Geophysical Journal International

Geophys. J. Int. (2020) **221,** 1802–1819
Advance Access publication 2020 March 4
GJI Seismology

doi: 10.1093/gji/ggaa106

Multimode multioffset phase analysis of surface waves, a new approach to extend MOPA to higher modes

I. Barone,[1] C. Strobbia[2] and G. Cassiani[1]

[1]*Dipartimento di Geoscienze, Università degli Studi di Padova, 35131 Padova, Italy. E-mail: ilaria.barone@phd.unipd.it*
[2]*RealTimeSeismic, 64000 Pau, France*

Accepted 2020 March 1. Received 2020 February 25; in original form 2019 June 26

SUMMARY

Multioffset phase analysis (MOPA) is a fairly recent technique for evaluating seismic surface wave dispersion and estimating the presence of lateral variations. The main limitation of MOPA is that it is based on the assumption of one predominant mode, usually the fundamental mode, in the wave propagation. However, MOPA can be extended (at least) to the two-mode case: this new technique will be called multimode MOPA (MMMOPA). The method employs both amplitude and phase spectral information. The analysis is performed for each frequency independently. The presence of two modes causes the amplitude to have an oscillating behaviour as a function of offset (beats): the spatial period of the oscillating amplitude is identified, amplitude maxima and minima are extracted, and the local wavenumber is computed via linear regression. The resulting multimodal dispersion curve is consequently derived. Model uncertainties can be estimated by propagating the experimental phase and amplitude error variances through the different steps of the analysis all the way to the final phase velocities. An algorithm running the process in an automatic way has been implemented and tested on both synthetic and real data, with success. This is the base for future developments that, in the MOPA framework, can take into account rapid lateral velocity variations within the same acquisition window and estimate the modal absorption, for the estimation of the damping ratio, even in the presence of multimode surface wave propagation.

Key words: Fourier analysis; Seismic attenuation; Surface waves and free oscillations.

1 INTRODUCTION

Seismic surface waves have been considered noise for many years, especially in the context of reflection surveys. The first attempts to extract useful information from these waves date back to the 1950s (Jones 1958, 1962; Ballard 1964), when a very simple approach, the steady-state Rayleigh method, was proposed for engineering site characterization purposes. In the early 1980s, spectral methods for surface wave analysis were introduced and applied in the context of geotechnical studies, leading to the development of the well-known spectral analysis of surface waves (SASW; Nazarian & Stokoe II 1984), but the true impulse to surface wave methods (SWM) was given by the historical progress in electronics and computers, which made it possible to acquire and process a large amount of data even in geotechnical applications: the consequence has been the rise of multichannel techniques (McMechan & Yedlin 1981; Park *et al.* 1999a; Xia *et al.* 1999), which revealed to be more robust than SASW for many types of applications at very different spatial scales, including the cases where multimodal propagation is critical.

The main purpose of SWM is to characterize the near surface in terms of shear elastic properties by studying the dispersive characteristics of the medium and by solving the related inverse problem. The whole analysis is made in the spectral domain which, depending on each specific technique, could be frequency–wavenumber, frequency–phase velocity, frequency–offset or others.

Both active and passive techniques have been developed over the years. Among the active techniques, the multichannel analysis of surface waves (MASW; Park *et al.* 1999a; Xia *et al.* 1999) is one of the most widely used, both for small-scale (environmental, engineering-geotechnical) and large-scale (oil and gas and mineral exploration) applications. The MASW method aims at retrieving dispersion curves by picking amplitude maxima in one of the aforementioned spectral domains, either manually or by the use of automatic algorithms. In some cases, also higher modes are picked and included in the inversion process (Gabriels *et al.* 1987; Park *et al.* 1999b), since these modes are known to bring independent additional information, helping constrain the inversion problem and increase both investigation depth and model resolution (Xia *et al.* 2003; Socco & Strobbia 2004; Maraschini *et al.* 2010).

The most critical aspect of the multimodal analysis is still the mode numbering identification. Normally, the most energetic mode is taken as the fundamental one, but this is not always the case, especially in the presence of inversely dispersive media or strong velocity contrasts (Tokimatsu *et al.* 1992; Maraschini *et al.* 2010;

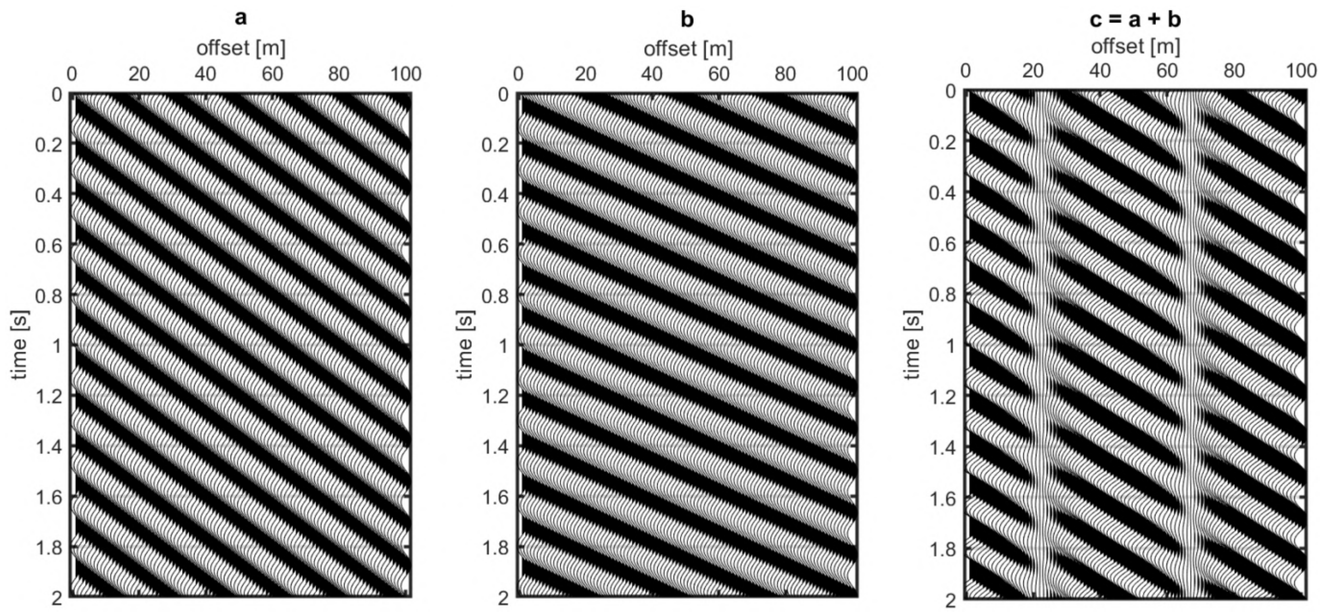


Figure 1. Two monochromatic waves, (a) and (b), having the same frequency $f = 6$ Hz and amplitude $A = 5$, but propagating in the same direction with different apparent velocities $c_a = 100$ m s^{-1} and $c_b = 160$ m s^{-1}. Their interference produces seismogram (c).

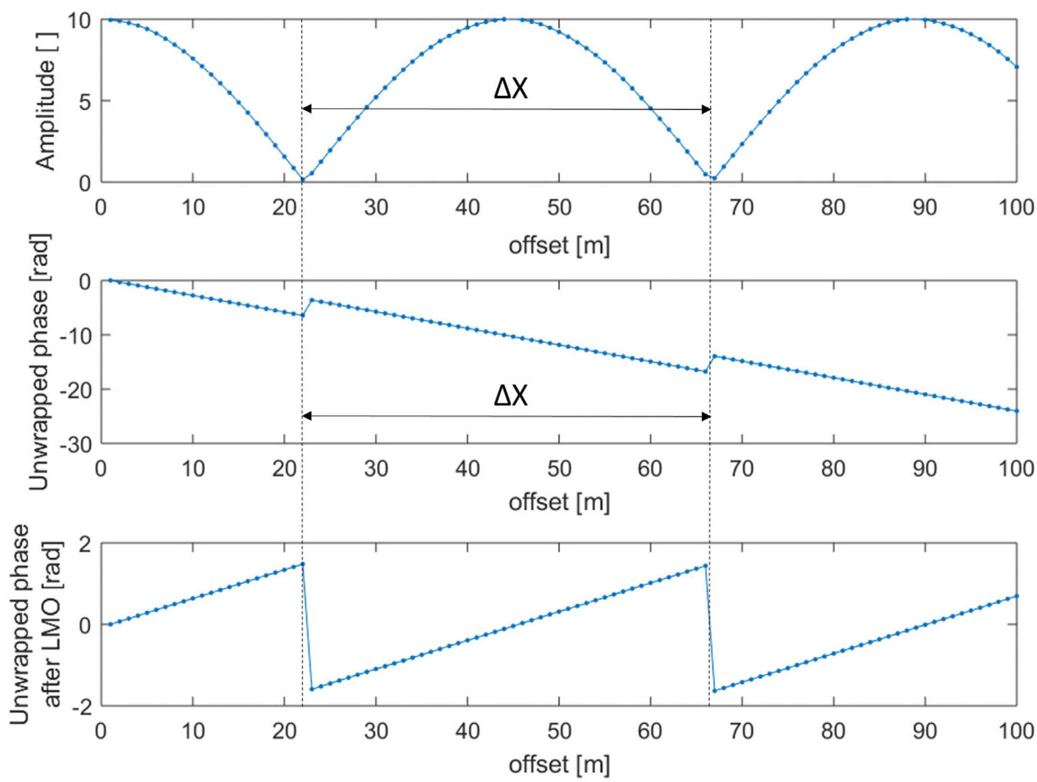


Figure 2. From top to bottom: amplitude, unwrapped phase and unwrapped phase after LMO correction for the seismogram in Fig. 1, third panel. The analysed frequency is 6 Hz and the velocity used for the LMO is 100 m s^{-1}.

Boaga *et al.* 2014). Mode misidentification is often caused by a bad choice of acquisition parameters: long, multicomponent and fine-sampled arrays permit a better resolution of the different modes (Foti *et al.* 2002) . This is particularly true in the case of the so-called 'osculation' of different modes, thoroughly discussed by Boaga *et al.* (2013), who also propose a solution based on multicomponent data acquisition. However, strong limitations are imposed by field conditions and instrument availability. Luo *et al.* (2008) propose the application of a high-resolution linear Radon transform to surface wave data in order to obtain *f*–*v* spectra with increased mode resolution with no need for larger arrays lengths.

Correctly picking a multimodal dispersion curve is, however, not only a matter of spectral resolution. Low-frequency linear noise

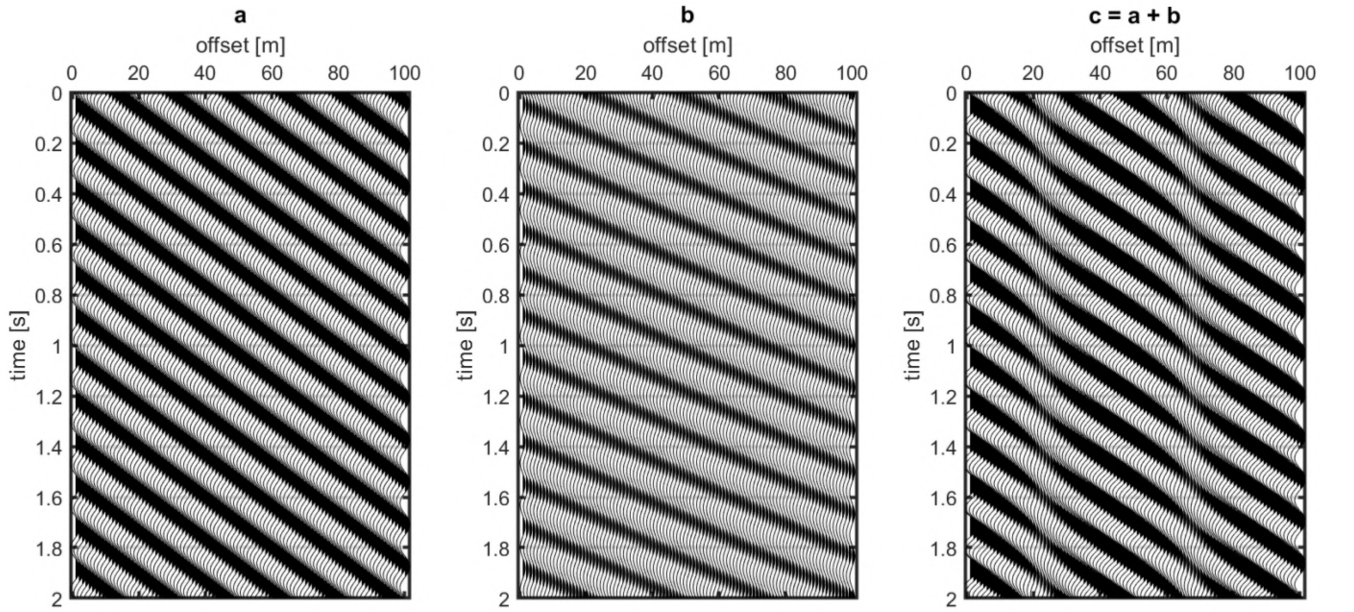


Figure 3. Two monochromatic waves, (a) and (b), having the same frequency $f = 6$ Hz, but different amplitudes $A_a = 10$ and $A_b = 5$, and propagating in the same direction with different apparent velocities $c_a = 100$ m s^{-1} and $c_b = 160$ m s^{-1}. Their interference produces seismogram (c).

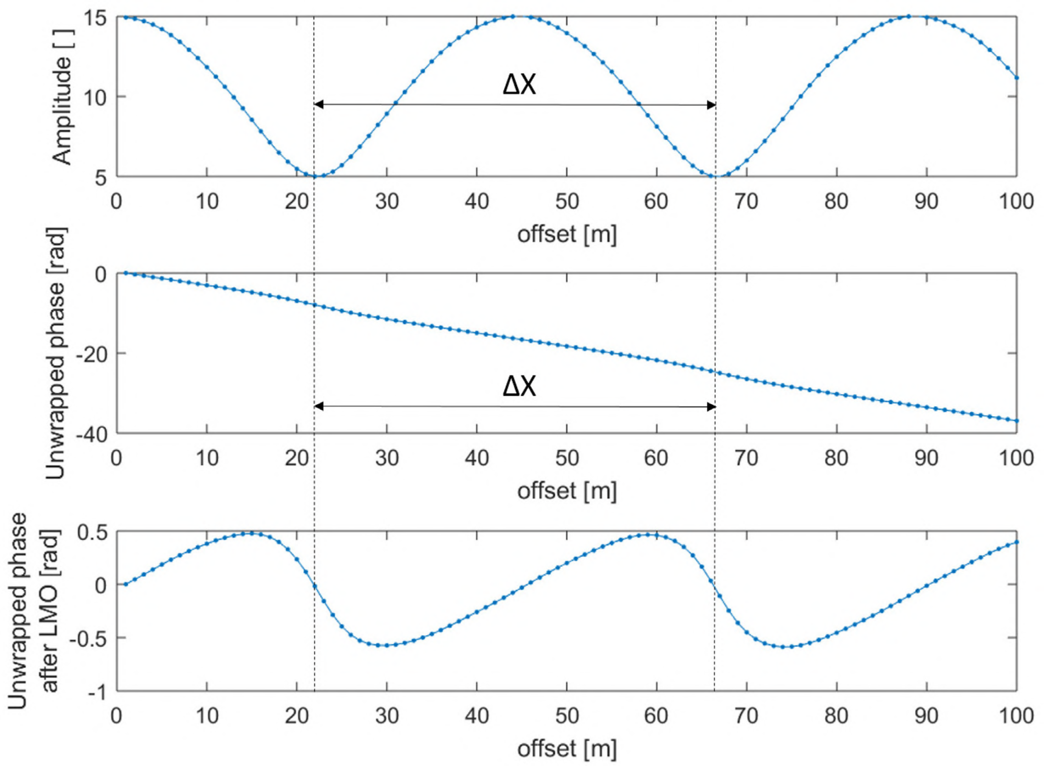


Figure 4. From top to bottom: amplitude, unwrapped phase and unwrapped phase after LMO correction for the seismogram in Fig. 3, third panel. The analysed frequency is 6 Hz and the velocity used for the LMO is 100 m s^{-1}.

and/or ignored lateral velocity variations may pollute the spectrum. While the first issue can be resolved with multiple repetitions of the same shot, the latter issue is subtler: Boaga *et al.* (2014) show the compresence of different branches for each mode in the *f–k* spectrum, which makes mode picking and identification challenging and potentially misleading. This is why the MASW method, in spite of its robustness, is not suitable in the presence of lateral heterogeneities. Adaptation of the method to laterally varying contexts

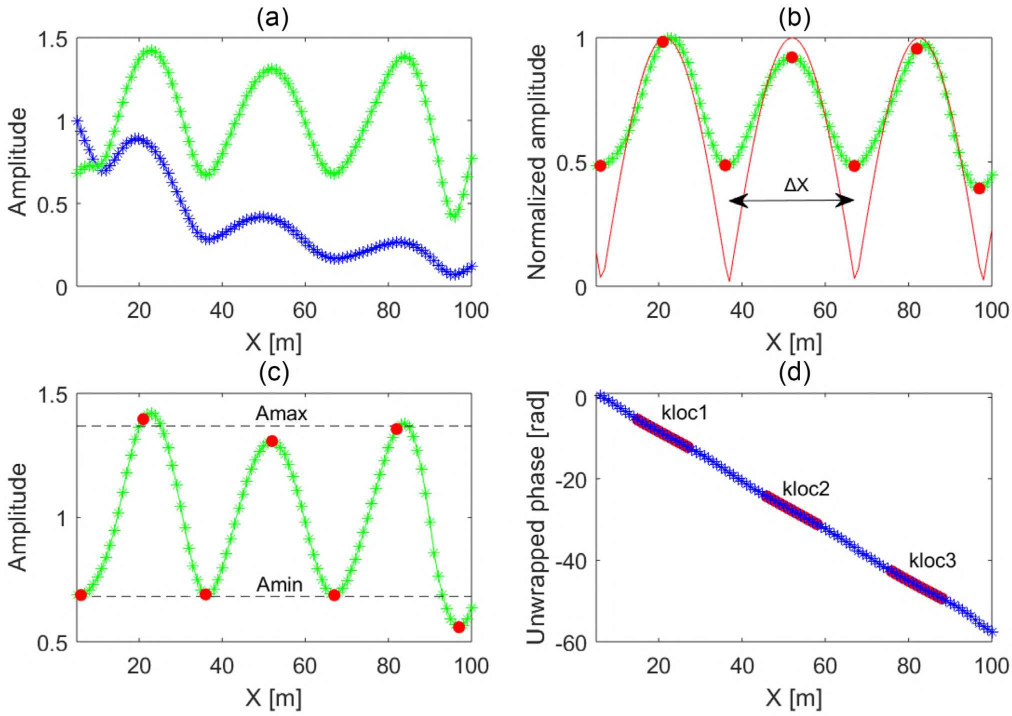


Figure 5. Steps of MMMOPA. (a) Raw amplitude (blue line) and amplitude corrected for attenuation (green line). (b) Normalized corrected amplitude (green), fitting sinusoid (red line) and position of local maxima and minima (red points). (c) Average amplitude maximum and minimum (black dotted lines) computed from all local maxima and minima (red points). (d) Unwrapped raw phase (blue line) and intervals selected for local linear regressions (red points). Example taken from a synthetic data set.

has been attempted, but this often consists only of moving window averaging (Boiero & Socco 2011; Bergamo 2012).

More recently, a technique based on the analysis of the phase versus offset has been developed by Strobbia & Foti (2006): this procedure, called multioffset phase analysis (MOPA), gives an estimate of phase velocities and their uncertainties, preserving lateral resolution, even in the presence of abrupt lateral variations (Vignoli & Cassiani 2010). When multiple shots are available, a statistical approach (sMOPA; Vignoli *et al.* 2011), taking advantage of data redundancy, is advisable. In this case a finer estimate of phase uncertainties is obtained, which may be used in turn to weight the linear regression of phase versus offset.

Despite the fact that MOPA is a robust technique, it has a critical limitation: it assumes that a single mode is dominant in the propagation at each frequency, and that the signal phase is a modal phase. If this condition is not satisfied, MOPA can still be used if the different modes are well separated: a preliminary processing step is required, which consists in isolating each mode, for example, by filtering the data in the *f*–*k* domain. Otherwise, when the different modes are close to each other and/or lateral variations are quite important, isolating modes prior to the analysis might be not applicable. Note however that, in the presence of lateral variations, the use of the *f*–*k* filtering does not take into account that energy maxima are the results of superposition coming from different parts of the survey lines, where modes are seemingly different. This may cause the identification of apparent modes that correspond to no real mode in no place of the profile.

We present here a new approach, extending MOPA to the two-mode case, and possibly to more modes. The resulting multimode MOPA (MMMOPA) is based on the concurrent analysis of both amplitude and phase.

In the sequel, we first explain the theoretical idea behind the method. Then, we spell out the criteria of applicability of the method and describe the processing algorithm with its main steps of application. Three successful examples are illustrated, one with a synthetic data set and two with real data. Finally, a brief discussion shows pros and cons of the method and future perspectives.

2 THEORETICAL BACKGROUND

MOPA is a very powerful technique to extract phase information from multichannel data, particularly in the presence of lateral variations. MOPA takes advantage of the linear dependence between the phase of the signal at a specific frequency and the offset, valid for the propagation of one single mode. In this case, the slope of the line is linked to the wavenumber, that is, to the phase velocity: by performing a frequency by frequency linear regression of the phase versus offset, the dispersion relation is inferred. Yet, MOPA has a drawback: when the underlying assumption, that most energy is focused on a single mode, is not fulfilled, multichannel filtering, such as *f*–*k* or *f*–*x* dip, is needed to isolate one mode (normally, the fundamental, or at least the one assumed to be such) from the others. This filtering operation could be, in some cases, quite complex, as mentioned in Section 1.

We present a method that aims to overcome the limits of the single mode propagation as assumed by standard MOPA. The proposed method (MMMOPA) is capable of retrieving dispersion curves of two different modes at the same time, and potentially more modes, with no specific need of filtering in the *f*–*k* domain. The main requirements for the method are that both modes are energetic and that the phase spectrum is affected by the mode superposition. In

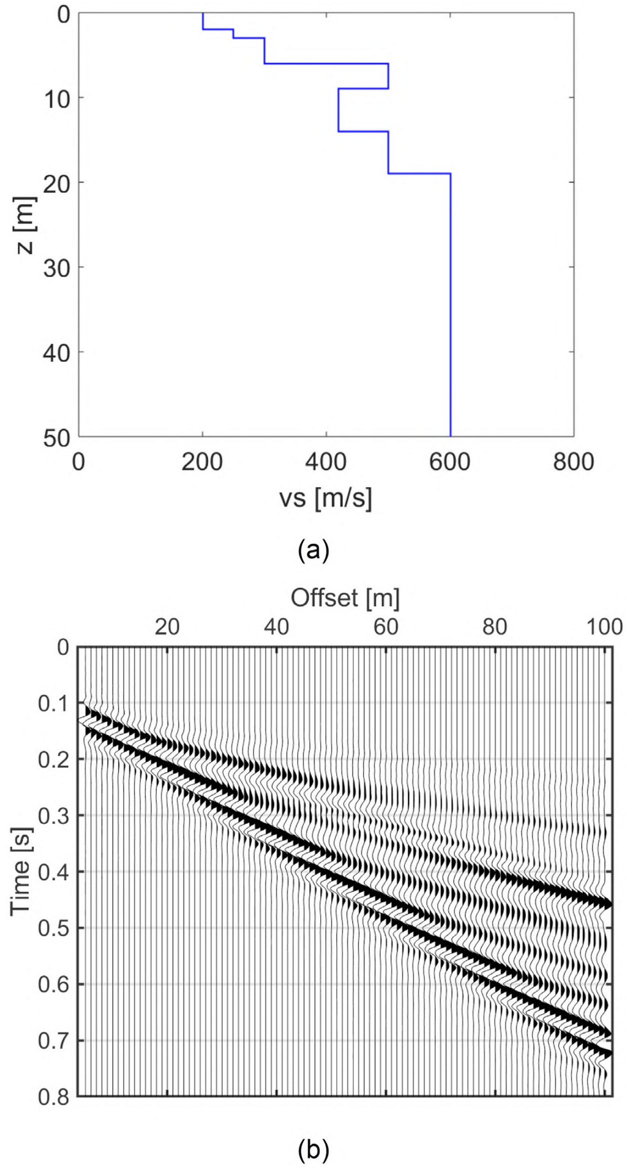


Figure 6. (a) 1-D Shear wave velocity model used to generate the synthetic data set and (b) resulting synthetic seismogram shown with trace by trace amplitude normalization.

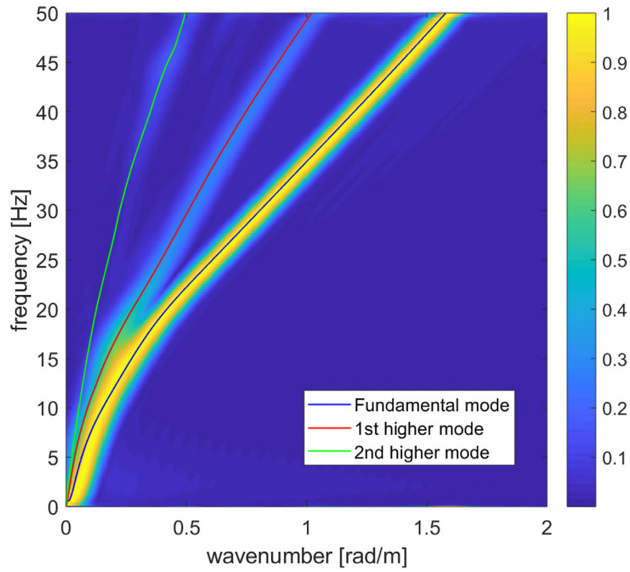


Figure 7. Normalized *f–k* spectrum for the seismogram in Fig. 6. The three first theoretical modes are also displayed.

addition, no other coherent signal should interfere. In the proposed method, both phase and amplitude need to be used. The analysis and use of the spectral amplitude simultaneously for multiple modes, as discussed later, also overcome the main limitations of the standard attenuation assessment techniques.

For illustration purposes, let us start with a very simple case. Let us consider the simple interference between two sinusoidal signals, a and b, both with frequency $f = 6$ Hz and amplitude $A = 5$, and yet propagating with different apparent velocities $c_a = 100\ \mathrm{m\,s^{-1}}$ and $c_b = 160\ \mathrm{m\,s^{-1}}$. Two synthetic seismograms with 100 traces, a receiver spacing of 1 m, a record length of 4 s and a sampling interval of 10 ms were generated both for a and b, and combined together by a simple sum: let us call c the combination of a and b (Fig. 1). Note that no specific reference is here made to surface waves, but only to the interference between two monochromatic waves. The amplitude, as expected, presents a regular oscillating pattern with offset due to positive and negative interferences between the two waves (beats): the spectral amplitude oscillates between $2A$ and 0, corresponding to perfect constructive and destructive interferences, respectively. Conversely, the phase presents regular jumps of π in conjunction with amplitude minima, at the positions of phase opposition (Fig. 2).

Let us consider now the same example, but varying the relative amplitude between the two sinusoids: in this second case, the amplitude of a is double that of b. The combined seismogram, c (Fig. 3), presents a pattern that resembles the previous case but here amplitudes are not going to zero at any location. Similarly, phase jumps are not as abrupt as in the previous case: the phase slope has smooth variations also in correspondence of amplitude minima, while it is constant for amplitude maxima (Fig. 4). Note that in both Fig. 2 and Fig. 4, the bottom panel shows phases with a linear-moveout correction (LMO) applied. LMO shifts seismic records back in time, where the shift Δt linearly depends on the source–receiver distance x. In formula:

$$\Delta t = \frac{x}{V_{\mathrm{Lmo}}}, \tag{1}$$

where V_{Lmo} is the velocity used for the correction. This time shift in the time domain translates to a phase shift in the spectral domain. LMO correction is here used for displaying purposes only, to enhance the small-scale phase variations produced by modes interference: to this end, the velocity used for the correction is the one better describing the large-scale phase trend.

The examples above show that the combination of two different monochromatic signals having the same frequency produces observable patterns on seismograms. In particular:

1. The oscillating spatial period of amplitude and phase, let us call it ΔX, is linked to the difference of wavenumbers (i.e. to the difference in velocity) of the two signals.
2. Amplitude maxima and minima give information about the sum (constructive or in-phase interference) and the difference (destructive or out-of-phase interference) of the individual amplitudes of the two sinusoids.
3. The local slope of the phase versus offset, corresponding to the position of amplitude maxima, is constant and always corresponds to a value in-between the wavenumbers $k_a = 2\pi f/c_a$ and $k_b = 2\pi f/c_b$ of the two sinusoids.

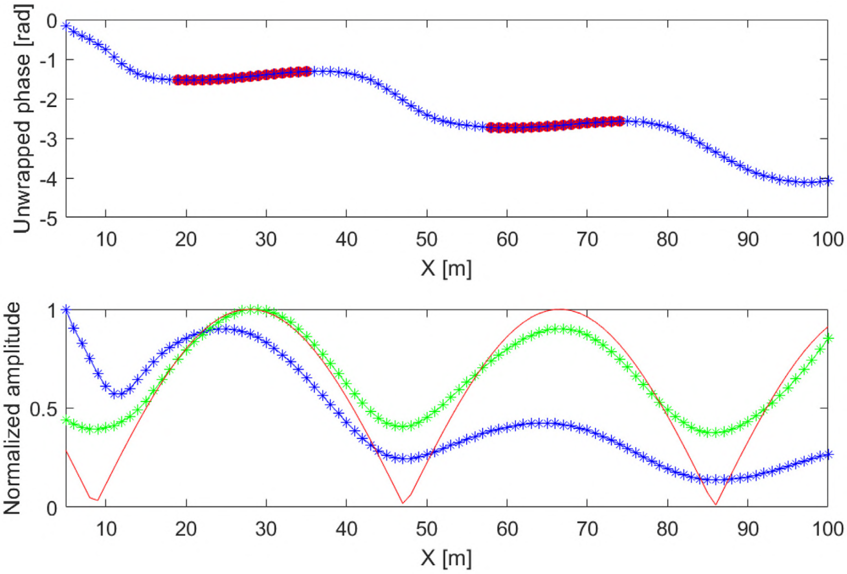


Figure 8. Top panel: unwrapped phase (blue line) with values automatically selected by the algorithm for each independent k_{loc_i} computation (red circles). Data have been LMO corrected with a velocity of 300 m s^{-1}. Bottom panel: normalized amplitude (blue line), normalized amplitude after correction for attenuation (green line), fitting sinusoid (red line). For both plots frequency is 22 Hz. Example from the synthetic data set.

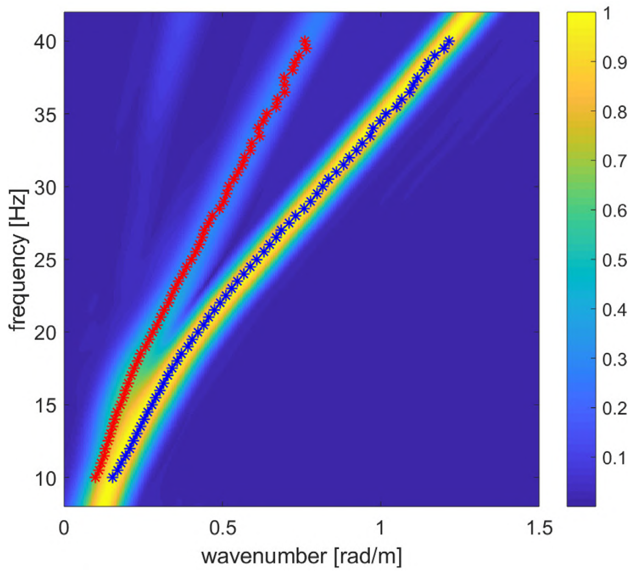


Figure 9. Final multimodal dispersion curve overlapped to the normalized *f*–*k* spectrum: fundamental mode is in blue, first higher mode in red. Example from the synthetic data set.

Regarding the first point above, ΔX may be seen as the spatial interval needed to produce the difference of one cycle between the two sinusoids:

$$k_a \Delta X - k_b \Delta X = 2\pi, \tag{2}$$

that is,

$$\Delta X = \frac{2\pi}{k_a - k_b}. \tag{3}$$

As far as amplitude and phase are concerned (second and third points of the list above), their behaviour can be mathematically described as follows. The wavefield resulting from the summation of the two sinusoidal signals may be expressed as

$$S_c(x) = A_c \mathrm{e}^{\mathrm{i}(\omega t - k_c x)} = A_a \mathbf{e}^{\mathrm{i}(\omega t - k_a x)} + A_b \mathrm{e}^{\mathrm{i}(\omega t - k_b x)}, \tag{4}$$

where $\omega = 2\pi f$ is the angular frequency, t is time and x is the offset coordinate. Let us first consider the amplitude A_c, which can be expressed as (see Appendix A for details):

$$A_c = \sqrt{A_a^2 + A_b^2 + 2A_a A_b \cos[(k_a - k_b)x]}. \tag{5}$$

Eq. (5) describes the oscillating behaviour of the amplitude: maxima and minima take place when $\cos[(k_a - k_b)\, x]$ is equal to 1 and -1, respectively. In formula:

$$A_{\max} = \sqrt{A_a^2 + A_b^2 + 2A_a A_b} = A_a + A_b \tag{6}$$

$$A_{\min} = \sqrt{A_a^2 + A_b^2 - 2A_a A_b} = \mid A_a - A_b \mid . \tag{7}$$

Note that eq. (7) has two possible solutions depending on the relative power of the two individual waves, so that

$$A_a = \frac{A_{\max} + A_{\min}}{2},\ A_b = \frac{A_{\max} - A_{\min}}{2}, \tag{8}$$

when A_a is greater than A_b, and

$$A_a = \frac{A_{\max} - A_{\min}}{2},\ A_b = \frac{A_{\max} + A_{\min}}{2}, \tag{9}$$

when A_b is greater than A_a. This means that the mere observation of the amplitude distribution with offset produces an intrinsic ambiguity about the attribution of the computed amplitude values to sinusoids *a* or *b*, unless we dispose of additional information about which of the two waves is more energetic. Let us now analyse the phase of the wavefield in eq. (4), which can be expressed as:

$$-k_c x = \arctan\left\{\frac{A_b \sin[(k_a - k_b)x]}{A_a + A_b \cos[(k_a - k_b)x]}\right\} - k_a x. \tag{10}$$

If we differentiate eq. (10) with respect to x, we obtain the expression of k_c as a function of x which, in the specific case of $\cos[(k_a - k_b)\, x] = 1$ and $\sin[(k_a - k_b)\, x] = 0$ (amplitude maxima), reduces to (see

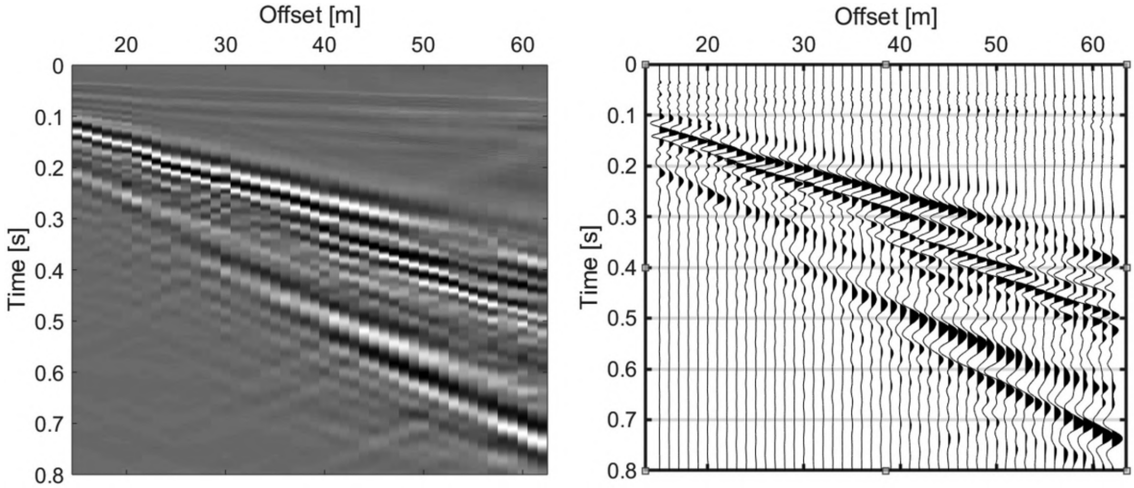


Figure 10. Stack of seismograms from 10 repeated shots, for the first shot location, shown with trace by trace normalization—Mirandola data set.

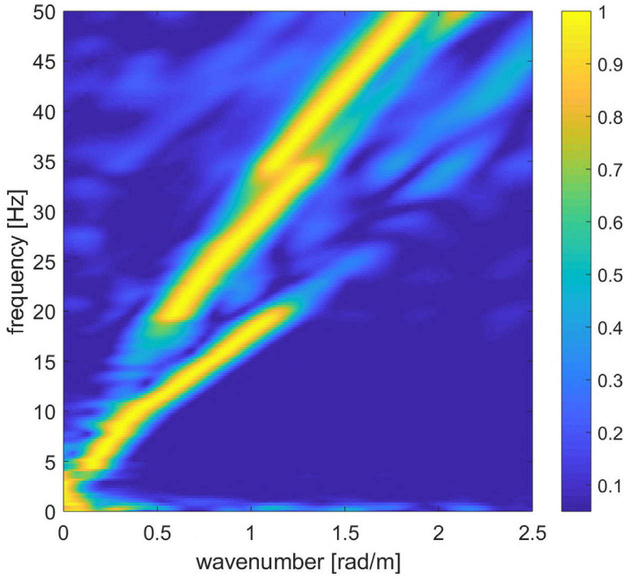


Figure 11. Normalized *f–k* spectrum for the seismogram in Fig. 10.

Appendix A for further details):

$$k_{\mathrm{loc}} = \frac{A_a k_a + A_b k_b}{A_a + A_b}. \tag{11}$$

Eq. (11) states that the local phase slope k_{loc}, corresponding to the amplitude maxima, is equal to the weighted mean of the two individual slopes k_a and k_b, where the weights are given by the sinusoid amplitudes.

Supposing we have some *a priori* information about which of the two waves is the most energetic, we are able to compute individual wave amplitudes, A_a and A_b, from either eq. (8) or (9). Finally, we can retrieve the phase slopes, or wavenumbers, from the solution of the system of eqs (3) and (11):

$$k_b = k_{\mathrm{loc}} - \frac{2\pi}{\Delta X}\frac{A_a}{A_a + A_b}. \tag{12}$$

$$k_a = k_{\mathrm{loc}} + \frac{2\pi}{\Delta X}\frac{A_b}{A_a + A_b}. \tag{13}$$

The concepts above, illustrated in the simple case of interference between two monochromatic waves propagating (in the same direction) with different velocities, can be easily extended to surface waves and, specifically, to the interference between different modes.

The wavefield associated to surface wave propagation into a laterally homogeneous medium can be seen as the superposition of different modal contributions (Aki & Richards 2002):

$$S(\omega, x) = \sum_n A_m(\omega, x)\mathrm{e}^{\mathrm{i}(\omega t - k_m(\omega)x + \phi_S(\omega))}, \tag{14}$$

where ω is angular frequency, x is offset, t is time, $A_m(\omega, x)$ and $k_m(\omega)$ are respectively the amplitude spectrum and the wavenumber of the mth mode and ϕ_S is the source phase. As a consequence of wave dispersion, all terms in eq. (14) depend on frequency. If we look at one single frequency, $\omega = \varpi$, and at the simplest case of a single mode propagation, eq. (14) reduces to:

$$S(\varpi, x) = A_0(\varpi, x)\mathrm{e}^{\mathrm{i}[\varpi t - k(\varpi)x + \phi_S(\varpi)]}, \tag{15}$$

where the modal phase is

$$\phi(\varpi, x) = -k(\varpi)x + \phi_S(\varpi). \tag{16}$$

This is the base for the classical MOPA approach, which performs an independent linear regression of the phase versus offset for each frequency. The final objective is to find the wavenumber (then converted to velocity) distribution with frequency, that is, the dispersion curve. Note that this method ignores the amplitude information and only concentrate on the phase of the signal.

In the case of coexistence of two modes, let us number them 0 and 1, eq. (14) becomes:

$$S(\varpi, x) = \mathrm{e}^{\mathrm{i}\phi_S(\varpi)}\left\{A_0(\varpi, x)\mathrm{e}^{\mathrm{i}[\varpi t - k_0(\varpi)x]} + A_1(\varpi, x)\mathrm{e}^{\mathrm{i}[\varpi t - k_1(\varpi)x]}\right\}, \tag{17}$$

where the first exponential term is a constant and the expression in brackets resembles eq. (4), with the only difference of the amplitude dependence on offset. Therefore, if we analyse the complex wavefield *S(ω, x)* produced by the interference of two (and only two) different modes, and we do it frequency by frequency, we apply the theoretical concepts and equations derived for the case of the two sinusoids. In this way, we can extract the wavenumber information of the two modes for each analysed frequency, $k_0(\varpi)$ and $k_1(\varpi)$, which consists in retrieving the multimodal dispersion curve. This

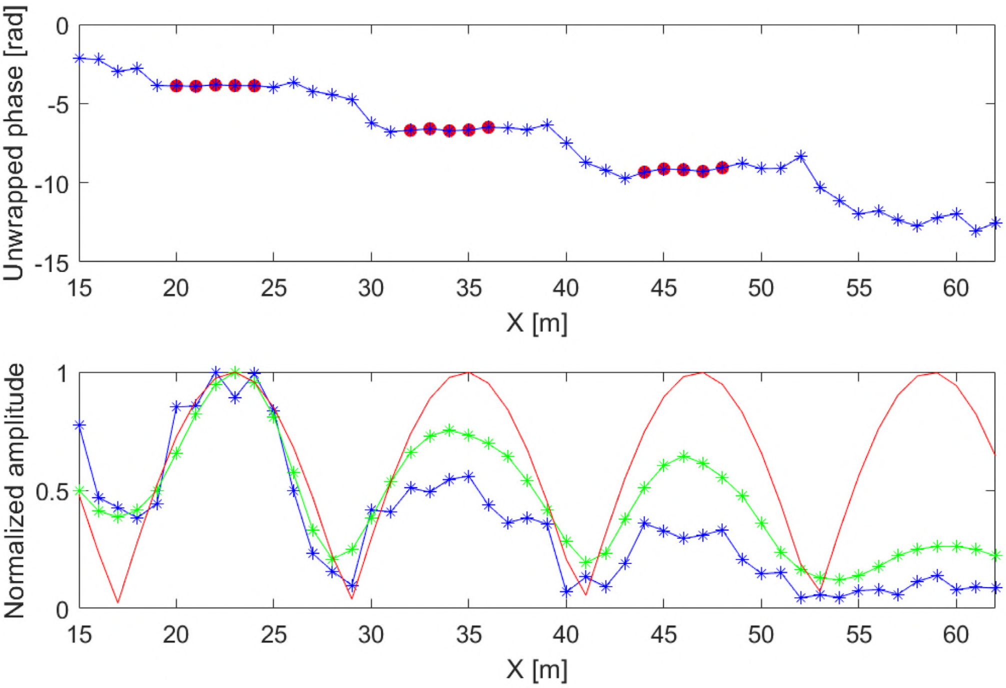


Figure 12. Top panel: unwrapped phase (blue line) with values automatically selected by the algorithm for each independent k_{loc_i} computation (red circles). Data have been LMO corrected with a velocity of 140 m s^{-1}. Bottom panel: normalized amplitude (blue line), normalized smoothed amplitude after correction for attenuation (green line), fitting sinusoid (red line). For both plots the frequency is 19 Hz—Mirandola data set.

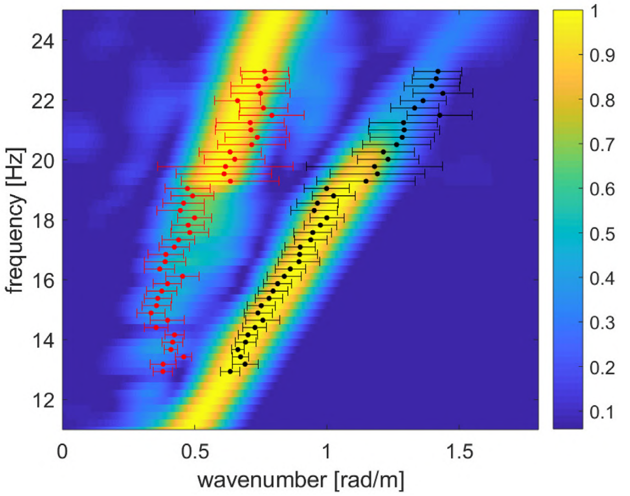


Figure 13. Final multimodal dispersion curve, with relative error bars, overlapped to the normalized *f*–*k* spectrum: fundamental mode is in black, first higher mode in red. Mirandola data set.

is the base for what we call MMMOPA. In this case, both phase and amplitude information have to be extracted and analysed.

3 PROPOSED ALGORITHM

When analysing surface waves and especially real data, things get more complicated compared to the case with two sinusoids. The additional problems we encounter are:

(i) We need to ensure that only two modes interfere at the same time, with comparable amplitudes.

(ii) Amplitude is not constant with offset, being affected by geometrical and intrinsic attenuation.

(iii) Real data are affected by noise which distorts phase information and affects true amplitude recovery.

(iv) The investigated area could be affected by lateral velocity variations.

Given the reasons above, we show here a robust and effective algorithm to perform MMMOPA which can be applied to real data under certain conditions, which are the compresence of two and only two energetic modes and the 1-D character of the medium within each analysis window. When the first condition is not met we can, in some cases, pre-process the data in order to make them still suitable for the analysis. Concerning the 1-D condition within the analysis window, this is a required assumption due to the simplified formulation of the method: however, as it is done with most surface wave methods including the standard MOPA, the method could be applied either piecewise over neighbouring intervals, or continuously on moving windows, in order to highlight smooth lateral heterogeneities. Further developments to catch rapid velocity changes are feasible in the MMMOPA framework and will be illustrated in Section 6.

To better guide the reader in the comprehension of the method, we list here the different steps of MMMOPA (Fig. 5):

1. Identification of a proper frequency range for the analysis, and of the most energetic mode for each frequency within that range.
2. Amplitude correction for attenuation.
3. Identification of the amplitude/phase spatial period and of the position of amplitude maxima and minima.
4. Computation of the individual modes amplitudes.
5. Phase linear regression and computation of an average local wavenumber.

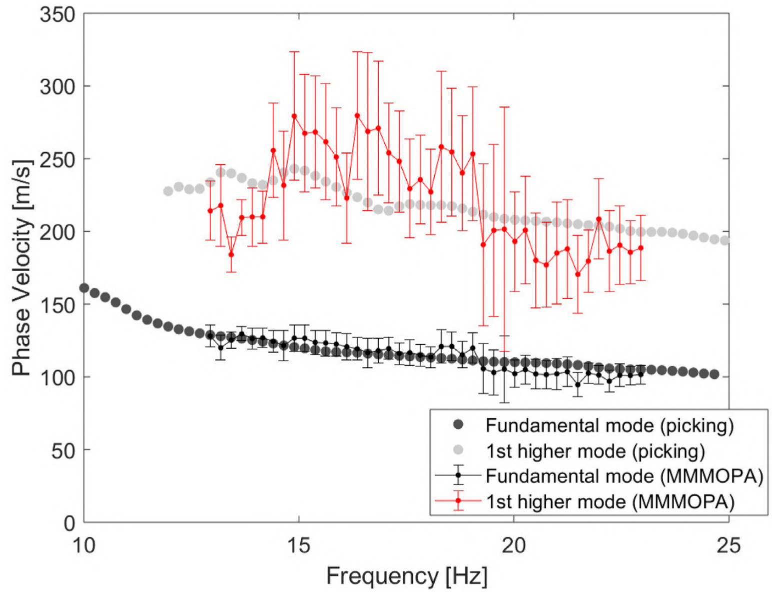


Figure 14. Final multimodal dispersion curve from MMMOPA over the same curve picked with an automatic algorithm. Mirandola data set.

Step 1 consists in analysing the *f–k* spectrum in order to identify possible frequency ranges where two different modes coexist with comparable levels of energy. Within them, frequency sub-ranges need to be defined, depending on which is the most energetic mode. If further modes exist, we should first analyse their impact in terms of amplitude (i.e. verify whether their energy is non-negligible and thus can produce further interference, compromising the two-mode analysis presented in this paper). When necessary, we could filter them out in the *f–k* domain. Of course, this is not always straightforward: in order to preserve the wanted information, we should apply a filter only when higher-orders modes velocities are far enough from the fundamental and first-order modes, going back to the problem faced when trying to isolate the fundamental mode only. However it should be noted that very rarely energy of Rayleigh wave is distributed, to a measurable extent, onto more than two modes.

Steps 2–5 describe the automatic analysis performed by the algorithm. This is done frequency by frequency. Step 2 consists in correcting the raw measured amplitude for attenuation in space: the algorithm first corrects for geometrical (cylindrical) spreading ($1/\sqrt{x}$, where x is the source–receiver distance), then performs a nonlinear regression to compute a rough estimate of the average (in space) apparent attenuation coefficient α_R, which describes the decay of the recorded amplitude as a combination of the joint propagation of the two modes. In other words, the algorithm estimates the overall amplitude loss ignoring oscillations due to modes superposition, and corrects for it (Fig. 5a). When considering both geometrical and intrinsic attenuation effects, and by ignoring the oscillating behaviour produced by modal interaction ($\cos[(k_0 - k_1)x] = 0$), eq. (5) becomes

$$A\frac{1}{\sqrt{x}}\mathrm{e}^{(-\alpha_R x)} = \sqrt{\frac{A_0^2}{x}\mathrm{e}^{-2\alpha_0 x} + \frac{A_1^2}{x}\mathrm{e}^{-2\alpha_1 x}}, \tag{18}$$

where α_0 and α_1 are the attenuation coefficients of the two modes. Note that for $x = 0$, $A = A_0 + A_1$. That said, we have

$$\mathrm{e}^{(-2\alpha_R x)} = \left(\frac{A_0}{A_0 + A_1}\right)^2 \mathrm{e}^{-2\alpha_0 x} + \left(\frac{A_1}{A_0 + A_1}\right)^2 \mathrm{e}^{-2\alpha_1 x} \tag{19}$$

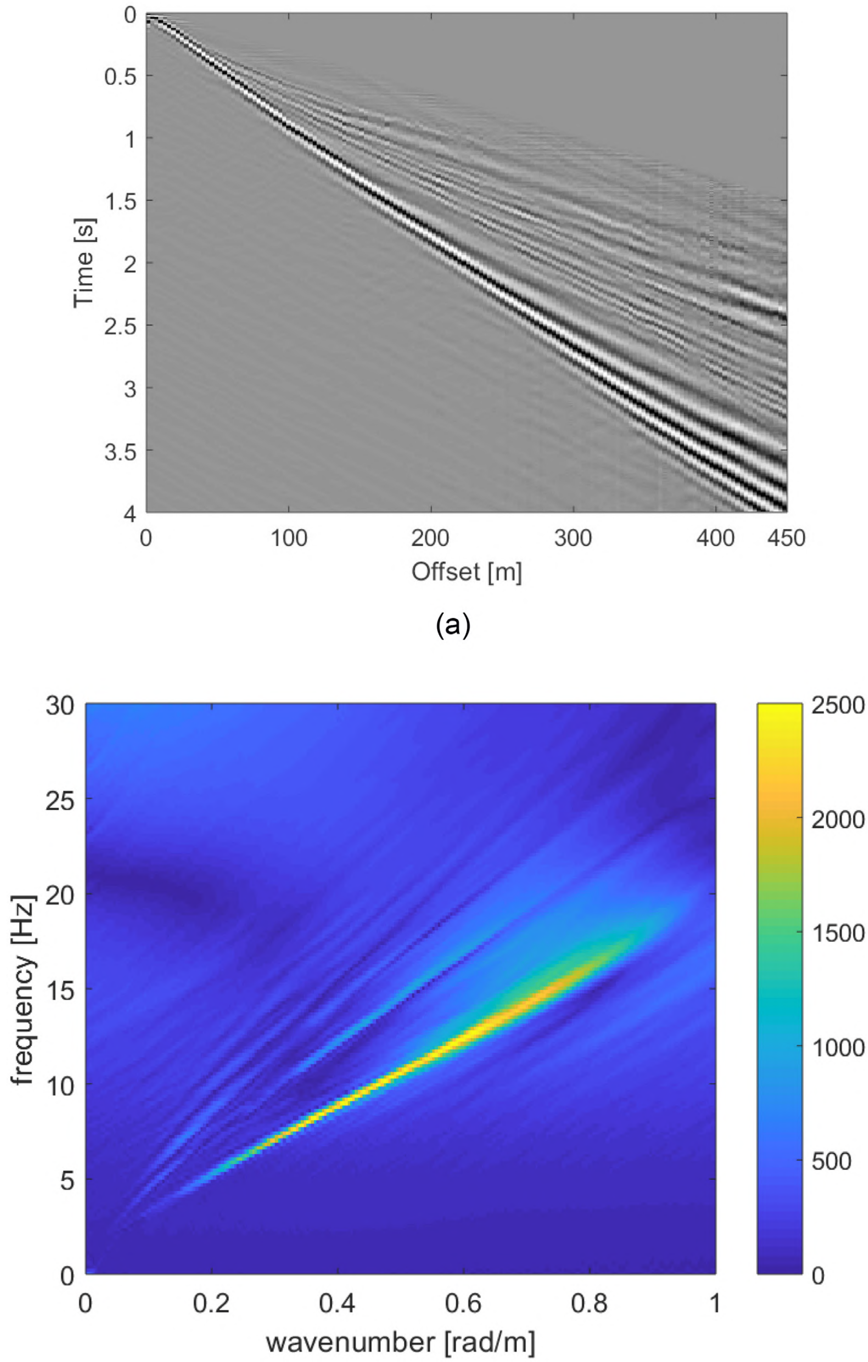
from which

$$\alpha_R = -\frac{\ln\left(\left(\frac{A_0}{A_0+A_1}\right)^2 \mathrm{e}^{-2\alpha_0 x} + \left(\frac{A_1}{A_0+A_1}\right)^2 \mathrm{e}^{-2\alpha_1 x}\right)}{2x}. \tag{20}$$

Eq. (20) describes the dependency of α_R on distance. However, a unique apparent attenuation coefficient is here computed through a nonlinear regression of amplitudes, using the left side of eq. (18): the resulting apparent attenuation coefficient represents an average value in space (Lai 1998). This procedure is based on the strong assumption of a homogeneous medium (Rix *et al.* 2000). However, as the goal here is to recover amplitude loss in an approximate sense (practically, using an exponential loss), and not to provide a precise estimate of attenuation, this approach is viable. More detailed analysis of attenuation may be performed after the dispersion information is recovered: individual modes attenuation coefficients could be easily extracted making use of all information acquired during the MMMOPA analysis (see Appendix B).

Once amplitude is corrected for attenuation, we come to step 3, where the oscillating period of amplitude versus offset (ΔX) is identified for each frequency. To this end, the algorithm searches for a positive periodic function (more specifically, the norm of a sinusoid) which approximates the observed data (Fig. 5b): the period and initial phase of the best-fitting cosine are found by minimizing a specific cost-function, which is based on the sign of the amplitude first and second derivatives. If the amplitude is particularly noisy, a smoothing step is necessary before fitting. From this first step we obtain, for each frequency, the value of ΔX and the offset locations of amplitude maxima and minima.

In step 4, the two individual modes amplitudes are extracted for each frequency. The algorithm computes amplitude minima and maxima as the average of the different local minima and maxima found in the previous step (Fig. 5c). Consider that data are affected by larger uncertainties at greater distance, so it is advisable to weight the average with respect to offset. Another possibility consists in selecting only a certain number of amplitude (and phase) oscillations, excluding from the whole analysis the ones at greater distance. This is required when the receiver array is very long and/or the signal-to-noise (S/N) ratio is low, and permits to adapt



(a)

(b)

Figure 15. (a) Raw seismogram for the selected shot, with trace by trace normalization. (b) Raw *f*–*k* spectrum corresponding to the seismogram in (a). UK data set.

the analysis window independently for each frequency. Then, the two individual amplitudes are computed using either eqs (8) or (9), depending on which of the two modes is the most energetic for that frequency.

In step 5, the local slope (or wavenumber) information, which is a combination of wavenumbers of the two modes, is extracted. The developed algorithm takes, for each frequency, the unwrapped (in offset) phase in a neighbourhood of each amplitude maximum *i* (Fig. 5d), and computes a linear regression as in the classic MOPA approach (Strobbia & Foti 2006). In formula:

$$\Phi = G \cdot M, \tag{21}$$

where G is the data kernel matrix, containing the offset information, Φ the phase vector and M the vector containing the unknown polynomial coefficients, $-k_{\mathrm{loc}_i}$ and $\phi_{\mathrm{S}i}$. Using a least-squares approach

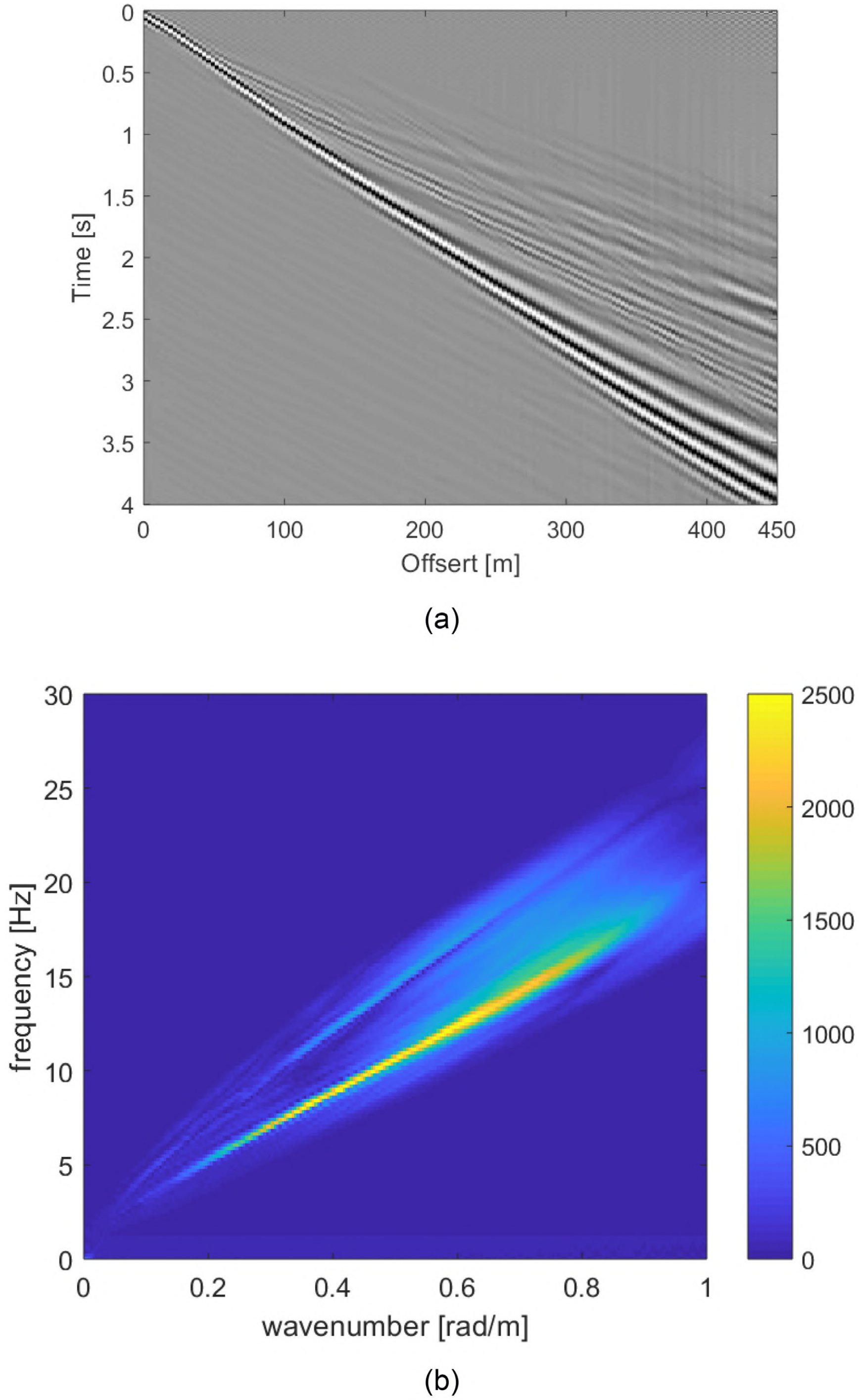


Figure 16. (a) Seismogram for the selected shot after *f*–*k* filtering, with trace by trace normalization. (b) Filtered *f*–*k* spectrum corresponding to the seismogram in (a). UK data set.

we obtain M as:

$$M = G^{-g} \cdot \Phi, \tag{22}$$

being $G^{-g}=(G^{T}G)^{-1}$, G^{T} the pseudo-inverse of G. The value of the final local wavenumber k_{loc} for each frequency is then computed as the weighted (with respect to offset, as done for amplitude) average of all the k_{loc_i} values, or of a selection of them. This is possible since we are performing a 1-D analysis (one-dimensional velocity profile with depth). In the presence of lateral variations of velocity, each k_{loc_i} would have a different value, and by averaging them we would obtain a result reflecting the average geological conditions within the computation window. If we would like to perform, in such cases, a proper 2-D analysis, aimed to highlight the wavenumber (i.e. the velocity) variability within the same analysis window, we should avoid this averaging step but work with the single $\mathbf{k_{loc_i}}$ values. The adaptation of the proposed method to the 2-D case will be described in Section 6.

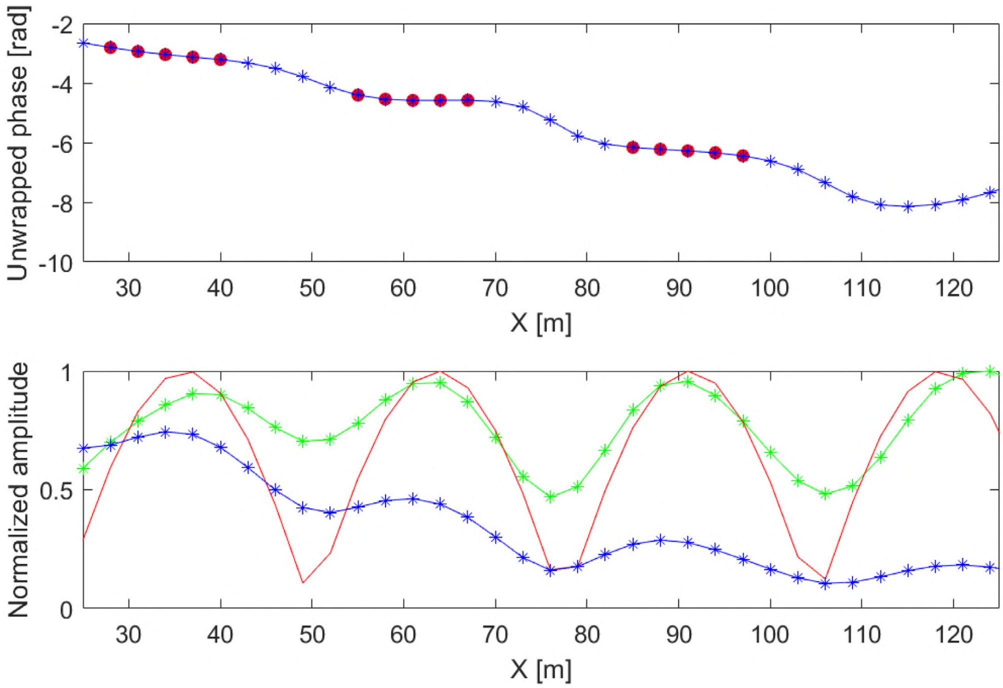


Figure 17. Top panel: Unwrapped phase (blue line) with values automatically selected by the algorithm for each independent k_{loc_i} computation (red circles). Data have been LMO corrected with a velocity of 140 m s^{-1}. Bottom panel: normalized amplitude (blue line), normalized smoothed amplitude after correction for attenuation (green line), fitting sinusoid (red line). For both plots the frequency is 14 Hz—UK data set.

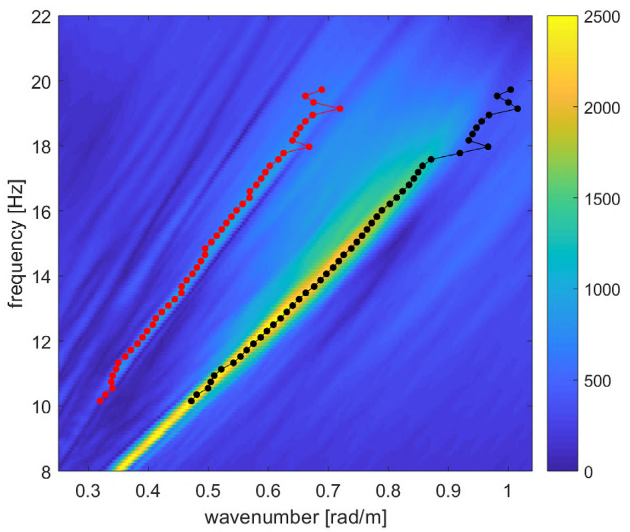


Figure 18. Final multimodal dispersion curve overlapped to the original (non-filtered) *f–k* spectrum: fundamental mode is in black, first higher mode in red. UK data set.

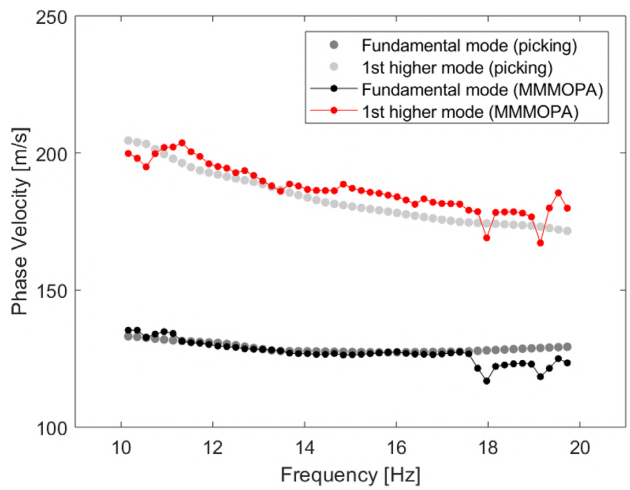


Figure 19. Final multimodal dispersion curve from MMMOPA over the same curve picked with an automatic algorithm. UK data set.

Finally, single modes wavenumbers are calculated for each frequency using eqs (12) and (13). The obtained values compose the multimodal dispersion curve.

3.1 Evaluating model uncertainty

In the section above, we illustrate a method to retrieve multimodal dispersion curves from a single shot. We now want to quantify the error on the final model, that is, on the estimated velocities. In order to improve the S/N ratio, 2-D seismic acquisitions normally provide redundant data, repeating the same shot and/or shooting at different locations along the receiver line. In the first case (vertical stack) it is possible to compute the phase variance, which may be used in turn to weight the linear regression of phase versus offset. Also amplitude variance may be computed, which will be certainly small with good source repeatability (e.g. controlled sources) and high with impulsive uncontrolled sources (e.g. sledgehammer). In the second case, with many shots at different locations, a statistical approach (such as sMOPA, Vignoli *et al.* 2011) is advisable. At any rate, when dealing with real data, it is always important to estimate the uncertainties related to the model, that is, the uncertainties of the estimated phase velocities for the two modes. For this purpose, we need to propagate phase and amplitude variance using a cascade approach.

Let us focus on the case of a single location with repeated shots. Selecting one frequency, we can easily calculate the variance of the recorded phase at each receiver location, $\sigma^2_{\phi_i}$. Then, phase variance can be used, when solving the linear system in eq. (21), to build the pseudo inverse matrix G^{-g} in a weighted manner (Strobbia & Foti 2006). In formula:

$$G_W^{-g} = (G^T W^T W G)^{-1} G^T W^T W, \tag{23}$$

where $W^T W$ is the diagonal matrix containing the inverse of the phase variances. This represents an improvement with respect to the non-weighted formulation, because data quality is taken into account.

We now need to estimate the error of each k_{loc_i}. Under the assumption of normally distributed data, the error on the polynomial coefficients vector M, $\sigma_M = [\sigma_{k_{\mathrm{loci}}}, \sigma_\phi S_{\mathrm{loci}}]$, is given by Santamarina & Fratta (1998) (in matricial notation):

$$\sigma_M = \sqrt{G2_W^{-g} \cdot \sigma^2_\Phi}, \tag{24}$$

where $G2_W^{-g}$ is the matrix containing all G_W^{-g} elements squared.

According to the law of error propagation (Mandel 1964), the uncertainty of the final k_{loc} value, being the latter a weighted average of the N different k_{loci}, is

$$\sigma_{k_{\mathrm{loc}}} = \sqrt{\sum_{i=1}^{N} (w_i^2 \cdot \sigma^2_{k_{\mathrm{loci}}})}, \tag{25}$$

where w_i are the weights used for the average, which sum to 1.

As for amplitudes, being $\sigma^2_{A_{\mathrm{maxi}}}$ and $\sigma^2_{A_{\mathrm{mini}}}$ the amplitude variances for a selected frequency at each maximum and minimum amplitude spatial location, the uncertainties of $A_{\max}$ and $A_{\min}$ can be calculated respectively as the weighted root mean square of the N $A_{\max_i}$ and of the M $A_{\min_i}$ values:

$$\sigma_{A_{\max}} = \sqrt{\frac{\sum_{i=1}^{N}[(A_{\max_i} - \overline{A_{\max}})^2 \cdot w_i']}{\sum_{i=1}^{N} w_i'}}, \tag{26}$$

$$\sigma_{A_{\min}} = \sqrt{\frac{\sum_{i=1}^{M}[(A_{\min_i} - \overline{A_{\min}})^2 \cdot w_i']}{\sum_{i=1}^{M} w_i'}}, \tag{27}$$

being w_i' the composed weights given by the product of w_i (the weights on distance, the same used in eq. 25) and the inverse of the amplitude variance ($1/\sigma^2_{A_{\mathrm{maxi}}}$ or $1/\sigma^2_{A_{\mathrm{mini}}}$). If we further propagate the error through the computation of the single modes amplitudes A_0 and A_1 (eqs 8 or 9), we obtain:

$$\sigma_{A_0} = \sigma_{A_1} = \frac{1}{2}\sqrt{\sigma^2_{A_{\max}} + \sigma^2_{A_{\min}}}. \tag{28}$$

Uncertainties on the two modes wavenumbers K_0 and k_1 will then be a combination of both amplitude and phase error propagation:

$$\sigma_{k_0} = \sigma_{k_1} = \sqrt{\sigma^2_{k_{loc}} + \left[-\frac{2\pi A_1}{\Delta X (A_0 + A_1)^2}\right]^2 \sigma^2_{A_0} + \left[\frac{2\pi A_0}{\Delta X (A_0 + A_1)^2}\right]^2 \sigma^2_{A_1}}. \tag{29}$$

Finally, the error on the estimated phase velocities of the two modes, c_0 and c_1, can be estimated as:

$$\sigma_{c_0} = \frac{2\pi f}{k_0^2} \cdot \sigma_{k_0} \tag{30}$$

$$\sigma_{c_1} = \frac{2\pi f}{k_1^2} \cdot \sigma_{k_1}. \tag{31}$$

4 RESULTS WITH SYNTHETIC DATA

In this section, we present the results of an application of MMMOPA to a 1-D synthetic data set where the first higher mode is quite strong in a specific frequency range, and it is measurable together with the fundamental mode.

The synthetic seismogram was generated using PUNCH (Kausel 1989), a powerful tool for the computation of Green's functions in a layered medium due to a variety of sources. The program is based on a finite element solution in the direction of layering, that is, it is based on the 'thin layer method' or TLM (Kausel 1981). For this reason, PUNCH can handle only 1-D vertical velocity models, which is sufficient for the required demonstration.

In the simulation we applied a single shot using a point-load source at the surface. The receiver array is composed of 96 vertical geophones, spaced 1 m along a line with minimum and maximum offsets of 5 and 100 m. We adopted a 1-D shear wave velocity profile presenting different velocity layers and a velocity inversion at a depth of around 9 m (Fig. 6a). The presence of a velocity inversion moves part of the surface wave energy to higher modes (Boaga *et al.* 2014): the frequency normalized *f*–*k* spectrum (Fig. 7) shows that, in a wide frequency range, most of the Rayleigh wave energy is distributed between the fundamental (always the most energetic) and the first higher mode. In the case presented herein, a suitable frequency range (10–40 Hz) can be identified easily: within this range the fundamental and first higher modes are well separated, their energy is comparable and clearly larger than the second higher mode. Therefore we proceeded with the application of the MMMOPA for each frequency within this range. The analysis was performed in a fully automatic fashion using the algorithm described in the previous section, specifically conceived for surface wave data. Fig. 8 shows an example of the algorithm performances for a selected frequency (22 Hz): amplitudes are corrected for attenuation and fitted with a periodic function to retrieve the spatial period and position of maxima and minima, while the unwrapped phase, showing a local linear trend in conjunction with amplitude maxima, undergoes linear fitting. Note that, due to the absence of noise, amplitudes need not be smoothed before fitting them with a positive sinusoids. For the same reason, uncertainties on the final results were not computed.

The application of the method to our synthetic data set gives promising results. By overlapping the final dispersion curves to the normalized *f*–*k* spectrum we observe a perfect match between the retrieved dispersion curves and the peaks of maximal energy (Fig. 9), already proved to correspond to the fundamental and first-order Rayleigh modes (Fig. 7). It is important to note that for low frequencies, where the two modes are very close to each other in terms of wavenumber, the two modes are perfectly recovered. If we had applied the standard MOPA we should have separated fundamental and first higher modes with the use of a properly designed filter, which is obviously not trivial. One could also note that the quality of the result slightly decreases with increasing frequency: the reason has to be found in the presence of a third mode, gaining more and more energy. This is why, for frequencies higher than 40 Hz, the analysis starts to fail: applicability conditions are no longer valid.

5 RESULTS WITH REAL DATA

5.1 Mirandola case study

We tested the proposed MMMOPA method also on real data. Since lateral variations are out of the scope of this work (in Section 7, we will disccuss the possibility of extending the method to that case), we chose a first test data set with laterally uniform conditions: geology consists of nearly 100 m of soft soils on the top of a stiff bedrock. The site is located in Mirandola, Italy, close to the epicentre of the 2012 earthquake in the Emilia region, and it is one of the three test sites where data were acquired in the context of the InterPACIFIC project (Garofalo *et al.* 2016). Many data sets concerning this case study, each referring to a specific acquisition scheme, are available online. Our analysis was focused on the 1 m spaced 48 channels active acquisition and, in particular, on the first shot location (energized 10 times). We focused our attention on the Rayleigh waves as recorded by 4.5 Hz natural frequency vertical geophones. Records are sampled at 0.25 ms and their length is 2 s. The first recorded offset is at 15 m, the last at 62 m (Fig. 10).

The first step of the analysis was the observation of the *f*–*k* spectra in order to identify possible frequency ranges where the method could be applied (two modes interfering). Between 10 and 23 Hz the fundamental and the first higher modes are both present and quite energetic, whereas for frequencies higher than 23 Hz the fundamental mode dramatically looses energy, showing spectral amplitudes comparable to the background noise. The *f*–*k* spectrum also shows that the first mode is the most energetic up to 19 Hz, while the second is stronger between 19 and 23 Hz (Fig. 11). This information is needed to constrain the final result and remove the intrinsic ambiguity about the amplitude already discussed above.

Then, we proceeded running the algorithm within the identified frequency interval. The main differences between this case and the synthetic one are the needs of smoothing the amplitude before fitting it with the positive sinusoid and of selecting a finite number of oscillations for the analysis, which was set to three, beyond which amplitude and phase are too noisy (Fig. 12). Both operations are performed in order to increase the robustness of the analysis. Moreover, with 10 repeated shots available, the average phase and amplitude spectra were used, local linear regressions were weighted using phase variance information and model uncertainties, both for fundamental and first higher mode, were computed as described above.

The algorithm performed successfully between 13 and 23 Hz (Fig. 13): at lower frequencies the two modes are not sufficiently separated, having very similar velocities, which reflects in a very large spatial period ΔX, exceeding the length of the acquisition array.

Even though the Mirandola data set has been processed by many scientific groups all around the world, no result on higher modes has been published. In order to have a benchmark for our results, we performed an automatic picking of the two most energetic maxima on the *f*–*k* spectrum for the frequency interval of interest. Manual or automatic picking in some spectral domain is the core of most state-of-the-art mode separation techniques.

Fig. 14 shows the comparison between the two results. The fundamental mode from MMMOPA shows a good match with the picked curve. As for the higher mode, MMMOPA result is apparently less precise, showing an average error of about 10 per cent, and certainly less smooth. Similarly, the computed uncertainties on the phase velocity distributions are acceptable for the fundamental mode, while rather large for the first higher mode, being the latter less energetic and faster than the first. Uncertainties also have a peculiar distribution in frequency up to 20 Hz, being for both modes higher for higher frequencies. This reflects the distribution of the input phase uncertainty, which is affected by a large amount of noise at high frequency and large offsets.

5.2 UK case study

Another case study is here presented. The data set consists of a 2-D seismic line taken from a wider acquisition for mining exploration in the UK. The acquisition scheme includes more than 200 shot positions, each of them energized once using explosive (150 g of dynamite at 1 m depth). For each shot the active spread is composed of 201 vertical component autonomous nodes with natural frequency of 10 Hz and spacing equal to 3 m. Traces were sampled at 10 ms and their length is 4 s. In order to test the MMMOPA algorithm we selected the record corresponding to the first shot position, with offsets between 0 and 450 m (fully developed surface wave trend, Fig. 15a). This data set is very different from the Mirandola one, both in terms of size of the acquisition and data quality (very powerful source, industrial equipment, etc.). As far as geology is concerned, former studies in the same area show a quite homogeneous near-surface, so that the medium could be assumed to be 1-D. For this reason, even if the large number of receivers composing the active spread would allow to apply MMMOPA on moving windows, we performed the measurement only once and retrieved one single multimodal dispersion curve.

We first computed the *f*–*k* spectrum for the selected seismogram and analysed it (Fig. 15b). In this case, different higher order modes are present and clearly visible in the spectrum, which is a quite rare condition. However, the most energetic ones are the first two, which we assume to be the fundamental and the first higher mode. In order to focus our analysis exclusively on them, some pre-processing is necessary. We manually designed a filter in the *f*–*k* domain and applied it to the spectrum. Consequently, we applied an inverse Fourier transform to project data back to the space-time domain and retrieve the filtered seismogram. The filtered spectrum and the corresponding seismogram are showed in Fig. 16.

After filtering the data, we were able to define a frequency range to test the MMMOPA algorithm: the two modes are both quite energetic between 10 and 18 Hz, with an overall prevalence of the fundamental mode. However, in order to push the analysis to its limits and see how the method behaves where the applicability conditions are not fully satisfied, we decided to extend the frequency range up to 20 Hz.

Since the recorded offset interval is extremely large, limiting the number of amplitude and phase oscillations included in the analysis, which is equivalent to perform a frequency-dependent offset selection, becomes an essential operation. In this case, to avoid both near- and far-offset effects, we excluded the first amplitude maximum and selected the next three. As for the amplitudes, they were slightly smoothed before the fitting step, as in the previous case study (Fig. 17).

In this example, however, only one single source was fired in each location: for this reason we were not be able to estimate the uncertainties on the final results, as for the Mirandola case.

Dispersion curves computed with MMMOPA show a good match with energy maxima on the *f*–*k* spectrum up to 17.6 Hz (Fig. 18). Beyond this frequency the method still gives a reasonable estimate of wavenumber for both modes, but the quality of the result is much degraded. As already done for the Mirandola data set, we compared our results with the curves obtained by picking energy maxima on the *f*–*k* spectrum: in this case, since the lack of energy for higher frequencies caused the failure of automatic picking, we opted for a manual picking. The analysis of the results must be separated: between 10 and 17.6 Hz the maximum error was about 2 per cent for the fundamental mode and 4 per cent for the higher mode, whereas for frequencies higher than 17.6 Hz errors raise to

10 per cent for the fundamental mode (where velocities are slightly overestimated) and 7 per cent for the higher mode (Fig. 19). This change in the MMMOPA performances around 18 Hz is certainly caused by:

(i) The lack of energy on both modes, which reflects in noisier phase;
(ii) The increasing distance (difference of phase velocity) between the two modes with frequency, which causes a decrease of the spatial period ΔX (see eq. 3) and a poor sampling of the local phases. This critical issue will be discussed in Section 6.

These two examples using real data show how the MMMOPA can correctly perform when the applicability conditions are met. In both cases phase velocity estimation is more accurate for the fundamental mode than for higher mode.

6 DISCUSSION

The MMMOPA method presents a valid alternative to most state-of-the-art methods for surface waves analysis in multimodal propagation. The real novelty brought by MMMOPA is that it is based on the analysis of the interaction between two different modes, from which it infers their individual characteristics. This could open new possibilities to extend the method to an undefined number of modes, simply by studying their (more complex) interaction. Another important difference with respect to existing surface wave methods is the concurrent analysis of phase and amplitude: both surface wave dispersion and absorption are considered and potentially assessed, as they are expressions of the same propagation phenomenon.

As for the current formulation of the method, it fails when Rayleigh wave propagation is dominated by more than two modes having comparable energy. Note that such situations are extremely rare. In these cases, as already mentioned, an *f*–*k* filter removing higher modes could be necessary. One could wonder which is the interest of applying an *f*–*k* filter to isolate two modes and apply MMMOPA instead of isolating one mode at a time and apply classic MOPA. The answer is plural: when dealing with lateral variations of velocity the resulting *f*–*k* spectrum for one shot, being the sum of the different spatial contributions, can be complex and misleading. In this case designing an *f*–*k* filter to isolate each single mode could be far from simple and the greatest risk is to destroy part of the wanted information: getting rid of higher order modes, preserving the two most energetic, represent a more prudent approach. The same argument is valid when the two most energetic modes have similar velocities within a certain frequency band. Moreover, MMMOPA is capable of retrieving dispersive information about two modes in one step, which may be used to constrain the surface wave velocity inversion. Last but not least, the MMMOPA analysis gives, together with the dispersion relation, also information about the relative amplitude or the two modes, which can be used to infer the attenuation coefficients of the two modes (see Appendix B).

It is important to point out that, although the examples shown here deal with the interference of fundamental and first higher-order modes, the method could be applied to any pair of modes of any order, with the condition of being the most energetic in the propagation, and provided they can ultimately be recognized in their true nature (for inversion purposes). Thus one could conceive also an iterative MMMOPA approach that analyses pairs of modes when energetic modes are more than two.

The main limitation of MMMOPA, indeed, consists in the need of a previous computation and analysis of the *f*–*k* spectrum, in order to define a suitable frequency range for its application and to determine which is the dominant mode for each frequency in that range. This operation suffers for user-dependency and could affect the success of the method. Future developments will focus on automatizing this part of the analysis as much as possible, including an autonomous identification of the most energetic modes involved in the Rayleigh wave propagation.

Another limitation consists in the 1-D character of the approach: as most surface wave methods do, MMMOPA retrieves one single (multimodal) dispersion curve for each measuring window, in which the medium is assumed rather homogeneous. A simple way to check whether the 1-D condition is met is based on the analysis of *f*–*k* spectra computed from reciprocal shots. If these shots are not consistent with each other, the analysis will return an average dispersion curve for that area, with more weight on shorter offsets. For this reason, as we move away from the 1-D condition, results of the MMMOPA analysis for reciprocal shots will differ from each other: in this case, dispersion uncertainties could be higher than those estimated as discussed in Section 3.1. As any other technique, MMMOPA may be also applied on sliding windows, in order to retrieve smooth lateral velocity variations. However, the size of each computation window should respect some minimum requirement to permit discriminating between two different modes when their velocities are similar. We call $\Delta X_{\max}$ the maximum detectable spatial period and make it coincide with the window length L. Starting from eq. (3) we get

$$\Delta k_{\min} = \frac{2\pi}{\Delta X_{\max}} = \frac{2\pi}{L}, \tag{32}$$

where $\Delta k_{\min}$ is the minimum detectable distance, in wavenumber, between two different modes. This limitation mostly affects low frequencies: an example is given by the first real case presented, where the method fails below 13 Hz. In the same way, the spatial interval between receivers (dr) has an impact on the minimum detectable period $\Delta X_{\min}$, therefore on the maximum detectable wavenumber distance between modes $\Delta k_{\max}$. In this case, it is important to take into account a minimum number of phase measurements for the linear regression, which depends on frequency and on the noise level. If we set it to 5 and take a regression window of 1/3 ΔX, we get:

$$\Delta X_{\min} = 12 \cdot \mathrm{dr}, \tag{33}$$

$$\Delta k_{\max} = \frac{\pi}{6\mathrm{dr}}. \tag{34}$$

If this condition is not met, the local phase extracted could be unreliable, and so the velocities for the two modes. A similar condition takes place in the second real case presented, where the phase velocities computed at higher frequencies are much less accurate. Nevertheless, it is important to keep in mind that the MMMOPA method is a direct derivation of the standard MOPA, which is very effective in targeting smooth to rapid lateral variations, and this is due to the possibility of selecting very short analysis windows for the phase linear regression. This is still valid for its multimode version, but in this case the measurements are sparse, because they are linked to the position of the amplitude maxima. In this first simplified formulation we assume 1-D conditions within the same spatial window, so that all information coming from different oscillations are averaged together. Therefore, a way to capture the lateral variability of phase velocity within the same analysis window is to treat each amplitude maximum as a single measurement point, avoiding averaging in space both amplitude maxima/minima and local wavenumbers. The resulting coarse sampling in space and the potential loss of robustness can only be compensated by data redundancy, which means using multishot data: different shot locations reflect in different amplitude maxima locations. Future developments of MMMOPA will be also oriented towards a generalization of the method to the true 2-D case, making use of multishot data.

Finally, although this work mainly focuses on the analysis of seismic dispersion, the estimation of the modal attenuation is also a potentially valuable approach. The global spectral decay is affected by the mode superposition and the presented formulation can provide the modal wavenumber and the modal attenuation for two interfering modes, giving a more complete picture of the shallow subsoil properties.

7 CONCLUSIONS AND OUTLOOK

In this paper, we present a new approach that makes it possible to apply the MOPA to two-mode situations: we called it MMMOPA. The method is based on the analysis of the interaction between modes producing specific patterns in the spectral domain. The analysis of both phase and amplitude spectra represents a main step forward with respect to the classic MOPA, but also to other commonly used techniques such as *f*–*k* analysis: modal dispersion is also taken into account.

MMMOPA was tested on both synthetic and real data sets, with success. When redundant data are available, uncertainties can be estimated by propagating the error from the phase and amplitude variances to the inverted model estimates.

This first formulation of the method is based on the assumption of 1-D vertical velocity profiles for each window used for the analysis, but it is propaedeutical to the application to 2-D media. As a start, MMMOPA could be applied on sliding spatial windows. To make the analysis even more focused on rapid 2-D lateral variations, a different formulation of the method, making use of multishot data, can be implemented.

In conclusion, MMMOPA is an extension of the MOPA technique when surface wave propagation is dominated by two modes. Depending on the situation, MMMOPA could be preferred to MOPA or vice versa: a combined approach, which makes use of one technique or the other depending on the energy distribution in each frequency band, is probably the best choice.

ACKNOWLEDGEMENTS

Authors thank Cariparo foundation, who financed Ilaria Barone PhD fellowship. We also thank all authors of the InterPacific project, who made Mirandola data available, and RealTimeSeismic for providing the UK dataset.

REFERENCES

Aki, K. & Richards, P.G., 2002. *Quantitative Seismology,* 2nd edn, University Science Books.

Ballard, R.F., 1964. Determination of soil shear moduli at depths by in-situ vibratory techniques, Miscellaneous Paper no. 4-691, Waterways Experiment Station, Vicksburg, Mississippi.

Bergamo, P., 2012. Surface wave analysis in laterally varying media, *PhD thesis*, Politecnico di Torino.

Boaga, J., Cassiani, G., Strobbia, C. & Vignoli, G., 2013. Mode misidentification in Rayleigh waves: ellipticity as a cause and a cure, *Geophysics,* **78**(4), EN17–EN28.

Boaga, J., Vignoli, G., Deiana, R. & Cassiani, G., 2014. The influence of subsoil structure and acquisition parameters in MASW mode mis-identification, *J. Environ. Eng. Geophys.,* **19**(2), 87–99.

Boiero, D. & Socco, L.V., 2011. The meaning of surface wave dispersion curves in weakly laterally varying structures, *Near Surf. Geophys.,* **9,** 561–570.

Foti, S., Sambuelli, L., Socco, L.V. & Strobbia, C., 2002. Spatial sampling issues in fk analysis of surface waves, in *Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems,* doi:10.4133/1.2927181

Gabriels, P., Snieder, R. & Nolet, G., 1987. In situ measurements of shear-wave velocity in sediments with higher-mode Rayleigh waves, *Geophys. Prospect.,* **35**(2), 187–196.

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

Jones, R.B., 1958. In-situ measurement of the dynamic properties of soil by vibration methods, *Gèotechnique,* **8**(1), 1–21.

Jones, R.B., 1962. Surface wave technique for measuring the elastic properties and thickness of roads: theoretical development, *Br. J. Appl. Phys.,* **13**(1), 21–29.

Kausel, E., 1981. An explicit solution for the Green functions for dynamic loads in layered media, Research Report R81-13, Massachusetts Institute of Technology, Department of Civil Engineering, Cambridge, Massachusetts.

Kausel, E., 1989. *Punch: Program for the Dynamic Analysis of Layered Soils, Version 3.0,* Massachusetts Institute of Technology.

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

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**(5), 903–922.

Mandel, J., 1964. *The Statistical Analysis of Experimental Data,* Dover Publications, Inc.

Maraschini, M., Ernst, F., Foti, S. & Socco, V., 2010. A new misfit function for multimodal inversion of surface waves, *Geophysics,* **75**(4), G31–G43.

McMechan, G.A. & Yedlin, M.J., 1981. Analysis of dispersive waves by wave field transformation, *Geophysics,* **46**(6), 869–874.

Nazarian, S.,Stokoe & , K.H., II, 1984. In situ shear wave velocity from spectral analysis of surface waves, in *Proceeding of the 8th Conference on Earthquake Engineering,* Vol. 3, San Francisco, pp. 31–38.

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

Park, C.B., Miller, R.D. & Xia, J., 1999b. Multimodal analysis of high frequency surface waves, in *Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems (SAGEEP)*, pp. 115–122, doi:10.4133/1.2922596.

Rix, G.J., Lai, C.G. & Spang, A.W., 2000. In situ measurement of damping ratio using surface waves, *J. Geotech. Geoenviron. Eng.,* **126**(5), 472–480.

Santamarina, C. & Fratta, D., 1998. *Introduction to Discrete Signals and Inverse Problems in Civil Engineering,* ASCE Press.

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

Strobbia, C. & Foti, S., 2006. Multi-offset phase analysis of surface wave data (MOPA), *J. Appl. Geophys.,* **59**(4), 300–313.

Tokimatsu, K., Tamura, S. & Kojima, H., 1992. Effects of multiple modes on Rayleigh wave dispersion characteristics, *J. Geotech. Eng.,* **118**(10), 1529–1543.

Vignoli, G. & Cassiani, G., 2010. Identification of lateral discontinuities via multi-offset phase analysis of surface wave data, *Geophys. Prospect.,* **58,** 389–413.

Vignoli, G., Strobbia, C., Cassiani, G. & Vermeer, P., 2011. Statistical multioffset phase analysis for surface-wave processing in laterally varying media, *Geophysics,* **76**(2), U1–U11.

Xia, J., Miller, R.D. & Park, C.B., 1999. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves, *Geophysics,* **64**(3), 691–700.

Xia, J., Miller, R.D., Park, C.B. & Tian, G., 2003. Inversion of high frequency surface waves with fundamental and higher modes, *J. Appl. Geophys.,* **52**(1), 45–57 .

APPENDIX A: WAVEFIELD EQUATIONS

The complex wavefield associated with the interference of two monochromatic signals is described by eq. (4). In the complex plane, we can represent the two sinusoids a and b as two vectors with moduli A_a and A_b and angle $-k_a x$ and $-k_b x$, respectively, and sinusoid c as their vectorial summation. The amplitude A_c is therefore given by

$$A_c = \sqrt{[A_a \cos(-k_a x) + A_b \cos(-k_b x)]^2 + [A_a \sin(-k_a x) + A_b \sin(-k_b x)]^2}, \tag{A1}$$

By solving eq. (A1), we come to

$$A_c = \sqrt{A_a^2 + A_b^2 + 2A_a A_b[\cos(-k_a x)\cos(-k_b x) + \sin(-k_a x)\sin(-k_b x)]}. \tag{A2}$$

If we rotate the complex plane in order to align the real axis to the direction of a, the new phases become

$$-k'_a x = 0 \tag{A3}$$

and

$$-k'_b x = -k_b x - (-k_a x) = (k_a - k_b)x, \tag{A4}$$

eq. (A2) therefore reduces to

$$\begin{aligned} A_c &= \sqrt{A_a^2 + A_b^2 + 2A_a A_b[\cos(-k'_a x)\cos(-k'_b x) + \sin(-k'_a x)\sin(-k'_b x)]} \\ &= \sqrt{A_a^2 + A_b^2 + 2A_a A_b \cos[(k_a - k_b)x]}, \end{aligned} \tag{A5}$$

which is equivalent to eq. (5).

Similarly, the new phase of c in the rotated system can be expressed as

$$-k'_c x = \arctan\left\{\frac{A_b \sin(-k'_b x)}{A_a + A_b \cos(-k'_b x)} s\right\} \tag{A6}$$

and by substituting eqs (A3) and (A4) into eq. (A6) we obtain

$$-k'_c x = \arctan\left\{\frac{A_b \sin[(k_a - k_b)x]}{A_a + A_b \cos[(k_a - k_b)x]}\right\}. \tag{A7}$$

Finally, the phase of c is given by

$$-k_c x = -k'_c x - k_a x = \arctan\left\{\frac{A_b \sin[(k_a - k_b)x]}{A_a + A_b \cos[(k_a - k_b)x]}\right\} - k_a x, \tag{A8}$$

which is the same as eq. (10). We are now interested in the slope of the phase. Let us compute the derivative of eq. (10) with respect to x:

$$\begin{aligned} \frac{\mathrm{d}(-k_c x)}{\mathrm{d}x} = -k_c = \frac{1}{1 + \left\{\frac{A_b \sin[(k_a - k_b)x]}{A_a + A_b \cos[(k_a - k_b)x]}\right\}^2} \cdot \\ \frac{A_b \cos[(k_a - k_b)x](k_a - k_b)\{A_a + A_b \cos[(k_a - k_b)x]\} + A_b^2\{\sin[(k_a - k_b)x]\}^2(k_a - k_b)}{\{A_a + A_b \cos[(k_a - k_b)x]\}^2} - k_a. \end{aligned} \tag{A9}$$

In the specific case when $\cos[(k_a - k_b)\, x] = 1$ and $\sin[(k_a - k_b)\, x] = 0$, eq. (A9) simplifies to

$$-k_{\mathrm{loc}} = \frac{A_b(k_a - k_b)}{A_a + A_b} - k_a = -\frac{A_a k_a + A_b k_b}{A_a + A_b}. \tag{A10}$$

This is the local value of k_c corresponding to all amplitude maxima positions, already described in eq. (11).

APPENDIX B: ATTENUATION ANALYSIS

The amplitude decay of surface waves for one mode of propagation is described by

$$A(\omega, x) = A(\omega, 0)\frac{1}{\sqrt{x}}\mathrm{e}^{-\alpha(\omega)x}, \tag{B1}$$

where ω is the angular frequency, x is the source–receiver distance and $\alpha(\omega)$ is the frequency-dependent attenuation coefficient. When dealing with two modes of propagation, the amplitude decay can be derived including eq. (B1) in eq. (5). We have:

$$A(\omega, x) = \sqrt{\frac{A_0(\omega)^2}{x}\mathrm{e}^{-2\alpha_0(\omega)x} + \frac{A_1(\omega)^2}{x}\mathrm{e}^{-2\alpha_1(\omega)x} + 2\frac{A_0(\omega)A_1(\omega)}{x}\mathrm{e}^{-(\alpha_0(\omega)+\alpha_1(\omega))x}\cos[(k_0(\omega) - k_1(\omega))x]}. \tag{B2}$$

If we correct amplitudes for geometrical spreading only, eq. (B2) becomes

$$A'(\omega, x) = \sqrt{A_0(\omega)^2 e^{-2\alpha_0(\omega)x} + A_1(\omega)^2 e^{-2\alpha_1(\omega)x} + \\ + 2A_0(\omega)A_1(\omega)e^{-(\alpha_0(\omega)+\alpha_1(\omega))x} \cos[(k_0(\omega) - k_1(\omega))x]}. \tag{B3}$$

The MMMOPA technique permits to retrieve both the $k_0(\omega)$ and $k_1(\omega)$ distributions and the amplitudes of the two modes after compensating for the apparent attenuation, let us call them $a_0(\omega)$ and $a_1(\omega)$ to distinguish them from the true individual mode amplitudes, $A_0(\omega)$ and $A_1(\omega)$. Let us express the latter as $A_0(\omega) = \frac{a_0(\omega)}{a_0(\omega)+a_1(\omega)}\overline{A(\omega)}$ and $A_1(\omega) = \frac{a_1(\omega)}{a_0(\omega)+a_1(\omega)}\overline{A(\omega)}$, where $\overline{A(\omega)} = A_0(\omega) + A_1(\omega)$. We obtain:

$$A'(\omega, x) = \sqrt{\frac{a_0(\omega)^2}{(a_0(\omega)+a_1(\omega))^2}\overline{A(\omega)}^2 e^{-2\alpha_0(\omega)x} + \frac{a_1(\omega)^2}{(a_0(\omega)+a_1(\omega))^2}\overline{A(\omega)}^2 e^{-2\alpha_1(\omega)x} + \\ + 2\frac{a_0(\omega)^2 a_1(\omega)^2}{(a_0(\omega)+a_1(\omega))^2}\overline{A(\omega)}^2 e^{-(\alpha_0(\omega)+\alpha_1(\omega))x} \cos[(k_0(\omega) - k_1(\omega))x]}. \tag{B4}$$

The individual mode attenuation coefficients can be derived from eq. (B4) through a nonlinear regression: for each frequency, the recorded displacement, compensated for geometrical spreading only, is fitted and $\overline{A}$, α_0 and α_1 are recovered.