Comparison of speaker recognition approaches for real applications

Geophysical Journal International

Geophys. J. Int. (2024) **237,** 506–524
Advance Access publication 2024 February 8
GJI Applied and Marine Geophysics

https://doi.org/10.1093/gji/ggae051

Novel techniques for *in situ* estimation of shear-wave velocity and damping ratio through MASW testing – I: a beamforming procedure for extracting Rayleigh-wave phase velocity and phase attenuation

Mauro Aimar [1], Sebastiano Foti[1] and Brady R. Cox[2]

[1]*Department of Structural, Building and Geotechnical Engineering (DISEG), Politecnico di Torino, Corso Duca degli Abruzzi* 24, 10129 *Torino, Italy.*
E-mail: mauro.aimar@polito.it
[2]*Department of Civil and Environmental Engineering, Utah State University,* 4110 *Old Main Hill, Logan, UT* 84321-4110*, USA*

Accepted 2024 February 1. Received 2024 January 10; in original form 2023 March 24

SUMMARY

A robust, *in situ* estimate of shear-wave velocity V_S and the small-strain damping ratio D_S (or equivalently, the quality factor Q_S) is crucial for the design of buildings and geotechnical systems subjected to vibrations or earthquake ground shaking. A promising technique for simultaneously obtaining both V_S and D_S relies on the Multichannel Analysis of Surface Waves (MASW) method. MASW can be used to extract the Rayleigh wave phase velocity and phase attenuation data from active-source seismic traces recorded along linear arrays. Then, these data can be inverted to obtain V_S and D_S profiles. This paper introduces two novel methodologies for extracting the phase velocity and attenuation data. These new approaches are based on an extension of the beamforming technique which can be combined with a modal filter to isolate different Rayleigh propagation modes. Thus, the techniques return reliable phase velocity and attenuation estimates even in the presence of a multimode wavefield, which is typical of complex stratigraphic conditions. The reliability and effectiveness of the proposed approaches are assessed on a suite of synthetic wavefields and on experimental data collected at the Garner Valley Downhole Array and Mirandola sites. The results reveal that, under proper modelling of wavefield conditions, accurate estimates of Rayleigh wave phase velocity and attenuation can be extracted from active-source MASW wavefields over a broad frequency range. Eventually, the estimation of soil mechanical parameters also requires a robust inversion procedure to map the experimental Rayleigh wave parameters into soil models describing V_S and D_S with depth. The simultaneous inversion of phase velocity and attenuation data is discussed in detail in the companion paper.

Key words: Elasticity and anelasticity; Fourier analysis; Seismic attenuation; Wave propagation.

1 INTRODUCTION

The characterization of soil response to dynamic loading has great relevance in geotechnical earthquake engineering. Key parameters to describe the stress-strain response of soils are the small-strain shear-wave velocity V_S and damping ratio D_S, which quantify respectively the stiffness and the internal energy dissipation by the soil at low strains (e.g. Foti *et al.* 2021). In geophysics and seismology, D_S is commonly replaced by the quality factor Q_S, defined as $Q_S = 1/(2D_S)$. However, for the sake of simplicity, the remainder of this paper will refer only to D_S. A promising approach to derive *in situ* estimates of these quantities relies on the Multichannel Analysis of Surface Waves (MASW; Foti 2000). Indeed, MASW can provide a reliable estimate of mechanical parameters for design purposes, as it investigates the soil behaviour in undisturbed conditions at a spatial scale compatible with the geotechnical application of interest (e.g. Comina *et al.* 2011; Passeri 2019).

MASW relies on the dispersive behaviour of Rayleigh waves (or *R*-waves) in layered media, for which the corresponding phase velocity V_R (i.e. the propagation speed) and the phase attenuation α_R (i.e. the spatial amplitude decay) exhibit a frequency dependence. The frequency-dependence of propagation parameters is a combined effect of geometric dispersion, which results from the variation of soil mechanical properties with depth, and intrinsic dispersion, due to the constitutive behaviour of linear viscoelastic media. The frequency-dependence is described by means of dispersion and attenuation data, hereafter labelled as $V_R(\omega)$ and $\alpha_R(\omega)$. Therefore, the MASW-based estimate of V_S and D_S usually refers to

the measurement of the spatial phase lag and attenuation of Rayleigh waves along linear arrays with active sources, from which $V_R(\omega)$ and $\alpha_R(\omega)$ are retrieved (Foti *et al.* 2014). Then, the V_S and the D_S profiles with depth are jointly estimated through an inversion scheme, where a theoretical soil model is calibrated to match the experimental $V_R(\omega)$ and $\alpha_R(\omega)$.

This paper deals with the estimation of Rayleigh-wave propagation parameters from the recorded waveform, whereas the derivation of ground models from experimental data is discussed in the companion paper. Manyfold procedures are available for the multimode estimation of $V_R(\omega)$ (e.g. McMechan & Yedlin 1981; Park *et al.* 1998; Forbriger 2003; Zywicki & Rix 2005; Luo *et al.* 2008; Askari & Ferguson 2012). Instead, the combined estimation of $V_R(\omega)$ and $\alpha_R(\omega)$ is a non-trivial task, especially in the presence of complex stratigraphy, and large uncertainties affect the estimated values. Indeed, only a few methods for jointly deriving $V_R(\omega)$ and $\alpha_R(\omega)$ are available. A common procedure infers *R*-wave propagation parameters through a regression of recorded traces (e.g. Rix *et al.* 2000, 2001; Lai *et al.* 2002; Foti 2003). However, these techniques assume that the wavefield is dominated by a single Rayleigh mode of propagation. Therefore, the result is an estimate of apparent Rayleigh phase dispersion and attenuation curves rather than modal values, that can be affected by modal superposition when multiple propagation modes are relevant (Foti *et al.* 2014). Furthermore, some procedures simplify the *R*-wave as planar, based on the far-field approximation of the displacement field induced by Rayleigh waves (e.g. Rix *et al.* 2001; Lai *et al.* 2002; Foti 2004; Askari & Ferguson 2012; Gao *et al.* 2018). However, this neglects the actual cylindrical shape of the wave front. This hypothesis negatively affects the quality of the estimated wave parameters at low frequencies, that is, at large wavelengths.

Alternatively, some recent approaches generalize modal identification techniques used in structural engineering to the characterization of *R*-waves (e.g. Badsar *et al.* 2010 and Verachtert *et al.* 2017). For instance, the Circle Fit Method (CFM; Verachtert *et al.* 2017) interprets the Nyquist plot of the *f*–*k* spectrum of measured displacement data to estimate modal *R*-wave dispersion and attenuation data. The transform is calculated through a Hankel transformation, to account for the cylindrical shape of the wavefront (Forbriger 2003). However, the estimated $V_R(\omega)$ and $\alpha_R(\omega)$ depend on two control parameters, which are to be calibrated based on *in situ* conditions. Furthermore, both methods by Badsar *et al.* (2010) and Verachtert *et al.* (2017) tend to overestimate $\alpha_R(\omega)$ at low frequencies (less than about 15 Hz), due to leakage.

Finally, the Wavefield Decomposition (WaveDec, hereafter labelled as 'WD') technique interprets measured displacement data to provide a joint maximum likelihood estimation of the Rayleigh wave propagation parameters (Maranò *et al.* 2017; Bergamo *et al.* 2018, 2019, 2023). This approach accommodates for the presence of multiple modes of propagation in the recorded wavefield by performing multiple fittings of the experimental wavefield, assuming different number of modes in the predicted displacement and selecting the most suitable one based on the Bayesian Information Criterion (BIC; Schwartz 1978). The method models the cylindrical shape of the Rayleigh wave front, thus mitigating near-field effects and ensuring good fitting quality also at low frequencies.

In summary, the available techniques rely on limiting assumptions about the composition and the shape of the recorded waveform, or they depend on processing parameters that might require a site-specific calibration for a proper application. These hypotheses may negatively affect the quality of the estimated wave parameters at low frequencies (that is, at long wavelengths) and in the presence of multiple propagation modes.

This paper introduces two novel methodologies, namely the Frequency Domain BeamForming Attenuation (FDBFa) and the Cylindrical Frequency Domain BeamForming Attenuation (CFDBFa), that aim at extending the framework of dispersion estimation techniques to obtain the *R*-wave phase attenuation. The principle of these approaches consists in applying a transformation to the wavefield, whose resulting function may be interpreted as a pseudo-wave. It is demonstrated that the phase attenuation can be derived through the dispersion analysis of the obtained pseudo-wavefield. In addition, a new modal filtering scheme is proposed, to isolate the contribution of each Rayleigh propagation mode and achieve reliable estimates of modal wave parameters. In this way, the quality and robustness of the modal dispersion and attenuation estimates can be improved.

The paper starts by providing a detailed description of the proposed methods. Then, the inclusion of the modal filtering technique is reported. The second part of the work tests the reliability and effectiveness of the proposed techniques to retrieve modal wave parameters, with reference to both synthetic waveforms and real surface wave data sets. Part of real data refer to the Garner Valley Downhole Array (GVDA) site, as the presence of information on both the velocity and damping structure provides an excellent benchmark on the validity of the inverted models. This study also includes experimental data measured at the Mirandola (MIR) site, since the complexity of the recorded data allows to assess the performance of the proposed algorithm in the presence of multimode wavefields. Specifically, this paper addresses the estimation of the phase velocity and attenuation data and the validity of the proposed technique is demonstrated by comparing the propagation estimates with values derived through alternative methods. The inversion and the comparison with alternative estimates are fully investigated in the companion paper (Aimar *et al*. 2024).

2 PROPOSED ALGORITHM

The Frequency-Domain BeamForming—Attenuation (FDBFa) is a transform-based approach that estimates α_R under the assumption that the recorded wavefield is composed by a single planar wave (e.g. a Rayleigh wave with a dominant propagation mode, recorded in the far field). The principle of this approach consists in applying a transformation to the recorded wavefield, such that the resulting function may be interpreted as a pseudo-wavefield, with dispersion characteristics corresponding to the phase attenuation of the original one. Thus, α_R can be derived by estimating the wavenumber of the pseudo-wavefield, for which a broad variety of tools is currently available. In this study, the wavenumber estimate is carried out through the Frequency-Domain BeamForming (FDBF; Lacoss *et al.* 1969) technique, hence the FDBFa may be interpreted as a generalization of the FDBF for the attenuation estimate. Indeed, the FDBF scheme is computationally fast and robust. Furthermore, the FDBF allows an immediate generalization to non-planar waves, as the inclusion of geometric effects due to the cylindric shape of the Rayleigh wave front is straightforward.

2.1 Frequency-domain beamforming—attenuation (FDBFa)

The FDBFa method is based on the following wavefield transformation:

$$v = [\bullet]^{\mathrm{i}}, \tag{1}$$

where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit. If the recorded wavefield is a plane wave inducing the displacement field $u_p(r) = e^{-\alpha r}e^{-\mathrm{i}kr}$ (where k is the wavenumber, α is the phase attenuation, and r is the spatial coordinate), it can be demonstrated that the resulting function $v(r)$ can be interpreted as a planar pseudo-wave, the wavenumber of which corresponds to the attenuation of $u_p(r)$. Indeed, the application of the transformation (eq. 1) to $u_p(r)$ returns the following function:

$$v\left(r\right) = \left[u_p\left(r\right)\right]^{\mathrm{i}} = e^{kr}e^{-\mathrm{i}\alpha r}. \tag{2}$$

Comparing eq. (2) with the expression of $u_p(r)$, $v(r)$ can be interpreted as a plane wave, whose displacement amplitude varies across space as e^{kr}, with harmonic oscillations according to $e^{-\mathrm{i}\alpha r}$. Therefore, α is the wavenumber of the 'pseudo'-wavefield $v(r)$, whereas k controls the spatial variation of the particle displacement amplitude (Figs 1a–d). Thus, α of the original planar wave $u_p(r)$ can be retrieved by searching for the wavenumber of $v(r)$. In this scheme, the FDBF technique is adopted to carry out the wavenumber analysis of $v(r)$. Therefore, $v(r)$ data across all the receivers are first combined into the spatio-spectral correlation matrix $\mathbf{R}$ (Zywicki 1999), which is a Hermitian-symmetric matrix where each element $R_{j,k}$ is the cross-power spectrum between the jth and the kth sensors for j, $k = 1$ to S (with $S =$ number of measurement points), defined as follows:

$$R_{j,k} = v\left(r_j\right)v^*\left(r_k\right). \tag{3}$$

Then, the procedure computes the pseudo-power $P_{BFa}(\alpha_t)$:

$$P_{BFa}\left(\alpha_t\right) = \mathbf{e}^H\left(\alpha_t\right)\mathbf{Re}\left(\alpha_t\right), \tag{4}$$

where H denotes the Hermitian transpose, and $\mathbf{e}(\alpha_t)$ is the planar steering vector, that is the mathematical representation of the linear phase shift, as a function of the trial attenuation value α_t:

$$\mathbf{e}\left(\alpha_t\right) = \left[\exp\left(-\mathrm{i}\alpha_t r_1\right), \ldots, \exp\left(-\mathrm{i}\alpha_t r_S\right)\right]^{\mathrm{T}}. \tag{5}$$

Ideally, $P_{BFa}(\alpha_t)$ exhibits a single peak, with location α_e (i.e. the estimated attenuation) corresponding to the actual attenuation α of the recorded wave (Fig. 1e).

However, the resulting pseudospectrum often exhibits side lobes together with the main peak. Differently from FDBF, the presence of multiple waves in the recorded wavefield does not result in multiple local maxima in $P_{BFa}(\alpha_t)$ corresponding to modal attenuation values, as it will be addressed in the following sections. Instead, side lobes may partially be the effect of spectral leakage, induced by windowing of the pseudo-wave. This perturbation can be mitigated by applying an appropriate tapering to $v(r)$ prior to the computation of $P_{BFa}(\alpha_t)$, for example by means of a Hanning window. Furthermore, side peaks are artefacts induced by the spatial variation of the amplitude of $v(r)$. Indeed, such amplitude usually assumes a sawtooth-like shape, where each segment increases according to an exponential function, that derives from the transformed wrapped phase of $u_p(r)$. It can be demonstrated that the presence of these discontinuities in the amplitude of $v(r)$ maps into a set of equally spaced side lobes inside $P_{BFa}(\alpha_t)$, with spacing equal to the wavenumber k (Aimar 2022). To avoid ambiguities in the identification of the peak of $P_{BFa}(\alpha_t)$, a possible strategy consists in normalizing $v(r)$ by its amplitude, so that the discontinuities cancel out, prior to the tapering. Alternatively, the transformation (eq. 1) may be applied only to $|u(r)|$, discarding the phase contribution. In both cases, the discontinuities in $v(r)$ cancel out, and $P_{BFa}(\alpha_t)$ exhibits a single peak.

The above formulation refers to active data recorded on 1-D arrays; however, it can be straightforwardly modified for passive data measured on 2-D setups (i.e. ambient vibration analysis).

2.2 Cylindrical frequency-domain beamforming—attenuation (CFDBFa)

Rayleigh waves generated by a point source propagate according to a cylindrical wave front. Therefore, the application of the FDBFa method to retrieve the phase attenuation might lead to biased estimates. For this reason, the FDBFa method can be modified to introduce explicit modelling of the cylindrical shape of the wave front. The proposed scheme is hereafter termed as cylindrical FDBFa (CFDBFa), which can be seen as a generalization of the cylindrical FDBF technique (Zywicki & Rix 2005).

Given a cylindrical wave inducing the displacement field $u_c(r) = H_0^{(2)}[(k - \mathrm{i}\alpha)r]$ (with $H_0^{(2)}$ being the Hankel function of second kind and zeroth order), the application of the transformation (eq. 1) returns a function $v(r)$ representing a pseudo-wavefield, whose phase variations reflect spatial changes of the amplitude of the Hankel function. Indeed, this situation is analogous to the application of eq. (1) to the plane wave $u_p(r)$. Therefore, the attenuation α can be retrieved by exploiting the phase information carried in $v(r)$. For this purpose, the CFDBFa scheme follows a procedure similar to the cylindrical FDBF (Zywicki & Rix 2005). Specifically, it computes the pseudospectrum $P_{CBFa}(\mathrm{K}_t)$ from the spatio-spectral correlation matrix $\mathbf{R}$ of $v(r)$ as:

$$P_{CBFa}\left(\mathrm{K}_t\right) = \mathbf{a}^H\left(\mathrm{K}_t\right)\mathbf{Ra}\left(\mathrm{K}_t\right). \tag{6}$$

The steering vector $\mathbf{a}(\mathrm{K}_t)$ is defined as follows:

$$\mathbf{a}\left(\mathrm{K}_t\right) = \left[\exp\left\{-\mathrm{i}\arg\left[h_0\left(\mathrm{K}_t r_1\right)\right]\right\}, \ldots, \exp\left\{-\mathrm{i}\arg\left[h_0\left(\mathrm{K}_t r_S\right)\right]\right\}\right]^{\mathrm{T}}, \tag{7}$$

where the function h_0 is the power of the Hankel function to the imaginary unit:

$$h_0\left(\bullet\right) = \left[H_0^{(2)}\left(\bullet\right)\right]^{\mathrm{i}}. \tag{8}$$

The steering vector $\mathbf{a}(\mathrm{K}_t)$ stretches the pseudo-wavefield according to phase variations of the power of a Hankel function to the imaginary unit, thus accounting for spatial changes in the phase of $v(r)$ in a proper way. In this way, the pseudo-cylindrical wavefield is converted into an equivalent plane wavefield, and the attenuation of the cylindrical wave is properly estimated. Differently from FDBFa, the argument of $\mathbf{a}(\mathrm{K}_t)$ is a trial complex wavenumber K_t which, by definition, includes both the real wavenumber and the phase attenuation term. Indeed, the estimate of the attenuation for a cylindrical wave should also account for the real wavenumber, as it affects spatial variations of the wave amplitude. As CFDBFa searches for α, the real wavenumber has to be fixed to a value k_{ref}, whereas trial values α_t for the attenuation are adopted. However, being the estimated attenuation α_e quite sensitive to k_{ref} (Aimar 2022), the choice of k_{ref} is not arbitrary, as it should be as close as possible to the actual wavenumber k characterizing the measured wavefield. In plane waves, instead, there is no influence of the wavenumber characteristics on the attenuation estimate, as the amplitude variations only depend on the attenuation itself. Therefore, CFDBFa should be combined with a robust method for estimating k prior to the derivation of α. For this purpose, CFDBF represents an effective strategy because of its accuracy and robustness. Furthermore, combining CFDBF and CFDBFa provides a physically consistent approach for estimating wave parameters, as both of them model the propagation of the wavefield according to a cylindrical scheme.

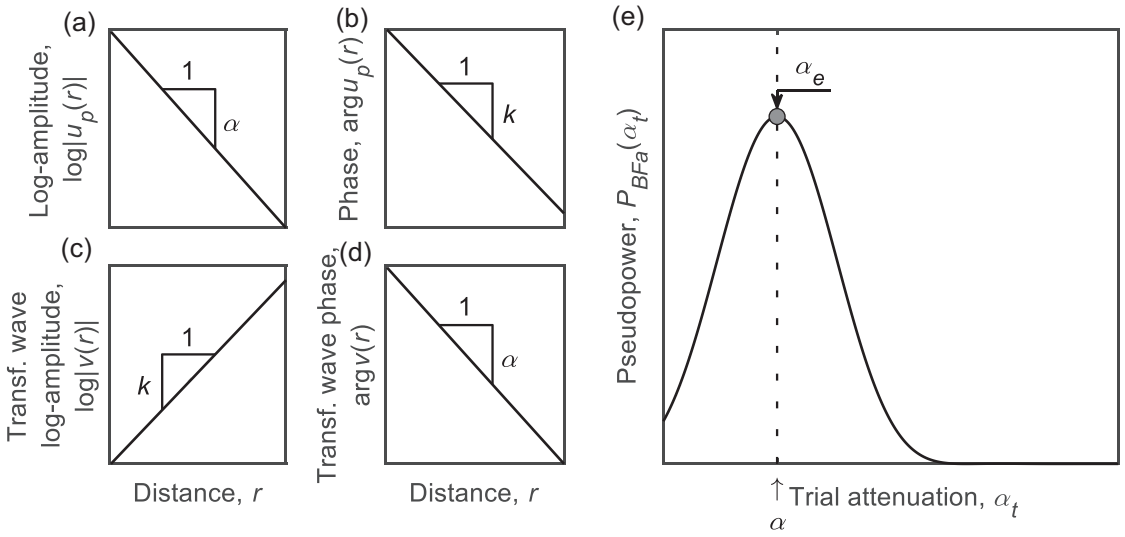

Figure 1. (a–b) Spatial variation of the displacement field $u_p(r)$ induced by the original planar wave, in terms of (a) log-amplitude and (b) phase; (c–d) spatial variation of the transformed wave $v(r)$, in terms of (c) log-amplitude and (d) phase; (e) pseudospectrum $P_{BFa}(\alpha_t)$ of the transformed wave, where the location of the spectral maximum α_e is compared with α.

Under these conditions, $P_{CBFa}(\mathrm{K}_t)$ ideally exhibits a single peak with location α_e coincident with α. However, actual data may provide side lobes in the pseudospectrum, which may corrupt the correct identification of the peak. Their presence can be mitigated by normalizing the pseudo-wavefield by its amplitude and applying proper tapering, as in FDBFa.

To demonstrate the validity of CFDBFa in modelling the propagation of cylindrical waves, Table 1 compares the estimated wave parameters for a specific wavefield, according to FDBFa (and FDBF for the wavenumber) and CFDBFa. The displacement field $u_c(r)$ is computed at 100 equally spaced locations, with spacing equal to 1 m, by setting k equal to 0.1 rad m^{-1} and α equal to 0.0015 rad m^{-1}. The simulated wavefield may be representative of the vertical displacement field induced by Rayleigh waves due to a monochromatic vertical point force, according to the far-field solution of the Lamb problem for a homogeneous half-space (e.g. Lai & Rix 1998). For instance, this may correspond to a R-wave with V_R approximately equal to 315 m s^{-1} and phase damping ratio D_R [defined as $D_R \approx \alpha_R \times V_R/(2\pi f)$–Misbah & Strobbia 2014] equal to 0.015, at a frequency of 5 Hz—this R-wave may be propagating in a homogeneous medium with $V_S \approx 335$ m s^{-1} and $D_S \approx 0.015$. As the FDBFa method does not rigorously model the spatial variation of a cylindrical wave, a potential solution to mitigate this problem consists in processing recorded traces after having scaled them according to $r^{1/2}$. Indeed, cylindrical waves asymptotically tend to be described as a spatially harmonic function (i.e. plane waves), but with amplitude scaled down by $r^{1/2}$ (e.g. Foti *et al.* 2014).

Explicit modelling of the cylindrical wavefield leads to an almost exact estimate of the wave parameters, whereas the use of a planar scheme results in some discrepancies between theoretical and estimated values. On the one side, FDBFa tends to slightly overestimate k—although the error is less than 1 per cent. This bias is consistent with various studies on surface wave testing, which claimed that a planar propagation model often underestimates dispersion data (i.e. the resulting wavenumber is excessively large), especially at low frequencies (e.g. Zywicki & Rix 2005). Instead, α is significantly underestimated when using FDBFa, with a relative difference of around 15 per cent. The error in both k and α estimates rises from the simplifying assumptions introduced by the planar model, which ignores the actual shape of the wave front, which is characterized by a non-linear change in both the displacement amplitude and phase with the distance, and this effect is more relevant at low k (e.g. low-frequency R-waves).

Thus, modelling the propagation of R-waves as plane waves would introduce a bias in the estimated phase wavenumber and the attenuation, especially for the usual source offsets adopted in active-source surface wave testing. This may dramatically alter the resulting V_S and D_S profiles, particularly at greater depths. Therefore, accurate modelling of the geometric spreading is crucial to achieving robust and reliable estimates of dispersion and attenuation data, as recommended by several Authors (e.g. Rahimi *et al.* 2021). For this reason, CFDBFa is a superior technique with respect to the use of a planar model in estimating Rayleigh wave parameters, and the remaining part of this paper will focus on it.

2.3 Application of CFDBFa for surface wave analysis

The application of CFDBFa for retrieving Rayleigh wave phase velocity and attenuation parameters from the interpretation of MASW data is quite straightforward. Indeed, although this method has been introduced with reference to monochromatic planar or cylindrical waves, the generalization to non-harmonic signals can be easily achieved through the Fourier decomposition of the wavefield. In this way, each frequency component of the recorded waveform is separated and it may be processed according to this technique. The resulting wave parameters are then combined frequency by frequency to obtain the experimental dispersion and attenuation curves.

As CFDBFa relies on explicit modelling of the cylindrical wavefield, the coupling between dispersion and attenuation estimates becomes relevant. For this reason, the adopted algorithm first estimates dispersion curves through the CFDBF approach. Then, the attenuation analysis is carried out using the CFDBFa method. In this step, wavenumber data returned by CFDBF are plugged into the steering vectors used in CFDBFa, frequency by frequency. The main steps of this algorithm are listed in Algorithm 1 (the notation follows the conventions used in computer science; e.g. Cormen *et al.* 2001). The algorithm assumes that the recorded motion is

Table 1 Estimated wavenumber and phase attenuation from $u_c(r)$, according to FDBFa and CFDBFa, compared with the corresponding true value.

	FDBFa	CFDBFa	True value
Estimated wavenumber k_e (rad m^{-1})	0.1009	0.1000	$k_R = 0.1$ rad m^{-1}
Estimated attenuation α_e (rad m^{-1})	0.0013	0.0015	$\alpha_R = 0.0015$ rad m^{-1}

Algorithm 1 Cylindrical Frequency-Domain BeamForming—Attenuation (CFDBFa algorithm) for MASW processing

Input:	$\{u_z(r_s, t_m)\}_{s=1,m=1}^{S,M}$: particle displacement recorded at S sensors with offsets r_s, at M time samples t_m
1:	Compute frequency spectra $\{u_z(r_s, \omega_w)\}_{s=1,w=1}^{S,W}$, for W frequencies ω_w, through Fourier transform
2:	**for** $w = 1$ to W **do**
3:	Compute pseudo-power spectrum $P_{CBF}(k_t, \omega_w)$ through CFDBF (Zywicki & Rix 2005)
4:	Identify wavenumber $k_{Re}(\omega_w)$ of the dominant peak of $P_{CBF}(k_t,\omega_w)$: $k_{Re}(\omega_w) \leftarrow \text{argmax} P_{CBF}(k_t,\omega_w)$
5:	Calculate $v(r_s,\omega_w)$ from eq. 1
6:	Remove the amplitude term: $v(r_s,\omega_w) \leftarrow v(r_s,\omega_w)/\lvert v(r_s,\omega_w)\rvert$
7:	Apply the tapering window $w(r_s,\omega_w)$: $v(r_s,\omega_w) \leftarrow w(r_s,\omega_w)v(r_s,\omega_w)$
8:	Compute spatiospectral correlation matrix $R_{j,k}(\omega_w)$: $R_{j,k}(\omega_w) \leftarrow v_z(r_j,\omega_w)\,[v_z(r_k,\omega_w)]^*$
9:	Construct steering vector $\mathbf{a}(\mathrm{K}_t)$ with trial complex wavenumber K_t: $\mathbf{a}(\mathrm{K}_t) \leftarrow [\exp\{-\mathrm{i}\arg[h_0(\mathrm{K}_t r_1)]\}, \ldots, \exp\{-\mathrm{i}\arg[h_0(\mathrm{K}_t r_S)]\}]^{\mathrm{T}}$, with $\mathrm{K}_t = k_{Re}(\omega_w) - \mathrm{i}\alpha_t$
10:	Calculate pseudospectrum $P_{CBFa}(\mathrm{K}_t,\omega_w)$ for varying α_t: $P_{CBFa}(\mathrm{K}_t,\omega_w) \leftarrow \mathbf{a}^{\mathrm{H}}(\mathrm{K}_t)\mathbf{R}(\omega_w)\mathbf{a}(\mathrm{K}_t)$
11:	Peak picking of $P_{CBFa}(\mathrm{K}_t,\omega_w)$: $\alpha_{Re}(\omega_w) \leftarrow \text{argmax} P_{CBFa}(\mathrm{K}_t,\omega_w)$
12:	**end for**
Output:	Estimate of Rayleigh wave dispersion curve $V_{Re}(\omega_w) = \omega_w/k_{Re}(\omega_w)$ and attenuation curve $\alpha_{Re}(\omega_w)$

the particle displacement in the vertical direction. Note that usual acquisition schemes rely on geophones, which return time histories particle velocities, and they should be converted into displacement data through integration in the time domain or multiplication in the frequency domain. Besides, $H_0^{(2)}(\cdot)$ should be replaced by $H_1^{(2)}(\cdot)$ when processing radial motion.

2.4 Dealing with multiple modes: modal filtering

A fundamental assumption of CFDBFa is that the recorded wavefield should consist of a single wave or, in surface wave analysis, that a single Rayleigh mode is dominant (generally, the fundamental mode). When this hypothesis is violated, it tends to return dominant wave components, that are representative of the effective dispersion and attenuation curve and may not coincide with any modal curve. This aspect is crucial in site characterization studies as the Rayleigh wavefield in layered media is the composition of several Rayleigh modes. Therefore, ignoring the influence of this mutual interaction may lead to phase velocity and attenuation estimates that, when misinterpreted as modal values, might return biased estimates of the V_S and D_S profile in the inversion stage. For this reason, CFDBFa is modified to account for the presence of multiple propagation modes, by including an additional step. This further step aims at isolating each R-wave propagation mode by means of a novel filtering technique, that exploits basic principles of digital signal processing. In this way, the main hypothesis of CFDBFa is satisfied, and this technique can be effectively applied to obtain reliable estimates of the Rayleigh wave phase velocity and attenuation parameters.

2.4.1 Proposed filtering technique

The proposed filtering technique is synthetically described in the scheme provided in Fig. 2. The method first identifies the Rayleigh mode under investigation (that is, the corresponding wavenumber, labelled as k_{tg}), for instance through the spectral analysis of the recorded displacement field $u(r,\omega)$. Then, the procedure constructs a sequence $h[n] = [h_1, h_2, \ldots, h_{N-1}]$ (where n denotes the discrete sample index), composed by $N - 1$ complex-valued coefficients h_n. This represents the filter, which is defined as a function of the wavenumber k_{tg} of the target mode. Then, the algorithm isolates the contribution of the desired Rayleigh mode by means of a spatial convolution between the vector of the frequency traces $u(r,\omega)$ and $h[n]$. The result is a new displacement field containing only the contribution of the reference mode, for which modal parameters can be extracted through the CFDBFa method.

This section describes the main features of $h[n]$ and how it is constructed, although a more detailed discussion on this and the involved terminology is provided in the online Supporting Information.

The adopted filter is a complex, bandpass and linear-phase finite-impulse-response (FIR) filter. Being a bandpass filter, $h[n]$ isolates the wavenumber component corresponding to the target Rayleigh mode, thus limiting the contribution of additional components to the wavefield. The chosen filter is a FIR-type system, because they are stable and they represent the optimal choice when the linearity in the phase response is crucial (Mitra 2006; Bruekers 2009). In this case, the phase-linearity is strongly desired as any distortion in the target wave component may affect the estimated attenuation. Finally, $h[n]$ has complex-valued coefficients because it operates with complex-valued wavefield traces, for which the wavenumber spectrum is not an even function. In these conditions, a complex FIR filter with passband centred at the target wavenumber is guaranteed to select and favour only the investigated wave component.

The design of $h[n]$ (hence, the estimation of filter coefficients h_n) relies on the assignment of two quantities:

(i) Magnitude k-response specifications: the k-response H is the spatial Fourier transform of $h[n]$, and it represents the filter response in the wavenumber domain. Being $h[n]$ a bandpass filter, H has to be symmetric around the target wavenumber k_{tg}. Its definition should involve the specification of an adequate passband together with an adequate transition band to control the decay in the response

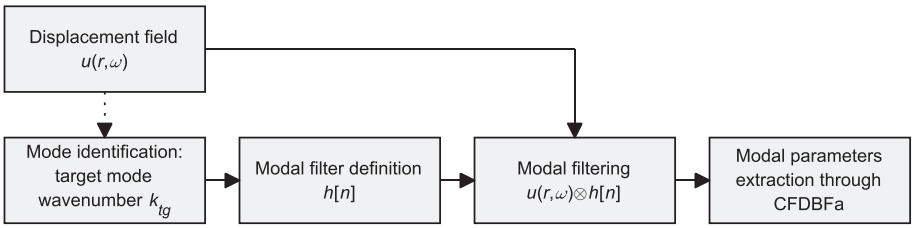

Figure 2. Scheme of the modal filtering procedure.

magnitude at $k \neq k_{tg}$. The boundaries of the transition band are labelled as k_p (i.e. the passband wavenumber) and k_s (i.e. the stopband wavenumber), respectively (Fig. 3a).

(ii) Filter order N: it is an integer corresponding to the number of coefficients involved in $h[n]$ and it controls the matching between the effective k-response H and the magnitude k-response specifications.

The definition of k-response specifications and N are not two independent steps. Indeed, narrow bandpass filters are effective in isolating the desired component of the wavefield, although such behaviour can be achieved only with high-order filters as low-N filters introduce broader transition bands, entailing a partial loss in the filtering capability (Fig. 3a). On the other hand, low-order filters are preferred when dealing with spatially sampled data. Indeed, only a portion of the output signal corresponds to the desired wave component and it includes only $S - N$ samples (with $S =$ number of spatial samples of the input), and N should be adequately less than S to allow processing of filtered values (see the online Supporting Information for further details). This issue may be critical in geophysical applications. Indeed, each spatial sample corresponds to a physical sensor and usual acquisition setups employ a limited number of sensors because of logistical issues, hence recorded signals only include a few tens of samples. Therefore, N has to be bounded within a proper maximum value, so that the number of samples in the desired output signal is large enough to allow a robust estimate of the wave parameters.

A calibration study on synthetic wavefields identified the following recommended values of the filter design parameters (i.e. k_p, k_s and N) to provide the best-quality estimates of modal R-wave parameters (Aimar 2022):

(i) Passband wavenumber: $k_p = k_{\text{res}}$;
(ii) Stopband wavenumber: $k_s = 2k_{\text{res}}$;
(iii) Filter order: $N = (1/2–2/3)S$ (rounded to the nearest integer),
where S is the number of receivers.

The term $k_{\text{res}} = 2\pi/D$ is the wavenumber resolution of the sensor array (where D is the array length), which describes the width of the spectral mainlobe of a waveform composed by a single propagating wave (Zywicki 1999). Thus, by setting the passband equal to k_{res}, the filter tends not to alter spectral information related to the target wave. Besides, filters with order $N = 1/2–2/3$ times the number of receivers may be considered as a valid reference for applying the modal filter in various site conditions. However, using rather low N returns an ineffective filter, which is not capable of well isolating the propagation features of each Rayleigh mode. Therefore, the modal filter effectively extracts information on the target mode when the number of sensors is greater than 20–24, as it guarantees an adequately large usable filter order.

From the computational viewpoint, filter coefficients h_n are obtained through the modulation method, starting from a real filter $g[n] = [g_1, g_2, \ldots, g_{N-1}]$ as follows:

$$h_n = g_n \exp\left(\mathrm{i}\vartheta_{tg} n\right), \ \vartheta_{tg} = \frac{2\pi}{2k_{Nyq}}\left(2k_{Nyq} - k_{tg}\right). \tag{9}$$

The system $g[n]$ is a lowpass real filter whose k-response is an even function, and it exhibits a passband centred at the zero wavenumber. The factor $\exp(\mathrm{i}\vartheta_{tg} n)$ is a modulation term, that shifts the passband of the filter in correspondence of k_{tg} and it depends on the Nyquist wavenumber k_{Nyq}. In this way, the lowpass real filter $g[n]$ is converted into the bandpass complex filter $h[n]$ (Fig. 3b). The modulation method allows to shift the design problem of a complex filter into the design of a real filter $g[n]$. Indeed, the estimation of the filter coefficients g_n can rely on computationally fast and stable algorithms. This task is carried out with the frequency sampling approach (Gold & Jordan 1969), which derives $g[n]$ as a function of the design parameters (i.e. k_p, k_s and N) for $h[n]$. For instance, this approach is implemented in the MATLAB function `fir2` and in the Python function `firwin2` (available in the open-source *Scipy* library).

2.4.2 *Implementation of modal filtering into the CFDBFa*

The inclusion of the modal filtering step inside the CFDBFa algorithm is quite straightforward. The principle of the updated workflow consists in performing a preliminary dispersion analysis. This stage identifies different Rayleigh modes contributing into the recorded wavefield from the peaks of the f–k spectrum, computed through CFDBF. Then, frequency by frequency, each mode is isolated and extracted from the waveform, by applying the modal filter to the recorded displacement. The filter is defined by eq. (9), by setting k_{tg} equal to the wavenumber of the investigated mode. The application of the filter is followed by the removal of the first N samples of the output and a correction of the offsets, that are shifted by $N/2$. These operations compensate for the transient effects and the delay introduced by the filter itself and obtain the desired wave component (see the online Supporting Information for further details). In this way, the recorded wavefield is transformed into a set of displacement data, each representative of a single Rayleigh mode. Thus, the application of CFDBFa to filtered data will return the modal phase velocity and the modal phase attenuation, for each considered propagation mode. Note that a preliminary estimate of the phase velocity is available from the initial stage of identification of Rayleigh modes, however this quantity is updated posterior to the filtering, also to assess the effectiveness of the modal extraction.

The CFDBFa algorithm with Modal Filtering will be hereafter labelled as CFDBFaMF, and a complete description is provided in Algorithm 2 (the notation follows the conventions used in computer science; e.g. Cormen *et al.* 2001). The algorithm assumes that the recorded motion is the particle displacement in the vertical direction. Note that $H_0^{(2)}(\cdot)$ should be replaced by $H_1^{(2)}(\cdot)$ when processing radial motion. However, it should be remarked that this technique is not compatible with waveforms measured in spatially irregular

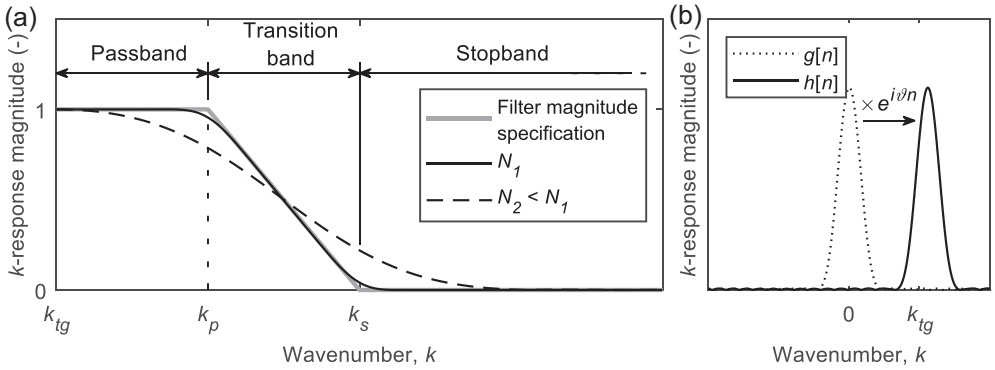

Figure 3. (a) Magnitude k-response specifications, with an explanation of parameters k_p and k_s and a visualization of the influence of the filter N on the fitting of the desired response and (b) scheme of the modulation algorithm for constructing a bandpass complex FIR filter, with the transformation of the real lowpass FIR filter $g[n]$ into the bandpass complex filter $h[n]$, represented in terms of k-response.

Algorithm 2 Cylindrical Frequency-Domain BeamForming—Attenuation with Modal Filtering (CFDBFaMF).

Input:	$\{u_z(r_s,t_m)\}_{s=1,m=1}^{S,M}$: particle displacement recorded at S sensors with offset r_s, at M time samples t_m
1:	Compute frequency spectra $\{u_z(r_s,\omega_w)\}_{s=1,w=1}^{S,W}$, for W frequencies ω_w, through Fourier transform
3:	**for** $w = 1$ to W **do**
4:	Compute pseudopower spectrum $P_{CBF}(k_t,\omega_w)$ through CFDBF (Zywicki & Rix 2005)
5:	Identify peaks $\{k_{Re,r}(\omega_w)\}_{r=1}^{R(\omega_w)}$ of $P_{CBF}(k_t,\omega_w)$, for $R(\omega_w)$ modes
6:	**for** $r = 1$ to $R(\omega_w)$ **do**
7:	Define filter $g[n]$ with MATLAB function `fir2` or Python function `firwin2`, based on the desired passband, transition band and N
8:	Define filter $h[n]$ through eq. 9, with $k_{tg} = k_{Re,r}(\omega_w)$
9:	Mode extraction: $\{u_{z,r}(r_s,\omega_w)\}_{s=1}^{S} \leftarrow \{u_z(r_s,\omega_w)\}_{s=1}^{S} \otimes h[n]$
10:	Removal of filter transient effects: $\{u_{z,r}(r_s,\omega_w)\}_{s=1}^{S-N} \leftarrow \{u_{z,r}(r_s,\omega_w)\}_{s=N+1}^{S}$
11:	Update of sensor offsets to compensate filter delay: $\{r_{s,r}\}_{s=1}^{S-N} \leftarrow \{r_s\}_{s=N/2+1}^{S-N/2}$
12:	Compute pseudopower spectrum $P_{CBF}(k_t,\omega_w)$ through CFDBF from $\{u_{z,m}(r_{s,r},\omega_w)\}_{s=1}^{S-N}$
13:	Peak picking of $P_{CBF}(k_t,\omega_w)$: $k_{Re,r}(\omega_w) \leftarrow \text{argmax}P_{CBF}(k_t,\omega_w)$
14:	Calculate $v_r(r_{s,r},\omega_w)$ from $\{u_{z,m}(r_{s,r},\omega_w)\}_{s=1}^{S-N}$ through eq. 1
15:	Remove the amplitude term: $v_r(r_{s,r},\omega_w) \leftarrow v_r(r_{s,r},\omega_w)/\lvert v_r(r_{s,r},\omega_w)\rvert$
16:	Apply the tapering window $w(r_{s,r},\omega_w)$: $v_r(r_{s,r},\omega_w) \leftarrow w(r_{s,r},\omega_w)v_r(r_{s,r},\omega_w)$
17:	Compute spatiospectral correlation matrix $R_{j,k}(\omega_w)$: $R_{j,k}(\omega_w) \leftarrow v_r(r_{j,r},\omega_w)\,[v_r(r_{k,r},\omega_w)]^*$
18:	Construct steering vector $\mathbf{a}(\mathrm{K}_t)$ with trial complex wavenumber K_t: $\mathbf{a}(\mathrm{K}_t) \leftarrow [\exp\{-\mathrm{i}\arg[h_0(\mathrm{K}_t r_1)]\},\ldots,\exp\{-\mathrm{i}\arg[h_0(\mathrm{K}_t r_S)]\}]^{\mathrm{T}}$, with $\mathrm{K}_t = k_{Re}(\omega_i) - i\alpha_t$
19:	Calculate pseudospectrum $P_{CBFa}(\mathrm{K}_t,\omega_w)$ for varying α_t: $P_{CBFa}(\mathrm{K}_t,\omega_w) \leftarrow \mathbf{a}^{\mathrm{H}}(\mathrm{K}_t)\mathbf{R}(\omega_w)\mathbf{a}(\mathrm{K}_t)$
20:	Peak picking of $P_{CBFa}(\mathrm{K}_t,\omega_w)$: $\alpha_{Re,r}(\omega_w) \leftarrow \text{argmax}P_{CBFa}(\mathrm{K}_t,\omega_w)$
21:	**end for**
22:	**end for**
Output:	Estimates of Rayleigh wave modal dispersion curves $V_{Re,r}(\omega_w) = \omega_w/k_{Re,r}(\omega_w)$ and modal attenuation curves $\alpha_{Re,r}(\omega_w)$ for $r = 1$ to $R(\omega_w)$ modes

Table 2 Ground model parameters adopted to generate the synthetic wavefield SW1.

Thickness (m)	S-wave velocity, V_S (m s^{-1})	S-wave damping ratio, D_S (per cent)
5	200	3.5
10	300	3
10	400	2.5
–	500	2

arrays, as the method used in the derivation of filter coefficients (i.e. the frequency sampling approach) can only be applied to uniformly sampled data.

3 TESTS ON SYNTHETIC WAVEFORMS

The reliability of the proposed approaches is tested with reference to a set of synthetic surface waveforms, that simulate results of MASW surveys carried out on idealized soil models. Selected soil models are compatible with the stratigraphy of typical soil deposits in engineering practice and the generated wavefields are characterized by a different degree of complexity.

3.1 Description of waveforms

The synthetic wavefield SW1 simulates results of a MASW survey carried out on a normally dispersive soil deposit, wherein V_S gradually increases with depth (Table 2; Figs 4a and b). In this case, the

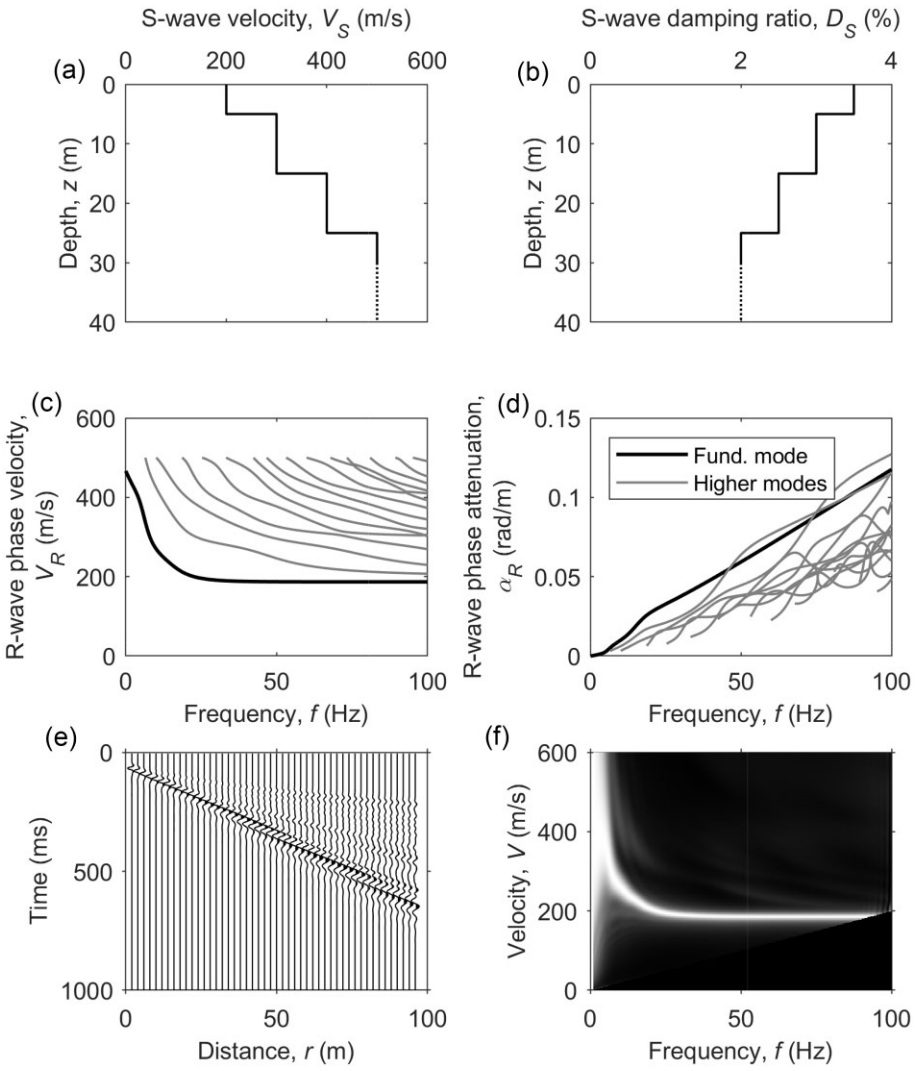

Figure 4. Main characteristics of the synthetic wavefield SW1: reference ground model, described in terms of the *S*-wave velocity (a) and damping ratio (b); *R*-wave phase velocity curves (c) and phase attenuation curves (d); time-domain traces (e) and pseudo-power spectrum in the frequency-velocity domain, where each mode is identified by the white patterns (f).

Rayleigh wavefield is multimodal, although the fundamental mode of propagation is dominant for a broad range of frequencies. For this reason, SW1 allows to investigate the performance of the proposed methods in the presence of a wavefield mostly composed by a single mode. The other synthetic example (labelled as SW2) is obtained from the simulation of a MASW survey for an inversely dispersive profile, characterized by a low-velocity zone (Table 3; Figs 5a and b). In this case, higher modes significantly contribute to the simulated wavefield. Therefore, the performance of the novel approaches in the presence of a multimodal waveform can be addressed. In both models, constant values of Poisson's ratio $\nu = 0.33$ and mass density $\rho = 1800\ \mathrm{kg\,m^{-3}}$ are kept throughout the layers, whereas the *P*-wave damping ratio D_P is assumed as equal to D_S.

SW1 and SW2 are computed assuming an impulsive source by means of the ElastoDynamics Toolbox (EDT; Schevenels *et al.* 2009), which implements the stiffness matrix method (Kausel & Roësset 1981). SW1 and SW2 include vertical displacement data computed at 48 evenly spaced locations on the surface of the model, with spacing equal to 2 m. Thus, the offset from the source ranges between 2 and 94 m. Both the number of receivers and the receiver spacing are consistent with the usual acquisition layout of MASW surveys for near-surface site characterization (Foti *et al.* 2018). The investigated frequency range spans between 1 and 100 Hz.

The resulting waveforms are plotted in Figs 4 and 5, both in terms of seismic traces defined in the time domain and of the pseudo-power spectrum in the frequency–velocity domain, computed through the FDBF approach (Lacoss *et al.* 1969). As expected, the wave energy of SW1 is mostly carried out by a single mode (Fig. 4f), whereas SW2 exhibits a remarkably strong multimodal propagation, with the dominant mode shifting with increasing frequency (Fig. 5f). Figs 4(c), (d) and 5(c), (d) also report the theoretical phase velocity curves $V_R(\omega)$ and phase attenuation curves $\alpha_R(\omega)$ corresponding to the models, as computed through EDT.

3.2 Synthetic wavefield SW1

Fig. 6 compares resulting dispersion and attenuation estimates obtained through the CFDBFaMF algorithm for the first three Rayleigh propagation modes for SW1, labelled as R0, R1 and R2, respectively. In addition, results from the CFDBFa scheme are included.

Table 3 Ground model parameters adopted to generate the synthetic wavefield SW2.

Thickness (m)	*S*-wave velocity, V_S (m s^{-1})	*S*-wave damping ratio, D_S (per cent)
5	250	2.5
3	150	4
–	350	1

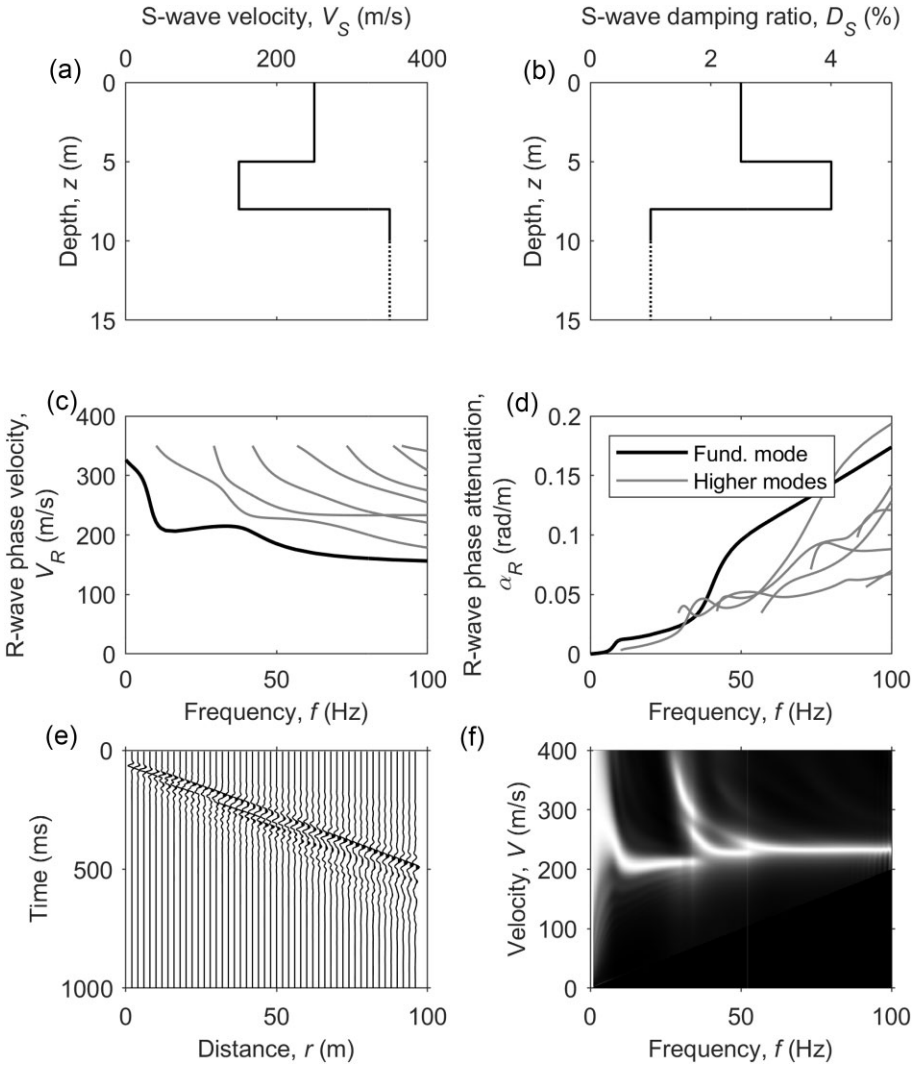

Figure 5. Main characteristics of the synthetic wavefield SW2: reference ground model, described in terms of the *S*-wave velocity (a) and damping ratio (b); *R*-wave phase velocity curves (c) and phase attenuation curves (d); time-domain traces (e) and pseudo-power spectrum in the frequency–velocity domain, where each mode is identified by the white patterns (f).

Note that the CFDBFa-based estimates are only reported in the R0 panel (i.e. Figs 6a and b), as they are expected to be representative of the fundamental mode, since SW1 is dominated by this propagation mode (see Fig. 4f). Reported data are sampled with a sampling frequency of 1 Hz, across the frequency band ranging between 5 and 100 Hz, covering a wavelength range compatible with the resolution limits of the ideal array used for generating SW1 (Foti *et al.* 2014). For comparison purposes, the theoretical modal curves are included.

As for CFDBFa, the estimated dispersion curve matches the fundamental mode, with some discrepancy at low frequencies (Fig. 6a). Indeed, the Rayleigh wavefield in normally dispersive media is typically governed by the fundamental mode, as highlighted in Fig. 4(f). However, the resulting attenuation curve equals the fundamental mode only at intermediate frequencies, whereas it tends to underestimate it at frequencies larger than 50 Hz (Fig. 6b). At this frequency, higher modes start to contribute to the wavefield, thus the estimated α_R is partially affected by them.

Instead, the CFDBFaMF algorithm returns reliable estimates of $\alpha_R(\omega)$ for R0 especially at high frequencies, where the CFDBFa-based attenuation estimate diverges from modal values. In addition, the estimated $V_R(\omega)$ well matches the R0 dispersion curve. This result is not surprising, as the fundamental mode is the dominant component of SW1, whereas higher modes provide a small contribution. Therefore, isolating the fundamental mode by removing other components is an easy task. The converse occurs when dealing with higher modes, because of the higher difficulties in extracting the desired mode. For instance, the retrieval of R1 wave parameters is

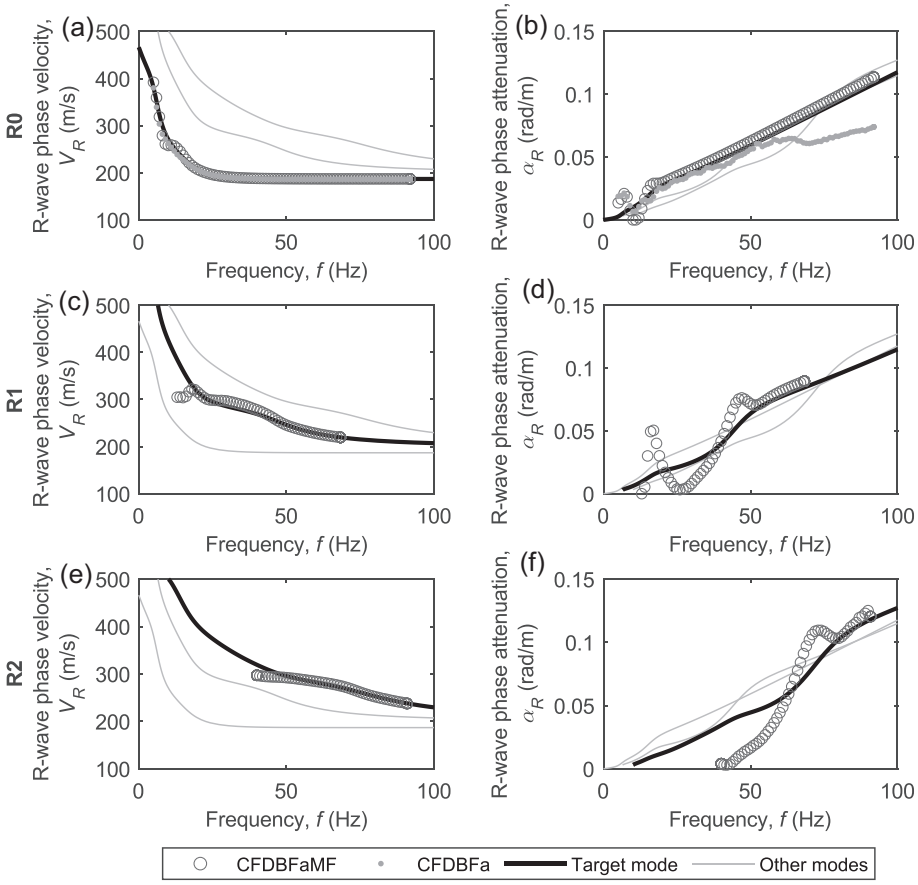

Figure 6. Application of the CFDBFaMF algorithm on SW1, with a focus on the first three modes: (a–b) estimated dispersion curves (a) and attenuation curves (b) for the fundamental mode, labelled as R0; (c–d) estimated dispersion curves (c) and attenuation curves (d) for the first higher mode, labelled as R1; (e–f) estimated dispersion curves (e) and attenuation curves (f) for the second higher mode, labelled as R2. Results of CFDBFa are also reported, in superimposition with R0 data, in subplots (a–b), for comparison purposes.

rather challenging at frequencies less than 20 Hz because the modal filtering returns a waveform where the fundamental mode is still dominant, as its contribution cannot be easily removed. Therefore, the estimated $V_R(\omega)$ and $\alpha_R(\omega)$ do not tend to match the corresponding modal values, except at higher frequencies. A similar result is observed for R2, where the degree of fit is improved at high frequencies (Figs 6e and f). Besides, due to the high difficulties in extracting these wave components, the sensitivity to filter parameters is more relevant than for R0 (Aimar 2022).

3.3 Synthetic wavefield SW2

Fig. 7 compares resulting dispersion and attenuation estimates obtained through the CFDBFaMF algorithm for the first three Rayleigh propagation modes for SW2, labelled as R0, R1 and R2, respectively. In addition, results from the CFDBFa scheme are included. Differently from SW1, CFDBFa-based estimates are reported in all the panels. Indeed, SW2 is a combination of several Rayleigh modes, each one providing a different degree of contribution as a function of the frequency (see Fig. 5f). Therefore, such estimates are expected to drift between different modal values, as a function of the frequency.

As expected, CFDBFa returns an estimated dispersion curve gradually shifting from the fundamental mode up to the second higher mode (Fig. 7a). Consistently, the estimated attenuation curve tends to follow the corresponding modal ones, and each transition occurs at the same frequency of the passages in the dispersion data (Fig. 7b, Fig. 7d, and Fig. 7f).

When applying CFDBFaMF, the estimated $V_R(\omega)$ and $\alpha_R(\omega)$ consistently match the corresponding modal values, especially for R0 and R2. However, this result is valid mainly in the frequency ranges where each mode dominates the wavefield. Instead, at other frequencies, the filter has to isolate a wave component which is not dominant and quite close to other modes, as the relative distance in the wavenumber domain is quite small. Therefore, the estimated $V_R(\omega)$ and $\alpha_R(\omega)$ do not tend to match the corresponding modal curves, entailing that the filter does not completely remove undesired wave components. This problem is evident for R1, where the strong interference due to R0 and R2 prevents an effective estimation of the wave parameter at many frequencies.

Finally, estimated $V_R(\omega)$ and $\alpha_R(\omega)$ for SW1 and SW2 systematically differ from the corresponding modal values at low frequencies, for both CFDBFa and CFDBFaMF. On the one hand, CFDBFaMF underestimates $\alpha_R(\omega)$ close to 10 Hz for SW1. This may be partly due to some influence of the first higher mode, which starts to contribute to the wavefield at 6 Hz, and whose effect on the displacement field may not have been completely removed by the modal filter. At lower frequencies, both CFDBFa and CFDBFaMF tend to underestimate $V_R(\omega)$ (especially for SW2) and overpredict $\alpha_R(\omega)$, with a relative error which is on average much greater than the corresponding error in the dispersion data. Such drift might be an artefact introduced by other wave components (e.g. body waves),

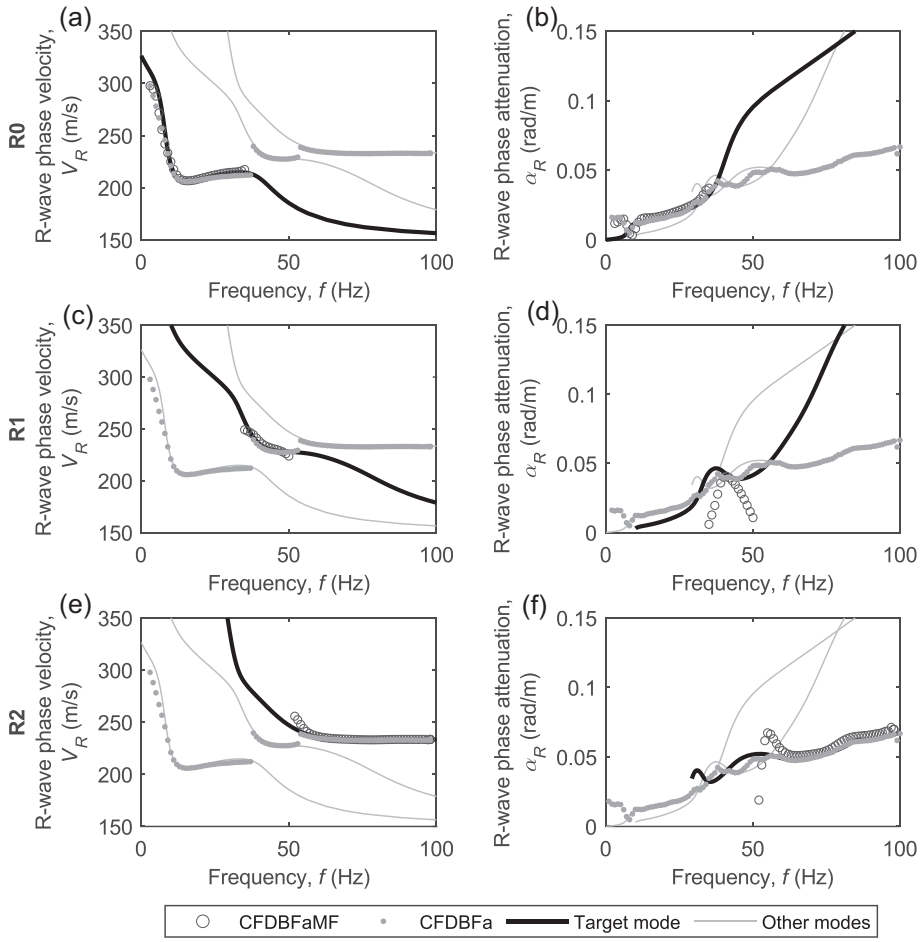

Figure 7. Application of the CFDBFaMF algorithm on SW2, with a focus on the first three modes: (a–b) estimated dispersion curves (a) and attenuation curves (b) for the fundamental mode, labelled as R0; (c–d) estimated dispersion curves (c) and attenuation curves (d) for the first higher mode, labelled as R1; (e–f) estimated dispersion curves (e) and attenuation curves (f) for the second higher mode, labelled as R2. Results of CFDBFa are also reported, for comparison purposes.

whose contribution is usually significant in this range (Rahimi *et al.* 2021).

3.4 Discussion

The comparisons reported in Figs 6 and 7 highlight a moderate discrepancy between the estimated modal R-wave parameters and the corresponding theoretical values for both SW1 and SW2, especially when dealing with high-order modes. However, a rigorous statement on the effectiveness of the proposed technique needs a comparative study with the performance of alternative estimation methods, in terms of both the retrieved frequency range and the estimation error.

To this aim, the considered processing techniques are herein the Circle Fit Method (CFM; Verachtert *et al.* 2017) and the Wavefield Decomposition (WD) technique (Maranò *et al.* 2017; Bergamo *et al.* 2018, 2019, 2023), as they include an explicit modelling of multiple propagation modes, hence they ideally provide estimates of the modal dispersion and attenuation curves. The default processing parameters are used for the CFM (Verachtert *et al.* 2017), whereas the WD is set to return a maximum likelihood estimate.

For each approach, the reliability is quantified by measuring the relative error between the estimated dispersion and attenuation data—$V_{R,e}(\omega)$ and $\alpha_{R,e}(\omega)$, respectively—and the theoretical $V_R(\omega)$ and $\alpha_R(\omega)$ curves, for each synthetic soil profile. A quantitative, global measure of such error is provided in terms of average relative error $\bar{\varepsilon}_{V\alpha}$:

$$\bar{\varepsilon}_{V\alpha} = \frac{1}{W}\sum_{w=1}^{W}\left(\frac{|V_{R,e} - V_R|}{V_R} + \frac{|\alpha_{R,e} - \alpha_R|}{\alpha_R}\right), \tag{10}$$

where W is the number of samples. This quantity is computed over the frequency range wherein all the considered techniques return wave parameters.

Fig. 8 summarizes the estimation error $\bar{\varepsilon}_{V\alpha}$ for each investigated mode of the considered synthetic wavefields. The last column of the figure also contains the global error for all the modes, given as the mean of $\bar{\varepsilon}_{V\alpha}$ for R0, R1 and R2, to provide insight on the overall performance of the processing techniques under examination. The aim of the figure is to provide an insight on the capability of the CFDBFaMF in retrieving both $V_R(\omega)$ and $\alpha_R(\omega)$ focusing on single modes and multiple modes, compared with state-of-the-art methods.

As for SW1, the estimation error for the CFDBFaMF tends to increase while moving from R0 to higher-order modes, as the weaker contribution of R1 and R2 in SW1 (see Fig. 4f) results in a drop in the quality of the corresponding estimated data. The proposed approach is slightly outperformed by the WD and the CFM in the estimation of the R0 and R1 parameters, respectively. However, it should be noted

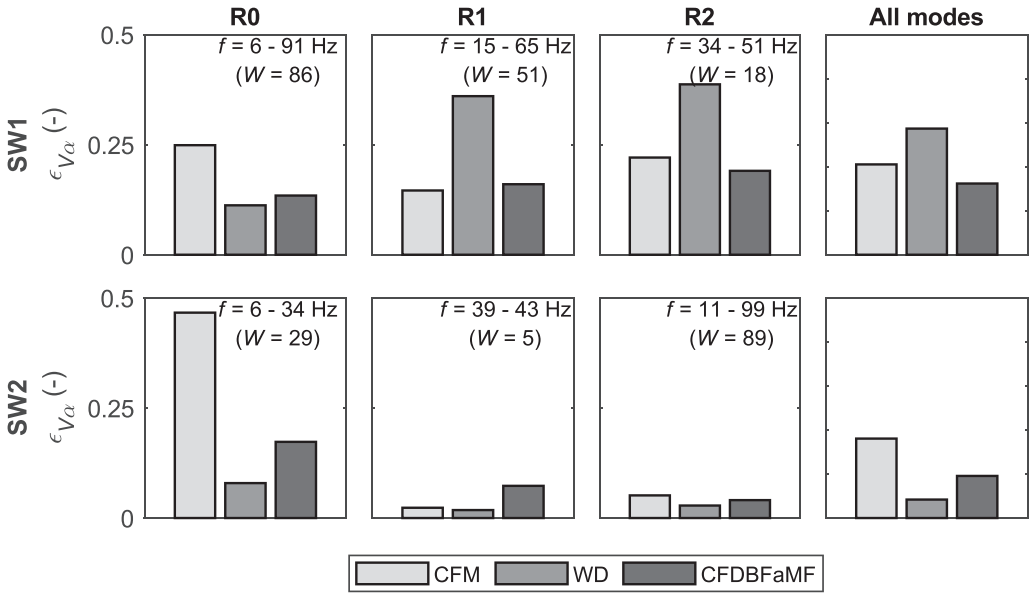

Figure 8. Average error $\bar{\varepsilon}_{V\alpha}$ for the combined estimated phase velocity and phase attenuation for the circle fit method (CFM), the wavefield decomposition approach (WD), and the cylindrical frequency-domain beamforming-attenuation with modal filtering (CFDBFaMF). Errors are computed with reference to the lowest-order three propagation modes (labelled as R0, R1 and R2, respectively) of the synthetic wavefields SW1 and SW2. For each panel, the frequency band and the number of samples involved in the error computation are reported.

that the corresponding $\bar{\varepsilon}_{V\alpha}$ for the CFDBFaMF shares the same order of magnitude of the other techniques. Furthermore, the WD approach does not effectively retrieve R1 and R2 wave parameters, whereas the estimation quality for the CFM drops when focusing on R0. Instead, $\bar{\varepsilon}_{V\alpha}$ for the CFDBFaMF is always smaller than 0.20—that is, the estimation error is on average less than 20 per cent. As a result, the corresponding average $\bar{\varepsilon}_{V\alpha}$ is smaller compared with the considered methods (see the last column of Fig. 8), entailing that the performance of the CFDBFaMF is comparable or even superior.

As for SW2, the performance of the CFDBFaMF is intermediate compared with alternative techniques. The proposed approach faces the most relevant issues in the identification of R1 parameters, as highlighted by the larger $\bar{\varepsilon}_{V\alpha}$. Indeed, R1 is dominant across a narrow frequency range, where the contribution of R0 and R2 is also significant (see Fig. 5f). Therefore, the filtering procedure implemented in the CFDBFaMF does not effectively remove the contribution by R0 and R2, thus returning a corrected wave with amplitude variations not reflecting those linked with R1. Similar considerations are valid for R2 at low frequencies, due to the interference by R1, although the better performance at high frequencies mitigates this issue. On the other hand, the estimation error for R0 is compatible with the one of other methods, also because the CFM-based estimate is affected by quite large $\bar{\varepsilon}_{V\alpha}$. Instead, the WD approach tends to globally outperform the other methods for this synthetic data set. However, albeit larger compared with the WD, it should be remarked that the average $\bar{\varepsilon}_{V\alpha}$ for the CFDBFaMF is still compatible with it and less than 0.2, entailing that the discrepancy between theoretical and estimated attenuation and dispersion data is smaller than 20 per cent, on average.

Furthermore, a better understanding on the relative performance of each processing technique is provided in Fig. 9, for R0, R1 and R2, with reference to SW1. This figure represents the frequency range at which each method returned dispersion and attenuation data, at the boundaries of which are included the wavelength limits. Each bar includes a pseudo-colour scheme according to the estimation error in terms of relative error ε_V and ε_α, for the phase velocity and phase attenuation, respectively:

$$\varepsilon_V = \frac{V_{R,e} - V_R}{V_R},\ \varepsilon_\alpha = \frac{\alpha_{R,e} - \alpha_R}{\alpha_R}. \tag{11}$$

This error metric is consistent with the one described in eq. (10) and, according to this, a positive value denotes an overestimation of the modal value. For simplicity, only results for SW1 are reported, as similar considerations apply for SW2.

As for R0, the resulting dispersion data match the theoretical one, for almost each approach. However, the CFDBFaMF returns dispersion data even at lower f values, down to 4 Hz, with the error magnitude bounded between 5 and 10 per cent. On the other hand, all the methods tend to overestimate the fundamental-mode $\alpha_R(\omega)$ at low frequencies, with a much larger drift than what is observed for dispersion data. This drift in attenuation data may be linked with near-field effects, that are not modelled in an explicit way, especially in terms of the body-wave contribution. Focusing on higher modes, all the methods successfully identify the R1 dispersion curve, at moderately high frequencies. However, both the CFM and the CFDBFaMF are affected by bias at low frequencies, and only the WD provides almost correct estimates in this range. Specifically, the CFDBFaMF underestimates $V_R(\omega)$, due to the inability of the modal filter in effectively separating the weak target mode from the dominant one (i.e. R0; see Fig. 6c). However, the error magnitude significantly reduces at around 20 Hz. As a consequence of the bias in the estimated velocity, the CFDBFaMF underestimates $\alpha_R(\omega)$ in the same frequency range. On the other hand, it achieves a rather good level of compatibility with the theoretical $\alpha_R(\omega)$ at $f >$ 50 Hz, whereas the estimated value is far from the true one for the other methods. For instance, although the WD perfectly matches the theoretical dispersion curve for R1, it fails in extracting the corresponding attenuation over a broad frequency range. This justifies the large $\bar{\varepsilon}_{V\alpha}$ reported in Fig. 8. Similar results apply for R2. In this case, the CFDBFaMF is quite effective in characterizing the corresponding modal parameters, returning estimates for a quite broad frequency range, in which WD faces some issues, perhaps due to the low energy associated with this mode. Specifically, the CFDBFaMF

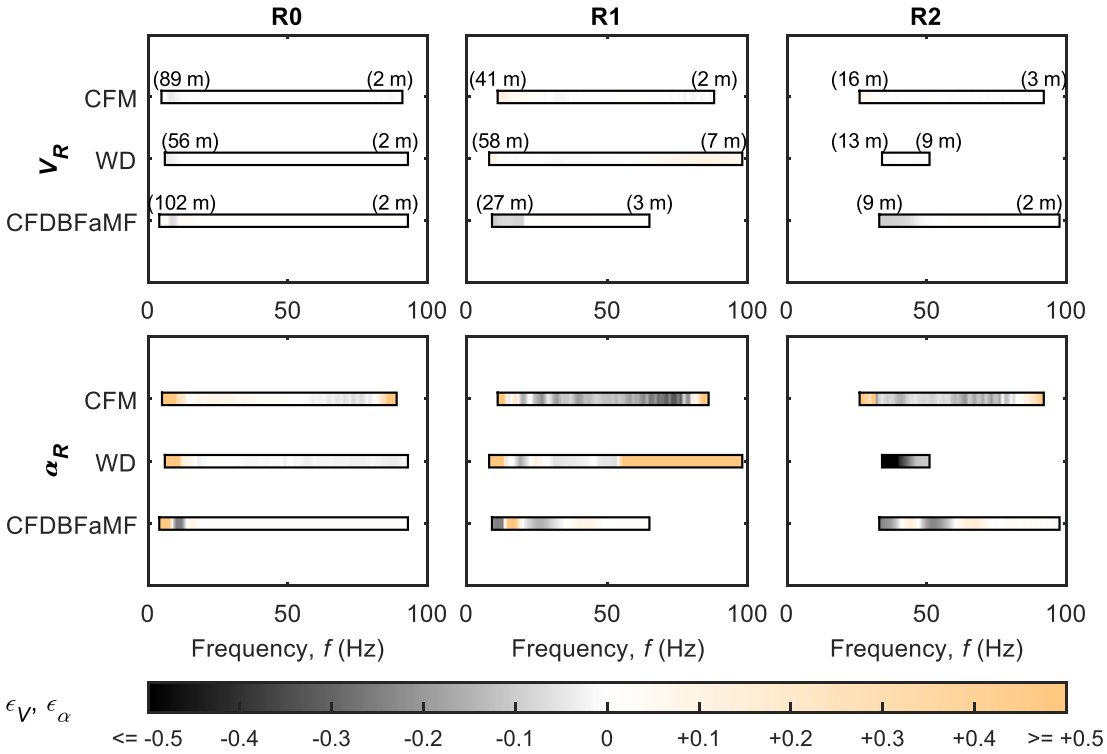

Figure 9. Assessment of the relative performance of the CFM, WD and CFDBFaMF methods for the *R*-wave dispersion and attenuation curves, with reference to SW1. Each plot contains bars spanning along the frequency range at which each method returned estimates of *R*-wave dispersion data (top row) and attenuation data (bottom row). The colour scale is a function of the magnitude of the relative error for the phase velocity ε_V or for the phase attenuation ε_α, for the fundamental mode R0 (left-hand column), first higher mode R1 (central column), and second higher mode R2 (right-hand column). The numeric values in brackets denote the maximum and the minimum detected wavelength.

well captures $V_R(\omega)$ above 50 Hz, whereas a good level of fitting with $\alpha_R(\omega)$ is achieved above 70 Hz, which is the frequency range where $V_R(\omega)$ is exactly matched.

In summary, the considered attenuation processing techniques are affected by an average estimation error ranging between 12 and 25 per cent, depending on the soil deposit characteristics. These moderately high error values can be partly attributed to problems in extracting wave parameters in multimode wavefields. This is due to difficulties in effectively separating the contribution of each propagation mode, especially when the target mode is not dominant. An additional source of error is due to near-field effects, which typically affect the estimated fundamental mode parameters at low frequencies, with slight underestimation of the phase velocity (Rahimi *et al.* 2021) and an overestimation of the phase attenuation (Yoon 2005; Aimar 2022). If not accounted for, the bias in the experimental data may affect the inversion process, resulting in erroneous estimates of velocity and damping ratio profiles, especially at large depths. Thus, reducing errors in attenuation estimates is important and deserves further research. A strategy to improve the quality of the estimates could be a processing scheme based on the interpretation of frequency-dependent, moving windows of the array. This could allow for mitigating the influence of near-field effects, by selectively removing the waveforms close to the source.

While the attenuation estimation error for the CFDBFaMF is generally compatible with those of other techniques (i.e. CFM and WD), the proposed technique can provide estimates of the *R*-wave parameters on a generally broader frequency range, with even smaller errors in some scenarios (e.g. normally dispersive medium), as demonstrated in Fig. 8. Overall, the CFDBFaMF is the only method to return robust estimates of the phase attenuation, especially at high frequencies, for all the propagation modes. However, it should be remarked that this study is only deemed to present the novel approach, showing its reliability with a simple, yet not complete, comparison with some state-of-the-art approaches. Indeed, both the limited number of synthetic data and the simplified error metrics do not allow for a rigorous statement on the relative performance of one method with respect to another. A thorough assessment requires additional analyses, which are beyond the aim of the present contribution.

4 TESTS ON REAL SURFACE WAVE DATA

The assessment of the performance of the proposed methods is also based on two surface wave data sets collected at different sites. The available data allow for a comprehensive investigation of the capability of the CFDBFa and CFDBFaMF methods in the presence of wavefields with different degrees of complexity.

4.1 Description of the surface wave data sets

The first data set was collected at the Garner Valley Down-Hole Array (GVDA, 33°40.127′N, 116°40.427′W—in the WGS84 Datum), located in a narrow valley in Southern California, USA. The near-surface site stratigraphy is characterized by two main geological units. The top unit is composed of soft silty and sandy alluvial soil, interbedded with some clay layers and lenses (Hill 1981). The water table oscillates between the ground surface down to 1–3 m depth. The alluvium transitions into a layer of gravelly sand resulting from

decomposed granite at about 20 m depth, although the depth of this interface is variable in space.

The GVDA site has been characterized by several geotechnical and geophysical tests. Furthermore, this site is instrumented with a seismic monitoring system, including surface and borehole accelerometers. For this reason, GVDA is an excellent candidate for the joint characterization of the velocity and damping ratio profiles through surface wave testing. Specifically, this section addresses the estimation of the phase velocity and attenuation data, whereas the inversion and the comparison with alternative estimates is addressed in the companion paper.

A MASW survey was carried out at the GVDA site, involving a 2-D array initially aimed at developing a 3-D model of the soil deposit (Aimar 2022; Vantassel *et al.* 2023). Herein, we only process one of the linear arrays that comprise the larger 2-D array. Specifically, we process the NS13 array located along the southeast edge, as shown in Fig. 10, which also shows that the investigated area is close to the instrumented boreholes.

The acquisition setup includes 14 geophones, with an inter-receiver distance equal to 5 m. The receivers were Magseis Fairfield Nodal ZLand 3C nodes, which contain a three-component geophone, with passband corner frequency equal to 5 Hz. Waveforms were generated at two shot points off each end of the array using a vibrating source, namely the NHERI@UTexas Thumper vibroseis truck (Stokoe II *et al.* 2020). The vibrating source generated a 12-s long sweep signal, from 5 to 30 Hz. The source-offsets of road-side shot points are 4.5 and 35.5 m respectively, whereas the ones on the parking side are 2.5 and 33.5 m from the closest sensor (Fig. 10). At each shot point, three repetitions were performed, whereas the number was increased to ten at the farthest source-offsets.

The second data set was collected at Mirandola (MIR, 44°52.621′N, 11°03.668′E—in the WGS84 Datum), located in the Po Plain in Italy. The stratigraphy of the site is an alluvial deposit, characterized by alternating sandy and silty-clayey layers, which transition to rock-like formations at depths between 50 and 150 m.

The MIR site was investigated in an extensive characterization campaign in the framework of the InterPACIFIC ('Inter-comparison of methods for site PArameter and veloCIty proFIle Characterization') project (Garofalo *et al.* 2016a;b), which compared the performance of different geophysical techniques (i.e. surface waves and invasive methods) in estimating the stiffness profiles. In this study, we process data from a linear array of 48 vertical 4.5-Hz geophones, with an inter-receiver spacing of 1 m. Waveforms were generated at six shot points from each end of the array using a sledgehammer. The source-offsets of the shot points are 2.5, 3 and 15 m. The data set also includes recordings from sources located within the array, but these data are not processed in this study for simplicity. At each shot point, ten repetitions were performed. Further details on the site description and the acquisition setup can be found in Garofalo *et al.* (2016b).

Fig. 11(a) reports the *f*–*k* representation of the recorded wavefield at the GVDA, obtained by applying the spatial Fourier transform to recorded data. Notably, the recorded wavefield is dominated by a single propagation mode in the whole investigated frequency range, which is expected to be the *R*-wave fundamental mode, labelled as R0. Instead, the MIR data set is characterized by a remarkably multimode wavefield, as highlighted by the multiple, distinct spectral peaks in Fig. 11(b). In this case, three dominant modes can be identified in the considered frequency range. The highest wavenumber mode (which may correspond to the *R*-wave fundamental mode, labelled as R0) tends to disappear at $f > 30$ Hz. At higher frequencies, two additional components of the wavefield become relevant, and they both contribute to the wavefield with similar amplitudes. These may be representative of the first two higher modes, labelled as R1 and R2, respectively. In this sense, in the MIR data set, the contribution of higher propagation modes and the modal superposition are significant over a wide range of frequencies. Thanks to the multimodal nature of the wavefield, the MIR site is an appealing case study to apply the CFDBFaMF technique and to assess the effectiveness of the modal filtering under real site conditions.

Thus, the GVDA and the MIR data sets provide a real benchmark to evaluate the performance of the proposed processing algorithms, in comparison with state-of-the-art methods, for both a single mode and a multimode scenario.

4.2 Results from GVDA

The CFDBFa algorithm is adopted to extract $V_R(\omega)$ and $\alpha_R(\omega)$ from the recorded seismic traces at GVDA, for two reasons. On the one hand, the array includes a moderately small number of sensors, which precludes a successful application of modal filtering techniques to well isolate the propagation features of each Rayleigh mode. Indeed, the modal filter effectively extracts information on the target mode when the number of sensors is greater than 20–24. Therefore, results from the CFDBFaMF are not included, as they would be identical to CFDBFa. Furthermore, the recorded wavefield is dominated by a single propagation mode, as stated above (Fig. 11a). Therefore, it is expected the CFDBFa to return effective wave parameters compatible with the modal ones, in this case.

Fig. 12 represents the estimated $V_R(\omega)$ and $\alpha_R(\omega)$ obtained through CFDBFa. As processing includes data from multiple shots, statistics are computed by combining results from different source offsets, in consistency with the multi-offset approach (Wood & Cox 2012; Vantassel & Cox 2022). Results are represented in terms of lognormal statistics (i.e. median and logarithmic standard deviation) of $V_R(\omega)$ and $\alpha_R(\omega)$ (Aimar 2022), computed from the elementary *R*-wave parameters corresponding to the 4 shot locations. Furthermore, CFDBFa results are superimposed with the estimated *R*-wave parameters according to the CFM and the WD.

The CFDBFa identifies *R*-wave parameters in the frequency range between 5 and 35 Hz. Specifically, the resulting $V_R(\omega)$ matches well with those provided by other techniques. However, it is affected by lower variability on average (across the full frequency bandwidth), as $\sigma_{\ln V}$ varies from 0.02 at high frequencies, up to 0.03–0.05 at longer wavelengths. As for the estimated $\alpha_R(\omega)$, the median trend is compatible with the one returned by alternative processing techniques. Notably, $\sigma_{\ln\alpha}$ is an order of magnitude greater than $\sigma_{\ln V}$, and the variability is larger at lower frequencies, with a variation from 0.3 and 0.5. Interestingly, CFDBFa manages to well constrain $\alpha_R(\omega)$ at low frequencies, at a range where the CFM is affected by larger variability both on $V_R(\omega)$ and $\alpha_R(\omega)$, and where the WD method does not provide any estimates of attenuation.

This result demonstrates the potential of the new CFDBFa method in retrieving dispersion and attenuation parameters, when the wavefield is mostly controlled by a single propagation mode. Indeed, given the strong agreement between the CFDBFa-based estimated $V_R(\omega)$ and $\alpha_R(\omega)$ with results from the WD and CFM, that return modal *R*-wave parameters, it can be stated that such estimates are representative of the *R*-wave fundamental mode. In other words, both modal techniques and those returning effective wave parameters provide comparable results, in this case. Furthermore, the proposed CFDBFa method is superior in terms of the effectiveness, as it can provide attenuation estimates over a broad frequency range,

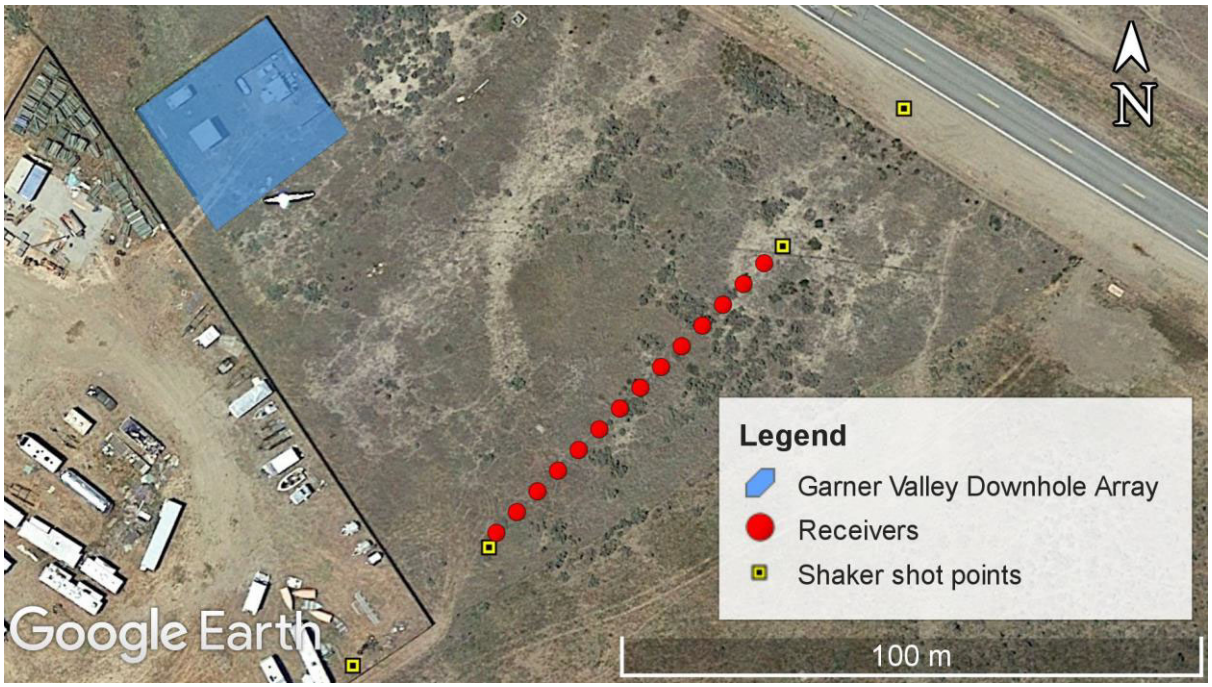

Figure 10. MASW array setup, with the NS13 array. The blue area identifies the location of the instrumented boreholes. Source: Google Earth Map data: © 2019 Google.

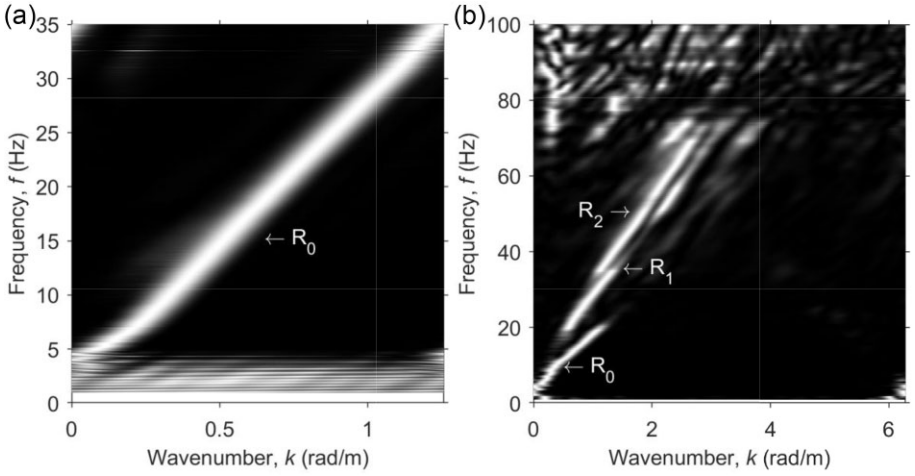

Figure 11. (a) Frequency–wavenumber representation of the wavefield recorded at GVDA. Data refer to the source-to-receiver distance equal to 5 m; (b) frequency–wavenumber representation of the wavefield recorded at MIR. Data refer to the source-to-receiver distance equal to 15 m. The brightest areas denote peaks of the frequency–wavenumber representation, each one corresponding to a *R*-wave propagation mode, labelled with R0, R1 and R2.

with lower variability compared to other techniques. The presence of well-constrained attenuation data at low frequencies is crucial for a reliable estimate of the damping ratio at greater depths.

4.3 Results from MIR

In this case, the estimation of $V_R(\omega)$ and $\alpha_R(\omega)$ is carried out through the CFDBFaMF algorithm. Indeed, due to the multimodal nature of the wavefield and the large number of receivers in the array, the MIR site is suitable for the application of this technique. Instead, the CFDBFa is expected to provide effective wave parameters that are different from the modal ones, hence it is not considered. The performance of the CFDBFaMF is compared to the CFM and the WD approaches. Specifically, Fig. 13 compares the median of the estimated $V_R(\omega)$ and $\alpha_R(\omega)$, computed from the elementary *R*-wave parameters corresponding to the 6 shot locations, for the three considered modes (i.e. R0, R1 and R2). Instead, Fig. 14 shows the corresponding logarithmic standard deviation.

For all the considered modes, the CFDBFaMF provides $V_R(\omega)$ and $\alpha_R(\omega)$ estimates generally compatible with the other processing techniques, over the same frequency ranges. Specifically, the resulting $V_R(\omega)$ agrees well with results provided by other techniques, although the CFM tends to return larger velocity values than the WD and the CFDBFaMF for R2. Instead, slightly larger differences characterize the estimated attenuation data. For instance, some discrepancies among the considered methods are observed for R0 at

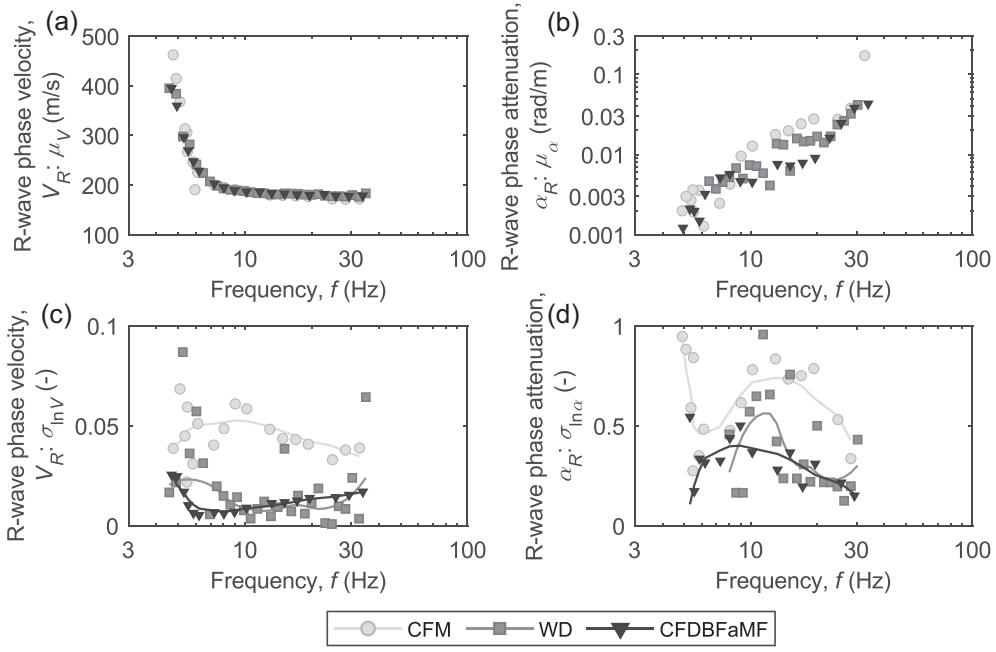

Figure 12. Estimated fundamental mode *R*-wave dispersion and attenuation curve according to the CFDBFa, WD and CFM methods, with reference to GVDA: (a) median estimated dispersion curves; (b) median estimated attenuation curves; (c) logarithmic standard deviation of the estimated dispersion curves; (d) logarithmic standard deviation of the estimated attenuation curves. Given the erratic trend of the standard deviation with the frequency, scatter plots in (c) and (d) are integrated with a continuous curve showing the trend for each approach, computed according to a moving average procedure, to facilitate data interpretation.

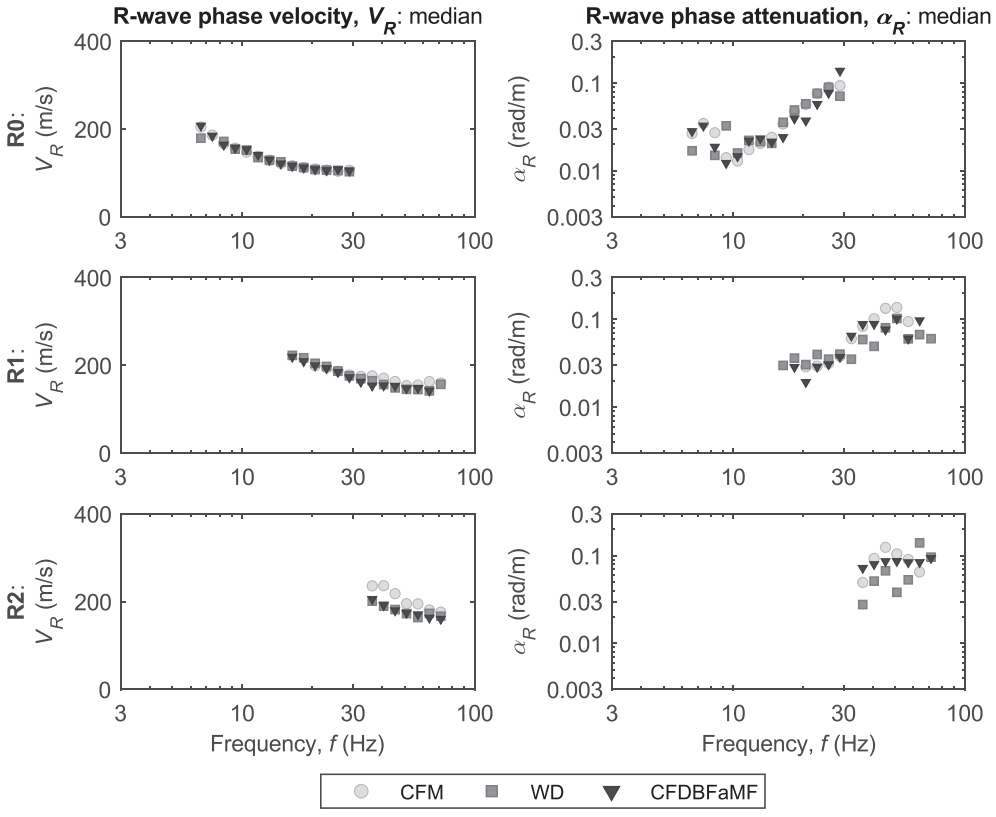

Figure 13. Median trends of the estimated modal *R*-wave dispersion curve (left-hand column) and attenuation curve (right-hand column) according to the CFDBFaMF, WD and CFM methods, with reference to MIR. Data refer to the first three *R*-wave modes, labelled as R0 (top row), R1 (middle row) and R2 (bottom row), respectively.

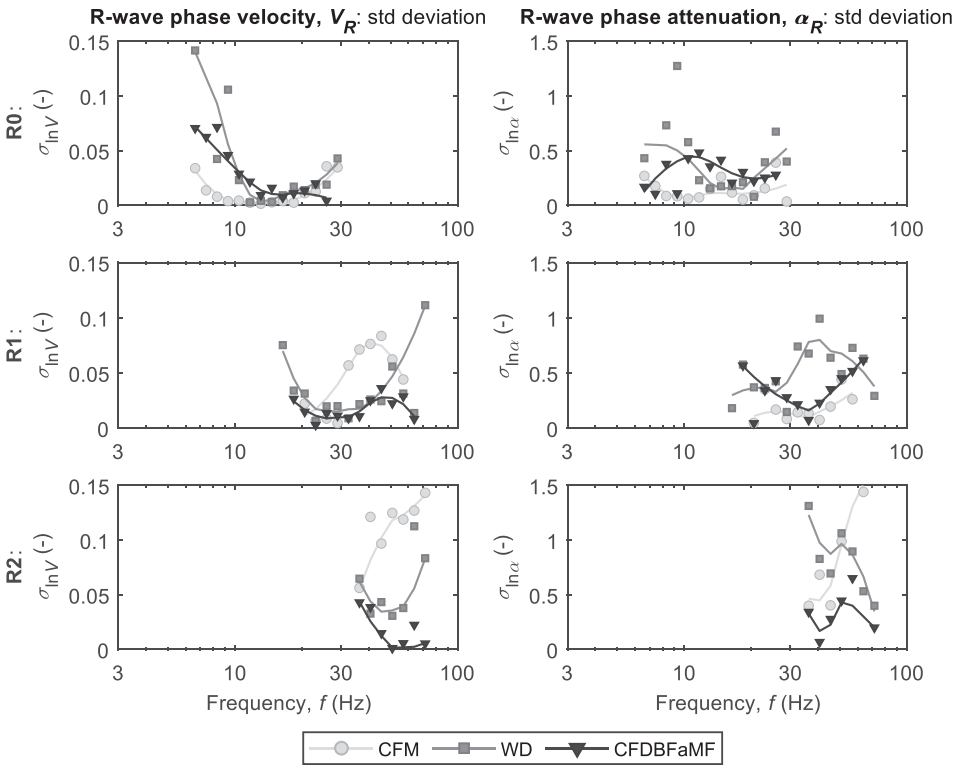

Figure 14. Trends of the logarithmic standard deviation of the estimated modal *R*-wave dispersion curve (left-hand column) and attenuation curve (right-hand column) according to the CFDBFaMF, WD and CFM methods, with reference to MIR. Data refer to the first three *R*-wave modes, labelled as R0 (top row), R1 (middle row) and R2 (bottom row), respectively. Given the erratic trend of the standard deviation with the frequency, scatter plots are integrated with a continuous curve showing the trend for each approach, computed according to a moving average procedure, to facilitate data interpretation.

low frequencies. In this range, $\alpha_R(\omega)$ tends to increase slightly, perhaps due to the influence of body waves. Furthermore, for R2, the CFM and the WD seem to give a more erratic trend with frequency, compared to the CFDBFaMF.

In terms of variability, the estimated phase attenuation is generally 5–10 times more variable than the corresponding phase velocity, for all the investigated modes and considered methods, similar to what was observed for GVDA. Specifically, for R0, the $V_R(\omega)$ variability is comparable for all the methods in the intermediate-to-high frequency band, with $\sigma_{\ln V}$ close to 0.02 at frequencies between 15 Hz and 25 Hz and slightly increasing to 0.05 at 30 Hz. At lower frequencies, however, $\sigma_{\ln V}$ tends to increase up to 0.07 for the CFDBFaMF. In this range, the $\sigma_{\ln V}$ is even larger for the WD ($\sigma_{\ln V} = 0.12$), whereas the CFM seems to better constrain the phase velocity estimates ($\sigma_{\ln V} = 0.04$). The variability of the estimated $\alpha_R(\omega)$ exhibits a similar trend, as $\sigma_{\ln \alpha}$ oscillates around 0.3 at high frequencies, whereas it increases up to 0.5 at low frequencies for the CFDBFaMF, which is an intermediate value between the CFM ($\sigma_{\ln \alpha} = 0.3$) and the WD ($\sigma_{\ln \alpha} = 0.6$). When focusing on higher modes, the average degree of uncertainty tends to increase for both $V_R(\omega)$ and $\alpha_R(\omega)$, especially for R2. The significantly larger variability for R2 may be a consequence of the strong influence of R1 on the frequency band in which R2 contributes to the recorded wavefield, which affects the quality of the estimated wave parameters for all the methods. Interestingly, although the CFM tends to better constrain R0 wave parameters compared to the other techniques, the estimation quality drops when focusing on higher modes for this site, especially in terms of variability (partially for R1 and more significantly for R2). Instead, the CFDBFaMF-based estimates have a larger variability than the CFM for R0, although it should be noted that the order of magnitude of $\sigma_{\ln V}$ and $\sigma_{\ln \alpha}$ is similar. Furthermore, when focusing on higher modes, the CFDBFaMF tends to outperform the other methods in terms of the variability of the $V_R(\omega)$ and $\alpha_R(\omega)$ estimates, especially for R2.

This result positively contributes to demonstrate the potential of the new CFDBFaMF method in retrieving dispersion and attenuation parameters, also when the wavefield is a combination of multiple propagating modes with comparable energy levels.

5 CONCLUSIONS

This paper presented a novel family of techniques, namely the FDBFa and the CFDBFa, for the estimation of the Rayleigh wave dispersion and attenuation data from seismic traces recorded in MASW surveys. Both techniques rely on a transformation of the wavefield, that allows to obtain the phase attenuation by carrying out a dispersion analysis on the transformed data. When the recorded wavefield consists of a pure, cylindrical *R*-wave, the cylindrical beamformer (hence the CFDBFa algorithm) perfectly matches the desired wave parameters. For this reason, the CFDBFa algorithm is the preferred choice to extract *R*-wave propagation parameters from active tests.

Furthermore, the proposed algorithm includes an additional step to extract modal wave parameters from multimode wavefields. The modified algorithm is named CFDBFaMF and it extracts each propagation mode from the recorded seismic traces by applying a band-pass filter, that preserves the mode of interest and removes additional components.

In general, the proposed algorithm is reliable and accurate on both synthetic waveforms and measured data from *in situ* surveys.

Indeed, a good level of compatibility of the results with alternative processing techniques is also demonstrated. Furthermore, the proposed procedure returns low-variability estimates over a broad frequency range, covering lower frequencies than those retrieved by other techniques. This advantage can be crucial when the purpose of the characterization is the definition of the damping structure at large depths.

A significant advantage of the proposed wavefield transformation is given by the possible extension to other processing techniques. For instance, the proposed technique can be straightforwardly applied to ambient vibration analysis on 2-D passive arrays. This would allow for improved coverage of experimental data in the low-frequency range, thus constraining estimated damping ratios at greater depths. Besides, the proposed filtering scheme can be generally applied for dispersion analysis, to improve the quality of the phase velocity estimates from multimode wavefields.

The companion paper deals with the solution of the Rayleigh inverse problem to estimate the shear-wave velocity and small-strain damping ratio from the experimental dispersion and attenuation data.

ACKNOWLEDGMENTS

The authors are grateful to Mauro Francavilla, who helped in processing data measured at the Garner Valley Downhole Array. The study has been partially supported by the ReLUIS 3 project, funded by the Italian Civil Protection Agency. Data at the Garner Valley Downhole Array were extracted from the NSF Project 'Collaborative Research: 3D Ambient Noise Tomography (3D ANT) for Natural Hazards Engineering' grant CMMI-1931162. However, any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The seismic instruments were provided by the Incorporated Research Institutions for Seismology (IRIS) through the PASSCAL Instrument Center at New Mexico Tech. Data collected will be available through the IRIS Data Management Center. The facilities of the IRIS Consortium are supported by the National Science Foundation's Seismological Facilities for the Advancement of Geoscience (SAGE) Award under Cooperative Support Agreement EAR-1851048.

DATA AVAILABILITY

Synthetic waveform data and codes are available on request from the authors. Data at the Garner Valley Downhole Array were extracted from the NSF Project 'Collaborative Research: 3D Ambient Noise Tomography (3D ANT) for Natural Hazards Engineering' grant CMMI-1931162, and they are available at http://service.iris.edu/ph5ws/. Data at the Mirandola site were extracted from the InterPacific Project and are available at http://interpacific.geopsy.org/.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Supp-Info

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.

REFERENCES

Aimar, M., 2022. Uncertainties in the estimation of the shear-wave velocity and the small-strain damping ratio from surface wave analysis, *PhD thesis*, Politecnico di Torino, Torino, Italy. https://hdl.handle.net/11583/2972200.

Aimar, M., Foti, S. & Cox,, B.R. 2024. Novel techniques for in-situ estimation of shear-wave velocity and damping ratio through MASW testing - II: A Monte Carlo algorithm for the joint inversion of phase velocity and phase attenuation. *Geophys. J. Int.,* https://doi.org/10.1093/gji/ggae050.

Askari, R. & Ferguson, R.J., 2012. Dispersion and the dissipative characteristics of surface waves in the generalized S-transform domain. *Geophysics,* **77,** V11–V20.

Badsar, S.A., Schevenels, M., Haegeman, W. & Degrande, G., 2010. Determination of the material damping ratio in the soil from SASW tests using the half-power bandwidth method. *Geophys. J. Int.,* **182,** 1493–1508.

Bergamo, P., Maranò, S. & Fäh, D., 2023. Joint estimation of s-wave velocity and damping ratio of the near-surface from active Rayleigh wave surveys processed with a Wavefield decomposition approach. *Geophys. J. Int.,* **233,** 1560–1579.

Bergamo, P., Maranò, S., Imperatori, W. & Fäh, D., 2018. Wavedec code: an application to the joint estimation of shear modulus and dissipative properties of the near-surface from multi-component, active surface-wave surveys, in *Proceedings of the 36th General Assembly of the European Seismological Commission,* ESC2018, Valletta, Malta, 2–7 September 2018.

Bergamo, P., Maranò, S., Imperatori, W., Hobiger, M. & Fäh, D., 2019. Wavefield decomposition technique applied to active surface wave surveys: towards joint estimation of shear modulus and dissipative properties of the near-surface, in *Proceedings of the EPOS@ SERA "Strong Motion Site Characterization" Workshop*.

Bruekers, A.A.M.L., 2009. *Symmetry and Efficiency in Complex FIR Filters,* Technische Universiteit Eindhoven.

Comina, C., Foti, S., Boiero, D. & Socco, L.V., 2011. Reliability of VS,30 evaluation from surface-wave tests. *J. Geotech. Geoenviron. Eng.,* **137,** 579–586.

Cormen, T.H., Leiserson, C.E., Rivest, R.L. & Stein, C., 2001. *Introduction to Algorithms,* MIT Press.

Forbriger, T., 2003. Inversion of shallow-seismic wavefields: I. Wavefield transformation. *Geophys. J. Int.,* **153,** 719–734.

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

Foti, S., 2000. Multistation methods for geotechnical characterization using surface waves, *PhD thesis*, Politecnico di Torino, Torino, Italy.

Foti, S., 2003. Small-strain stiffness and damping ratio of Pisa clay from surface wave tests. *Géotechnique,* **53,** 455–461.

Foti, S., 2004. Using transfer function for estimating dissipative properties of soils from surface-wave data. *Near Surf. Geophys.,* **2,** 231–240.

Foti, S., Aimar, M. & Ciancimino, A., 2021. Uncertainties in small-strain damping ratio evaluation and their influence on seismic ground response analyses, in *Latest Developments in Geotechnical Earthquake Engineering and Soil Dynamics,* Springer.

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

Gao, L., Pan, Y., Tian, G. & Xia, J., 2018. Estimating Q factor from multimode shallow-seismic surface waves. *Pure appl. Geophys.,* **175,** 2609–2622.

Garofalo, F. *et al.* 2016a. InterPACIFIC project: comparison of invasive and non-invasive methods for seismic site characterization. Part II: intercomparison between surface-wave and borehole methods. *Soil Dyn. Earthq. Eng.,* **82,** 241–254.

Garofalo, F. *et al.* 2016b. InterPACIFIC project: comparison of invasive and non-invasive methods for seismic site characterization. Part I: intra-comparison of surface wave methods. *Soil Dyn. Earthq. Eng.,* **82,** 222–240.

Gold, B. & Jordan, K.L., 1969. A direct search procedure for designing finite-duration impulse response filters. *IEEE Trans. Audio Electroacoust.,* **17,** 33–36.

Hill, R.I., 1981. Geology of Garner Valley and vicinity, in *Geology of the San Jacinto Mountains, Annual Field Trip Guidebook No. 9,* eds Brown, A.R. & Ruff, R.W., South Coast Geological Society.

Kausel, E. & Roësset, J.M., 1981. Stiffness matrices for layered soils. *Bull. seism. Soc. Am.,* **71,** 1743–1761.

Lacoss, R.T., Kelly, E.J. & Toksoz, M.N., 1969. Estimation of seismic noise structure using arrays. *Geophysics,* **34,** 21–38.

Lai, C.G. & Rix, G.J., 1998. Simultaneous inversion of Rayleigh phase velocity and attenuation for near-surface site characterization, *PhD thesis*, Georgia Institute of Technology, Atlanta, GA, USA.

Lai, C.G., Rix, G.J., Foti, S. & Roma, V., 2002. Simultaneous measurement and inversion of surface wave dispersion and attenuation curves. *Soil Dyn. Earthq. Eng.,* **22,** 923–930.

Luo, Y., Xia, J., Miller, R.D., Xu, Y., Liu, J. & Liu, Q., 2008. Rayleigh-wave dispersive energy imaging using a high-resolution linear Radon transform. *Pure appl. Geophys.,* **165,** 903–922.

Maranò, S., Hobiger, M., Bergamo, P. & Fäh, D., 2017. Analysis of Rayleigh waves with circular wavefront: a maximum likelihood approach. *Geophys. J. Int.,* **210,** 1570–1580.

McMechan, G.A. & Yedlin, M.J., 1981. Analysis of dispersive waves by wave field transformation. *Geophysics,* **46,** 869–874.

Misbah, A.S. & Strobbia, C., 2014. Joint estimation of modal attenuation and velocity from multichannel surface wave data. *Geophysics,* **79,** EN25–EN38.

Mitra, S.K., 2006. *Digital Signal Processing—A Computer-Based Approach,* McGraw-Hill.

Park, C.B., Miller, R.D. & Xia, J., 1998. Imaging dispersion curves of surface waves on multi-channel record, in *Proceedings of the SEG: 68th Annual International Meeting,* pp. 1377–1380, Society of Exploration Geophysicists.

Passeri, F., 2019. Development of advanced geostatistical models of shear wave velocity profiles to manage uncertainties and variabilities in ground response analyses, *PhD thesis*, Politecnico di Torino, Torino, Italy.

Rahimi, S., Wood, C.M. & Teague, D.P., 2021. Performance of different transformation techniques for MASW data processing considering various site conditions, near-field effects, and modal separation. *Surv. Geophys.,* **42,** 1197–1225.

Rix, G.J., Lai, C.G. & Foti, S., 2001. Simultaneous measurement of surface wave dispersion and attenuation curves. *Geotech. Test. J.,* **24,** 350–358.

Rix, G.J., Lai, C.G. & Wesley Spang, A., Jr., 2000. In situ measurement of damping ratio using surface waves. *J. Geotech. Geoenviron. Eng.,* **126,** 472–480.

Schevenels, M., Degrande, G. & François, S., 2009. EDT: an elastodynamics toolbox for MATLAB. *Comput. Geosci.,* **35**(8), 1752–1754.

Schwartz, G.E., 1978. Estimating the dimension of a model. *Ann. Stat.,* **35,** 461–464.

Stokoe, K.H., II, Cox, B.R., Clayton, P.M. & Menq, F.-Y., 2020. NHERI@UTexas Experimental facility with large-scale mobile shakers for field studies. *Front. Built Environ.,* **6,** 575973.

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

Vantassel, J.P., Crocker, J.A., Cox, B.R. & Tran, K.T., 2023. Subsurface imaging dataset acquired at the Garner Valley downhole array site using a dense network of three-component nodal stations, *Earthq. Spectra,* **40**(1), doi:10.1177/87552930231209734.

Verachtert, R., Lombaert, G. & Degrande, G., 2017. Multimodal determination of Rayleigh dispersion and attenuation curves using the circle fit method. *Geophys. J. Int.,* **212,** 2143–2158.

Wood, C.M. & Cox, B.R., 2012. A comparison of MASW dispersion uncertainty and bias for impact and harmonic sources, in *Proceedings of the GeoCongress 2012: State of the Art and Practice in Geotechnical Engineering,* Oakland, CA, 25–29 March 2012, pp. 2756–2765.

Yoon, S., 2005. Array-based measurements of surface wave dispersion and attenuation using frequency-wavenumber analysis, *PhD thesis*, Georgia Institute of Technology, Atlanta, GA, USA.

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

Zywicki, D.J., 1999. Advanced signal processing methods applied to engineering analysis of seismic surface waves, *PhD thesis*, Georgia Institute of Technology, Atlanta, GA, USA.