Geophysical Journal International

Geophys. J. Int. (2024) **237,** 1751–1764
Advance Access publication 2024 April 11
GJI Seismology

https://doi.org/10.1093/gji/ggae134

Urban subsurface exploration improved by denoising of virtual shot gathers from distributed acoustic sensing ambient noise

Leila Ehsaninezhad [1,2] Christopher Wollin,[1] Verónica Rodríguez Tribaldos,[1] Benjamin Schwarz [3] and Charlotte M. Krawczyk[1,2]

[1] *German Research Centre for Geosciences (GFZ),* 14473 *Potsdam, Germany. E-mail: leila@gfz-potsdam.de*
[2] *Institute of Applied Geosciences, Technical University of Berlin,* 10623 *Berlin, Germany*
[3] *Fraunhofer Institute for Wind Energy Systems (IWES),* 27572 *Bremen, Germany*

Accepted 2024 April 2. Received 2024 March 16; in original form 2023 November 17

SUMMARY

Ambient noise tomography on the basis of distributed acoustic sensing (DAS) deployed on existing telecommunication networks provides an opportunity to image the urban subsurface at regional scales and high-resolution. This capability has important implications in the assessment of the urban subsurface's potential for sustainable and safe utilization, such as geothermal development. However, extracting coherent seismic signals from the DAS ambient wavefield in urban environments at low cost remains a challenge. One obstacle is the presence of complex sources of noise in urban environments, which may not be homogeneously distributed. Consequently, long recordings are required for the calculation of high-quality virtual shot gathers, which necessitates significant time and computational cost. In this paper, we present the analysis of 15 d of DAS data recorded on a pre-existing fibre optic cable (dark fibres), running along an 11-km-long major road in urban Berlin (Germany), hosting heavy traffic including vehicles and trains. To retrieve virtual shot gathers, we apply interferometric analysis based on the cross-correlation approach where we exclude low-quality virtual shot gathers to increase the signal-to-noise ratio of the stacked gathers. Moreover, we modify the conventional ambient noise interferometry workflow by incorporating a coherence-based enhancement approach designed for wavefield data recorded with large-N arrays. We then conduct multichannel analysis of surface waves to retrieve 1-D velocity models for two exemplary fibre subsegments, and compare the results of the conventional and modified workflows. The resulting 1-D velocity models correspond well with available lithology information. The modified workflow yields improved dispersion spectra, particularly in the low-frequency band (<1 Hz) of the signal. This leads to an increased investigation depth along with lower uncertainties in the inversion result. Additionally, these improved results were achieved using significantly less data than required using conventional approaches, thus opening the opportunity for shortening required acquisition times and accordingly lowering costs.

Key words: Distributed acoustic sensing; Seismic interferometry; Seismic noise; Seismic tomography; Surface waves and free oscillations; Waveform inversion.

1 INTRODUCTION

Understanding subsurface structure in urban areas is necessary to assess the feasibility and viability of utilizing the subsurface for energy production such as geothermal development. A common method for subsurface exploration is based on seismic experiments using large receiver arrays, deployed either at the surface or down a borehole, and appropriate sources of excitation (Krawczyk *et al.* 2019b; Martuganova *et al.* 2022). However, large-scale seismic surveys for exploration purposes in urban areas are challenging due to the restrictions on the use of active seismic sources and receiver installations. To address these hurdles, passive seismic methods that utilize the ambient seismic field as the source of seismic energy have been used as a feasible alternative in urban areas. Moreover, the combination of the passive seismic with measurements using distributed acoustic sensing (DAS) on dark telecommunication fibres have attracted particular attention as its deployment is particularly cheap and effortless (Dou *et al.* 2017; Tribaldos *et al.* 2021; Yang *et al.* 2022).

Generally, ambient seismic noise is continuous and prevalent, and originates from either natural sources, such as the ocean microseism (Sabra *et al.* 2005; Bensen *et al.* 2008; Clayton 2020; Jiang &

© The Author(s) 2024. Published by Oxford University Press on behalf of The Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

Denolle 2022), or from anthropogenic activity (Asten 2006; Nakata *et al.* 2011). Over the last decade, several studies have successfully shown that ambient noise interferometry, that is the extraction of the Green's function between two stations from the recorded noise, can be utilized to retrieve shear-wave velocity models of the subsurface (Shapiro *et al.* 2005; Mordret *et al.* 2014; Ryberg *et al.* 2016; Zhang *et al.* 2020). Although ambient noise interferometry is capable of extracting body and surface wave components of the Green's function in theory, the extraction of surface waves is more achievable as they are prevalent on the surface and dominate the ambient-noise field (Behm *et al.* 2014; Nakata *et al.* 2015). Surface waves, and in particular Rayleigh waves, generated from low-frequency ocean microseismic noise (5 s and above) have been commonly used for large-scale tomographic imaging reaching depths beyond 10 km and yielding a spatial resolution in the order of several km (Sabra *et al.* 2005; Bensen *et al.* 2008; Clayton 2020; Jiang & Denolle 2022). In urban environments, however, the ambient seismic field is mostly dominated by surface waves originating from anthropogenic activity. Anthropogenic seismic noise is prominent in the vicinity of any urban agglomeration, and it is composed of frequencies higher than that of the ocean microseism (>1 Hz), thus allowing for higher resolution imaging at shallower depths (Nakata *et al.* 2011; Cheng *et al.* 2019).

Ambient-noise tomography at high spatial resolution also requires a dense array of receivers. DAS is a developing technology that uses fibre-optic cables, widely used for data transmission, as a sensor (Jousset *et al.* 2018; Krawczyk *et al.* 2019a; Lindsey *et al.* 2020). It offers the opportunity to convert vast optic telecommunication networks into large-N seismic arrays, that is into strings of densely spaced seismic receivers. DAS uses interrogator units (IUs) that probe the fibre cable by injecting laser pulses into one of its ends and analysing light backscattered all along its trajectory and recorded back at the IUs. By applying optical interferometry to the light originating from a small fibre interval referred to as the gauge-length, DAS renders this interval's strain-rate (Hartog 2017).

In a seismological setting where the fibre is coupled to the Earth, elastic strain changes are attributable to the seismic wavefield travelling along the fibre (Lindsey *et al.* 2017). The cost-effective deployment of the interrogator unit and the durability of the fibre in harsh environments offer the possibility to use DAS in a wide range of experiments and applications such as vertical seismic profiling (Götz *et al.* 2018; Henninges *et al.* 2021; Martuganova *et al.* 2022), monitoring of geothermal well completion (Lipus *et al.* 2021), volcanoes (Jousset *et al.* 2022) as well as glaciers (Klaasen *et al.* 2021) or ocean bottom characterization with submarine fibres (Taweesintananon *et al.* 2021; Cheng *et al.* 2021; Spica *et al.* 2022). In urban areas, where the deployment of large arrays of conventional sensors is challenging, the deployment of DAS on already existing, unused fibre-optic cables (known as 'dark fibre') is an attractive alternative for high-resolution imaging (Zhu & Stensrud 2019; Ajo-Franklin *et al.* 2019; Spica *et al.* 2020; Li *et al.* 2023).

The applicability of ambient noise interferometry to DAS recordings has been demonstrated in numerous recent studies (Dou *et al.* 2017; Tribaldos *et al.* 2021; Yang *et al.* 2022). Most investigations combine ambient noise interferometry with subsequent multichannel analysis of surface waves (MASWs) to map shallow structures(Dou *et al.* 2017; Song *et al.* 2021; Zhou *et al.* 2022), where MASW calculates the dispersion spectra of surface waves across the array, which can then be inverted to obtain shear-wave velocity models of the subsurface (Foti *et al.* 2018).

Most DAS urban ambient noise experiments using anthropogenic noise focus on the upper 100 m of the subsurface due to the limited low-frequency content of this noise field (Dou *et al.* 2017; Yang *et al.* 2023a). However, a few experiments have investigated the use of alternative sources with lower frequency content such as trains (Tribaldos *et al.* 2021), and the use of larger subarrays (Cheng *et al.* 2023) in order to increase the investigation depth. These techniques improved the retrieval of the phase velocities at lower frequencies and, accordingly, investigation depth. However, little attention has been given to the retrieval of low-frequency content (<1 Hz) from the DAS urban seismic noise, which is mainly dominated by high-frequency (>1 Hz) energy. Consequently, the impact of the complex, broadband urban seismic noise on the quality of the extracted virtual shot gathers, and the potential of using urban noise DAS records for investigating the subsurface down to depths of several 100 m to a few km depths, remains largely unexplored.

One approach to enhance the fidelity of low-frequency signals involves improving the signal-to-noise ratio (SNR) through denoising techniques (Chen *et al.* 2019). Recently, the use of denoising techniques on DAS data in seismology has gained attention due to the need for high-quality seismic signals (Williams *et al.* 2019; Atterholt *et al.* 2022; Yang *et al.* 2023b). Various studies have targeted different sources of noise, such as noise stemming from coupling conditions (Martuganova *et al.* 2021), noise originating from DAS instruments (Isken *et al.* 2022; Yang *et al.* 2023b; Lapins *et al.* 2024) and mixed noise stemming from DAS instrument and unwanted coherent recorded noise (e.g. traffic noise), prioritizing the mitigation of noise primarily originating from DAS instruments (van den Ende *et al.* 2021; Atterholt *et al.* 2022; Yang *et al.* 2023b). Most of these approaches primarily focus on enhancing earthquake signals detection in raw DAS recordings through deep learning techniques (van den Ende *et al.* 2021; Yang *et al.* 2023b) and transforming raw DAS recording into alternative representations (e.g. frequency–wavenumber, wavelet or curvelet transform; Williams *et al.* 2019; Atterholt *et al.* 2022). Deep learning approaches are known for their fast and data-driven nature; however, they require extensive training datasets and have limited adaptability across various datasets (Chen *et al.* 2019). Alternatively, some techniques involve transforming data into other representations. However, choosing a suitable threshold function poses a challenge (Yang *et al.* 2023b). Recently Schwarz (2019) introduced a denoising method based on wavefield decomposition, offering a fully automated solution capable of precisely targeting specific wavefield Chen *et al.* (2019). It is suitable for scenarios where preserving spatial coherence and acquiring accurate dispersion information are necessary, such as seismic velocity tomography utilizing surface waves.

In this study, we introduce a modified workflow for DAS ambient noise interferometry in an inner-city urban environment to improve the SNR of the virtual surface-wave gathers and accordingly to increase the investigation depth. We introduce a coherence-enhancement filter, based on wavefiled decomposition, a tool recently developed in seismic imaging (Schwarz 2019), which aims to separate coherent and random noise and to suppress the latter. We then apply MASW on the resulting virtual shot gathers and compare the dispersion spectra resulting from the modified and conventional workflows. We evaluate the amount of data required for stacking to obtain high-quality virtual shot gathers from the two workflows. Finally, we retrieve shear-wave velocity models for two subsegments of the interrogated fibre and compare them with the documented geological stratigraphy. The paper concludes with the discussion on the limitation of urban traffic noise on the maximum resolvable

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

frequency as well as the effectiveness of the proposed workflow for improving the quality of the obtained virtual shot-gathers and consequently reducing the uncertainty of the inverted velocity models and increasing the investigation depth.

2 STUDY SITE AND DATA

2.1 Geology and measurement setup

The study site is situated in the Adlershof district in the southeast of the city of Berlin, Germany. Geologically, Berlin is part of the North German Basin (NGB), which is considered one of the three major target sedimentary basin settings for geothermal applications in Germany (Noack *et al.* 2013). Consequently, subsurface investigations for geothermal energy development and utilization in this area are gaining increased attention. However, subsurface utilization for geothermal exploitation may be accompanied with hazards to the environment and infrastructure, for example water pollution following leakage due to the presence of a Rupelian structure in this area. This prominent structure in this setting is a Tertiary clay formation that originated from a marine ingression, with a thickness that ranges from a few decimetres to about 120 m across the NGB (Stackebrandt & Franke 2015). As a continuous layer, it separates the upper fresh water aquifer from the lower saline aquifer (Norden *et al.* 2010), acting as a hydraulic barrier which also has a potential impact on the thermal field (Noack *et al.* 2013). The upper bound of this layer in this region is at a depth between 150 and 190 m. Any disruption of the layer could cause saline water to enter the shallower freshwater horizons, causing a salinization of the groundwater aquifers. Therefore, before any manmade disruption, for example drilling for geothermal development, delineating the boundary of this layer is of great benefit. (Noack *et al.* 2013; Frick *et al.* 2016; Schilling *et al.* 2018; Frick *et al.* 2019).

We conduct our measurements along a major road in southeast Berlin which is characterized by major traffic of motorized vehicles throughout its length. An 11-km-long trajectory of the dark telecommunication fibre shown in Fig. 1 was used as a sensor in this study. The longest straight subsection of the fibre is ∼6 km long and runs parallel to 'Adlergestell' road and a railway track carrying suburban and cargo trains. Parts of this section in the northwest also carries trams. Before and after the 'Adlergestell' section, the fibre runs orthogonal to this major road. The fibre optic cable was rented from a local telecommunication company and had been deployed by blowing a strand of multiple fibres into a small conduit with pressurized air, where the small conduit itself lies within a larger cased conduit (also compare Ajo-Franklin *et al.* 2019). This setup is situated approximately 1 m below the ground and transverses numerous cable funnels where the fibre is coiled up in loops of slack.

The fibre was interrogated by a commercial Silixa iDAS unit recording strain-rate with a gauge length of 10 m. Measurements with a spatial sampling of 8 m and 1 kHz temporal frequency were conducted between March 17 and April 19, 2021. Accurate spatial localization of the fibre-optic cable for determining the geometry of the measurements was done by performing tap tests along the entire fibre trajectory using a hammer and a GPS device. The spatial localization error is between 8 and 10 m, and thus in the order of the gauge length.

In the following we analyse two approximately 1-km-long segments of the array (segments A and B, indicated in blue and magenta in Fig. 1, with boundaries shown in black dashed lines in Fig. 2). Segment A is used here to describe the workflow steps. The results for analysis of Segment B can be found in the Supplementary material. Segment A contains the channels at optical distances between 1700 and 2760 m, where the cable is mostly straight and without any crossings of major side roads, which could become local noise sources and potentially introducing spurious signals in the process of ambient noise interferometry (Cheng *et al.* 2023).

2.2 General characteristics of the recorded ambient noise field

Fig. 2(a) shows trace-normalized strain-rate for a 5-min-long DAS record along the 11 km-long fibre segment, where the data was decimated from 1000 to 125 Hz for plotting. The numerous diagonal features in this spatio-temporal recording correspond to seismic noise generated mainly by moving traffic such as cars, buses and trains travelling parallel to the fibre trajectory. Stationary noise sources such as construction yards and workshops are also present. The most dominant events are imprints of trains travelling in opposite directions, visible as zig-zag patterns. The discontinuity of this pattern at an optic distance of 3200 m shows the location of a train station (red dashed line in Fig. 2a).

We inspect the frequency content of our data by calculating the power spectral density (PSD) of hour-long recordings of each individual DAS channel. PSDs of each channel averaged over a time period of 2 d in Fig. 2(b) shows that below 50 Hz, the dominant spectral energy lies between 8 and 40 Hz. There are some mono-frequent signals which are visible across several neighbouring channels which could be caused by stationary works or machinery. For instance, 12 and 20 Hz (e.g. at distance of 11000 m), 18 Hz (e.g. at distance of 2500 m) and 22, 25, 28, 34 and 40 Hz (e.g. at distance of 6000 m) are shown with black arrows in Fig. 2(b).

Beyond a distance of 4000 m along the fibre, individual noise features seem less well-defined and amplitudes appear lower. This is likely due to a change in cable-ground coupling conditions or the exposure to a noisier environment. The blue dashed lines in Fig. 2(a) represent the locations where the cable branches off the main road. Large spectral energy throughout all frequencies indicates that the traffic crosses the cable there, saturating the signal.

To have a better understanding of the temporal pattern of the seismic noise recorded along the analysed Segment A, PSDs covering the entire 15 d of recording were averaged across all channels within this 1 km section of the cable. The resulting spectrogram (Fig. 3) clearly reveals a diurnal noise pattern, with increased and diminished levels of spectral energy during the day and night time, respectively. The two weekends on 9 and 16 of April can also be distinguished with intermediary levels of energy in both spectrogram (Fig. 3a) and temporal PSDs variation for some distinct frequencies (Fig. 3b).

Although the diurnal pattern of varying seismic energy can be distinguished throughout the inspected frequency band, particularly elevated levels of working days can be observed between 8 and 40 Hz during the daytime. The temporal pattern confirms that the seismic energy stems from anthropogenic activities. Similarly, we calculate PSDs covering the entire 15 d of recording were averaged across all channels within 1 km section of Segment B. The seismic energy in Fig. S8 shows the same temporal pattern as Segment A, but less predominant in frequency below 8 Hz.

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

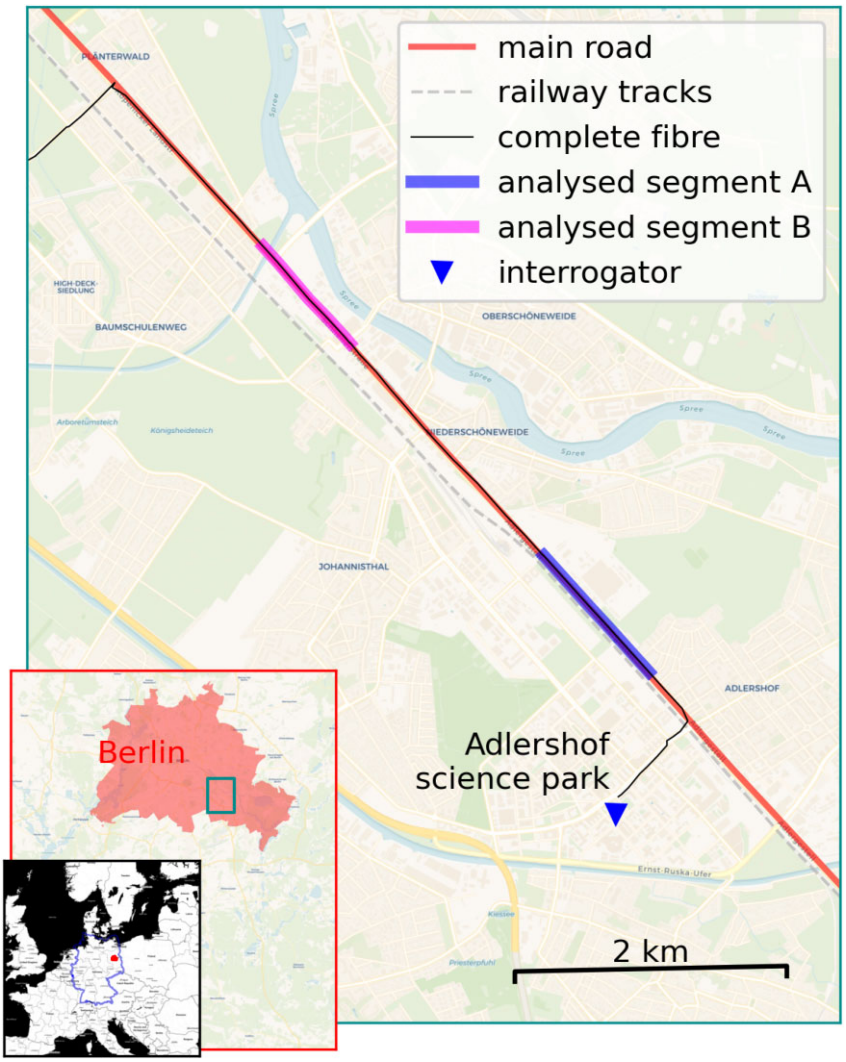

Figure 1. Map of the study area (data by OpenStreetMap contributors 2017). The insets in the lower left show the location of Adlershof (cyan frame) in greater suburban Berlin and the location of Berlin within Germany and central Europe (data by OpenStreetMap contributors 2017, map tiles by Stamen Design). The main plot indicates the locations of the interrogator (blue triangle), the fibre trajectory alongside the major road 'Adlergestell' (red), a strand of railway tracks running parallel to it (grey dashed line), and the 1-km-long segments of cable A and B analysed in this study (blue and magenta lines).

3 PROCESSING WORKFLOW

We present the entire workflow through the exemplary Segment A shown in blue in Fig. 1. We choose the southernmost channel of this 1-km-long segment (at optical distance of 1722 m) as the virtual source and perform the cross-correlation with all adjacent channels in northwest direction within a 1 km distance comprising 132 channels in total. Cross-correlation was performed for each 1-hr-long DAS record for the 15-d acquisition period thus yielding 360 virtual shot gathers (i.e. Green's functions) in total. To better assess our modified workflow, we apply it to a second independent fibre segment shown in magenta in Fig. 1. Segment B is at optical distance of 5025 m without any overlap with the first segment (Segment A).

3.1 Ambient noise interferometry

At first, continuous records sampled at 1000 Hz were decimated to 250 Hz using dedicated software incorporating an adequate low-pass FIR filter (Isken *et al.* 2021). Before applying the correlation analysis, hourly-long traces were pre-processed based on the standard workflow proposed by Bensen *et al.* (2007). This step involved trace-wise removal of the mean and linear trend as well as tapering followed by bandpass filtering between 0.2 and 40 Hz. Then a moving average window normalization in the time domain and spectral whitening were applied. The resulting whitened signals retained a large amplitude spike at the frequency of 35 Hz, which is also visible in the raw data spectrogram (Fig. 3) particularly during daytime. This feature suggests some anthropogenic activity along the fibre exposing the surroundings to a mono-frequent seismic signal. Thus, we further restrict the investigated frequency band by adding another band-pass filter between 0.5 and 28 Hz after the spectral whitening as a final step in the preparatory processing prior to cross-correlation.

3.2 Evaluation of the virtual shot gathers and stacking

For Segment A, approximately 212 out of 360 resulting virtual shot gathers are of good quality and display the expected v-shaped pattern. This results in the surface-wave moveout on both the causal

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

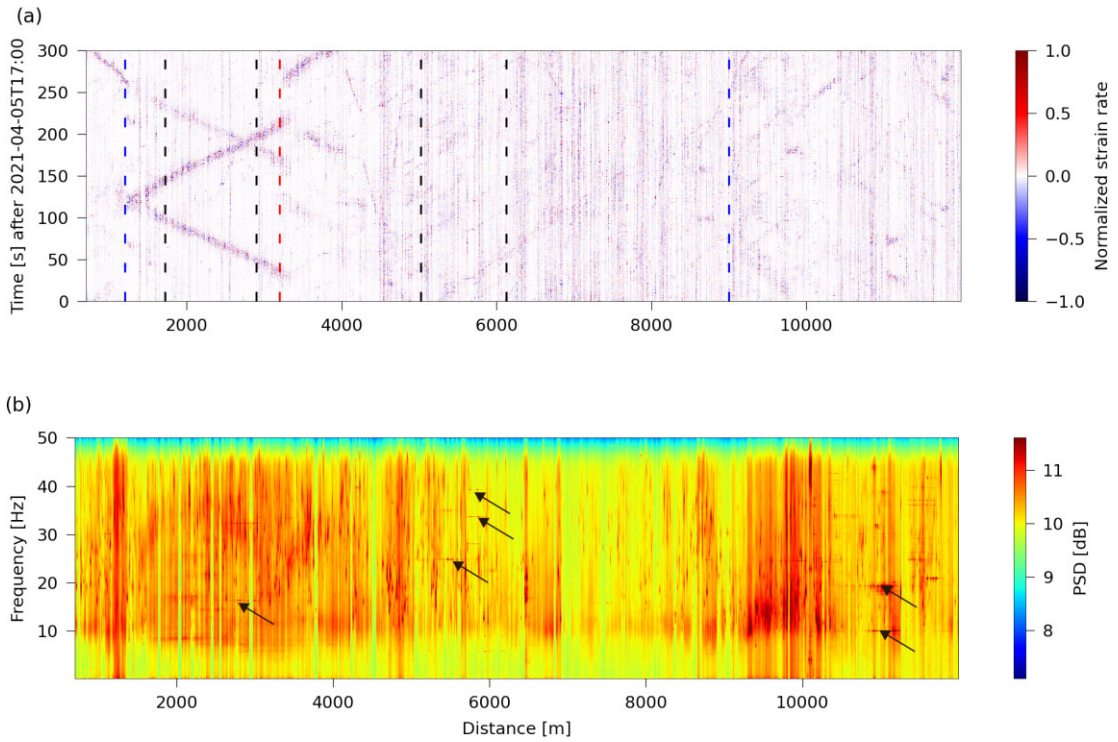

Figure 2. Amplitude and power spectral density (PSD) of DAS data recorded on the 11-km-long fibre optic cable along Adlergestell (a) 5-min-long DAS recording. The recordings were decimated to 125 Hz and normalized channel-wise for plotting. Zig-zag lines show the imprints of trains travelling along the cable. Black dashed lines show the boundaries of Segment A and Segment B (see supplementary material for Segment B) analysed in this study. Blue and red dashed lines show the location of bends and a train station, respectively. (b) Channel-wise PSD of hour-long recordings averaged over a time period of 2 d (black arrows indicate some examples of mono-frequent signals).

(positive) and acausal (negative) sides of the cross-correlation which is indicative of a homogeneous distribution of noise sources. Low quality gathers (Figs S1b and c) were discarded after visual inspection.

The first type of the excluded gathers displays an artefact that is repeatedly visible at the same fibre location. Trains of surface waves originate from a point in the immediate vicinity of the fibre (Fig. S1b). Therefore we call this source 'local noise' which likely stems from a workshop that could be identified at this optical distance of the array. This type of noise has been observed in a similar way in other studies too (Retailleau & Beroza 2021)

In contrast, the second type of noise contaminates the entire virtual shot gather with continuous mono-frequent wave trains (Fig. S1c). The expected moveout of the surface waves is visible only in the immediate vicinity of the virtual source, and it is likely that these signals are caused by various periodically working machinery in the wider area, for example pumps or power-transformers. In comparison with the included shot gathers, the spectra of the corresponding virtual shot gather displays a multitude of narrow frequency peaks (Fig. S1d). Accordingly, we call this source 'mono-frequent noise'.

After the exclusion of contaminated virtual shot gathers, the suitable gathers were linearly stacked to increase signal-to-noise ratio. This step yielded the final common virtual shot gather used for further processing (Fig. 4a). The virtual shot gather obtained from this selective stacking procedure shows consistency of the Rayleigh wavefield for both causal and acausal parts (Fig. 4a), suggesting a sufficiently uniform distribution of noise sources on each extremity and inline of the fibre. Similarly, for Segment B, 277 out of 360 virtual shot gathers are of good quality and display the expected surface-wave moveout and v-shaped pattern (Fig. S1a).

In order to evaluate the improvement in data quality obtained by stacking the selected gathers only, we calculate the *SNR* for each channel as the ratio of signal to noise energy using:

$$\mathrm{SNR} = 20 \log_{10} \frac{rms_{\mathrm{signal}}}{rms_{\mathrm{noise}}}, \tag{1}$$

where *rms* is root-mean-squared amplitude. The rms of the signal and noise were calculated within a time window around the first arrival given by two diverging linear functions (delineated in blue in Fig. S2) and the last two seconds of each trace (delineated in red in Fig. S2), respectively. The increased SNR of the selective stack compared to a stack including all single gathers in Fig. S3 confirms the effectiveness of the selective stacking in improving quality of the stacked virtual shot gather. Accordingly, the dispersion spectra resulting from selectively stacking virtual shot gathers produces a more continuous dispersion curve (Fig. S4). In contrast, stacking all virtual shot gathers leads to the presence of artefacts visible as local minima in the crest of dispersion mode.

In the further evaluation step, when working with DAS on dark fibres, recordings of virtual sensors located within loops of slack fibre must be removed from the record section. For this purpose, we searched the virtual shot gathers for spatial intervals with a discontinuous surface-wave moveout and faint amplitudes. Removing the corresponding traces yields a continues surface wave moveout and equidistantly spaced sensing nodes for the correct dispersion-spectra calculation. Exemplary shot gathers before and after removal of channels within loops of slack are shown in Figs S5(a) and (b), respectively.

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

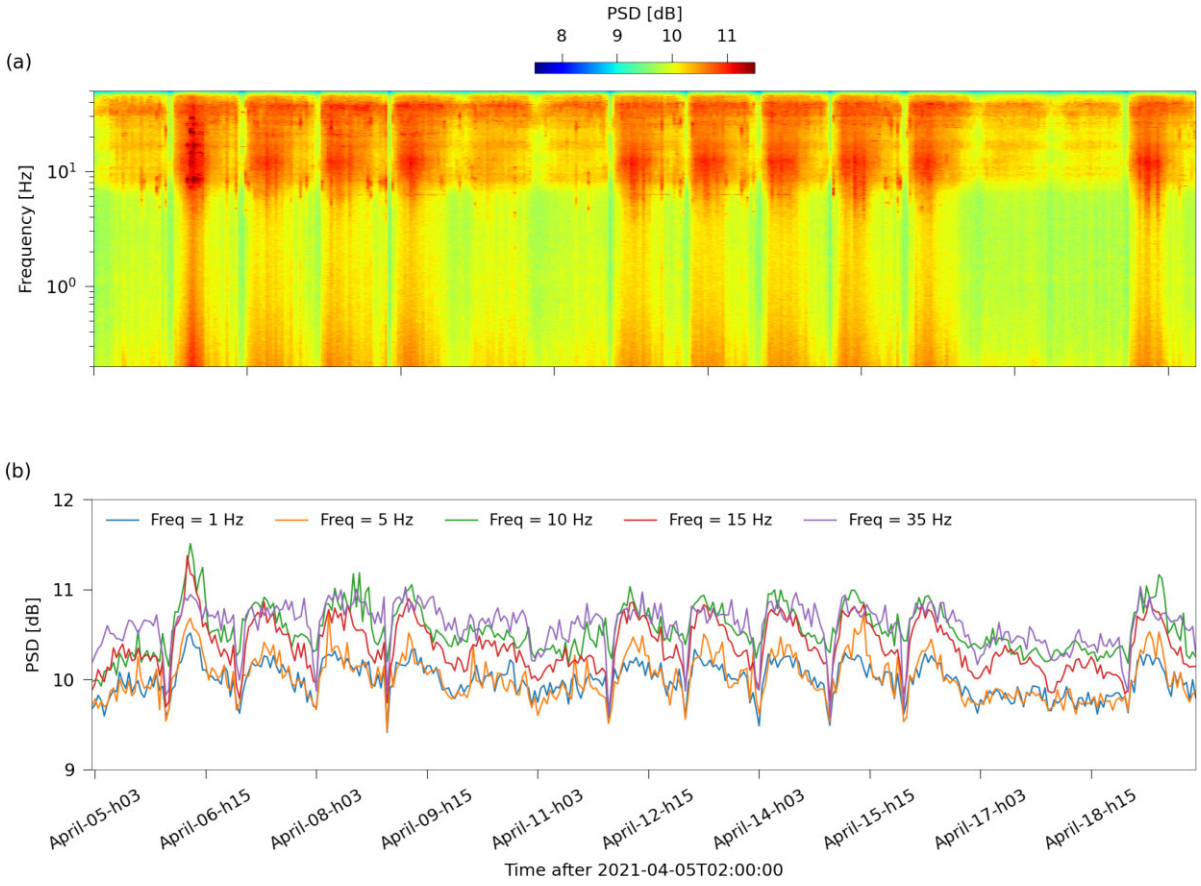

Figure 3. Power spectral density (PSD) for whole 15-d data (a) Spectrogram over a period of 15 d obtained from averaging PSDs of hour-long recordings across all channels of the targeted fibre segment (see Fig. 1). (b) The time-series of the PSDs variation for some distinct frequencies.

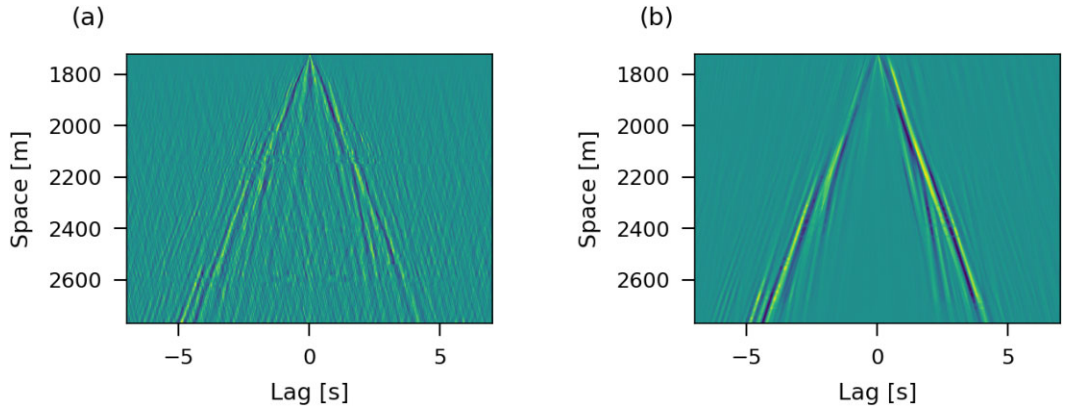

Figure 4. Virtual shot gathers (a) Stack of 212 selected virtual shot gathers after applying the conventional ambient-noise processing workflow of Bensen *et al.* (2007) to 1-hr-long ambient noise DAS recordings and removing the virtual sensors within loops of slack. (b) Result from subjecting (a) to coherence-based enhancement after (Schwarz 2019).

3.3 Improving the signal through coherence-based enhancement

At this point in the workflow, we introduce a coherence-based wavefield decomposition technique to further enhance the SNR of the stacked shot gather. Essentially, after decomposition of the wavefield into random noise and the spatially coherent wavefield, only the latter is kept for further processing (Schwarz 2019). The used method originates in exploration seismology where the systematic and repetitive use of controlled seismic sources and dense and regular receiver arrays allows for targeted data processing schemes that explicitly utilize waveform similarities. To quantify coherence in seismic records, typically, the semblance norm is used, which can be interpreted as a ratio of the coherent energy and the total energy contained in a wavefield (Neidell & Taner 1971). For fibre position x_0, recording time t_0, and the wavefield D, the semblance

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

S_k is computed for each local wave front slope p_k via:

$$S_k(x_0, t_0) = \frac{1}{N} \frac{\sum_{\delta t} \left[\sum_{i=1}^{N} D(x_0 + \Delta x_i, t_0 + p_k \Delta x_i) \right]^2}{\sum_{\delta t} \sum_{i=1}^{N} D^2(x_0 + \Delta x_i, t_0 + p_k \Delta x_i)}, \quad (2)$$

where the index i denotes a trace in the vicinity, Δx_i is its separation from the reference trace at location x_0, N is the total number of considered neighbouring traces, and δt is a centred time window whose extent should be in the order of the predominant signal period. The index k parametrizes the local wave front slope p and honours the fact that several wavefields with different inclination can intersect at data point (x_0, t_0).

If only a major wave train is of interest, the single dominant slope can be gained by locally maximizing eq. (2). If however, interference phenomena are important and need to be preserved, the semblance can be evaluated on a predefined horizontal slowness grid (Schwarz 2019; Jousset *et al.* 2022). As a result of the procedure, every data point (x_0, t_0) can be attributed with one or several semblance values, depending on whether the most dominant or a range of data slopes where investigated. Together with the slope field p_k, the coherent wavefield can be reconstructed by the semblance-weighted (multi-)directional stack:

$$D_{coher}(x_0, t_0) \propto \sum_k S(x_0, t_0) \sum_i D(x_0 + \Delta x_i, t_0 + p_k \Delta x_i). \quad (3)$$

When scattering phenomena are concerned, interference phenomena become important and the more general multidirectional reconstruction in eq. (3) should be considered (Jousset *et al.* 2022). If, however, surface-wave transmission (with a specific directionality) is at the centre of attention, the less dominant, conflicting terms in eq. (3) can be neglected, thus further boosting the enhancing effect on the outcomes. In the following, coherently enhanced results where gained following the second strategy.

3.4 MASW

MASW has been successfully used for the retrieval of shear-wave velocity models of the subsurface. The method has been tested in several constellations on the source side, that is both with active and passive experiments, as well as on the receiver side, that is with geophones (Park & Miller 2008; Foti *et al.* 2014; Cheng *et al.* 2016; Morton *et al.* 2021) and DAS (Dou *et al.* 2017; Tribaldos *et al.* 2021; Song *et al.* 2021). It involves three main processing steps, including dispersion spectra computation, picking of dispersive modes and inversion for a 1-D-velocity model through numerical optimization (Park *et al.* 1999).

Before calculating the dispersion spectra, the SNR of the virtual-shot gathers is further increased by averaging over causal (positive) and acausal (negative) parts of the extracted virtual shot gathers (Bensen *et al.* 2007). Slant-stacking of the phase shift is then used to map the virtual-shot gathers from the space–time to the frequency-phase velocity domain (Park *et al.* 1998; Xia *et al.* 2007; Dou *et al.* 2017). We then semi-automatically pick the maximum velocity for each frequency within a manually defined area around the fundamental mode of the resulting dispersion spectra.

The inversion of the picked dispersion curve to a 1-D shear-wave velocity model is done with the Geopsy software (Wathelet *et al.* 2020) which essentially performs a stochastic direct-search (Neighbourhood Algorithm; Sambridge 1999; Wathelet 2005) of velocity models whose layer thicknesses and velocities per layer are varied for a defined, constant number of discrete layers over a half-space. Through the forward modelling of velocity models and according dispersion curves, misfit values with respect to the measured dispersion curve are obtained. Except for the number of interfaces expected, the unknown model parameters during the inversion process include layer thickness, shear wave velocity, mass density and primary wave velocity.

Shear wave velocities and thicknesses are the most sensitive parameters (Foti *et al.* 2018; Dokht Dolatabadi Esfahani *et al.* 2020). Therefore, having some *a priori* knowledge of the subsurface can accelerate the inversion process by limiting the searched model space of each interface's depth and each layer's velocity. We use the geological model of the state of Brandenburg to extract a stratigraphic cross-section below our receiver array, that is the fibre trajectory along Adlergestell road (Schilling *et al.* 2018; Noack *et al.* 2013). The model is valid down to 200 m depth and includes the upper interface of the Rupelian clay. Boundaries of the main layers are located at depths of 60, 95 and 170 m (see Fig. 8c). Accordingly, we restrict the number of layers above the half-space to four. In addition, we choose a lower Poisson ratio value for the upper layers which are predominantly composed of sand. The boundaries for the tested shear wave velocities, layer thicknesses as well as the Poisson ratio for the common lithological structures of the site are presented in Table 1. The inversion process is affected by non-uniqueness, that is various models describe the same dispersion data with low misfits (Wathelet 2005). Following (Vantassel & Cox 2021; Mordret *et al.* 2014), we select the 1000 models with the lowest misfit out of 30 000 generated models, and use the distribution of layer thicknesses and velocities as a measure of confidence for our results (Fig. 7).

4 RESULTS

Virtual shot gathers result from the selective stacking and the same after application of the coherence-based enhancement after removal of all channels within the loops of slacks are shown in Figs 4(a) and (b). We calculate the channel-wise PSD of both the non-enhanced and enhanced virtual shot gathers (Fig. S6) and average across all channels (Fig. 5a). The results show that most of the energy retained by the coherence enhancement is concentrated below 4 Hz. We further analyse the frequency content of coherent and random wavefield components by using the same windows used for the SNR calculation described above (Fig. S2) to calculate the PSD over all channels for signal and noise separately. The average PSDs across all channels reveal that the coherent signal emerging after interferometric analysis is predominantly composed of frequencies below 4.5 Hz, whereas the random noise is characterized by elevated energy at higher frequencies (Fig. S7). Therefore, in the MASW process, we expect to resolve the dispersion curve in the frequency band of the signal window. The comparison of the channel-wise SNR of the stacked virtual shot gather before and after the application of the coherence-based enhancement in Fig. 5(b) shows the substantial improvement for all traces. Again the SNR was calculated following the procedure described earlier and eq. (1).

Similar results are obtained for fibre Segment B (Figs S9 and S10).

4.1 Dispersion analysis of convectional versus coherence-enhanced shot gathers

The dispersion spectra retrieved from virtual shot gathers without and with the applied coherence enhancement are illustrated in Figs 6(a) and (b), respectively. The dispersion spectra resulting from

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

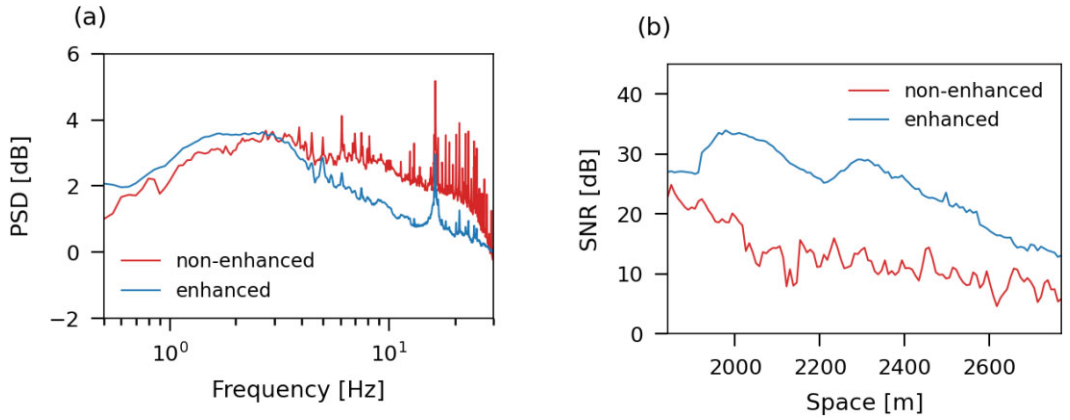

Figure 5. Comparison of enhanced and non-enhanced virtual shot gathers shown in Figs 4(a) and (b) (with and without application of the coherence-based enhancement procedure): (a) average PSD of all channels and (b) channel-wise SNR.

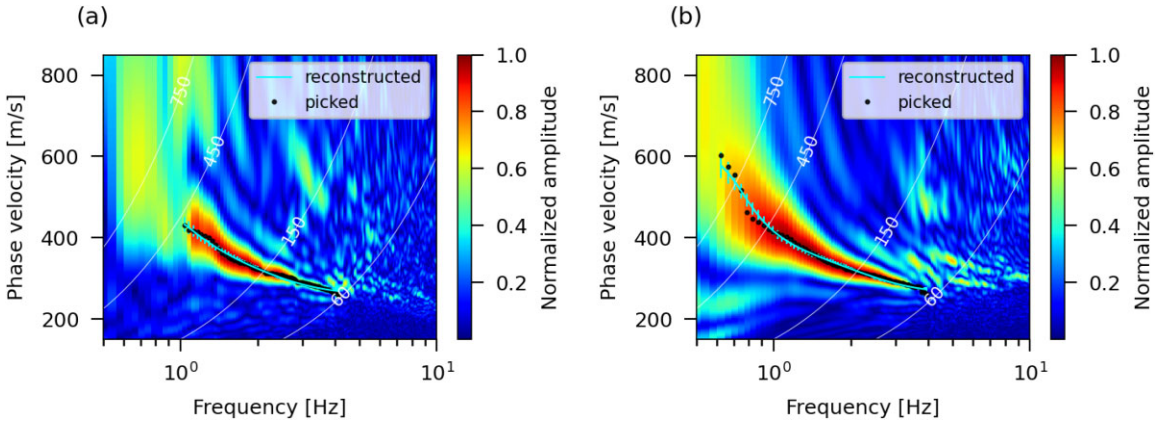

Figure 6. Rayleigh wave dispersion spectra resulting from slant stacking of (a) the non-enhanced virtual-shot gather and (b) the virtual-shot gather after coherence enhancement. Solid white lines represent contours of constant wavelength associated to respective pairs of velocity and frequency.

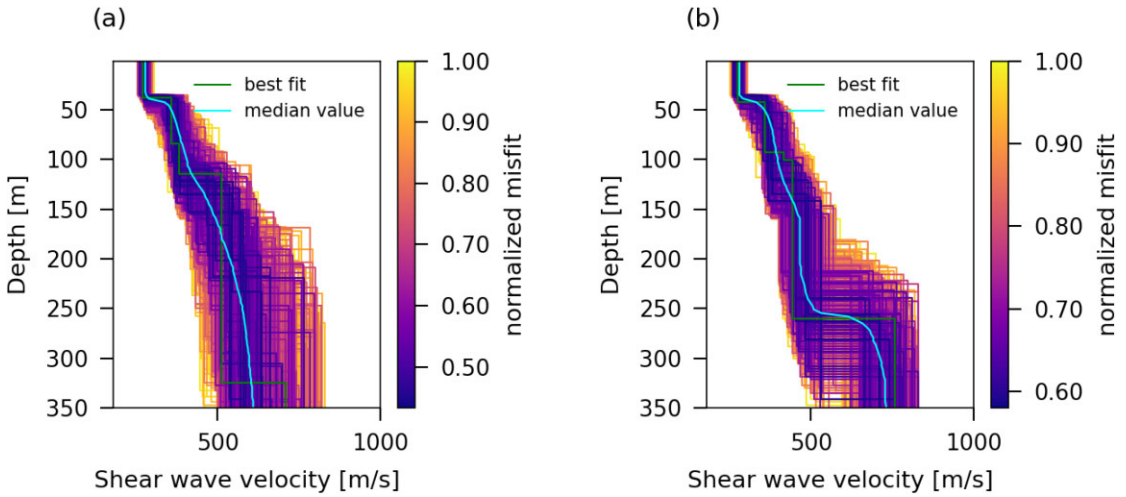

Figure 7. 1-D velocity models across the 1000 best-fitting models resulting from the inversion of the dispersion curves shown in Fig. 6, (a) from the non-enhanced and (b) for the coherence-enhanced shot gather.

the non-enhanced gather shows two regions of high amplitude energy. Between 1 and 4.5 Hz the fundamental mode can clearly be distinguished. Above this curve at higher velocities a possible higher dispersion mode can be distinguished as another streak of high amplitudes between 2.5 and 3.5 Hz. The dispersion spectra resulting after application of the coherence enhancement yields a continuous dispersion curve (Fig. 6b). The fundamental mode that was cut off at a lower bound of 1 Hz and 500 m s^{-1} in the non-enhanced result is now smoothly continued down to 0.67 Hz and 600 m s^{-1}. The comparison of the picked maxima of each fundamental mode is

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

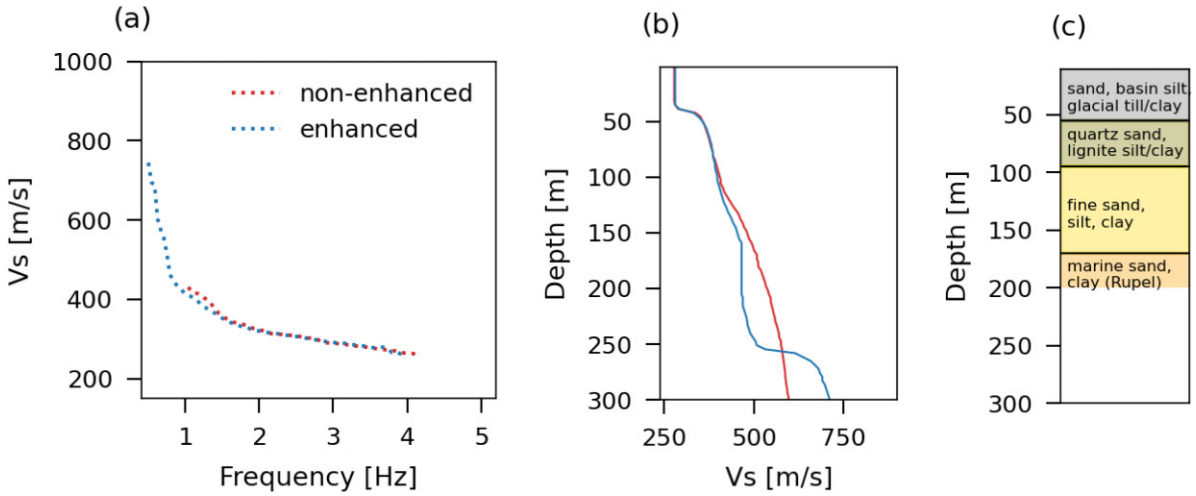

Figure 8. Comparison of the picked dispersion curves and velocity models with the geological profile: (a) Picked fundamental modes obtained from dispersion curves in Fig. 6. (b) Respective 1-D shear-wave velocity models (also see Fig. 7). (c) Geological profile based on the stratigraphic model of the geological survey of the state of Brandenburg (Schilling *et al.* 2018) and requested for the section along the optical fibre examined here. It models the first 200 m only.

Table 1. Search ranges of inversion parameters.

Layer	Interface depth (m)	*Vs* (m s^{-1})	Poisson ratio
1	[35–70]	[150–400]	[0.15–0.35]
2	[40–100]	[300–500]	[0.15–0.35]
3	[80–160]	[320–700]	[0.15–0.35]
4	[150–350]	[400–900]	[0.2–0.45]
5	half-space	[500–2000]	[0.2–0.5]

shown in Fig. 8(a). Solid white lines in Fig. 6 represent contours of constant wavelength associated to respective pairs of velocity and frequency. In terms of wavelength, the fundamental mode of the dispersion curve without enhancement has clearly readable maxima up to 450 m. In contrast, the enhanced curve is readable until 900 m. Likewise, we calculate dispersion spectra for the virtual shot gathers obtained for Segment B. The dispersion spectra in Fig. S11(b) result in continuous and improved fundamental mode for the enhanced workflow compared to the non-enhanced workflow Fig. S11(a).

4.2 Inverted *Vs* structure and comparison to geological data

The investigation depth depends on the maximum measured wavelength (conservatively, it corresponds to half of the maximum wavelength; Foti *et al.* 2018). With a maximum wavelength of 450 m for the non-enhanced case, this corresponds approximately to a depth of 225 m. The velocity profiles resulting from the inversion of the fundamental mode of the non-enhanced and enhanced dataset are shown in Fig. 7, revealing three distinct layers with increasing shear wave velocity from approximately 300–800 m s^{-1}. The green and cyan solid lines show the best-fitting and a median model, respectively, where the latter is calculated for every depth in 1 m steps from the median values of all velocities from the best 1000 models. The resulting velocity models (Fig. 8b) are in good agreement with boundaries delineated by the stratigraphic model of the state of Brandenburg (Fig. 8c).

We obtain very similar results for the Segment B presented in Fig. S12.

The distribution of the depths of each individual interfaces across the 1000 best accepted velocity models is shown in Figs 9(a) and (b). Vertical and horizontal whiskers and boxes shown on the right and top edge of each panel in Fig. 9 refer to the accordingly coloured distributions and indicate their extent as well as the extent of the second and third quartiles. As expected, the uncertainty of each interface increases with its depth for both of the resulting velocity models (without and with enhancement procedure). However, the depth uncertainty of the last layer is significantly reduced for the workflow including the coherence enhancement. This is expected, as the coherence-enhanced virtual shot gather constraints the fundamental dispersion mode significantly better at low-frequencies. The same holds true for the distribution of the last layer's velocity (Figs 9c and d). Similarly, for the example Segment B, the non-enhanced model produces a very wide velocity distribution with higher uncertainly (Fig. S13).

5 DISCUSSION

We applied a modified seismic interferometry workflow to urban ambient noise recordings acquired with DAS along a telecommunication fibre-optic cable. We present the results from two 1-km-long subsections of the fibre with 132 receivers. Here we discuss the maximum and minimum resolvable frequencies in the surface wave analysis and time duration required for stacking of cross-correlations.

5.1 Impact of coherence-based enhancement in stacking time duration

The time required for stacking of cross correlation function depends on the length of time series and the complexity of seismic noise being processed (Behm *et al.* 2014). In the presence of complex seismic noise such as urban traffic noise, a longer time series and more careful stacking are needed to obtain accurate results. As mentioned earlier for the example Segment A, the quality of the signal improves after selective stacking of 212 of 360 virtual shot gathers. We investigate the quality of dispersion spectra as a function of the number of stacked virtual shot gathers with and without applying

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

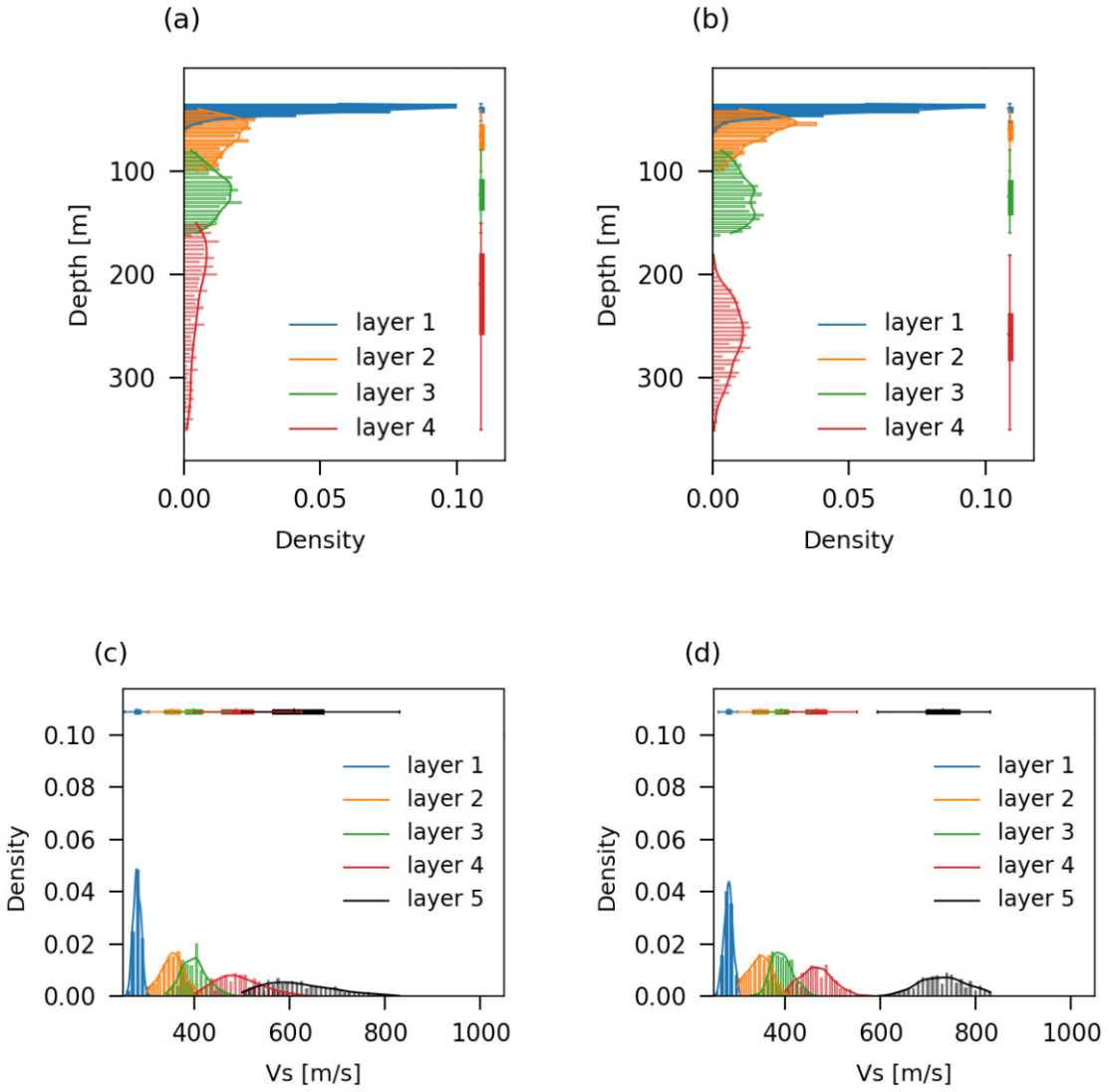

Figure 9. Distribution of the interface depths and the layer velocities across the 1000 best-fitting velocity models: (a) and (c) for the non-enhanced shot gather and (b) and (d) for the enhanced shot gather. Vertical and horizontal whiskers and boxes shown on the right and top edge of each panel refer to the accordingly coloured distributions and indicate their extent as well as the extent of the second and third quartiles.

the coherence enhancement to the final stack. The dispersion spectra obtained after stacking 6, 24, 48, 72, 96 and 150 hr without and with subsequent enhanced workflow are depicted in Fig. S14 (first and third as well as second and fourth row, respectively). Whereas the dispersion spectra without enhancement remains fragmented until the stacking of at least 72 gathers, a continuous dispersion curve can be achieved already after stacking 48 gathers if the coherence enhancement is applied. Without coherence enhancement, the extension of the fundamental mode to frequencies smaller than 1 Hz can only be distinguished after 150 stacks. To achieve the same result with coherence enhancement, 48 stacks are sufficient. Moreover, the application of coherence enhancement provides more prominent and continuous fundamental dispersion energy, allowing for more confidence in picking the dispersion curve.

To investigate this observation quantitatively, we use the spectral root-mean-square deviation (RMSD) between each iteration of stacking (i) as a measure of differences between two consecutive stacks. In particular, we analyse the reduction in RMSD with increasing stacking duration, where N is the number of channels, M is the number of frequency samples and $F(\omega_n, s_m)$ is the absolute energy of the virtual shot gather for the frequency sample m and channel n.

$$RMSD(i) = \sqrt{\frac{\sum_n^N \sum_m^M (F_i(\omega_n, s_m) - F_{i-1}(\omega_n, s_m))^2}{N \cdot M}}. \quad (4)$$

The convergence to a stable value of RMSD indicates that the quality of the stack cannot be substantially improved further by adding more virtual shot gathers to the stack. We compare the RMSD convergence of consecutive stacks without and with coherence enhancement in Fig. S15. While both RMSD series decrease with the increasing number of stacks, the enhanced scheme yields higher RMSD values and results into a faster convergence to a stable RMSD value. For the coherence-enhances scheme, a stable RMSD is obtained after approximately 48 stacks, whereas RMSD values keep decreasing after over 80 stacks for the non-enhanced scheme. As for the quality of the dispersion spectra described above and shown in Fig. S14, a faster RMSD convergence indicates that less data is required to obtain good quality virtual-shot gathers for the coherence-based enhancement approach compared to the non-enhancement scheme.

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

To further assess the efficiency of our newly developed workflow, we also investigate the impact of selecting high-quality virtual-shot gathers for stacking, as opposed to staking all available gathers. When comparing the quality of dispersion curves between the selective stack of 212 high-quality gathers (Fig. 6b) and the stack of all 360 available gathers after coherence enhancement (same as Fig. 6b), we observe striking similarities in the resultant dispersion spectra. This can be attributed to the notably higher proportion of good-quality gathers within the dataset. Thus, we systematically investigate the impact of selective stacking by comparing the stack of all gathers with that of selected ones over a consecutive and increasing time interval (24, 48, 72 and 96 hr, see Fig. S16). Through this, we simulate scenarios with limited data availability with an increasing ratio of high- to low-quality gathers over the growing recording period (10/14, 21/27, 39/33 and 61/35). Regardless of whether virtual shot gathers received enhancement, we observe that selective stacking consistently enhances the quality of dispersion spectra (Fig. S16). Although coherence enhancement consistently improves dispersion spectra quality, artefacts and discontinuities in dispersion spectra at shorter recording durations (characterized by a low ratio of good-to-bad quality gathers) impact the enhancement result and remain visible as discontinuities and fluctuations of the fundamental mode.

5.2 Influence of the acquisition, geometry and geology on the maximum resolvable frequency

It is fundamental to the DAS measurement principle that seismic wavelengths smaller than the gauge length cannot be resolved properly (Vantassel *et al.* 2022). Therefore, depending on the wavelengths of interest, DAS measurements with a sufficiently small gauge length must be considered. As the gauge length affects the efficiency of the optical interferometry during the analysis of the backscattered light, and thus the resulting SNR, a trade-of between the optimum value for the gauge length and a required sensitivity must be found (Dean *et al.* 2017). Generally, a gauge length smaller than half of the minimum wavelength is recommended for optimum measurements with no spatial aliasing.

In our case, the gauge-length is 10 m. Based on the velocity structure revealed by our analysis, our DAS set up should be able to resolve frequencies up to approximately 14 Hz. However, we observe a maximum resolvable frequency of the fundamental mode in the dispersion curve of approximately 4.2 Hz for both workflows (non-enhanced and enhanced). The channel spacing of 8 m is smaller than the gauge length and thus would also allow for the resolution of higher frequencies. The observed upper frequency bound cannot be explained by limitations of the acquisition system or the measurement geometry. Therefore, the upper bound of resolvable frequencies in the dispersion spectra is likely related to attenuation of the higher frequency components of the coherent seismic energy that contributes to the cross-correlations. In this area, the noise sources that mainly contribute to the coherent cross-correlation most likely originate from traffic surrounding the study area beyond the far ends of the cable. Although these noise sources are homogeneously distributed, their high-frequency content attenuates and we only observe their contribution to the low-frequency part of cross-correlation (<5 Hz). Additionally, the high-frequency noise sources close to the cable are more complex and not necessarily homogeneously distributed. Therefore, they do not coherently contribute to the cross-correlations.

This hypothesis is supported by the spatially averaged PSDs for the non-enhanced and enhanced stacks, calculated separately for signal and noise windows (Fig. S7). Compared to the noise, the frequency content of the signal can clearly be distinguished in the non-enhanced gather, as the spectral level is significantly elevated below a bound at slightly above 4 Hz. The same spectral analysis for the enhanced gather shows how the frequency band below and above this boundary are amplified and damped, respectively, thus confirming that the seismic wavefield at higher frequencies does not travel coherently in space.

5.3 Efficiency of the enhancement with respect to maximum depth of investigation and its application in urban geothermal operation

The application of the coherence-based enhancement had three positive effects. First, it improved the SNR significantly (Fig. S7) by separating signal from non-coherent noise components, despite the strongly diminished database of recordings due to selective stacking. This demonstrates the efficiency of our enhancement approach in boosting coherent signal and suppressing random noise. Secondly, the low-frequency information of the shot gather considerably improved, thus amplifying the dispersion curve in this frequency band and resulting in an increased depth of investigation.

These findings shed light on the shallow subsurface structure and provided valuable insights into the structure and composition of the subsurface. Therefore, the widespread availability of fibre optic cables below the city, and the use of urban ambient noise as a source of seismic energy along with the improvement of seismic imaging through our newly developed enhancement of low-frequency signals at a shorter time and lower cost, allow us to apply the method to other urban regions for preliminary mapping of subsurface conditions. This could be used in many aspects of urban planning and development, including infrastructure assessment and design, geothermal development and hazard mitigation.

6 CONCLUSION

Our study shows the feasibility of deriving a shear wave velocity structure down to a depth of 350 m from the ambient seismic wavefield recorded by a fibre-optic telecommunication cable parallel to a busy major road in a noisy urban area. After selective stacking, an increased investigation depth was achieved by denoising the stacked virtual-shot gather with a coherence-based enhancement method. The application of the proposed workflow on two different subsections of the array confirms the effectiveness of the workflow in improving the low-frequency part of dispersion spectra and accordingly increase investigation depth. In the process of selective stacking, numerous shot gathers had to be discarded due to spurious signals overprinting the expected surface wave moveout. Although this procedure resulted in a reduction of the collected data by 41 per cent, it improved the final SNR.

Our results suggests that, at our study site, traffic noise generated beyond the end of the array is sufficiently homogeneous to contribute to the cross-correlation and thus allow for MASW. These sources are mostly traffic noise emitting Rayleigh waves of broad frequency range. However, only Rayleigh waves with a frequency range below $\sim$4.2 Hz that coherently travel along the fibre are resolved, and the high-frequency components of the noise are attenuated. Our results demonstrate that our coherence-based denoising scheme decreases the amount of data required to obtain a high quality dispersion curve and thus a robust velocity model. Compared to the workflow

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

without coherence enhancement, the workflow including enhancement requires significantly less stacks of virtual shot gathers. These results indicate that the application of the coherence-enhancement scheme enables the possibility of significantly shortened campaign duration. This plays a significant role in decreasing both storage and computational expenses, which are critical issues for the successful and routine application of DAS for subsurface exploration. This reduction in the amount of data needed to obtain robust results also has important implications for time-lapse monitoring studies, since it will allow higher temporal resolution.

The resulting velocity models agree well with general geologic data for the study site. The obtained shear wave velocity model is capable of retrieving layers down to a depth of 350 m, where the deepest layer likely represents the Rupelian clay. Imaging this layer in the North German Basin is an important step with respect to the development of geothermal projects as it helps to estimate the required drilling depth as well as the potential impact on ground water resources.

Depiction of larger and smaller wavelengths which are sensitive to the deeper and shallower layers, respectively, could be further improved by better understanding the distribution and origin of noise sources. Further important insights could be gained from an extension of our study to further fibre segments thus yielding a continuous 2-D shear-wave velocity model (e.g. Cheng *et al.* 2023) as well as its applicability for large scale investigation. However, this step requires an automatic identification and exclusion of incoherent noisy patterns in either raw data or processed virtual shot gathers and will be addressed in a separate study.

ACKNOWLEDGMENTS

This study was supported by German Federal Ministry for Economic Affairs and Climate Action (BMWK) under grant number 03EE4009A as a part of project SENSE. We thank Ulf Kociemba (Senior Key Account Manager, Large Enterprises Nord/Ost) from 1&1 Versatel Deutschland GmbH, Aroser Allee 78, D-13407 Berlin (Ulf.Kociemba@1und1.net, www.1und1.net) for advice and support while choosing and interrogating the fibre, as well as for giving permission to publish the information related to the cable. We thank Maximilian Frick, Ben Norden and Ali Saadat from GFZ for their advice and support. The authors declare no competing interests.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

suppl_data

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.

DATA AVAILABILITY

The lithology data were extracted from http://berlin.geo-3d.de/berlin3d/portal. Python was used for data analysis and plotting Beyreuther *et al.* (2010) and Hunter (2007). Geopsy package 3.4.2 (https://www.geopsy.org) was used for the inversion process. The map and the insets in Figure1 are based on data from (OpenStreetMap contributors 2017) and were downloaded with contextily-python API (https://contextily.readthedocs.io) where the insets use map tiles by Stamen Design, under CC BY 4.0. Both stacked virtual shot gathers (enhanced and non-enhanced), all 360 virtual shot gathers (including those shown in Fig. S1), picked dispersion curves (Fig. 6), velocity models (Fig. 7) as well as all subsets of hourly decimated strain-rate recordings associated with Segment A shown in Fig. 1 and list of channels within loops of slack are available online (at https://doi.org/10.5880/GFZ.2.2.2023.003).

REFERENCES

Ajo-Franklin, J.B. *et al.*, 2019. Distributed acoustic sensing using dark fibre for near-surface characterization and broadband seismic event detection, *Sci. Rep.,* **9**(1), doi:10.1038/s41598-018-36675-8.

Asten, M.W., 2006. On bias and noise in passive seismic data from finite circular array data processed using SPAC methods, *Geophysics,* **71**(6), V153–V162.

Atterholt, J., Zhan, Z., Shen, Z. & Li, Z., 2022. A unified wavefield-partitioning approach for distributed acoustic sensing, *Geophys. J. Int.,* **228**(2), 1410–1418.

Behm, M., Leahy, G.M. & Snieder, R., 2014. Retrieval of local surface wave velocities from traffic noise–an example from the La Barge basin (Wyoming), *Geophys. Prospect.,* **62**(2), 223–243.

Bensen, G., Ritzwoller, M. & Shapiro, N.M., 2008. Broadband ambient noise surface wave tomography across the United States, *J. geophys. Res.,* **113**(B5), doi:10.1029/2007JB005248.

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**(3), 1239–1260.

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.

Chen, Y., Zhang, M., Bai, M. & Chen, W., 2019. Improving the signal-to-noise ratio of seismological datasets by unsupervised machine learning, *Seismol. Res. Lett.,* **90**(4), 1552–1564.

Cheng, F. *et al.*, 2016. Multichannel analysis of passive surface waves based on crosscorrelations, *Geophysics,* **81**(5), EN57–EN66.

Cheng, F., Xia, J., Behm, M., Hu, Y. & Pang, J., 2019. Automated data selection in the tau–p domain: application to passive surface wave imaging, *Surv. Geophys.,* **40**(5), 1211–1228.

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

Cheng, F., Ajo-Franklin, J.B., Nayak, A., Tribaldos, V.R., Mellors, R. & Dobson, P.& Team I. V. D.F., 2023. Using dark fibre and distributed acoustic sensing to characterize a geothermal system in the Imperial Valley, Southern California, *J. geophys. Res.,* **128**(3), e2022JB025240.

Clayton, R.W., 2020. Imaging the subsurface with ambient noise autocorrelations, *Seismol. Res. Lett.,* **91**(2A), 930–935.

Dean, T., Cuny, T. & Hartog, A.H., 2017. The effect of gauge length on axially incident P-waves measured using fibre optic distributed vibration sensing, *Geophys. Prospect.,* **65**(1), 184–193.

Dokht Dolatabadi Esfahani, R., Gholami, A. & Ohrnberger, M., 2020. An inexact augmented Lagrangian method for nonlinear dispersion-curve inversion using dix-type global linear approximation, *Geophysics,* **85**(5), EN77–EN85.

Dou, S. *et al.*, 2017. Distributed acoustic sensing for seismic monitoring of the near surface: a traffic-noise interferometry case study, *Sci. Rep.,* **7**(1), doi:10.1038/s41598-017-11986-4.

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

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.

Frick, M., Sippel, J., Cacace, M. & Scheck-Wenderoth, M., 2016. The geothermal field below the city of Berlin, Germany: results from structurally and parametrically improved 3D models, *Ener. Proc.,* **97,** 334–341. doi: 10.1016/j.egypro.2016.10.011.

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

Frick, M., Scheck-Wenderoth, M., Schneider, M. & Cacace, M., 2019. Surface to groundwater interactions beneath the city of Berlin: results from 3D models, *Geofluids,* **2019**(2), 1–22.

Götz, J., Lüth, S., Henninges, J. & Reinsch, T., 2018. Vertical seismic profiling using a daisy-chained deployment of fibre-optic cables in four wells simultaneously–case study at the ketzin carbon dioxide storage site, *Geophys. Prospect.,* **66**(6), 1201–1214.

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

Henninges, J., Martuganova, E., Stiller, M., Norden, B. & Krawczyk, C.M., 2021. Wireline distributed acoustic sensing allows 4.2 km deep vertical seismic profiling of the Rotliegend 150 c geothermal reservoir in the North German Basin, *Solid Earth,* **12**(2), 521–537.

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

Isken, M.P., Wollin, C., Heimann, S., Quinteros, J., Jäckel, K.-H. & Jousset, P., 2021. DAS Convert - Convert distributed acoustic sensing data. V. 1.0, GFZ Data Services, Potsdam. doi: 10.5880/GFZ.2.1.2021.005.

Isken, M.P., Vasyura-Bathke, H., Dahm, T. & Heimann, S., 2022. Denoising distributed acoustic sensing data using an adaptive frequency–wavenumber filter, *Geophys. J. Int.,* **231**(2), 944–949.

Jiang, C. & Denolle, M.A., 2022. Pronounced seismic anisotropy in Kanto sedimentary basin: a case study of using dense arrays, ambient noise seismology, and multi-modal surface-wave imaging, *J. geophys. Res.,* **127**(8), 1–28.

Jousset, P. *et al.*, 2018. Dynamic strain determination using fibre-optic cables allows imaging of seismological and structural features, *Nat. Commun.,* **9**(1), 1–11.

Jousset, P. *et al.*, 2022. Fibre optic distributed acoustic sensing of volcanic events, *Nat. Commun.,* **13**(1), 1753 , doi:10.1038/s41467-022-29184-w.

Klaasen, S., Paitz, P., Lindner, N., Dettmer, J. & Fichtner, A., 2021. Distributed acoustic sensing in volcano-glacial environments—Mount Meager, British Columbia, *J. geophys. Res.,* **126**(11), e2021JB022358, doi:10.1029/2021JB022358.

Krawczyk, C., Jousset, P. & Reinsch, T., 2019a. Fibre-optic strain sensing: Game changer for (urban) seismic surveying?, in *Proceedings of the 1st Conference on Geophysics for Infrastructure Planning Monitoring and BIM,* The Hague, Netherlands, 8–12 September 2019, doi:10.3997/2214-4609.201902547.

Krawczyk, C.M., Stiller, M., Bauer, K., Norden, B., Henninges, J., Ivanova, A. & Huenges, E., 2019b. 3-D seismic exploration across the deep geothermal research platform groß schönebeck north of berlin/germany, *Geotherm. Ener.,* **7**(1), 1–18.

Lapins, S., Butcher, A., Kendall, J.-M., Hudson, T. S., Stork, A. L., Werner, M. J., Gunning, J. & Brisbourne., A. M., 2024. Das-n2n: machine learning distributed acoustic sensing (das) signal denoisingwithout clean data, *Geophys. J. Int.,* **236**(2), 1026–1041.

Li, Y., Perton, M., Gaite, B., Ruiz-Barajas, S. & Spica, Z.J., 2023. Near-surface characterization using distributed acoustic sensing in an urban area: Granada, Spain, *Geophys. J. Int.,* **235**(2), 1849–1860.

Lindsey, N.J., Martin, E.R., Dreger, D.S., Freifeld, B., Cole, S., James, S.R., Biondi, B.L. & Ajo-Franklin, J.B., 2017. Fibre-optic network observations of earthquake wavefields, *Geophys. Res. Lett.,* **44**(23), 11 711–11 799.

Lindsey, N.J., Rademacher, H. & Ajo-Franklin, J.B., 2020. On the broadband instrument response of fibre-optic DAS arrays, *J. geophys. Res.,* **125**(2), e2019JB018145, doi:10.1029/2019JB018145.

Lipus, M.P., Reinsch, T., Weisenberger, T.B., Kragset, S., Stefánsson, A. & Bogason, S.G., 2021. Monitoring of a reverse cement job in a high-temperature geothermal environment, *Geotherm. Ener.,* **9**(1), 5 , doi:10.1186/s40517-021-00187-y.

Martuganova, E., Stiller, M., Bauer, K., Henninges, J. & Krawczyk, C.M., 2021. Cable reverberations during wireline distributed acoustic sensing measurements: their nature and methods for elimination, *Geophys. Prospect.,* **69**(5), 1034–1054.

Martuganova, E., Stiller, M., Norden, B., Henninges, J. & Krawczyk, C.M., 2022. 3D deep geothermal reservoir imaging with wireline distributed acoustic sensing in two boreholes, *Solid Earth,* **13**(8), 1291–1307.

Mordret, A., Landés, 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**(3), 1514–1525.

Morton, S.L., Ivanov, J., Peterie, S.L., Miller, R.D. & Livers-Douglas, A.J., 2021. Passive multichannel analysis of surface waves using 1D and 2D receiver arrays, *Geophysics,* **86**(6), doi:10.1190/geo2020-0104.1.

Nakata, N., Snieder, R., Tsuji, T., Larner, K. & Matsuoka, T., 2011. Shear wave imaging from traffic noise using seismic interferometry by cross-coherence, *Geophysics,* **76**(6), SA97–SA106.

Nakata, N., Chang, J.P., Lawrence, J.F. & Boué, P., 2015. Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry, *J. geophys. Res.,* **120**(2), 1159–1173.

Neidell, N.S. & Taner, M.T., 1971. Semblance and other coherency measures for multichannel data, *Geophysics,* **36**(3), 482–497.

Noack, V., Scheck-Wenderoth, M., Cacace, M. & Schneider, M., 2013. Influence of fluid flow on the regional thermal field: results from 3d numerical modelling for the area of Brandenburg (North German Basin), *Environ. Earth Sci.,* **70,** 3523–3544. doi: 10.1007/s12665-013-2438-4.

Norden, B., Förster, A., Vu-Hoang, D., Marcelis, F., Springer, N. & Le Nir, I., 2010. Lithological and petrophysical core-log interpretation in CO_2SINK, the European CO_2 onshore research storage and verification project, *SPE Res. Eval. Eng.,* **13**(02), 179–192.

OpenStreetMap contributors, 2017. *Planet dump retrieved from https://planet.osm.org*. https://www.openstreetmap.org.

Park, C.B. & Miller, R.D., 2008. Roadside passive multichannel analysis of surface waves (MASW), *J. Environ. Eng. Geophys.,* **13**(1), 1–11.

Park, C.B., Miller, R.D., Xia, J. & Survey, K.G., 1998. Imaging dispersion curves of surface waves on multi-channel record 1998, in *SEG Expanded Abstracts Imaging dispersion curves of surface waves 1998 SEG Expanded Abstracts,* Society of Exploration Geophysicists. https://doi.org/10.1190/1.1820161.

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

Retailleau, L. & Beroza, G.C., 2021. Towards structural imaging using seismic ambient field correlation artefacts, *Geophys. J. Int.,* **225**(2), 1453–1465.

Ryberg, T., Muksin, U. & Bauer, K., 2016. Ambient seismic noise tomography reveals a hidden caldera and its relation to the Tarutung pull-apart basin at the Sumatran Fault Zone, Indonesia, *J. Volc. Geotherm. Res.,* **321,** 73–84.

Sabra, K.G., Gerstoft, P., Roux, P., Kuperman, W. & Fehler, M.C., 2005. Surface wave tomography from microseisms in southern california, *Geophys. Res. Lett.,* **32**(14), doi:10.1029/2005GL023155.

Sambridge, M., 1999. Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space, *Geophys. J. Int.,* **138**(2), 479–494.

Schilling, M., Simon, A., Jahnke, C. & Höding, T., 2018. Brandenburg 3D – DAS geologische 3D modell brandenburgs im internet veröffentlicht, *Brandenburg. geowiss. Beitr.,* **25**(1/2), 39–46.

Schwarz, B., 2019. Coherent wavefield subtraction for diffraction separation, *Geophysics,* **84**(3), V157–V168.

Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science,* **307**(5715), 1615–1618.

Song, Z., Zeng, X., Xie, J., Bao, F. & Zhang, G., 2021. Sensing shallow structure and traffic noise with fibre-optic internet cables in an urban area, *Surv. Geophys.,* **42**(6), 1401–1423.

Spica, Z.J., Perton, M., Martin, E.R., Beroza, G.C. & Biondi, B., 2020. Urban seismic site characterization by fibre-optic seismology, *J. geophys. Res.,* **125**(3), e2019JB018656, doi:10.1029/2019JB018656.

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.

Stackebrandt, W. & Franke, D., 2015. *Geologie von Brandenburg,* Schweizerbart. ISBN 978-3-510-65295-2.

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

Taweesintananon, K., Landrø, M., Brenne, J.K. & Haukanes, A., 2021. Distributed acoustic sensing for near-surface imaging using submarine telecommunication cable: a case study in the Trondheimsfjord, Norway, *Geophysics,* **86**(5), doi:10.1190/geo2020-0834.1.

Tribaldos, V.R. *et al.*, 2021. Surface wave imaging using distributed acoustic sensing deployed on dark fibre: moving beyond high-frequency noise, in *Distributed Acoustic Sensing in Geophysics: Methods and Applications,* pp. 197–212, eds Li, Y., Karrenbach, M. & Ajo-Franklin, J. B., Geophysical Monograph Series, AGU.

van den Ende, M., Lior, I., Ampuero, J.-P., Sladen, A., Ferrari, A. & Richard, C., 2021. A self-supervised deep learning approach for blind denoising and waveform coherence enhancement in distributed acoustic sensing data, *IEEE Trans. Neural Networks Learn. Syst.,* **34**(7), 3371–3384.

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., 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.

Wathelet, M., 2005. *Array Recordings of Ambient Vibrations: Surface-Wave Inversion,* pp. 1–177, Liège University. https://www.geopsy.org/publi/Wathelet2005.pdf.

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.

Williams, E.F., Fernández-Ruiz, M.R., Magalhaes, R., Vanthillo, R., Zhan, Z., González-Herráez, M. & Martins, H.F., 2019. Distributed sensing of microseisms and teleseisms with submarine dark fibres, *Nat. Commun.,* **10**(1), 1–11.

Xia, J., Xu, Y. & Miller, R.D., 2007. Generating an image of dispersive energy by frequency decomposition and slant stacking, *Pure appl. Geophys.,* **164**(5), 941–956.

Yang, J., Zhou, J., Zhang, H., Xu, T., Deng, D. & Geng, J., 2023a. Revealing the shallow soil structure of the Yigong Lake in the Tibetan Plateau using a portable distributed acoustic sensing interrogator, *Front. Earth Sci.,* **10,** doi:10.3389/feart.2022.1018116.

Yang, L., Fomel, S., Wang, S., Chen, X., Chen, W., Saad, O.M. & Chen, Y., 2023b. Denoising of distributed acoustic sensing data using supervised deep learning, *Geophysics,* **88**(1), WA91–WA104.

Yang, Y., Atterholt, J.W., Shen, Z., Muir, J.B., Williams, E.F. & Zhan, Z., 2022. Sub-kilometer correlation between near-surface structure and ground motion measured with distributed acoustic sensing, *Geophys. Res. Lett.,* **49**(1), doi:10.1029/2021GL096503.

Zhang, K., Li, H., Wang, X. & Wang, K., 2020. Retrieval of shallow S-wave profiles from seismic reflection surveying and traffic-induced noise, *Geophysics,* **85**(6), EN105–EN117.

Zhou, W., Butcher, A., Brisbourne, A.M., Kufner, S.-K., Kendall, J.-M. & Stork, A.L., 2022. Seismic noise interferometry and distributed acoustic sensing (DAS): inverting for the firn layer S-velocity structure on rutford ice stream, Antarctica, *J. geophys. Res.,* **127**(12), e2022JF006917.

Zhu, T. & Stensrud, D.J., 2019. Characterizing thunder-induced ground motions using fibre-optic distributed acoustic sensing array, *J. geophys. Res.,* **124**(23), 12 810–12 823.

Downloaded from https://academic.oup.com/gji/article/237/3/1751/7644363 by GFZ Potsdam user on 12 July 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of The Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.