2.2 On the effect of the number of receivers

The number of receivers used in the slant-stack analysis impacts the spectral resolution of dispersion images. More specifically, surface wave modes cannot be properly separated if the array aperture is too small (e.g. Foti *et al*. 2015). In Fig. 4, we show the dispersion images computed from 1-bit CCFs by considering 200, 300, 400 and 500 receivers, which corresponds to array apertures of 1020, 1530, 2040 and 2550 m, respectively. Note that we change the virtual source location to ensure that the middle point of each set of CCFs remains the same (e.g. channel 3000 in Fig. 4). The resolution of dispersion images and the number of DCs significantly increase with respect to the number of receivers. For 200 channels, four coarse modes appear between 0.5 and 5 Hz. By broadening the array aperture, we are able to better separate the higher modes and retrieve sharper dispersive features with many local maxima up to 5 Hz.

In Fig. 4(b), we show the selected dispersion points for 300, 400 and 500 receivers. For velocities faster than 1 km s^{-1}, we observe that the modal content slightly increases between 300 and 400 receiver channels, but remains almost stable between 400 and 500 channels. This suggests that we converge towards a reliable representation of the deep structure, which generally exhibits less lateral variations than shallow layers. For velocities slower than 1 km s^{-1}, the broadening of the array aperture from 300 to 500 channels increases the modal content, but comes at a cost of smoothing potential lateral velocity variations (e.g. Foti *et al*. 2015). Therefore, we select 400 receivers as a trade-off between mode separation and lateral resolution.

3.3 On the variability of dispersion images along the array

Based on the analysis performed in Section 3.2, we compute dispersion images from 1-bit CCFs over 400 receivers. In Fig. 5, we show a subset of the dispersion images computed at 14 different positions along the cable. We observe significant differences between the dispersion images, especially in terms of modal content. Between channels 1500 and 4000, the modal content is particularly rich. However, spatial aliasing appears after virtual source 5000 and is the strongest at virtual source 5500 despite the fact that CCFs are computed with 1-bit normalization. Yet, clear modes can still be observed below 3 Hz for all the virtual sources and provides confidence in the coupling of the cable with the surrounding sediments. We further discuss the lateral variations of the dispersion image modal content in Section 4.3.

High-energy artefacts caused by spatial aliasing need to be removed from the selected local maximum energy points as they can bias the inverted velocity models. We reject selected points that increase with both velocity and frequency and show the effect of the selection process in Fig. S4 at three virtual sources. In Fig. 5, we only show the selected dispersion points used to perform the inversion. Note that despite the fact that spatial aliasing appears at some stations (i.e. virtual source 5500), clear selected dispersion points can be selected and used to better constrain the inversion.

3.4 Theoretical considerations: constant V_S gradient models and modal content

Two types of waves can propagate along a flat solid-fluid interface: leaky-Rayleigh and Scholte waves (Gusev *et al*. 1996; Zhu *et al*. 2004). The term ‘leaky’ is used as Rayleigh waves radiate energy not only in the solid but also in the fluid, which causes their fast attenuation with distance. For a hard solid medium (e.g. V_S in the solid is much higher than the acoustic fluid velocity), most of the energy is lost in leaky-Rayleigh waves and the presence of Scholte waves is limited to solid media presenting a gradient velocity (Glorieux *et al*. 2001). However, for a medium where V_S is much slower than the acoustic fluid velocity, leaky-Rayleigh waves disappear and Scholte waves can be used to accurately assess V_S in the solid

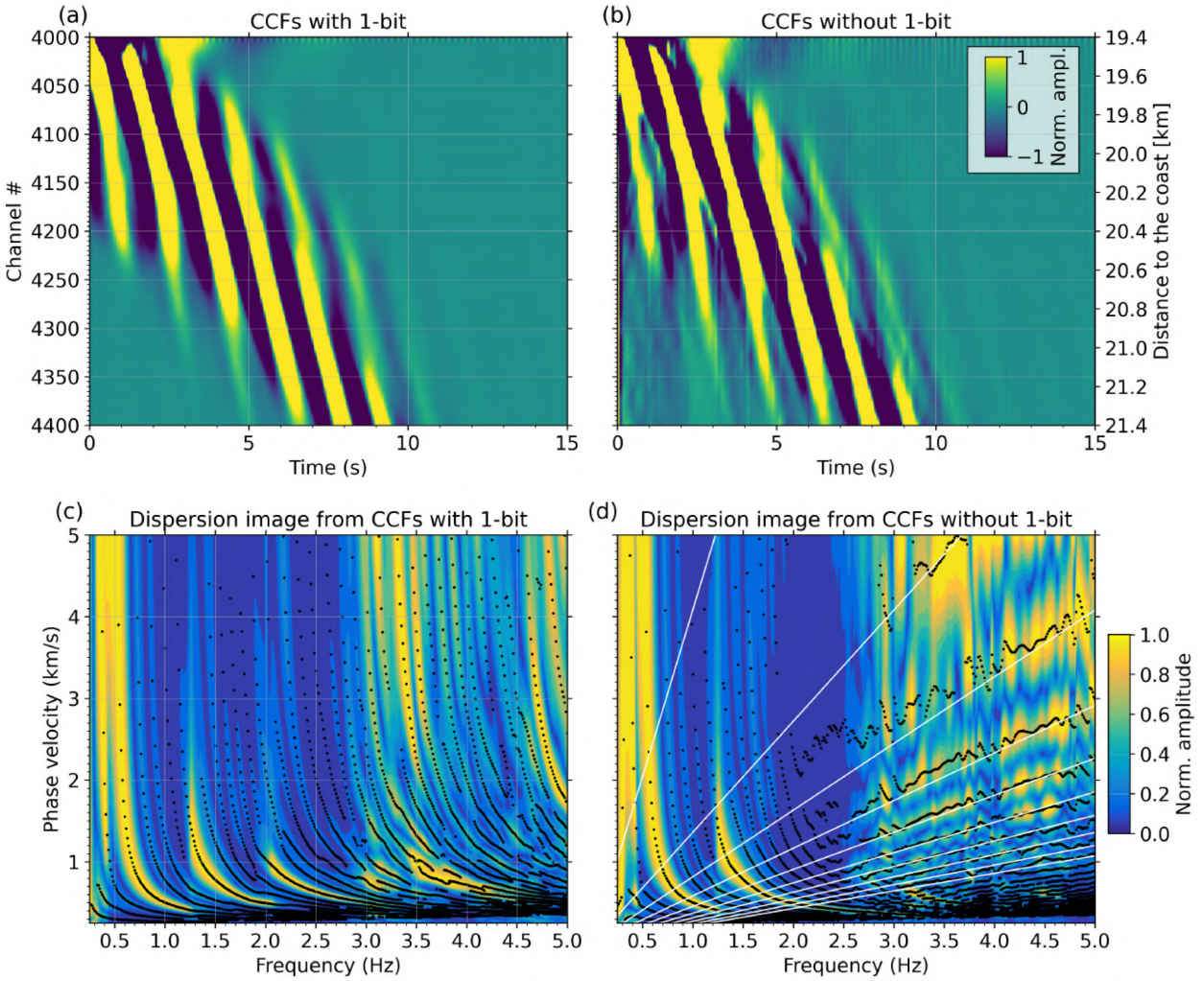

Figure 3. Stacked CCFs computed (a) with and (b) without applying 1-bit normalization to the continuous data recorded by all channels between virtual source 4000 and channel 4400. All the waveforms are bandpass filtered between 0.25 and 5 Hz. (c) Dispersion image obtained from the CCFs shown in (a) together with the selected phase dispersion points (black dots). (d) Same as (c) for the CCFs shown in (b). The white lines depicts the first 10 spatial aliasing lines (i.e. $i = 10$, eq. 2).

(Ali & Broadhead 1995). Sedimentary media, and more particularly marine sediments, are composed of soft materials which generally present a V_S gradient with depth and a near-null velocity at the top interface (Hamilton 1979). The effect of gradient media on the propagation of body waves has been studied in refraction seismology and body-wave traveltime studies (e.g. Stein & Wysession 2003). Yet, its effect on surface wave propagation is not as well constrained, especially in marine environments.

We present a simple theoretical case to illustrate the effect of constant V_S gradient media on the propagation of surface waves. While the nature of the subsurface generally differs from a constant velocity gradient, simulations can help us to better understand DAS observations and validate the approach that we propose. As we only consider media with very slow shear wave velocities in the shallow subsurface (i.e. $V_S \ll 1500$ m s^{-1}), we perform the simulations with a solid half-space layered gradient medium without a water layer. Nevertheless, we numerically confirm that the computed DCs are identical to those obtained with a solid half-space layered gradient medium overlaid by a water layer, but have a higher computational stability. This demonstrates that leaky-Rayleigh waves are not excited in the media considered for the simulations.

We consider two velocity models with different constant V_S gradients with depth (Fig. 6a). The constant V_S gradients are approximated by layered media in which each layer is homogeneous, which we refer to as stepwise linear gradient media in this study. The two stepwise linear gradients are defined as $\Delta V_S/\Delta z = 1$ s^{-1} and $\Delta V_S/\Delta z = 2$ s^{-1}, where ΔV_S and Δz are the S-wave velocity and depth differences between two contiguous layers, respectively. For the two stepwise linear gradient models, the layer thickness (h) is proportional to its depth (z) and is defined as $h \approx z/3$. The density and compressional-wave velocity (V_P) are obtained for each layer from V_S through empirical relationships (Berteussen 1977; Brocher 2005). Surface wave DCs for the two V_S stepwise linear gradient media are computed following the method introduced in Perton & Sánchez-Sesma (2016) and depict different behaviours (Fig. 6b). First, the number of modes is inversely proportional to the V_S gradient value of the stepwise linear models. Secondly, several parts of the DCs obtained with both gradients are superimposed. Thirdly, apparent DCs, which are constituted of a succession of osculation points from true DCs, appear in Fig. 6(b). We observe that the lower the gradient, the lower the velocity of the first apparent DC and the higher the number of apparent DCs for a fixed frequency range.

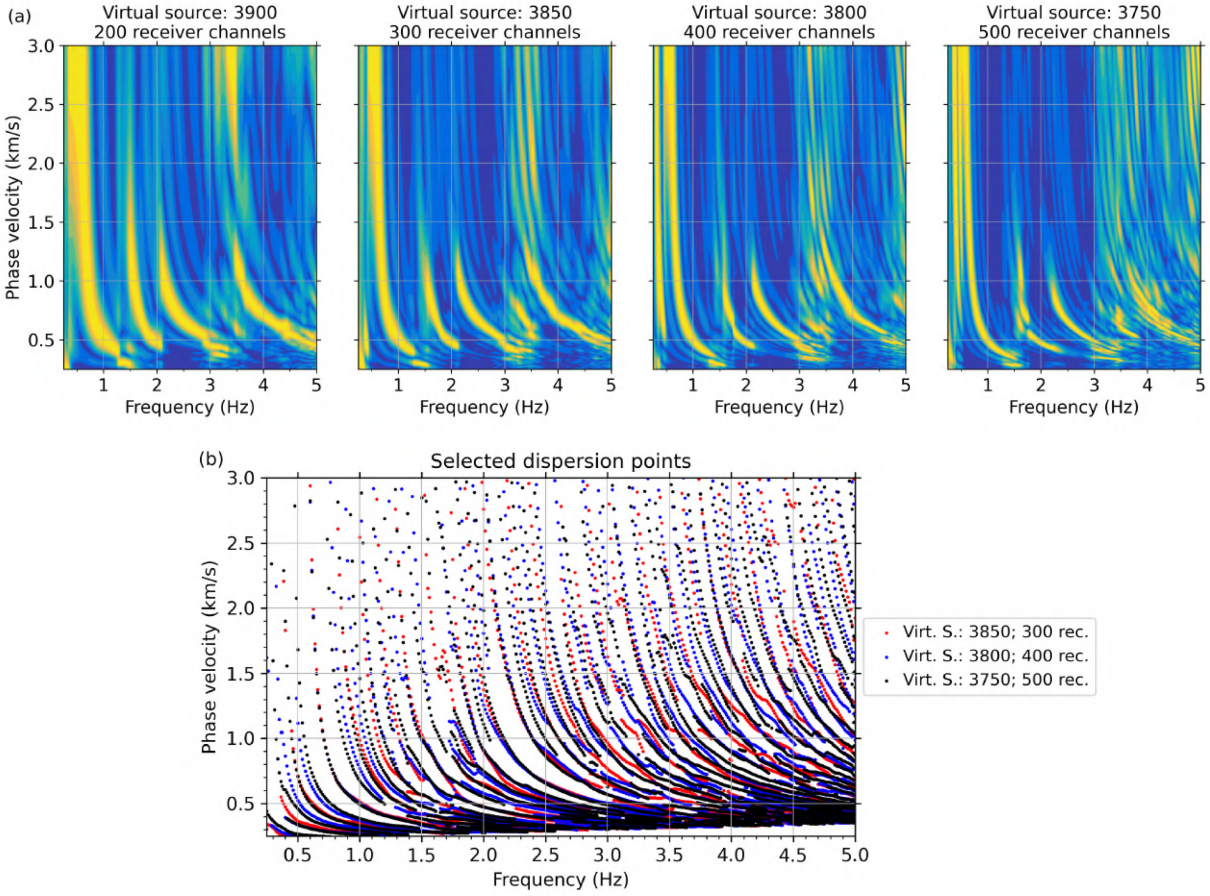

Figure 4. (a) Dispersion images computed from 1-bit stacked CCFs by considering 200, 300, 400 and 500 receivers. The virtual source is shifted so that the dispersion images focus on the same medium. (b) Selected dispersion points computed from the panels in (a) for 300, 400 and 500 receivers.

To better understand the partition of energy in dispersion images, we simulate a dispersion image from horizontal strain waveforms excited by a horizontal force (both along the axis of the cable) with a Discrete Wave Number method (Bouchon 2003) using a velocity gradient of 1 s^{-1} (Fig. 6c). This setup is equivalent to the wavefield recorded by DAS experiments with subhorizontal fibre-optic cables (Nakahara & Haney 2022). We observe that higher modes with high energy appear between 0.5 and 3 Hz and that apparent DCs do not hold any energy. The large number of modes combined with the complexity of dispersion images obtained from DAS data makes DC selection and identification difficult. Nevertheless, we minimize the risk of mode mis-identification by picking local energy maxima and not attributing the selected dispersion points to DCs.

We also investigate the energy partition in dispersion images computed from vertical displacement waveforms generated by a vertical source, which is the equivalent of a dispersion image that could be obtained from the vertical component of OBSs (Fig. S5). In this case, apparent DCs hold most of the energy and could easily be mistaken for true DCs. Therefore, special care is recommended when selecting DCs from CCFs computed from the vertical component of dense OBS networks, as mistaking apparent DCs for true DCs would undeniably bias inverted velocity models.

3.5 Multimode inversion scheme

3.5.1 Resolution of DAS dispersion images

The number of modes that can be retrieved from dispersion images depends on the array aperture (e.g. Section 3.2) and the velocity gradient in the medium (e.g. Section 3.4). Therefore, it is critical to observe a convergence of the number of modes in dispersion images to avoid finding a gradient value that is representative of the array aperture and not of the medium properties. Here, we observe a stable number of modes for phase velocities above 1 km s^{-1} by considering 400 receivers (Fig. 4b).

The slant-stack algorithm has a velocity resolution of $\Delta x/\Delta t$, where Δx and Δt are the channel spacing and temporal sampling of the CCFs, respectively (e.g. similar to the frequency–wavenumber (f–k) resolution described in Ventosa *et al*. 2012). In our experiment setting, the slowest velocity difference that can be resolved is 51 m s^{-1} as Δx and Δt are equal to 5.1 m and 0.1 s, respectively. In addition, we note that for a gradient velocity medium, the number of DCs significantly increases at slow velocities with increasing frequency (e.g. below 0.25 km s^{-1} in Fig. 6b). This leads to DCs that cannot be separated at high frequencies due to the limited resolution of the phase velocity discretization. To account for these limitations, we define a function to automatically reject dispersion

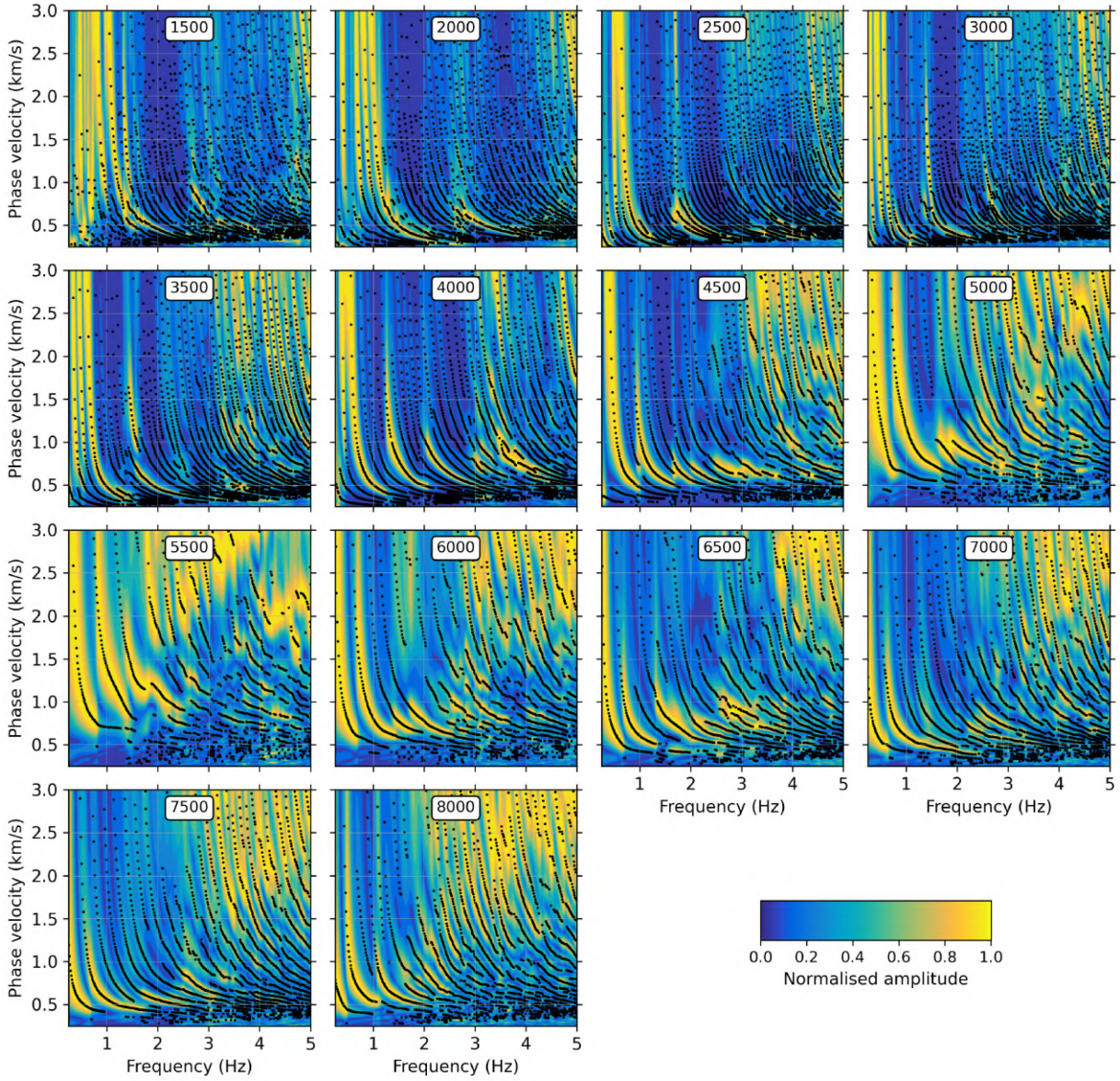

Figure 5. Dispersion images computed from 1-bit CCFs and by considering 400 receivers along the cable; every 500 virtual sources. Selected dispersion points, after rejecting spatial aliasing artefacts, are shown by the black dots.

points slower than a given phase velocity $c_{\min}$ as

$$c_{\min} = \begin{cases} 250 \text{ m s}^{-1} \text{ for } f <= 1 \text{ Hz} \\ (250 + 0.025(f-1)) \text{ m s}^{-1} \text{ for } f > 1 \text{ Hz}, \end{cases} \quad (3)$$

where f is the frequency. The effect of the data selection process is shown in Fig. 5, where no dispersion points are selected for velocities below $c_{\min}$.

3.5.2 *Objective function*

The selected dispersion points are considered independently as their identification into separated DCs is challenging, especially at high frequencies. A drawback of this approach is that the inversion of a large number of dispersion points can easily be biased. A popular misfit function used to perform multimode inversions is the (RMS) root mean square function (Perton *et al*. 2019). The RMS error (ϵ_{DC}) can be expressed as

$$\epsilon^2_{\text{DC}} = \frac{1}{J_{\max}} \sum_{f=f_{\min}}^{f_{\max}} \sum_{j=0}^{j_{\max}} \sum_{n=0}^{n_{\max}} G\left(\left|c_j^{\text{obs}}(f) - c_n^{\text{th}}(f)\right|\right)^2, \quad (4)$$

where n is the number of theoretical modes and j the number of observed dispersion points at a frequency f. $J_{\max}$ is the number of selected dispersion points in a specific frequency band [i.e. $J_{\max} = \sum_{f=f_{\min}}^{f_{\max}} j_{\max}(f)$]. c_n^{th} and c_j^{obs} are the phase velocities of the theoretical and observed DCs, respectively. G is a function that considers if a theoretical dispersion point matches a selected dispersion point and is defined as

$$G(x) = \begin{cases} x \text{ when } |x| \leq \delta \\ \delta \text{ when } |x| > \delta \end{cases}, \quad (5)$$

where δ is a threshold value that is equal to the mean velocity difference between the observed dispersion points. Yet, eq. (4) does not perform well for inverting a large number of selected dispersion

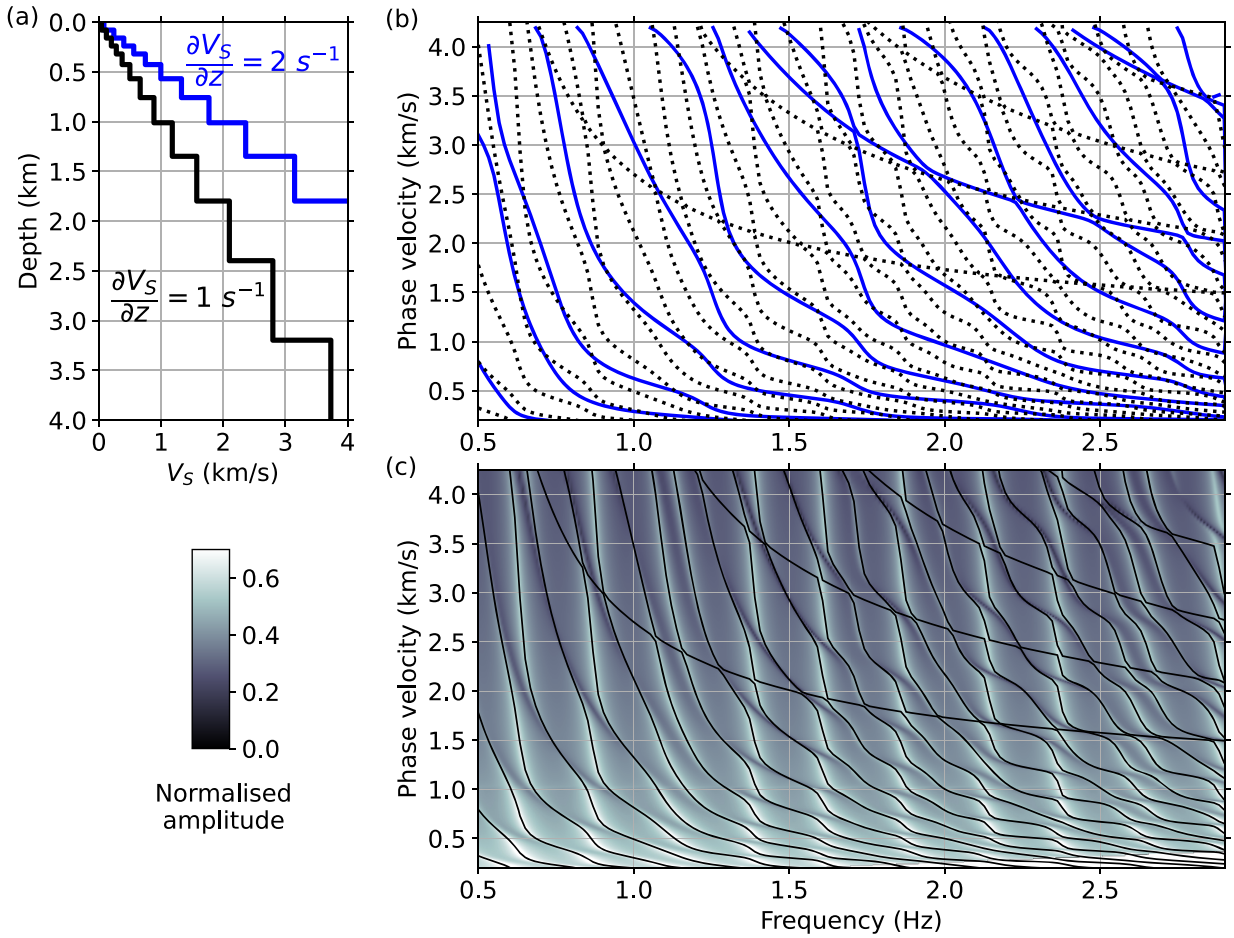

Figure 6. (a) Velocity models for two stepwise linear V_S gradients ($\Delta V_S/\Delta z = 1,\ 2\ \mathrm{s}^{-1}$). (b) Theoretical DCs for the two velocity models shown in (a) using the same colour code. (c) Theoretical dispersion image calculated from horizontal strain waveforms excited by an horizontal force and using the velocity model with a gradient of $\Delta V_S/\Delta z = 1\ \mathrm{s}^{-1}$. The amplitude of the energy is normalized between 0 and 1 and the DCs from (b) are shown by the black lines.

points. We systematically converge towards a low velocity gradient medium with a large number of inverted DCs that directly minimize the misfit function (Fig. S6).

To overcome this problem, we modify eq. (4) to penalize configurations that use more theoretical modes than the number of selected points at each frequency and in specific phase velocity ranges (e.g. $c_{\min}$ to $c_{\max}$). The phase velocity range is defined as the area where we confirmed the retrieval of all modes given the 2040 m array aperture. Note that $c_{\max}$ is fixed to 2 km s^{-1} as the slope of DCs above this value is too strong for dispersion points to be accurately resolved. We define $n_{\min}$ as the number of the first theoretical mode with a velocity above $c_{\min}$ and $n_{\max}$ as the number of the last theoretical mode with a phase velocity below $c_{\max}$. As the number of theoretical modes between $n_{\min}$ and $n_{\max}$ is $n_{\max} - n_{\min} + 1$ and the number of observed points is $j_{\max}$, we weight the misfit function with $|n_{\max} - n_{\min} + 1 - j_{\max}|$. The final misfit function is therefore defined as

$$\epsilon^2_{\mathrm{DC}} = \frac{1}{J_{\max}} \sum_{f=f_{\min}}^{f_{\max}} \sum_{j=0}^{j_{\max}} \sum_{n=0}^{n_{\max}} G \left(\left| c_j^{\mathrm{obs}}(f) - c_n^{\mathrm{th}}(f) \right| \right)^2 \left(1 + \frac{|n_{\max} - n_{\min} + 1 - j_{\max}|}{j_{\max}} \right). \quad (6)$$

This functional form adapts to the changing modal content and works well for all locations along the cable.

3.5.3 *Parametrization*

We start the inversion for the selected dispersion points at virtual source 3000. We use a constant gradient of $\Delta V_S/\Delta z = 0.8\ \mathrm{s}^{-1}$ as the starting velocity profile and consider the gradient value as the only free parameter. After obtaining the gradient value, we first invert for the thickness of each layer and then, separately, for their shear wave velocity. Since surface waves have a low sensitivity to density and V_P, we estimate them from V_S through empirical relationships (Berteussen 1977; Brocher 2005). Similarly to Perton *et al.* (2019), we use a constrained non-linear optimization procedure to minimize the misfit function (Byrd *et al.* 1999), where we impose the highest velocity to be in the half-space as it helps to compute more stable DCs. We then use the inverted velocity profile at channel 3000 as the input model for the neighbouring virtual sources and iteratively invert the selected dispersion points along the cable from virtual sources 1500 to 8000.

4 RESULTS AND DISCUSSION

4.1 On the reliability of the inversion

We show the 1-D inversion results at four virtual sources along the cable in Fig. 7. The inverted DCs fit well most of the selected points (Fig. 7a) and provide confidence in the inverted 1-D velocity

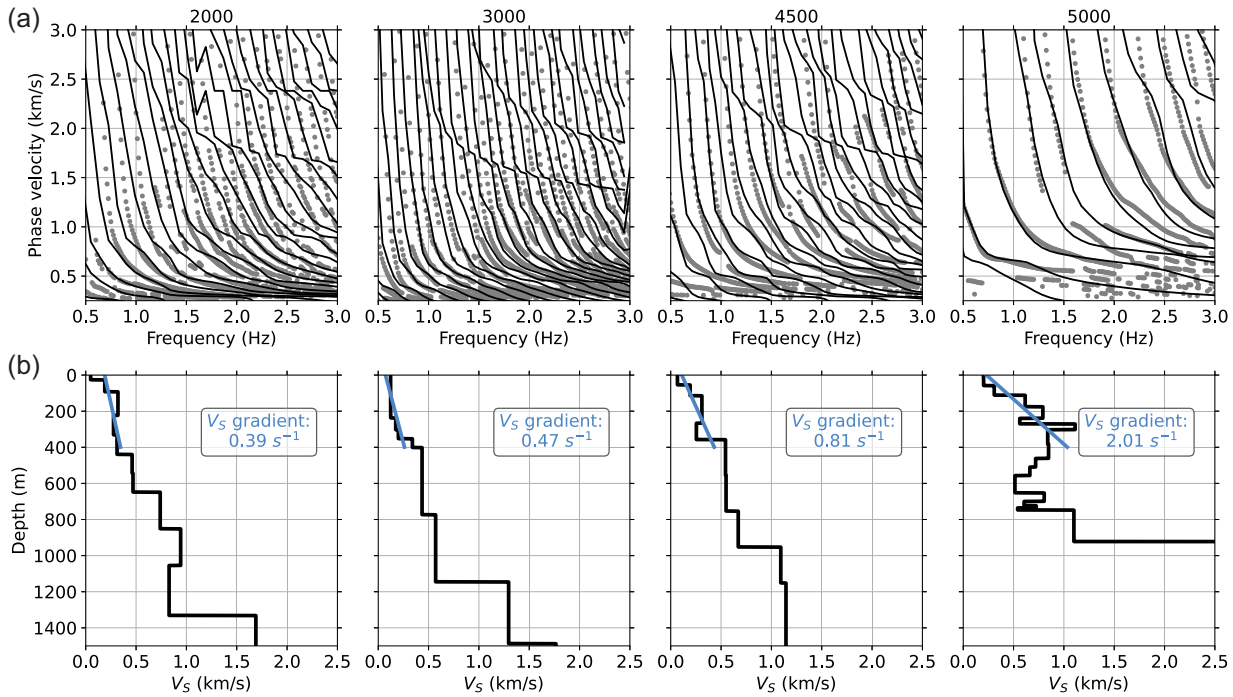

Figure 7. (a) Selected dispersion points (grey dots) and inverted DCs (black lines) for four sections along the cable. The virtual source number is indicated on top of each subplot. (b) Inverted V_S velocity model (black) for the four sections shown in (a). V_S gradient values and lines computed by fitting the inverted 1-D V_S models between the ocean floor and 400 m depth with a straight line are shown in blue.

models shown in Fig. 7(b). Apparent DCs appear in the fitted DCs for the four virtual sources, but are not visible in the selected dispersion points. This difference can be explained by the simulations performed in Section 3.4, where we show that apparent DCs do not hold much energy for DAS-like deformation dispersion images (Fig. 6c). Therefore, we do not expect the selected dispersion points to contain the signature of apparent DCs.

The misfit function minimized by eq. (6) does not allow us to estimate the reliability of the inversion along the array as the weights change with the number of fitted DCs. To provide misfit estimates that can be compared along the array, we compute misfit values (i.e. ϵ_{DC}) from the final inversion results using eq. (4) and show them in Fig. 8(a). We observe relatively constant misfit values along the cable with a mean value of 90 m s^{-1} and a one standard deviation to the mean of 23 m s^{-1}. The results from a few individual virtual sources have larger misfit values, but their inverted velocity models are consistent with the neighbouring ones, which demonstrates the stability of the velocity model along the array.

Finally, we also compute the Scholte-wave sensitivity kernels from the inverted 1-D velocity models at channels 3000 and 5000 (Fig. S7). For both locations, we observe some sensitivity down to a depth of approximately 2.5 km, with a better sensitivity to the shallow structure (e.g. $\leq$1 km). Therefore, we focus the rest of this study on the velocity structure shallower than 2.5 km from the ocean bottom.

4.2 Potential sources of bias

The CCFs used in this study to invert the velocity model are computed from 13 d of continuous data. While the computed CCFs are stable over 13 d (Figs S1 and S2) and the extracted surface waves likely travel along the axis of the cable due to the nature of their excitation (i.e. wave–wave interactions of ocean waves travelling perpendicularly to the coast), we cannot rule out the potential effect of persistent noise sources not aligned with the cable. In such a case, the surface wave arrivals of the different modes would exhibit apparent velocities different from those of the true surface waves, and lead to a biased V_S model. Nevertheless, this effect should be mitigated by the broadside insensitivity of DAS to seismic waves with particle motions oblique to the fibre axis. Indeed, Martin *et al.* (2021) showed that DAS measurements are primarily sensitive to Rayleigh waves travelling longitudinally along the fibre. This implies that the impact of noise sources not aligned with the axis of the cable should theoretically be reduced by the nature of DAS measurement. This could therefore be an advantage for noise tomography studies with linear cables (Shragge *et al.* 2021). Nevertheless, there are other factors contributing to the array sensitivity including the velocity structure, gauge length and analysed frequency range. Therefore, future work beyond the scope of this study should investigate the directivity of noise sources and their actual contribution to OBDAS CCFs computed along linear arrays.

Another source of bias could come from a non-optimal distribution of noise sources around the array leading to the summation of cross-modal terms that would introduce spurious modes in the CCFs (Halliday & Curtis 2008). Nevertheless, the stacking of the 20-min CCFs over 13 d should reduce the potential effect of cross-modal term (Snieder 2004) and noise sources are likely present all around the cable, which should lead to CCFs free of spurious modes (Kimman & Trampert 2010). Finally, Perton & Sánchez-Sesma (2016) showed that cross-terms theoretically cancel each other. Therefore, biases arising from the cross-correlation process should be limited.

The cable installation report indicates that the 35.7 km section of the cable analysed in this study is buried under 60–70 cm of sediments, which should result in a cable well coupled with the surrounding medium. While we retrieve clear DCs from the CCFs along the entire analysed section, there is no guarantee that the

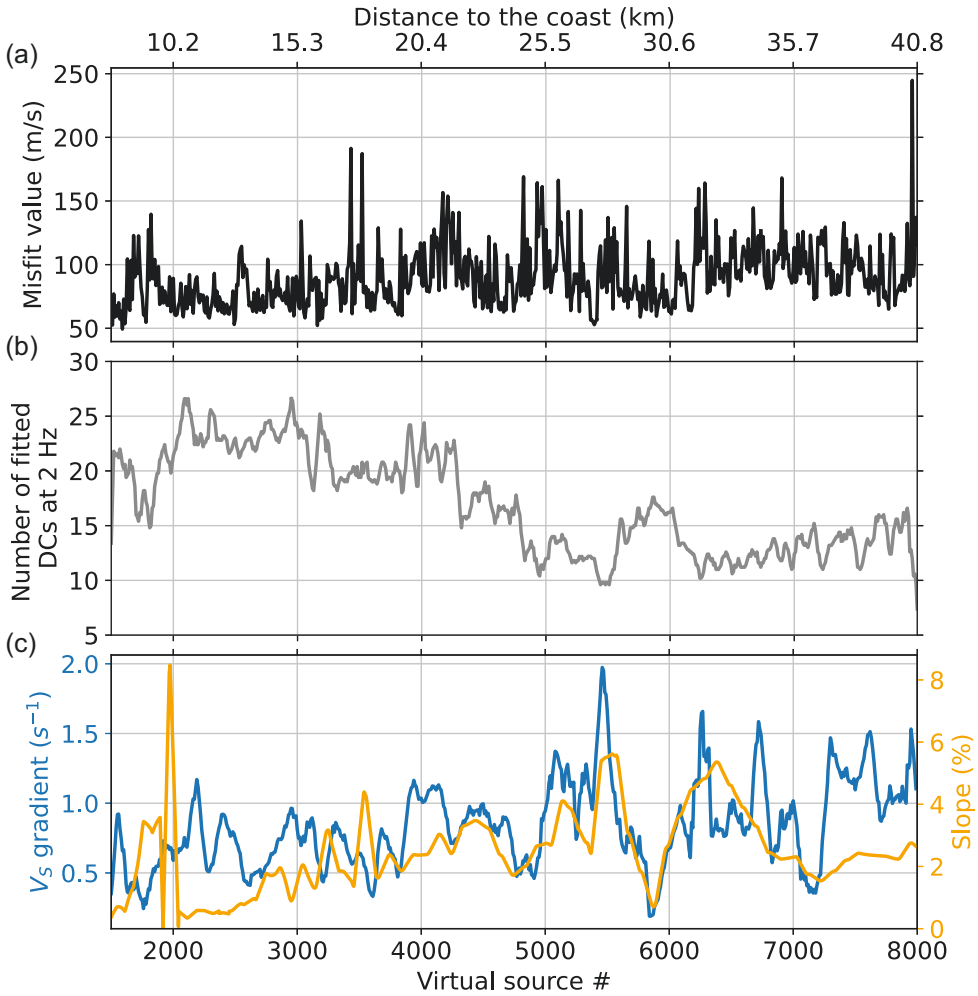

Figure 8. (a) Misfit value along the cable computed using eq. (4) from the final inverted DCs. (b) Number of fitted modes for each dispersion image at a frequency of 2 Hz. (c) V_S gradient value calculated between the ocean floor and 400 m depth (blue) and the slope of the bathymetry (orange).

cable is indeed buried and well coupled. As non-coupled fibres provide lower quality DAS data than buried fibres (Nap *et al*. 2020), future work should focus on determining the exact burial depth of the cable. For example, this could be performed by computing high-frequency autocorrelation functions of earthquake and/or ambient noise data to retrieve the *S* wave two-way traveltime between the cable and the seafloor.

4.3 2-D shear wave velocity model, bathymetry and their effect on the modal content

The final inverted 2-D V_S model is shown in Fig. 9(a). The model is smoothed using the median value of a 2-D sliding window (i.e. horizontal and vertical). The lengths of the smoothing window are 510 and 20 m in the horizontal and vertical directions, respectively. Along the array, V_S is generally relatively slow near the ocean bottom (e.g. $<$200 m s^{-1}) and increases with depth. A stiffer material with a velocity of 2500 m s^{-1} is observed at a depth of 1000–1500 m below the ocean floor.

The modal content of dispersion images significantly varies along the array. In Fig. 8(b), we show the number of fitted DCs at 2 Hz, which is a proxy for evaluating the modal content of dispersion images over the whole spectrum. We observe the largest number of fitted DCs between channels 2000 and 3000, with over 25 modes at some stations. The number of modes then decreases to approximately 10 modes near channel 5500. To better understand the impact of shallow layers on the modal content, we compute a V_S gradient value between the ocean floor and 400 m depth for each 1-D model. V_S gradients are defined as the slope of a straight line fitted to each V_S 1-D profile between the surface and 400 m depth. Examples of V_S gradient lines and values are shown in Fig. 7(b). In Fig. 8(c), we show V_S gradient values along the array after applying a lateral smoothing over ten 1-D models. V_S gradients vary between 0.3 and 1.2 s^{-1} between channels 1500 and 4500, peak to 2.0 s^{-1} near channel 5500, and finally vary between 0.3 and 1.6 s^{-1} until channel 8000. We observe an anticorrelation between the number of fitted DCs and V_S gradients. This shows that rapidly increasing velocities in the medium lead to dispersion images with less surface wave modes. This observation agrees with the theoretical results obtained in Section 3.4 for gradient media.

Changes of bathymetry can have a direct impact on the velocity structure, and therefore, on the modal content of dispersion images. To characterize this effect, we show the slope of the bathymetry in Fig. 8(c). The slope is defined as the ratio between the vertical

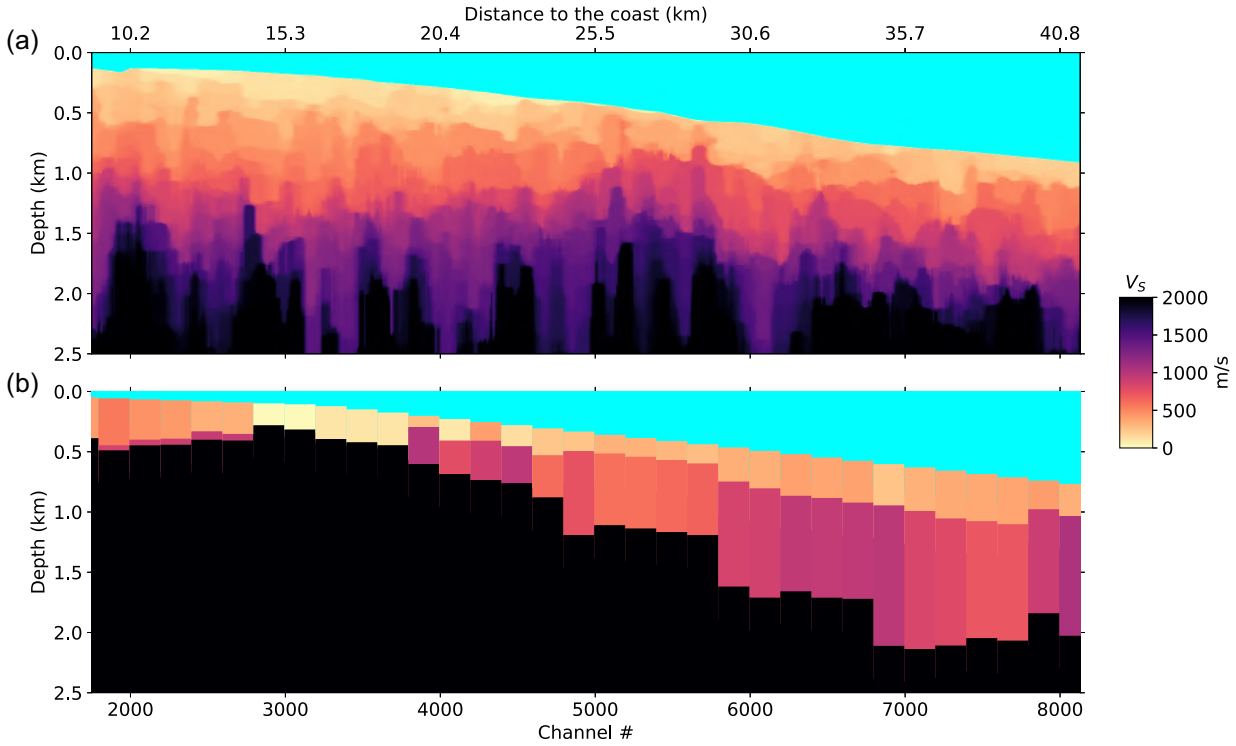

Figure 9. *S*-wave velocity model obtained (a) in this study and (b) by Spica *et al*. (2020). V_S is clipped at 2500 m s^{-1} for both models.

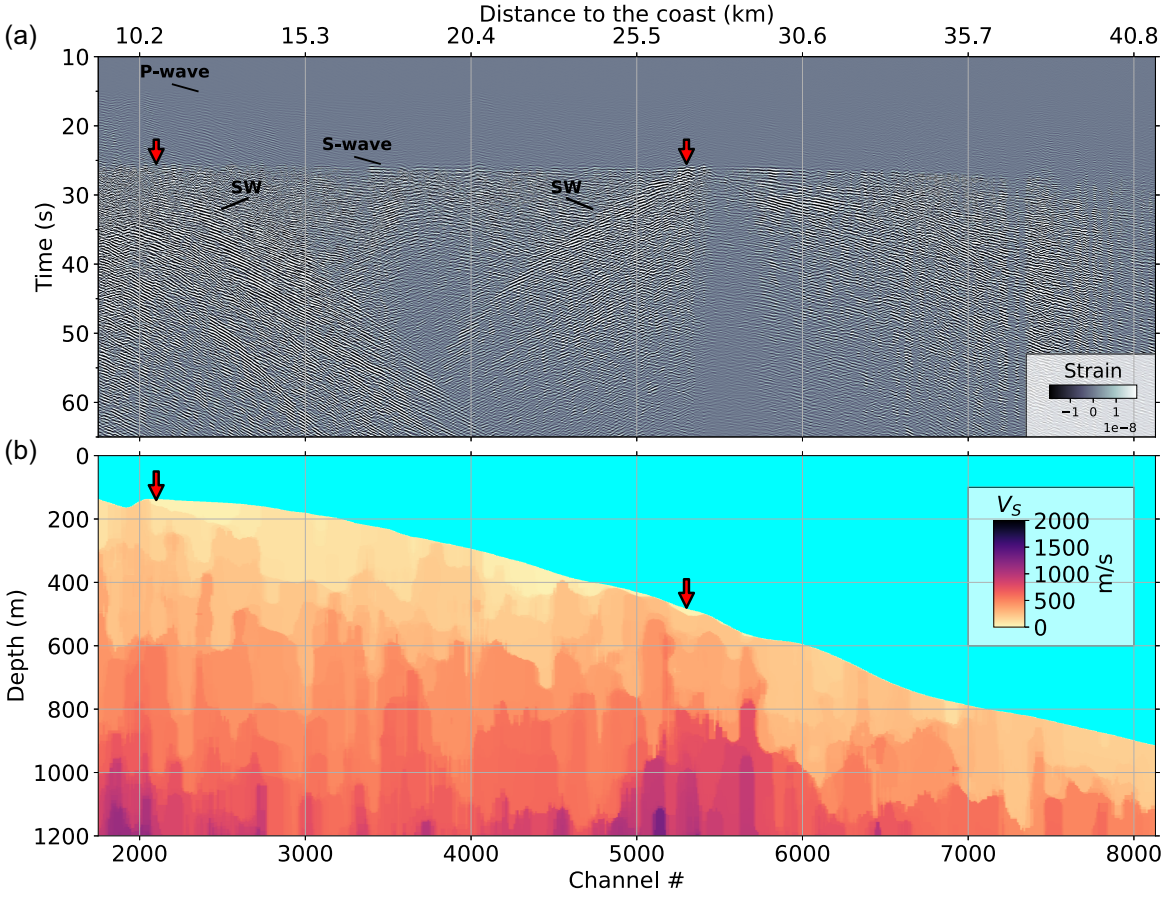

Figure 10. (a) Strain waveforms of a M_w 3.7 earthquake bandpass filtered between 2 and 8 Hz (location in Fig. 1). Direct *P* and *S* waves arrive around 15 and 25 s after the origin time, respectively. Regions where surface waves (SW) are locally generated are shown by the vertical red arrows. (b) Zoom on the shallow part of the inverted velocity model shown in Fig. 9(b). The red arrows correspond to the regions where surface waves are generated.

distance over a horizontal distance between two points, multiplied by 100 to make them percentages. In this study, the slope is computed along the array using a sliding horizontal distance of 408 m. We observe a good correlation between the slope of the bathymetry and V_S gradient values. Flat sections of the cable, where the slope is less than 4 per cent, generally coincide with slow V_S gradient values and more fitted DCs. Moreover, the region between channels 5000 and 5500, where spatial aliasing appears in the dispersion images, coincides with the region where the largest gradient values are observed. This demonstrates that the number of selected dispersion points and the presence/absence of spatial aliasing are closely related to the nature of the sediments in the shallow subsurface.

4.4 Comparison with another velocity model

In Fig. 9, we compare the inverted velocity model with that obtained by Spica *et al*. (2020), which was computed from data of the same fibre-optic cable. Both models exhibit soft ($V_S < 500$ m s^{-1}) sediment layers at shallow depth (<500 m below the ocean floor) and a stiff material ($V_S \geq 2000$ m s^{-1}) below 1–2 km from the ocean bottom. This first order observation agrees with the results from an active source reflection study conducted in the vicinity of the cable by Kodaira *et al*. (2017), which highlighted the presence of Pliocene–Pleistocene and Miocene sediments overlaying a Cretaceous bedrock at a depth of 1–2 km below the ocean floor. Between channels 5000 and 8400, both models show a similar shallow structure and a bedrock located approximately at the same depth (i.e. between 1 and 2 km below the ocean floor). Between channels 2500 and 4500, both models display a very slow V_S layer (∼200 m s^{-1}) at shallow depth (<400 m below the ocean floor). However, the main discrepancy between the two models resides in the depth of the bedrock between channels 1500 and 5000. While the model from this study shows a bedrock depth constantly deeper than 1 km below the ocean floor, the Spica *et al*. (2020) model displays a very shallow bedrock in this region. Nevertheless, the velocity model presented in this study is likely to be more representative of the structure below the cable for the reasons detailed below.

While the same fibre-optic cable is used by both studies, the two velocity models were obtained using different signal sources, processing steps and inversion schemes. The Spica *et al*. (2020) model was computed by inverting phase velocity DCs obtained from *f*–*k* power spectra sliding over the cable. Each *f*–*k* spectrum was computed over 600 channels (i.e. 3060 m), every 200 m, using 2 hr of continuous data, and DCs were extracted between 0.2 and 1 Hz. The first potential bias of the Spica *et al*. (2020) model comes from the fact that only 2 hr of data are used. Therefore, their extracted DCs are more likely to be sensitive to potential transient seismic signals not propagating along the axis of the fibre that would exhibit an apparent velocity. In our study, we use CCFs that favour the measurement of Scholte waves propagating perpendicularly to the coast (i.e. along the cable), which reduces the risk of measuring apparent velocities, as discussed in Section 4.2. Secondly, the larger aperture (i.e. 600 channels) considered in the Spica *et al*. (2020) study naturally provides a smoother velocity model with less lateral sensitivity compared to the 400 channel aperture used in our model. Thirdly, the inversion of the Spica *et al*. (2020) model was performed for the thickness and V_S of two layers overlying a half-space, and by only using the fundamental mode of Scholte waves between channels 1500 and 5000 and the fundamental and first-higher modes after channel 5000 at low frequencies (0.2 and 1 Hz). This inversion scheme provides another source of uncertainties as it is well established that surface wave DCs are sensitive to the absolute velocity in the medium and generally provide a non-unique solution of the layered structure (e.g. Scherbaum *et al*. 2003). Note that this is mainly caused by the broad sensitivity kernels of surface waves which sample a wide range of depths depending on their frequency. As surface wave inversions can be better constrained by adding different observables (e.g. Lin *et al*. 2012; Spica *et al*. 2017, 2018) or a larger number of higher modes sensitive to different depths (e.g. Aki & Richards 2002; Socco & Strobbia 2004; Perton *et al*. 2019), the large number of DCs inverted with the multimode inversion scheme presented in this study likely improves the accuracy of the inverted velocity model. More specifically, we retrieve and use a large number of surface wave modes with clear dispersive features up to phase velocities of 3 km s^{-1} before channel 5000, which provide better constraints on the velocity model down to depths around 2.5 km below the ocean floor. Finally, the similarity between the two velocity models after channel 5000 can be explained by the fact that both velocity structures are obtained by inverting multiple surface wave modes, which provides a better sensitivity at greater depths.

4.5 Effects on the earthquake wavefield

Numerous local and regional earthquakes were recorded during the 13-d period of the DAS experiment. In Fig. 10, we show the strain waveforms of a moment magnitude (M_w) 3.7 earthquake bandpass filtered between 2 and 8 Hz. The earthquake occurred on 21 November 2019, at 16:29:13 (UTC) at a depth of 33 km (National Research Institute for Earth Science and Disaster Prevention centroid moment tensor solution) and its location is shown in Fig. 1. The direct *P*-wave arrives around 15 s but is barely visible as it propagates nearly orthogonally to the cable (Martin *et al*. 2021). On the other hand, the direct *S* wave arrives between 25 and 30 s and can be observed at all channels. In addition, we also observe that strong surface waves are generated after the direct *S*-wave arrival near channels 2100 and 5200 and propagate along the array.

On the ocean bottom, surface waves can be excited locally by a variety of phenomena, including body-wave scattering caused by sharp changes of the bathymetry (Zheng *et al*. 2013), strong lateral heterogeneities such as fault zones (Sato *et al*. 2012) and water phase reverberations (Spica *et al*. 2022). Figs 10 and S8 show that surface waves are generally excited at the same two locations during earthquakes (i.e. near channels 2100 and 5200). These two locations are characterized by very soft and shallow V_S layers and relatively sharp bathymetry changes. Therefore, a combined effect of soft and shallow layers and bathymetry changes is likely responsible for the generation of surface waves in these two regions. We also note that the amplitude of later arrivals (i.e. after the direct *S* wave) are small near channel 5500 compared to that near channel 3000. This can be explained by the relatively high V_S layers in the shallow subsurface below channel 5500 (Fig. 10b), which do not have the potential to trap and amplify incoming high-frequency seismic waves. In contrast, the ground motions near channel 3000, where the structure is characterized by very slow V_S layers over the first 200 m, exhibit larger amplitudes that last beyond 50 s after the *P*-wave arrival. This strong and lasting wave amplification can be explained by seismic waves trapped in the shallow structure.

5 CONCLUSIONS

We retrieved surface waves by cross-correlating continuous strain signals recorded along a fibre-optic cable offshore the Sanriku

Coast, Japan. We first analysed the effect of data pre-processing on the retrieval of dispersion images and concluded that computing 1-bit CCFs and considering 400 receiver channels (i.e. 2040 m array aperture) offers a good trade-off between mode separation and lateral spatial resolution. We then presented a theoretical case to shade some lights on the complexity of dispersion images computed from DAS data using stepwise linear V_S gradient media. Our simulations confirmed that gradient media, which are representative of marine sediments, can generate a large number of surface wave modes. Based on data processing and theoretical considerations, we inverted surface wave DCs to provide a 2-D model of the V_S structure. We found that shallow low-velocity layers combine with bathymetry changes can greatly impact the modal content of surface waves retrieved from CCFs. We finally compared the inverted V_S model with that from Spica *et al*. (2020) and discussed the effect of the shallow structure on the propagation of earthquake seismic waves.

OBDAS provides a unique opportunity to image marine sediments with an unprecedented spatial resolution and to better understand the seismic wavefield. This can be critical for OBDAS experiments that are being considered to be part of earthquake early warning systems in subduction zones. While there is no doubt that earthquakes can be detected by OBDAS arrays, rapid magnitude estimation is likely to be more challenging due to the complexity of the recorded wavefield. Nevertheless, imaging the shallow subsurface beneath fibre-optic cables can help us reducing such uncertainties by providing better constraints on local seismic wave amplifications.

Finally, the methodology presented in this study is readily applicable to onshore metropolitan areas characterized by shallow and low V_S sediments that can significantly amplify earthquake ground motions, such as Jakarta, Los Angeles, Mexico City, Taipei and Tokyo. Nevertheless, future work should also include a full characterization of the distribution of ambient noise sources and analyse cable coupling conditions to reduce potential sources of bias on inverted velocity models.

ACKNOWLEDGMENTS

We thank Takeshi Akuhara for providing useful information about the targeted area and DAS measurements. We thank *Fujitsu* for cooperating with the Earthquake Research Institute (ERI), The University of Tokyo, for the DAS measurement campaigns. We are grateful to the Editor, Cornelis Weemstra and two anonymous reviewers for their constructive comments that helped us improve this study. All the figures are plotted with Matplotlib (Hunter 2007) or the Generic Mapping Tool (Wessel *et al*. 2019). Some of the data processing steps have been performed using ObsPy (Beyreuther *et al*. 2010).

Funding: This project was partly supported by the discretionary budget of the director of ERI. The observations were carried out as part of the Earthquake and Volcano Hazards Observation and Research Program by the Ministry of Education, Culture, Sports, Science, and Technology of Japan. LV was supported by NSF award EAR2022716 and MP is supported by CONACYT award CF-2019-G-6655. Evaluación de alta resolución de las propiedades sísmicas en la Ciudad de México. ZJS acknowledges support from the Air Force Research Laboratory grant FA9453-21-2-0018.

Competing interests: The authors declare that they have no competing interests.

DATA AVAILABILITY

The codes and data to reproduce most figures of the paper and the 2-D velocity model are available at https://github.com/lviens/2022_DAS_Tomography.

REFERENCES

Aki, K. & Richards, P., 2002. *Quantitative Seismology,* University Science Books.

Ali, B.H. & Broadhead, M.K., 1995. *Full Field Inversion Methods in Ocean- and Seismo-Acoustics,* Kluwer Academic.

Aoi, S. *et al.*, 2020. Mowlas: Nied observation network for earthquake, tsunami and volcano, *Earth, Planets Space,* **72**(1), 126 , doi:10.1186/s40623-020-01250-x.

Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.,* **169,** 1239–1260.

Berteussen, K.-A., 1977. Moho depth determinations based on spectral-ratio analysis of NORSAR long-period P waves, *Phys. Earth planet. Inter.,* **15**(1), 13–27.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y. & Wassermann, J., 2010. ObsPy: a Python toolbox for seismology, *Seismol. Res. Lett.,* **81**(3), 530–533 .

Bouchon, M., 2003. A review of the discrete wavenumber method, *Pure appl. Geophys.,* **160**(3), 445–465.

Brocher, T.M., 2005. Empirical relations between elastic wavespeeds and density in the Earth's crust, *Bull. seism. Soc. Am.,* **95**(6), 2081–2092.

Bussat, S. & Kugler, S., 2011. Offshore ambient-noise surface-wave tomography above 0.1 hz and its applications, *Leading Edge,* **30**(5), 514–524.

Byrd, R., Hribar, M. & Nocedal, J., 1999. An interior point algorithm for large-scale nonlinear programming, *SIAM J. Opt.,* **9**(4), 877–900.

Castellanos, J.C., Clayton, R.W. & Pérez-Campos, X., 2018. Imaging the eastern trans-Mexican volcanic belt with ambient seismic noise: evidence for a slab tear, *J. geophys. Res.,* **123,** 7741–7759.

Cedilnik, G., Lees, G., Schmidt, P., Herstrøm, S. & Geisler, T., 2019. Ultra-long reach fiber distributed acoustic sensing for power cable monitoring, in *Proceedings of the 10th International Conference on Insulated Power Cables,* held on 23-27 June, 2019 in Versailles, France.

Cheng, F., Chi, B., Lindsey, N.J., Dawe, T.C. & Ajo-Franklin, J.B., 2021. Utilizing distributed acoustic sensing and ocean bottom fiber optic cables for submarine structural characterization, *Sci. Rep.,* **11**(1), 1–14.

Dai, T., Hu, Y., Ning, L., Cheng, F. & Pang, J., 2018. Effects due to aliasing on surface-wave extraction and suppression in frequency-velocity domain, *J. appl. Geophys.,* **158,** 71–81 .

de Ridder, S. & Dellinger, J., 2011. Ambient seismic noise eikonal tomography for near-surface imaging at Valhall, *Leading Edge,* **30**(5), 506–512.

Foti, S., Lai, C., Rix, G. & Strobbia, C., 2015. *Surface Wave Methods for Near-Surface Site Characterization,* CRC Press.

Fukushima, S., Shinohara, M., Nishida, K., Takeo, A., Yamada, T. & Yomogida, K., 2022. Detailed s-wave velocity structure of sediment and crust off Sanriku, Japan by a new analysis method for distributed acoustic sensing data using a seafloor cable and seismic interferometry, *Earth, Planets Space,* **74**(1), 92 , doi:10.1186/s40623-022-01652-z.

Glorieux, C., Van de Rostyne, K., Nelson, K., Gao, W., Lauriks, W. & Thoen, J., 2001. On the character of acoustic waves at the interface between hard and soft solids and liquids, *J. acoust. Soc. Am.,* **110**(3), 1299–1306 .

Guerin, G., Rivet, D., van den Ende, M. P.A., Stutzmann, E., Sladen, A. & Ampuero, J.-P., 2022. Quantifying microseismic noise generation from coastal reflection of gravity waves recorded by seafloor DAS, *Geophys. J. Int.,* **231**(1), 394–407 .

Gusev, V., Desmet, C., Lauriks, W., Glorieux, C. & Thoen, J., 1996. Theory of Scholte, Leaky Rayleigh, and lateral wave excitation via the laser-induced thermoelastic effect, *J. acoust. Soc. Am.,* **100**(3), 1514–1528.

Halliday, D. & Curtis, A., 2008. Seismic interferometry, surface waves and source distribution, *Geophys. J. Int.,* **175**(3), 1067–1087 .

Hamilton, E.L., 1979. Sound velocity gradients in marine sediments, *J. acoust. Soc. Am.,* **65**(4), 909–922.

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

Hayes, G.P., Moore, G.L., Portner, D.E., Hearne, M., Flamme, H., Furtney, M. & Smoczyk, G.M., 2018. Slab2, a comprehensive subduction zone geometry model, *Science,* **362**(6410), 58–61.

Hunter, J.D., 2007. Matplotlib: a 2D graphics environment, *Comput. Sci. Eng.,* **9**(3), 90–95 .

Ide, S., Araki, E. & Matsumoto, H., 2021. Very broadband strain-rate measurements along a submarine fiber-optic cable off Cape Muroto, Nankai subduction zone, Japan, *Earth, Planets Space,* **73**(1), 1–10.

Kanazawa, T. & Hasegawa, A., 1997. Ocean-bottom observatory for earthquakes and tsunami off Sanriku, North-East Japan using submarine cable, in *Proceedings of the International Workshop on Scientific Use of Submarine Cables, Comm. for Sci. Use of Submarine Cables,* Okinawa, Japan, pp. 208–209.

Karrenbach, M., Ellwood, R., Yartsev, V., Araki, E., Kimura, T. & Matsumoto, H., 2020. Long-range das data acquisition on a submarine fiber-optic cable, in *EAGE Workshop on Fiber Optic Sensing for Energy Applications in Asia Pacific,* Vol. **2020,** pp. 1–5, European Association of Geoscientists & Engineers,.

Kimman, W.P. & Trampert, J., 2010. Approximations in seismic interferometry and their effects on surface waves, *Geophys. J. Int.,* **182**(1), 461–476 .

Kodaira, S. *et al.*, 2017. Depth-varying structural characters in the rupture zone of the 2011 Tohoku-oki earthquake, *Geosphere,* **13**(5), 1408–1424 .

Lin, F.-C., Ritzwoller, M.H., Townend, J., Bannister, S. & Savage, M.K., 2007. Ambient noise Rayleigh wave tomography of New Zealand, *Geophys. J. Int.,* **170**(2), 649–666 .

Lin, F.-C., Schmandt, B. & Tsai, V.C., 2012. Joint inversion of Rayleigh wave phase velocity and ellipticity using USArray: constraining velocity and density structure in the upper crust, *Geophys. Res. Lett.,* **39**(12), doi:10.1029/2012GL052196.

Lindsey, N.J., Dawe, T.C. & Ajo-Franklin, J.B., 2019. Illuminating seafloor faults and ocean dynamics with dark fiber distributed acoustic sensing, *Science,* **366**(6469), 1103–1107.

Lior, I. *et al.*, 2021. On the detection capabilities of underwater distributed acoustic sensing, *J. geophys. Res.,* **126**(3), e2020JB020925 , doi:10.1029/2020JB020925.

Lior, I., Mercerat, E.D., Rivet, D., Sladen, A. & Ampuero, J., 2022. Imaging an underwater basin and its resonance modes using optical fiber distributed acoustic sensing, *Seismol. Res. Lett.,* **93**(3), 1573–1584.

Longuet-Higgins, M.S., 1950. A theory of the origin of microseisms, *Phil. Trans. R. Soc. Lond., A,* **243,** 1–35.

Martin, E.R., Lindsey, N.J., Ajo-Franklin, J.B. & Biondi, B.L., 2021. *Introduction to Interferometry of Fiber-Optic Strain Measurements,* Chapter 9, pp. 111–129, American Geophysical Union (AGU).

Matsumoto, H. *et al.*, 2021. Detection of hydroacoustic signals on a fiber-optic submarine cable, *Sci. Rep.,* **11**(1), 1–12.

Mordret, A., Landès, M., Shapiro, N., Singh, S., Roux, P. & Barkved, O., 2013. Near-surface study at the valhall oil field from ambient noise surface wave tomography, *J. geophys. Int.,* **193**(3), 1627–1643.

Mordret, A., Landès, M., Shapiro, N.M., Singh, S.C. & Roux, P., 2014. Ambient noise surface wave tomography to determine the shallow shear velocity structure at Valhall: depth inversion with a Neighbourhood Algorithm, *Geophys. J. Int.,* **198,** 1514–1525 .

Nakahara, H. & Haney, M.M., 2022. Connection between the cross correlation and the Green's function: strain and rotation of surface waves, *Geophys. J. Int.,* **230**(2), 1166–1180.

Nakamura, T., Nakano, M., Hayashimoto, N., Takahashi, N., Takenaka, H., Okamoto, T., Araki, E. & Kaneda, Y., 2014. Anomalously large seismic amplifications in the seafloor area off the Kii peninsula, *Mar. Geophys. Res.,* **35**(3), 255–270 .

Nap, A., Edme, P., Schmelzbach, C., Paitz, P. & Robertsson, J., 2020. Surface seismic with distributed acoustic sensing: is trenching worthwhile?, in *NSG2020 26th European Meeting of Environmental and Engineering Geophysics,* December 2020, Vol. 2020, pp. 1–5, 1–5, European Association of Geoscientists & Engineers, doi:10.3997/2214-4609.202020145.

Nayak, A. & Ajo-Franklin, J., Team T. I. V. D.F., 2021. Measurement of surface-wave phase-velocity dispersion on mixed inertial seismometer – distributed acoustic sensing seismic noise cross-correlations, *Bull. seism. Soc. Am.,* **111**(6), 3432–3450 doi:10.1785/0120210028.

Nishida, K., Kawakatsu, H. & Obara, K., 2008. Three-dimensional crustal S wave velocity structure in Japan using microseismic data recorded by Hi-net tiltmeters, *J. geophys. Res.,* **113**(B10), doi:10.1029/2007JB005395.

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,* pp. 1377–1380, Society of Exploration Geophysicists.

Perton, M. & Sánchez-Sesma, F.J., 2016. Green's function calculation from equipartition theorem, *J. acoust. Soc. Am.,* **140**(2), 1309–1318.

Perton, M., Spica, Z.J., Clayton, R.W. & Beroza, G.C., 2019. Shear wave structure of a transect of the Los Angeles basin from multimode surface waves and H/V spectral ratio analysis, *Geophys. J. Int.,* **220**(1), 415–427 .

Rivet, D., de Cacqueray, B., Sladen, A., Roques, A. & Calbris, G., 2021. Preliminary assessment of ship detection and trajectory evaluation using distributed acoustic sensing on an optical fiber telecom cable, *J. acoust. Soc. Am.,* **149**(4), 2615–2627.

Sabra, K.G., Gerstoft, P., Roux, P., Kuperman, W.A. & Fehler, M.C., 2005. Extracting time-domain Green's function estimates from ambient seismic noise, *Geophys. Res. Lett.,* **32**(3), doi:10.1029/2004GL021862.

Sato, H., Fehler, M.C. & Maeda, T., 2012. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth,* Vol. **496,** Springer.

Scherbaum, F., Hinzen, K.-G. & Ohrnberger, M., 2003. Determination of shallow shear wave velocity profiles in the Cologne, Germany area using ambient vibrations, *Geophys. J. Int.,* **152**(3), 597–612 .

Schimmel, M. & Paulssen, H., 1997. Noise reduction and detection of weak, coherent signals through phase-weighted stacks, *Geophys. J. Int.,* **130,** 497–505 .

Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.,* **31**(7), doi:10.1029/2004GL019491.

Shinohara, M. *et al.*, 2019. Distributed acoustic sensing measurement by using seafloor optical fiber cable system off Sanriku for seismic observation, in *Proceedings of the Oceans 2019 MTS/IEEE Seattle,* 27–31 October 2019, Seattle, WA, USA, pp. 1–4, IEEE.

Shinohara, M., Yamada, T., Uehira, K., Sakai, S., Shiobara, H. & Kanazawa, T., 2021. Development and operation of an ocean bottom cable seismic and tsunami (OBCST) observation system in the source region of the Tohoku-oki earthquake, *Earth Space Science,* **8**(3), e2020EA001359 , doi:10.1029/2020EA001359.

Shinohara, M., Yamada, T., Akuhara, T., Mochizuki, K. & Sakai, S., 2022. Performance of seismic observation by distributed acoustic sensing technology using a seafloor cable off Sanriku, Japan, *Front. Mar. Sci.,* doi:10.3389/fmars.2022.844506.

Shragge, J., Yang, J., Issa, N., Roelens, M., Dentith, M. & Schediwy, S., 2021. Low-frequency ambient distributed acoustic sensing (DAS): case study from Perth, Australia, *Geophys. J. Int.,* **226**(1), 564–581.

Snieder, R., 2004. Extracting the green's function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E,* **69,** 046610 , doi:10.1103/PhysRevE.69.046610.

Socco, L. & Strobbia, C., 2004. Surface-wave method for near-surface characterization: a tutorial, *Near Surf. Geophys.,* **2**(4), 165–185.

Spica, Z., Perton, M., Calò, M., Legrand, D., Córdoba-Montiel, F. & Iglesias, A., 2016. 3-D shear wave velocity model of Mexico and South US: bridging seismic networks with ambient noise cross-correlations (C1) and correlation of coda of correlations (C3), *Geophys. J. Int.,* **206**(3), 1795–1813.

Spica, Z., Perton, M., Nakata, N., Liu, X. & Beroza, G.C., 2018. Shallow VS imaging of the Groningen area from joint inversion of multimode surface waves and H/V spectral ratios, *Seismol. Res. Lett.,* **89**(5), 1720–1729.

Spica, Z.J., Perton, M., Nakata, N., Liu, X. & Beroza, G.C., 2017. Site characterization at Groningen gas field area through joint surface-borehole H/V analysis, *Geophys. J. Int.,* **212**(1), 412–421.

Spica, Z.J., Nishida, K., Akuhara, T., Pétrélis, F., Shinohara, M. & Yamada, T., 2020. Marine sediment characterized by ocean-bottom fiber-optic seismology, *Geophys. Res. Lett.,* **47**(16), e2020GL088360 , doi:10.1029/2020GL088360.

Spica, Z.J., Castellanos, J.C., Viens, L., Nishida, K., Akuhara, T., Shinohara, M. & Yamada, T., 2022. Subsurface imaging with ocean-bottom distributed acoustic sensing and water phases reverberations, *Geophys. Res. Lett.,* **49**(2), e2021GL095287, doi:10.1029/2021GL095287.

Stehly, L., Fry, B., Campillo, M., Shapiro, N.M., Guilbert, J., Boschi, L. & Giardini, D., 2009. Tomography of the Alpine region from observations of seismic ambient noise, *Geophys. J. Int.,* **178**(1), 338–350 .

Stein, S. & Wysession, M., 2003. *Introduction to Seismology, Earthquakes, and Earth Structure,* Blackwell Publishing.

Takano, H., Shinohara, M., Nakata, R., Kurashimo, E., Ishiyama, T. & Mochizuki, K., 2021. Seismic reflection survey using seafloor optical cable and DAS measurement off sanriku, in *Proceedings of the Seismological Society of Japan Fall Meeting,* Vol. 2021, pp. ROMBUNNO.S02P-06, SSJ.

Tonegawa, T. *et al.*, 2022. Extraction of p wave from ambient seafloor noise observed by distributed acoustic sensing, *Geophys. Res. Lett.,* **49**(4), e2022GL098162, doi:10.1029/2022GL098162.

Ugalde, A., Becerril, C., Villaseñor, A., Ranero, C.R., Fernández-Ruiz, M.R., Martin-Lopez, S., González-Herráez, M. & Martins, H.F., 2021. Noise levels and signals observed on submarine fibers in the Canary Islands using DAS, *Seismol. Res. Lett.,* **93**(1), 351–363.

Ventosa, S., Simon, C. & Schimmel, M., 2012. Window length selection for optimum slowness resolution of the local-slant-stack transform, *Geophysics,* **77**(2), V31–V40 .

Wang, H.F., Zeng, X., Miller, D.E., Fratta, D., Feigl, K.L., Thurber, C.H. & Mellors, R.J., 2018. Ground motion response to an ML 4.3 earthquake using co-located distributed acoustic sensing and seismometer arrays, *Geophys. J. Int.,* **213**(3), 2020–2036 .

Wessel, P., Luis, J.F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H.F. & Tian, D., 2019. The generic mapping tools version 6, *Geochem., Geophys., Geosyst.,* **20**(11)), 5556–5564.

Williams, E.F., Fernández-Ruiz, M.R., Magalhaes, R., Vanthillo, R., Zhan, Z., González-Herráez, M. & Martins, H.F., 2021. Scholte wave inversion and passive source imaging with ocean-bottom DAS, *Leading Edge,* **40**(8), 576–583 .

Williams, E.F., Zhan, Z., Martins, H.F., Fernández-Ruiz, M.R., Martín-López, S., González-Herráez, M. & Callies, J., 2022. Surface gravity wave interferometry and ocean current monitoring with ocean-bottom DAS, *J. geophys. Res.,* **127**(5), e2021JC018375 , doi:10.1029/2021JC018375.

Xiao, H., Tanimoto, T., Spica, Z.J., Gaite, B., Ruiz-Barajas, S., Pan, M. & Viens, L., 2022. Locating the precise sources of high-frequency microseisms using distributed acoustic sensing, *Geophys. Res. Lett.,* **49**(17), e2022GL099292 , doi:10.1029/2022GL099292.

Yao, H., van Der Hilst, R.D. & de Hoop, M.V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis — I. Phase velocity maps, *Geophys. J. Int.,* **166**(2), 732–744 .

Zeng, X., Lancelle, C., Thurber, C., Fratta, D., Wang, H., Lord, N., Chalari, A. & Clarke, A., 2017. Properties of noise cross-correlation functions obtained from a distributed acoustic sensing array at Garner Valley, California, *Bull. seism. Soc. Am.,* **107**(2), 603–610 .

Zheng, Y., Fang, X., Liu, J. & Fehler, M.C., 2013. Scholte waves generated by seafloor topography, arXiv:1306.4383.

Zhu, J., Popovics, J.S. & Schubert, F., 2004. Leaky Rayleigh and Scholte waves at the fluid–solid interface subjected to transient point loading, *J. acoust. Soc. Am.,* **116**(4), 2101–2110.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. 20-min CCFs computed between channels 4000 (virtual source) and 4300 (receiver) over the 13 d of data. One-bit normalization is applied to the continuous data at both channels prior to computing 20-min CCFs and the 20-min CCFs are bandpass filtered between 0.25 and 5 Hz.

Figure S2. (a) 13-d stack of the 20-min CCFs computed between channel 4000 (virtual source) and the following 400 channels. (b) Same as (a) for a 3-d stack of the 20-min CCFs. One-bit normalization is applied to the continuous data before computing the 20-min CCFs and all the stacked CCFs are bandpass filtered between 0.25 and 5 Hz. (c) Selected dispersion points from the dispersion images computed from the CCFs illustrated in (a) and (b) are shown in red and black dots, respectively.

Figure S3. CCFs computed between channel 4000 (virtual source) and (a) 200, (b) 300 and (c) 400 receiver channels. The corresponding dispersion images are shown in (d–f). For each panel in (d–f), the spatial aliasing lines computed with eq. (2) of the main manuscript up to $i = 10$ are also shown.

Figure S4. All selected local maxima (black) versus refined selected dispersion points (red) for the dispersion images computed from virtual sources 5000, 5500 and 6000. The dispersion images, which are shown in Fig. 5 of the main manuscript, are calculated from 1-bit CCFs and 400 receiver channels.

Figure S5. (a) V_S (black) and V_P (blue) velocity profiles for a step gradient velocity profile of $\frac{\Delta V_S}{\Delta V_z} = 1\,s^{-1}$. (b) Theoretical dispersion image computed from vertical displacement waveforms generated by a vertical source using the velocity profile shown in (a). The energy of the dispersion image is normalized between 0 and 1 and the theoretical DCs are shown by black lines.

Figure S6. Fitted DCs with Equations (blue) 4 and (red) 5 of the main manuscript and selected dispersion points (black dots) for the dispersion image computed with virtual source 4000. (b) *S*-wave velocity profiles for the two inversion schemes.

Figure S7. (a) Scholte wave sensitivity kernels at frequencies of 0.5, 1, 1.5, 2.1, 2.6 and 3.1 Hz for the inverted 1-D V_S model using channel 3000 as the virtual source. (b) Same as (a) for the inverted 1-D V_S model using channel 5000 as the virtual source. (c) Sum of sensitivity of the first 25 modes for the inverted 1-D V_S model using channels 3000 (blue) and 5000 (orange) as virtual sources at different frequencies.

Figure S8. (a) Strain waveforms of a M_V 2.5 earthquake bandpass filtered between 1 and 3 Hz. *P* and *S* waves arrive around 5 and 10 s after the origin time, respectively. (b) Zoom on the shallow part of the inverted velocity model shown in Fig. 9(b).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.