Geophysical Journal International

Geophys. J. Int. (2019) **218,** 1873–1891
Advance Access publication 2019 June 21
GJI Seismology

doi: 10.1093/gji/ggz279

Full-waveform matching of VP and VS models from surface waves

Farbod Khosro Anjom,[1] Daniela Teodor,[2] Cesare Comina,[2] Romain Brossier,[3] Jean Virieux[3] and Laura Valentina Socco[1]

[1] *Politecnico di Torino, 10129 Torino, Italy. E-mail: khosro-anjom.farbod@polito.it*
[2] *Università di Torino, 10125 Torino, Italy*
[3] *Univ. Grenoble Alpes, ISTerre, 38000 Grenoble, France*

Accepted 2019 June 12. Received 2019 May 7; in original form 2018 October 11

SUMMARY

The analysis of surface wave dispersion curves (DCs) is widely used for near-surface *S*-wave velocity (VS) reconstruction. However, a comprehensive characterization of the near-surface requires also the estimation of *P*-wave velocity (VP). We focus on the estimation of both VS and VP models from surface waves using a direct data transform approach. We estimate a relationship between the wavelength of the fundamental mode of surface waves and the investigation depth and we use it to directly transform the DCs into VS and VP models in laterally varying sites. We apply the workflow to a real data set acquired on a known test site. The accuracy of such reconstruction is validated by a waveform comparison between field data and synthetic data obtained by performing elastic numerical simulations on the estimated VP and VS models. The uncertainties on the estimated velocity models are also computed.

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

INTRODUCTION

Analysis of surface wave dispersion-curve (DC) is a well-established technique to obtain local near-surface *S*-wave velocity (VS) models for stratified media (Socco *et al.* 2010). These local vertical stratified models could be used for building 2-D and 3-D velocity models through lateral interpolation. There are several established approaches available in literature to obtain DCs from the data (McMechan & Yedlin 1981; Bergamo *et al.* 2012). The DC inversion is an ill-posed problem and different approaches have been investigated with varying degrees of complexity, from deterministic to stochastic strategies for better assessing the non-uniqueness of this reconstruction problem (e.g. Xia *et al.* 1999; Beaty *et al.* 2002; Socco & Boiero 2008; Bodin *et al.* 2012).

More important is the local laterally invariant assumption of the shallow structures, which can be a limitation when significant velocity variations and complex topography are considered. This has been the motivation for considering the full-waveform inversion (FWI) which potentially alleviates this lateral invariance restriction (Brossier *et al.* 2009; Masoni *et al.* 2014; Pérez Solano *et al.* 2014; Yuan *et al.* 2015). However, this nonlinear imaging technique needs an initial guess of the medium properties. In spite of many attempts to start FWI from crude initial models (Warner & Guasch 2016; van Leeuwen & Herrmann 2013) in scheme of acoustic FWI crustal targets, workflows as traveltime tomography of first-arrival and/or reflected *P* waves are current working strategies when missing low-frequency data. For near-surface experiments, aside cycle-skipping problems of FWI induced by the selected initial model (Virieux & Operto 2009), separation of phases, needed for picking times/phases (to be used by tomography methods), is very challenging. Moreover, for near-surface targets, surface waves have high amplitudes and can be easily recorded in noisy environments. Their use in building the initial model will hence be preferable because the picking of first-arrival *P* times or the picking of reflections could be difficult (Badji *et al.* 2016).

Could one use such reconstructed models from the DC analysis of surface waves as initial models for FWI? This is the question we want to answer, and this work proposes a strategy for the reconstruction of both VS and *P*-wave velocity (VP) from surface waves. The DC exhibits strong sensitivity to VS while VP has a weaker effect and the density an even smaller influence on the DC (Foti & Strobbia 2002). However, recently, Socco *et al.* (2017) have shown that DCs have higher sensitivity to time-average *S*-wave velocity (VSZ) than to layered velocity models. Consequently, a method to directly transform DCs into VSZ models avoiding the inversion step has been proposed (Socco *et al.* 2017). The method requires a single specified VSZ reference profile along the seismic line together with its associated DC in order to obtain a relationship between the DC wavelength and the investigation depth (the so-called W/D relationship). The W/D relationship is obtained by searching the wavelength at which the surface-wave phase velocity is equal to the VSZ at a certain depth. The wavelength–depth pairs provide a relationship that represents the skin depth of the fundamental mode and can be applied to all the DCs in the data set to directly transform them into local VSZ profiles. Socco *et al.* (2017) applied the W/D method to two sets of synthetic models with different level of complexity and they estimated the VSZs with less than 5 per cent error compared to the true VSZs. In addition, they applied the method

to two experimental data sets and, similarly, less than 5 per cent error was encountered by comparing the results with benchmarks. Socco & Comina (2017) showed that the W/D relationship is highly sensitive to the Poisson ratio and they exploited this sensitivity to estimate also the time-average VP (VPZ) profile, using only the surface wave data. Socco & Comina (2017) performed three synthetic tests (on models with different level of complexity). The estimated VPZ showed an error of less than 4 per cent with respect to the true values. In addition, they applied the method to two experimental data sets and in both cases the difference of the estimations with the benchmarks was within 10 per cent.

This approach can be extremely efficient but, in case of strong lateral variations, Khosro Anjom *et al.* (2017) have shown that a further step is needed. Using a single reference model (a single W/D relationship) for the whole data set can lead to incorrect velocity profiles in some places of the medium where strong lateral variations occur. As proposed by Khosro Anjom *et al.* (2017), a hierarchical clustering algorithm allows to group similar DCs into consistent clusters for which the same W/D relationship can be used. Accordingly, instead of using a single W/D relationship for the whole seismic line, they proposed to use a reference W/D relationship for each DC cluster.

The time-average velocities estimated with this approach can be directly used in many applications such as seismic hazard (VS30) or static corrections for seismic reflection processing. However, wave propagation modelling (as required for FWI, e.g.) requires interval velocities which need to be deduced from time-average velocities.

In this paper, we make use of the above-mentioned methods to estimate the VSZ and the VPZ in a laterally varying site and we provide a formulation to transform the time-average velocities into interval VS and VP models, which requires differentiation regularization tools. The estimated VS and VP models can be used by seismic modelling method to generate synthetic seismograms. We use a spectral element-based method (Trinh *et al.* 2019) to check whether the synthetic data are immune to cycle-skipping issues when compared with the real data. We also propagate the experimental uncertainties of the DCs through the process and estimate the uncertainties on the final models. In the following, as for many methods of surface wave analysis, we assume that the DCs are related to the fundamental mode of Rayleigh waves and they are representative of a local vertical profile of the subsurface.

METHOD

Let us assume that a set of DCs has been obtained along a seismic line (2-D geometry) or over an acquisition area (3-D geometry). These DCs will be used for constructing 2-D or 3-D velocity models, following the workflow described in Fig. 1.

DCs are analysed using a specific algorithm for grouping them into clusters of rather homogeneous sets (a). For each cluster, the following different analyses are performed: A reference DC is selected and inverted with a Monte Carlo Inversion (MCI) algorithm to obtain VS and VSZ models as reference models (b). This estimated VSZ and the corresponding DC are used to compute a reference W/D relationship for this cluster (c). The reference W/D relationship is used to transform all DCs of this cluster into VSZ models (d). An apparent Poisson ratio for this cluster is estimated using the reference W/D relationship and the reference VS model (e). Using the apparent Poisson ratio of this cluster, each VSZ profile is transformed into a VPZ profile (f). All the reconstructed VSZ and VPZ profiles are then transformed into interval VS and VP profiles using specific rules of differentiation (g). Finally, laterally varying VS and VP 2-D or 3-D models can be interpolated from all these interval VS and VP profiles. These models can be used for computing seismograms which can be compared with the recorded ones. Let us now consider each step of this workflow for a more detailed description.

Dispersion curves clustering

Given a data set of DCs, a hierarchical clustering algorithm (Rokach & Maimon 2005) is applied to divide the data set into homogeneous sets of DCs (Khosro Anjom *et al.* 2017). Hierarchical clustering algorithms do not require prior information, regarding, for example, possible lateral velocity variations. Moreover, the number of clusters does not need to be defined and they are obtained from the algorithm. The distinction between the clusters is based on the dissimilarity between DCs, which is obtained by defining appropriate metrics and linkage criterion. Euclidean distance between two DCs is used as metrics and is expressed as

$$D\left(v_i, v_j\right) = \sqrt{(v_{i1} - v_{j1})^2 + (v_{i2} - v_{j2})^2 + \ldots + (v_{ip} - v_{jp})^2}\,, \quad (1)$$

where the quantity D is the Euclidian distance and where frequency-dependent phase velocities $\{v_i\}$ and $\{v_j\}$ are vectors of the two DCs. The linkage criterion is the distance between clusters, computed as the average distance between each component of one cluster to each component of the other clusters. The average distance between clusters allows all elements of each DC to contribute in the clustering process. The outcome of the hierarchical clustering is a dendrogram that is used to identify the clusters obtained for the DC set.

Reference models and W/D estimation

a reference VS and a reference VSZ models and their corresponding DC are needed for each cluster in the data transform process. For each cluster, the DC with the broadest frequency band is chosen as the reference DC. It is inverted using an MCI algorithm (Socco & Boiero 2008), to estimate a reference 1-D VSZ model. A wide model space is defined by selecting ranges for each model parameters (VS, thicknesses and the Poisson ratio for each layer). A large population of random models is generated, and their relevant fundamental modes are computed using the Haskell and Thompson forward modelling (Haskell 1953; Thomson 1950; Maraschini 2008). The MCI is optimized by applying scale properties between parameters (see Socco & Boiero 2008 for details). The final result of the MCI is a set of accepted VS models, according to the imposed level of confidence in a statistical one-tailed Fisher test. The accepted models are then transformed into their corresponding VSZ models by the equation

$$\mathrm{VSZ}(z) = \frac{\sum_n h_i}{\sum_n h_i / VS_i}\,, \quad (2)$$

where n is the number of layers down to depth z while the thickness of the ith layer is denoted by h_i and its S-wave velocity by VS_i. The reference VS and VSZ models are computed by averaging the accepted values at each depth. A W/D relationship is then computed, following the method of Socco *et al.* (2017).

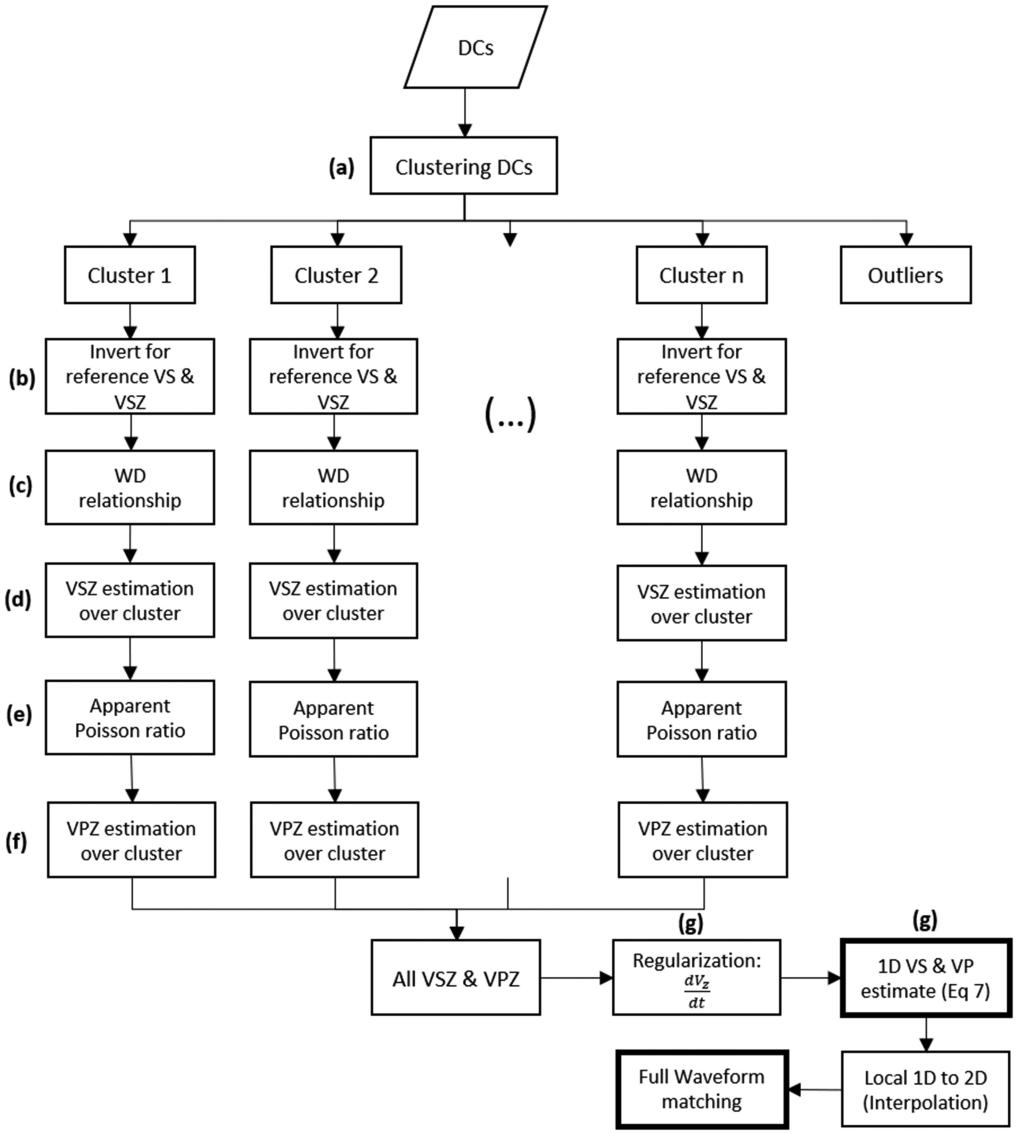

Figure 1. Complete workflow of the direct transformation of DCs of surface wave data into VS and VP models. The DCs are grouped into clusters. For each cluster, one inversion for a reference DC is performed leading to reference VS and VSZ profiles. A W/D relationship is deduced as well as a reference apparent Poisson ratio profile. VSZ and VPZ models are estimated over each cluster and they are used to construct interval VS and VP models. Through interpolation, 2-D or 3-D models are obtained, and they could be used as input for 2-D or 3-D wave propagation modelling.

VSZ estimation over a cluster

Socco *et al.* (2017) showed that given a homogenous set of DCs (where no strong lateral variations exist), the reference W/D relationship can be valid for all DCs to obtain the local VSZs at the position of each DC. The clustering mimics somehow the same configuration as for a stratified medium and the reference W/D relationship for this cluster can be used for all DCs of this cluster to retrieve the associated VSZ profiles.

Reference apparent Poisson ratio and VPZ estimation

Through synthetic and real examples, Socco & Comina (2017) proved that the W/D relationship is sensitive to apparent Poisson ratio υ_z given at the depth z by the relation

$$\upsilon_z = \frac{1}{2}\frac{\left(\frac{\mathrm{VPZ}}{\mathrm{VSZ}}\right)^2 - 2}{\left(\frac{\mathrm{VPZ}}{\mathrm{VSZ}}\right)^2 - 1}. \tag{3}$$

In order to find the apparent Poisson ratio, a set of DCs is computed using the reference VS model and different values of the Poisson ratio, which are the same for all the layers of each synthetic DC. Using the reference VSZ and the synthetic DCs, we compute a set of W/D relationships representing different Poisson ratios. By comparing the experimental W/D relationship and the synthetic W/D relationships, we obtain an apparent Poisson ratio for the reference VSZ model of this cluster. Assuming negligible variation of the apparent Poisson ratio inside each cluster, we extend the reference apparent Poisson ratio to all VSZs and we obtain the VPZs with the equation

$$\mathrm{VPZ} = \mathrm{VSZ}\sqrt{\frac{2\,(\upsilon_z - 1)}{2\upsilon_z - 1}}, \tag{4}$$

corresponding to the position of each DC. The result of such analysis is hence a set of VSZ and VPZ models: more details in Socco & Comina (2017).

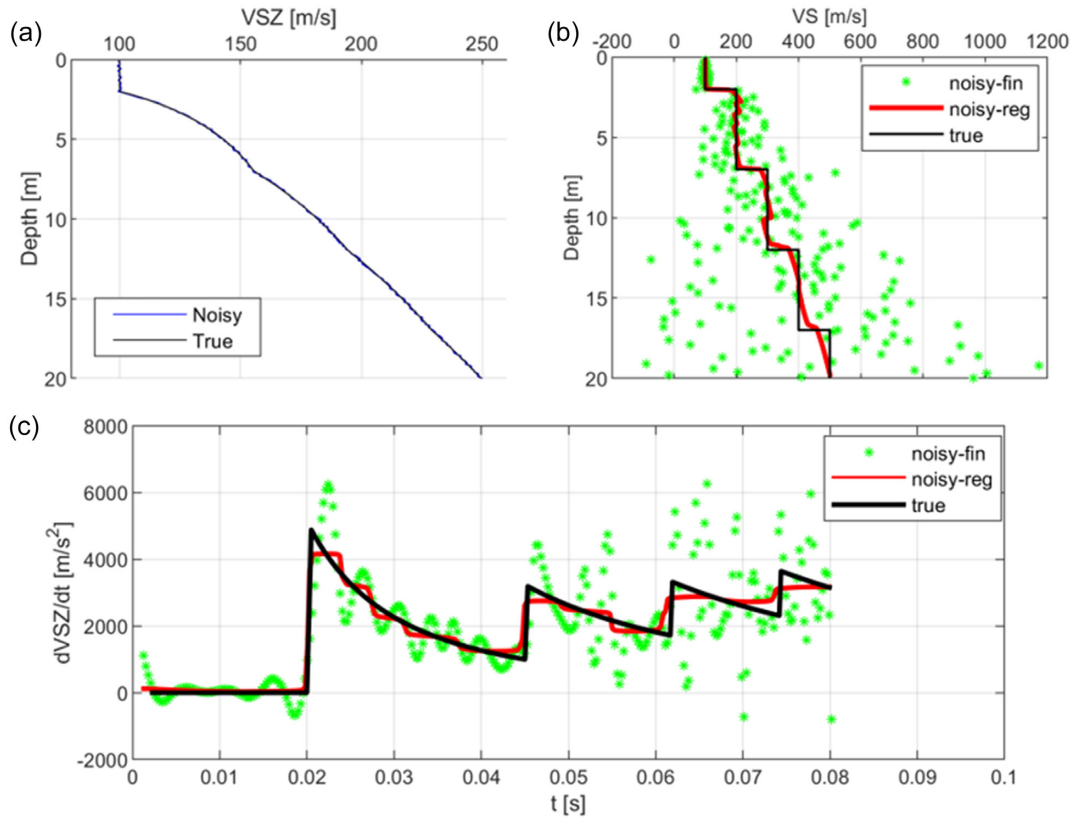

Figure 2. (a) Noise-free VSZ (True) in black and noisy VSZ in blue. (b) Estimated VS profiles using numerical stepwise finite differentiation (noisy-fin, green) and regularized differentiation (noisy-reg, red) compared to the true profile. (c) Estimated d(VSZ)/dt using numerical stepwise finite differentiation (noisy-fin, green) and regularized differentiation (noisy-reg, red) compared to the true derivative.

Estimation of interval velocities from time-average velocity profiles

Eq. (2) expresses the time-average velocity profile. We develop a stable approach for inverting this equation to transform time-average velocities into interval velocities and apply it to the VSZ and VPZ models relevant to all the DCs. Let us define the resident time Δt_i inside the ith interval. Considering $h_i = v_i\ \Delta t_i$, eq. (2) can be written as

$$V_z = \frac{\sum_{i=1}^{n}(v_i \Delta t_i)}{\sum_i^n \Delta t_i}, \tag{5}$$

and, by moving to a continuous depth description, the time-average velocity can be expressed under the integral form as

$$V_z = \frac{\int_0^t v_t\ \mathrm{d}t}{\int_0^t \mathrm{d}t} = \frac{\int_0^t v_t\ \mathrm{d}t}{t}, \tag{6}$$

where the quantity v_t is the interval velocity defined within each interval dt. In this paper, by interval velocity we mean the continuous local velocity profile and it should not be confused with the layered velocity. Eq. (6) is simply the continuous analogue of the discrete eq. (2): an illustration why the quantity V_z is called time-average velocity. This relation is unique between the time and depth variables. The differentiation of the eq. (6) gives $v_t \mathrm{d}t = \mathrm{d}(V_z.t)$ from which we can deduce the interval velocity (v_t) through the expression:

$$v_t = V_z + t\ \frac{\mathrm{d}V_z}{\mathrm{d}t}. \tag{7}$$

Based on eq. (7), the interval velocity v_t consists of a time-average velocity term and a derivative term. For all VSZ and VPZ models that sample the lateral variations along the seismic line, the difference dV_z can be discretized at each depth and divided by the stepwise Δt to get a numerical value of the derivative; however, small perturbations in the time-average velocity lead to large relative changes in the derivative estimation. Hence, a different approach must be applied for the derivative calculation, especially for noisy data.

Many methods are available in literature for such derivative estimation of noisy data, such as least-squares polynomial approximation, Tikhonov regularization and total-variation regularization (Knowles & Renka 2014). The estimated VSZ and VPZ are noisy, because they are directly obtained through data transform. On the other hand, abrupt vertical variations in the velocity profile produce discontinuities in the derivative trend. So, the method to calculate the derivative term should intrinsically allow discontinuities in the derivative estimation to avoid too strong smoothing. Total variation regularization seems to be a good option. Total variation regularization was initially introduced by Rudin *et al.* (1992) to remove the noise from images without smearing the edges. In this method, given a discrete set of data points V_z, the first derivative ($\frac{\mathrm{d}V_z}{\mathrm{d}t}$) is approximated by the solution u of the minimization of the function given by

$$F(u) = \alpha \int_0^T \left|u(t)'\right|\ \mathrm{d}t\ + \frac{1}{2}\ \int_0^T |Au(t) - V_z(t)|^2 \mathrm{d}t, \tag{8}$$

where the equation $Au\ (t) = \int_0^t u(\tau)\mathrm{d}\tau$ defines the smoothed V_z and the regularization parameter is denoted by the symbol α. The term $|Au(t) - V_z(t)|^2$ minimizes the distance between the smoothed

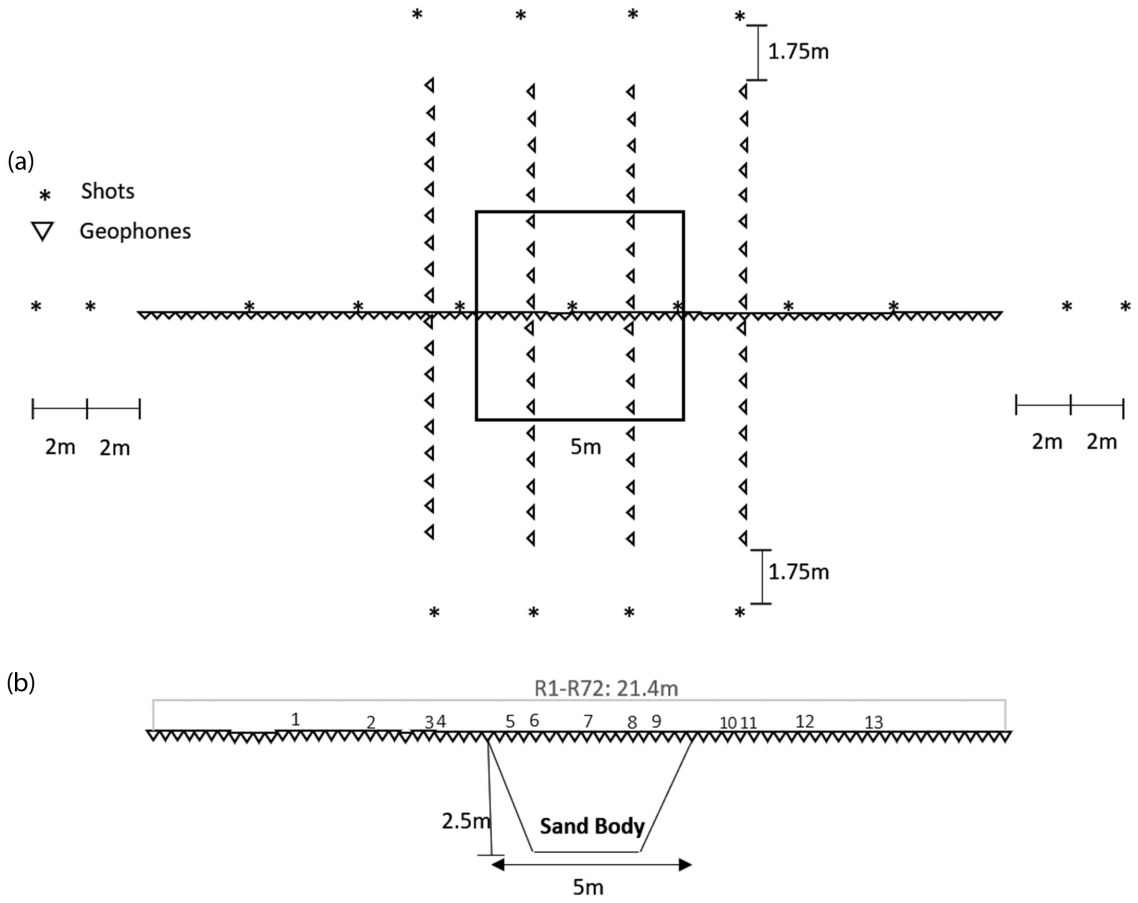

Figure 3. Map (a) and cross-section (b) of the acquisition outline over the known sand body. In (b), the position of the estimated DCs (numbers) is also reported.

function $Au(t)$ and the input function $V_z(t)$. On the other hand, $\int_0^T |u(t)'|\mathrm{d}t$ represents the total variation of the first derivative u. The regularization parameter controls the balance between the two terms. The numerical implementation in Matlab to be used in our application is provided by Chartrand (2011).

Let us consider a synthetic example for the illustration of the noise onto the adopted derivative solution. In Fig. 2 we show the application of eqs (7) and (8) to a noise-free case and to a noisy case. In Fig. 2(a) the noisy VSZ (noise level ± 2 m s^{-1}) is plotted in blue and the noise-free VSZ is represented in black. The difference between these two VSZ is almost unnotable given the low noise level. In Fig. 2(c), the derivative term ($\frac{\mathrm{d}V_z}{\mathrm{d}t}$) of noisy data is evaluated either by finite differentiation in green or by the total variation technique of eq. (8) in red, while Fig. 2(b) provides the wanted interval velocity: a minor amount of noise, present in the time-average velocity, has a great impact on the interval velocity estimation. The total variation regularization provides acceptable reconstruction of the true model (shown in red). This procedure is applied to all estimated VSZ and VPZ models from the seismic experiment to obtain interval VS and VP. These interval velocity profiles could be laterally interpolated for deducing 2-D or 3-D velocity models which can be used for waveform modelling.

The synthetic and real examples previously carried out (Socco & Comina 2017; Socco *et al.* 2017) have shown the validity of the workflow up to the estimation of the time-average VS and VP models. Teodor *et al.* (2018) successfully applied a similar workflow to a synthetic 3-D model to estimate the interval VS and VP. We illustrate the procedure of building interval velocity models on a real data set from a controlled test site which has the advantage of showing a real example of a known model. Reconstructed velocity models are then used for computing synthetic waveforms which are compared with the recorded ones for assessing the potential use of these models as viable initial models for FWI.

TEST SITE

The field data are acquired at CNR (National Research Council) headquarter in Turin, Italy (Fig. 3). This test site contains an artificial loose sand body embedded in more compacted geological formations (i.e. sand and gravel). The sand body occupies an area of 5 m $\times$ 5 m and the sand volume extends to 2.5 m depth.

The acquisition is carried out along a line crossing the sand body, using an 8 kg hammer source in 11 shot positions and 72 vertical 4.5 Hz geophones, evenly spaced every 0.3 m. The acquisition layout is reported in Fig. 3(a). For each source position 8–10 shots are stacked depending on the signal-to-noise ratio. In addition, four cross-line acquisitions are carried out using the same hammer source shooting externally to 18 vertical 4.5 Hz geophones, 0.5 m spaced, for each line. The cross-line acquisitions are carried out in the framework of future 3-D analyses and are here used to increase the density of the DCs along the main (2-D) seismic line. More details about the seismic acquisition can be found in Teodor *et al.* (2017).

RESULTS

The DCs are retrieved along the seismic line in different positions using a moving Gaussian window and the *f–k* processing (Bergamo *et al.* 2012; Fig. 3b). For each position of the window the spectra from the four sources external to the line are stacked to improve the signal-to-noise ratio before extracting the DC. The standard deviation of the adopted Gaussian windows is defined as

$$\sigma = \frac{N}{2\alpha}, \tag{9}$$

where σ is standard deviation of the Gaussian window in terms of number of receivers from the centre of the window, N is the number of receivers minus one and α is a parameter related to the width of the Gaussian window. The value of the parameter α is determined based on the minimum detectable wavenumber required in the *f–k* domain. Having 72 receivers for the experimental data CNR along the 2-D seismic line, we use an α number of 6 which corresponds to standard deviation of 1.8 m and wavenumber resolution of 0.3 rad. We acquire nine DCs along the seismic line each 1.8 m apart. In addition, to increase the density of the DCs along the seismic 2-D line, we compute four DCs using the cross-line acquisitions.

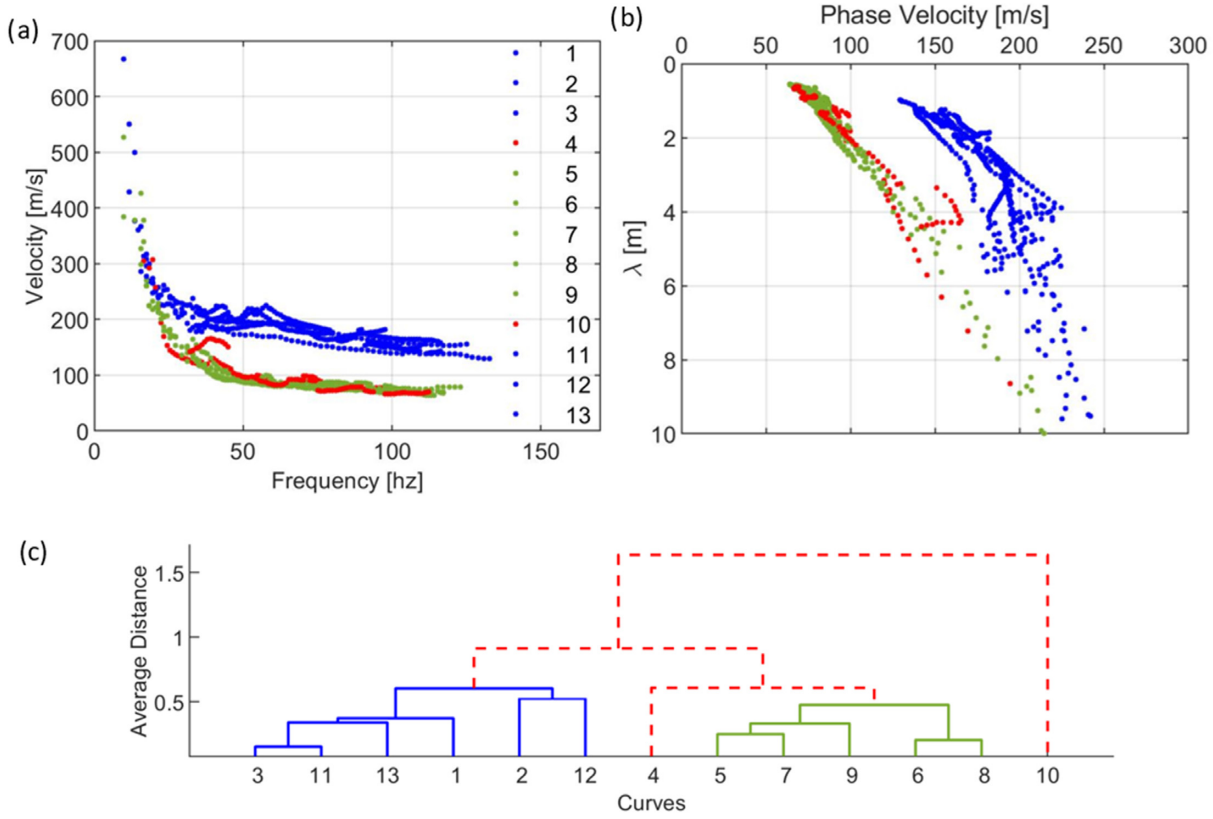

Figure 4. (a) EstimatedDCs in frequency, indicated with different colours based on clusters. (b) Estimated DCs in wavelength, indicated with different colours based on the clusters. (c) The dendrogram obtained through hierarchical clustering.

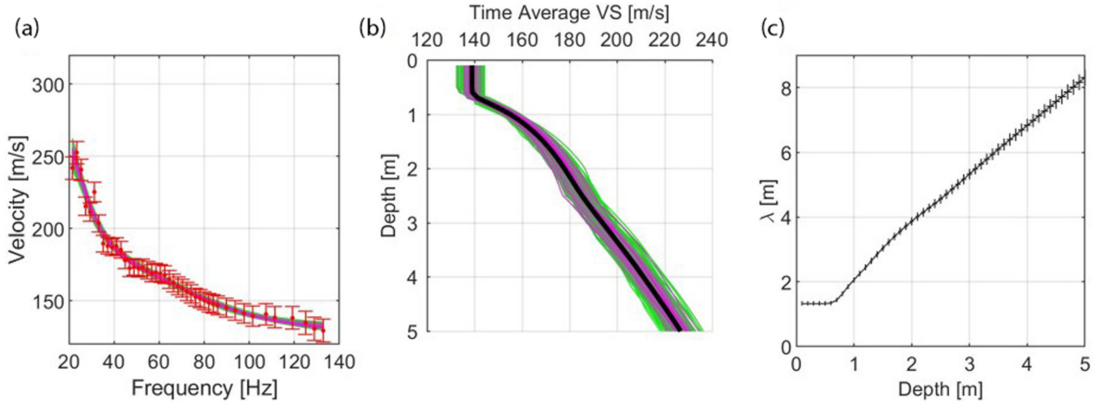

Figure 5. The resultsof the MCI for the reference model outside the sand body. (a) The experimental DC with uncertainties and the accepted DCs, (b) the accepted VSZs and the average VSZ used to retrieve the W/D and (c) the W/D of the accepted models with the horizontal and vertical error bars representing the standard deviation of depth and wavelength, respectively.

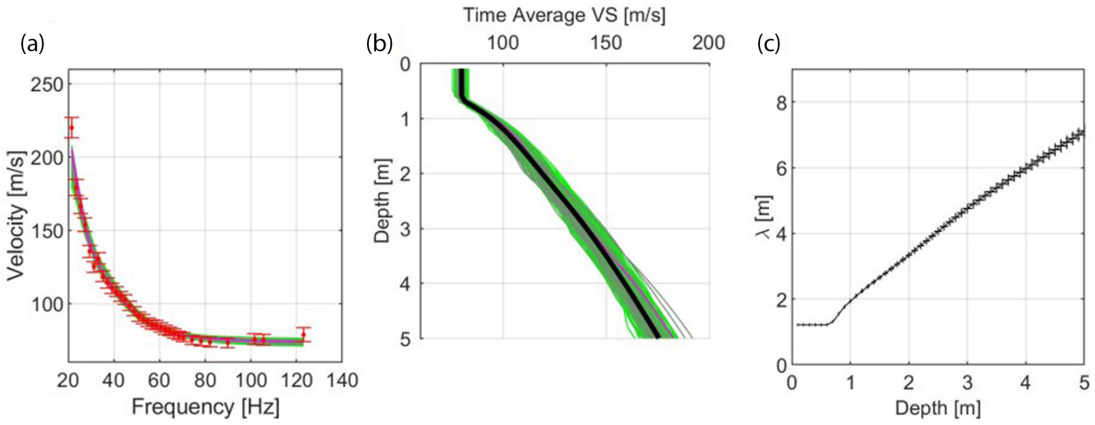

Figure 6. The results of the MCI for the reference model inside the sand body. (a) The experimental DC with uncertainties and the accepted DCs, (b) the accepted VSZs and the average VSZ used to retrieve the W/D and (c) the W/D of the accepted models with the horizontal and vertical error bars representing the standard deviation of depth and wavelength, respectively.

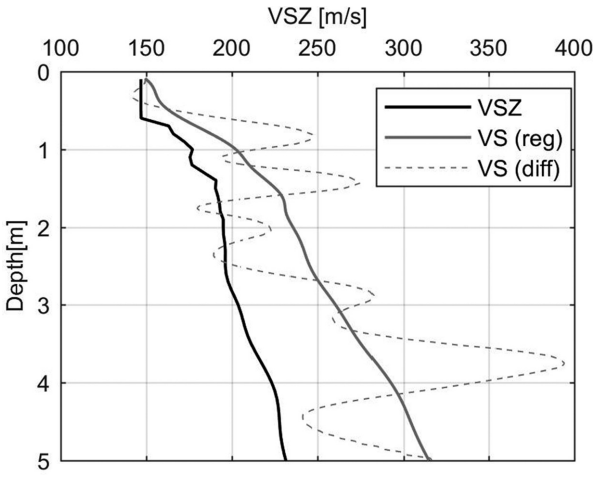

Figure 7. Transformation of the estimated VSZ from DC number 1 of CNR to interval VS with (reg) and without (diff) regularization.

The retrieved DCs (represented in terms of phase velocity versus frequency in Fig. 4a) are continuous and smooth and since the expected velocity model at the site does not contain strong velocity contrast or velocity inversions, we assume that the DCs correspond to the fundamental mode of Rayleigh wave propagation without any contamination by higher modes.

The extracted DCs are analysed by the hierarchical clustering strategy (Fig. 4c). Two main clusters are detected: six profiles for the blue cluster and five profiles for the green cluster. Two DCs are identified as outliers by the clustering process (indicated in red), they are therefore excluded from the workflow. It can be observed that the two selected clusters correspond to DCs located inside and outside the sand body and that the outliers are located at the sand body boundaries. The clustering is, therefore, effective in identifying positions over the survey line where sharp lateral variations occur. The same DCs in terms of wavelength are shown in Fig. 4(b) where the colour scale indicates the clusters and outliers corresponding to the dendrogram in Fig. 4(c).

A reference DC is chosen in each cluster based on the frequency band. The DC number 11 is selected as representative of the cluster blue related to the medium away from the sand body and the DC number 8 is chosen for the other cluster relevant to the sand body. In Figs 5 and 6, we show the results of the MCI to estimate VSZ reference models from the two selected DCs. DCs, VSZs profiles and W/D relationships are shown in Fig. 5 for the background medium and in Fig. 6 for the sand body. When comparing the results shown in Figs 5 and 6, the different velocity between loose sand and more compact surrounding can be seen. The experimental uncertainties associated to the DCs in Figs 5(a) and 6(a) have been determined from a recent work of Passeri (2019), who performed a consistent statistical analysis of DCs uncertainty on a wide data set (see the Appendix). The synthetic DCs shown in Figs 5(a) and 6(a) are relevant to the accepted VSZ models of Figs 5(b) and 6(b) where the black profile shows the average VSZ of the accepted models. Using the experimental DC and the average VSZ, the W/D relationship is deduced (Figs 5c and 6c). In Figs 5(c) and 6(c), we also show the uncertainties associated to the reference W/D relationships as depth standard deviation (horizontal error-bars) and wavelength standard deviation (vertical error-bars). The uncertainties associated to the estimated W/D relationship are estimated by uncertainty propagation analysis through the whole workflow as reported in the Appendix.

All DCs inside each cluster are transformed into VSZ profiles using the reference W/D relationship of the cluster. These VSZ profiles are converted into interval VS profiles by applying eqs (7) and (8). Fig. 7 is an illustration of the more robust estimation of interval velocities using the derivative regularization for the DC number 1 shown in solid grey with respect to finite differentiation shown in dashed grey.

From the 11 DCs, 11 VS profiles are estimated. These continuous depth-dependent VS models are linearly interpolated and a 2-D section of the *S*-wave velocity is obtained. In Fig. 8(a), we show the estimated interval VS with the sand body shape superimposed on the continuous velocity description. The estimated VS shows lower velocity inside the loose sand region, as expected. We use the reference W/D relationships (Figs 5c and 6c) to transform the DCs wavelength into depths. Being the adopted workflow a data transform, the DCs data distribution directly reflect the vertical and lateral resolution that can be attended in the 2-D section of the *S*-wave velocity. In Fig. 8(a), we also show the DCs data distribution in depth and along the profile (black dots). In Fig. 8(b), we show the normalized uncertainty distribution of the estimated 2-D section of the *S*-wave velocity (Fig. 8a), following the uncertainty propagation process described in the Appendix.

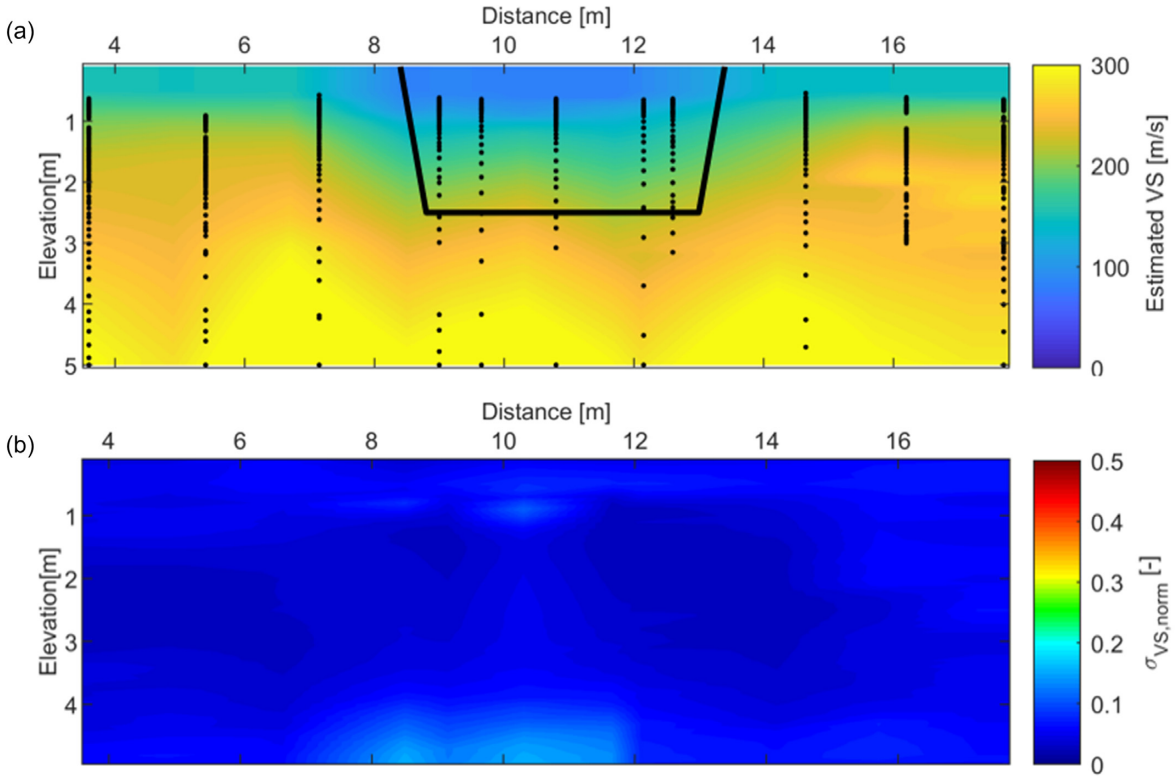

Figure 8. (a) The 2-D VS model extracted from the proposed workflow of DCs, as well as the sand body shape superimposed in black; the black dots represent the DC data distribution along the profile. (b) The normalized uncertainty of the estimated VS in Fig. 8(a).

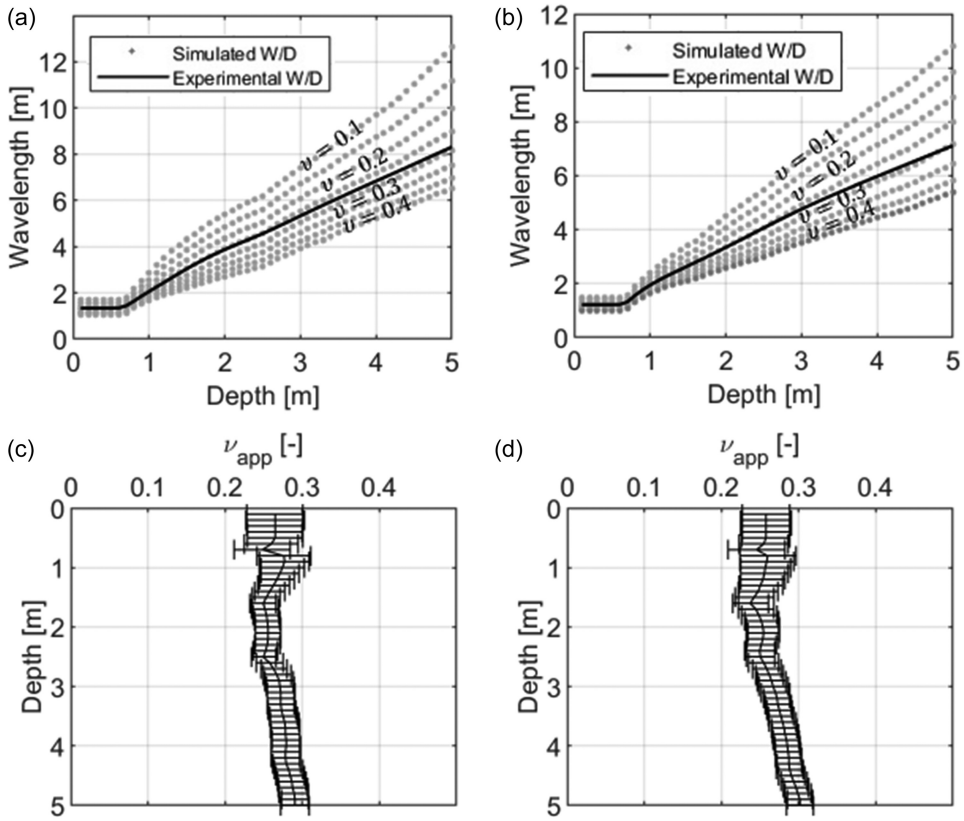

Figure 9. The Poisson ratio analysis ofthe W/D relationships in Figs 5 and 6. (a) Comparing the experimental and synthetic W/Ds for the reference model in Fig. 5. (b) Comparing the experimental and synthetic W/Ds for the reference model in Fig. 6. (c) The estimated apparent Poisson ratio of the reference model in Fig. 5 where the horizontal error bars are the standard deviations. (d) The estimated apparent Poisson ratio of the reference model in Fig. 6 where the horizontal error bars are the standard deviations.

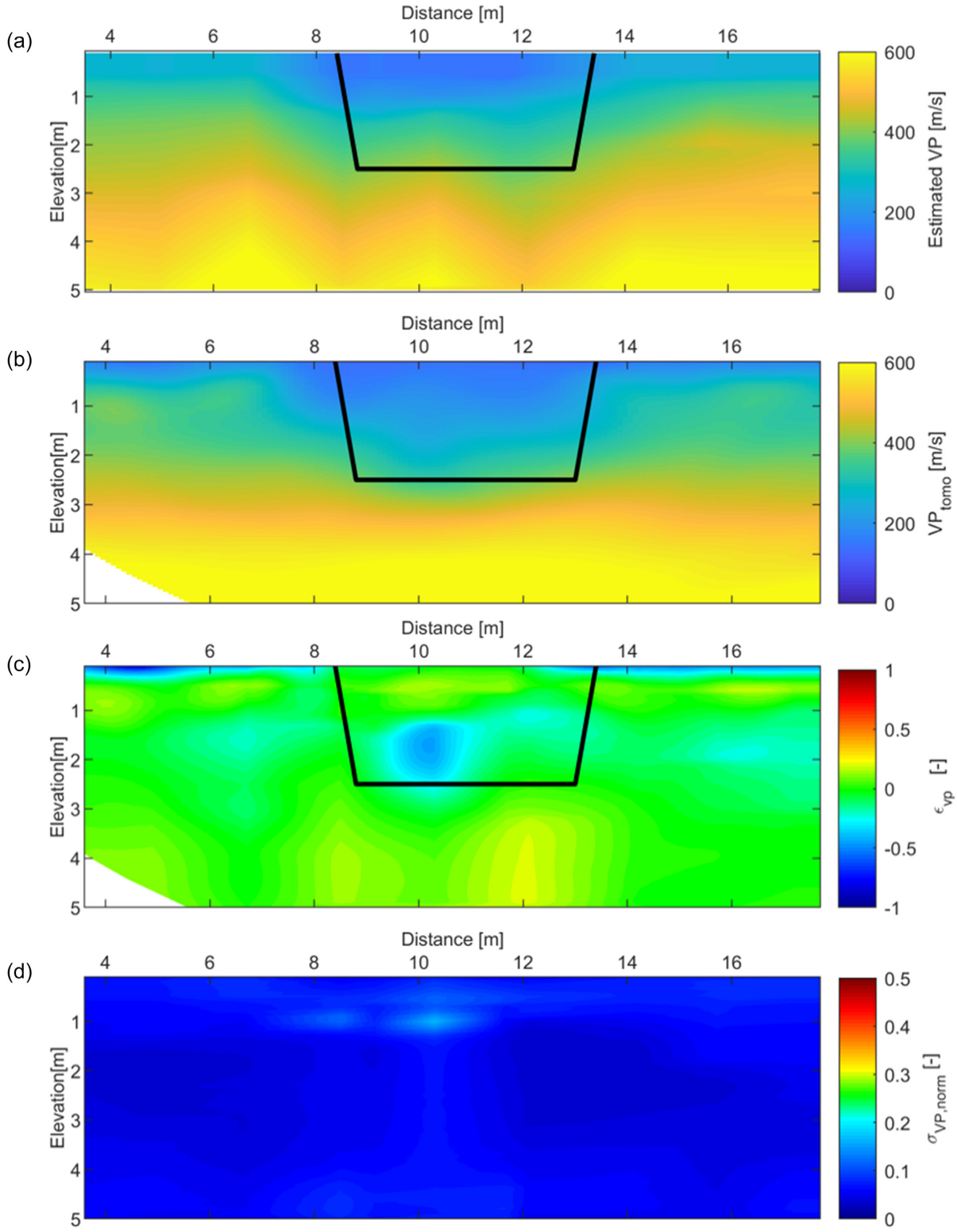

Figure 10. 2-D VP models andsand body shape superimposed: (a) Estimated VP. (b) Benchmark VP from the *P*-wave tomography. (c) The normalized difference between the estimation and the benchmark. (d) The normalized uncertainty of the estimated VP in (a).

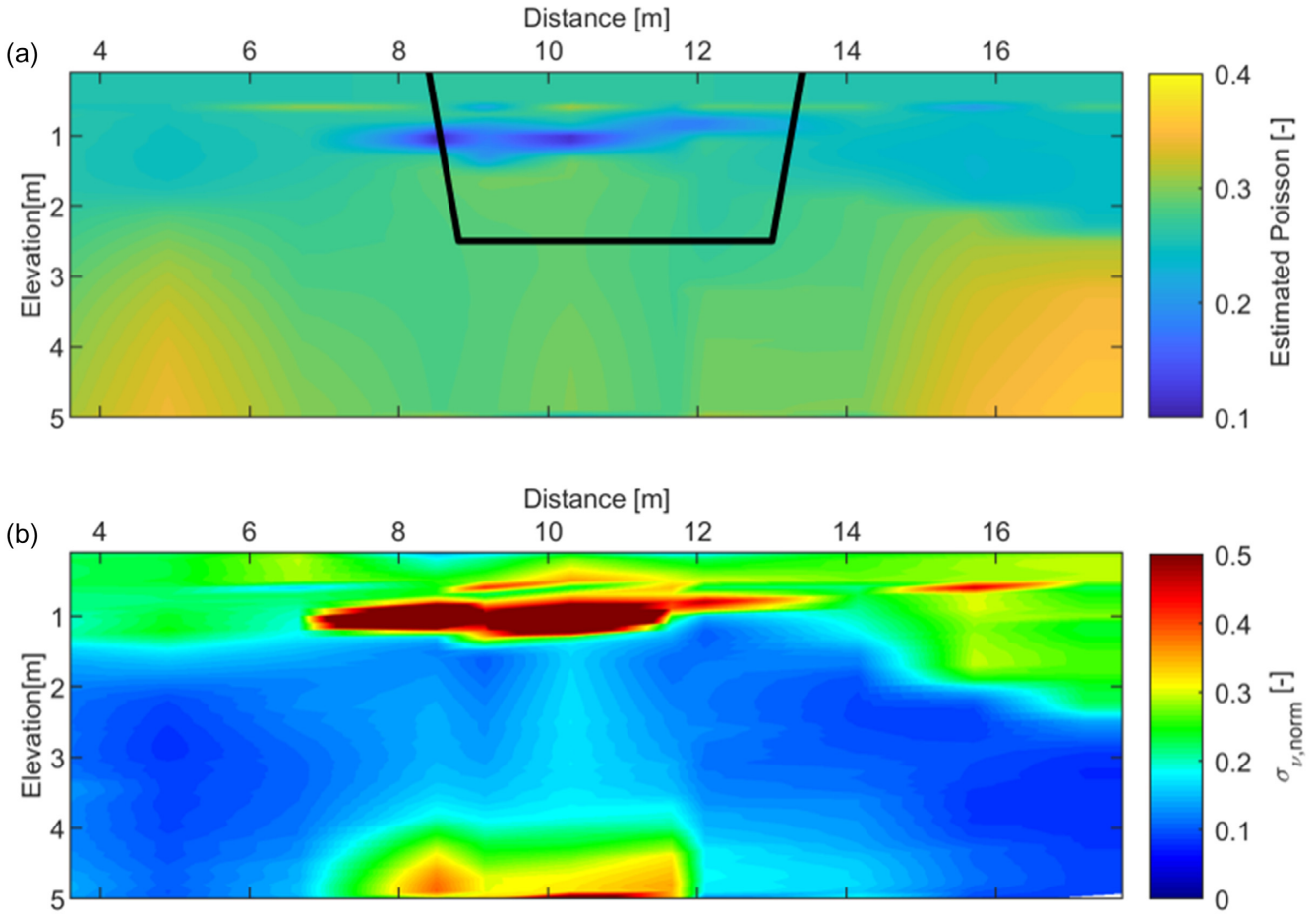

Figure 11. (a) The estimated Poisson ratio calculated from the estimated VP (Fig. 10a) and estimated VS (Fig. 8a). (b) The normalized uncertainty of the estimated Poisson ratio.

The reference W/D relationships (Figs 5c and 6c) and the reference VS (from the MCI) of each cluster are used to estimate the reference apparent Poisson ratio. In Figs 9(a) and (b), we show the experimental W/D relationships along with synthetic W/D relationships (each related to a constant Poisson ratio) corresponding to the reference models in Figs 5 and 6, respectively. By comparing the experimental W/D relationships with the synthetic ones, we obtained the apparent Poisson ratio profiles in Figs 9(c) and (d), respectively, where the error bars show the standard deviations. The process of obtaining these uncertainties is reported in the Appendix.

VPZ profiles are then obtained from the estimated VSZ models by using the reference apparent Poisson ratio in each cluster. Using eqs (7) and (8), these VPZ models are converted to interval VP and interpolated into a 2-D VP model (called VP_{est}) shown in Fig. 10(a). In Fig. 10(b), VP_{est} is compared with the VP model (called VP_{tom}) obtained from a *P*-wave traveltime tomography. At each point, the difference ε between these two models is calculated through the expression

$$\varepsilon = \frac{VP_{tom} - VP_{est}}{VP_{tom}} . \quad (10)$$

In Fig. 10(c), we show this normalized difference ε. In most of the regions of the 2-D section, the difference between the estimated VP and the benchmark is less than 5 per cent. The highest differences are registered at very shallow depths as well as at the bottom and below the sand body. The normalized uncertainty distribution of the estimated 2-D section of the *P*-wave velocity (Fig. 10a) following the uncertainty propagation process described in the Appendix is shown in Fig. 10(d).

The Poisson ratio is computed using the estimated VP and VS models . In Fig. 11 we show the estimated Poisson ratio. In most parts of the models, the estimated Poisson ratio shows values between 0.2 and 0.35 which is expected for the dry granular materials encountered at such site. The normalized uncertainty of the estimated Poisson ratio (Fig. 11a) is shown in Fig. 10(b). The process of retrieving this uncertainty is reported in the Appendix.

FULL-WAVEFORM SIMULATION

Using the estimated 2-D VP and VS models, we construct a full 3-D model, representative of the site conditions, to numerically simulate the full-waveform propagation from the position of the sources and store the recordings at the position of the 72 receivers along the 2-D acquisition line. These synthetic records are then compared with the observed (real) recordings.

Full-waveform numerical simulations are performed using a 3-D spectral-element method (Trinh *et al.* 2019). To construct the 3-D simulation domain, we extend laterally the estimated 2-D models to a 3-D structure, by symmetrically replicating the 2-D section along the cross-line direction (*y*-direction). The extent of the sand body in the *y*-direction is kept equal to the one observed in the *x*-direction.

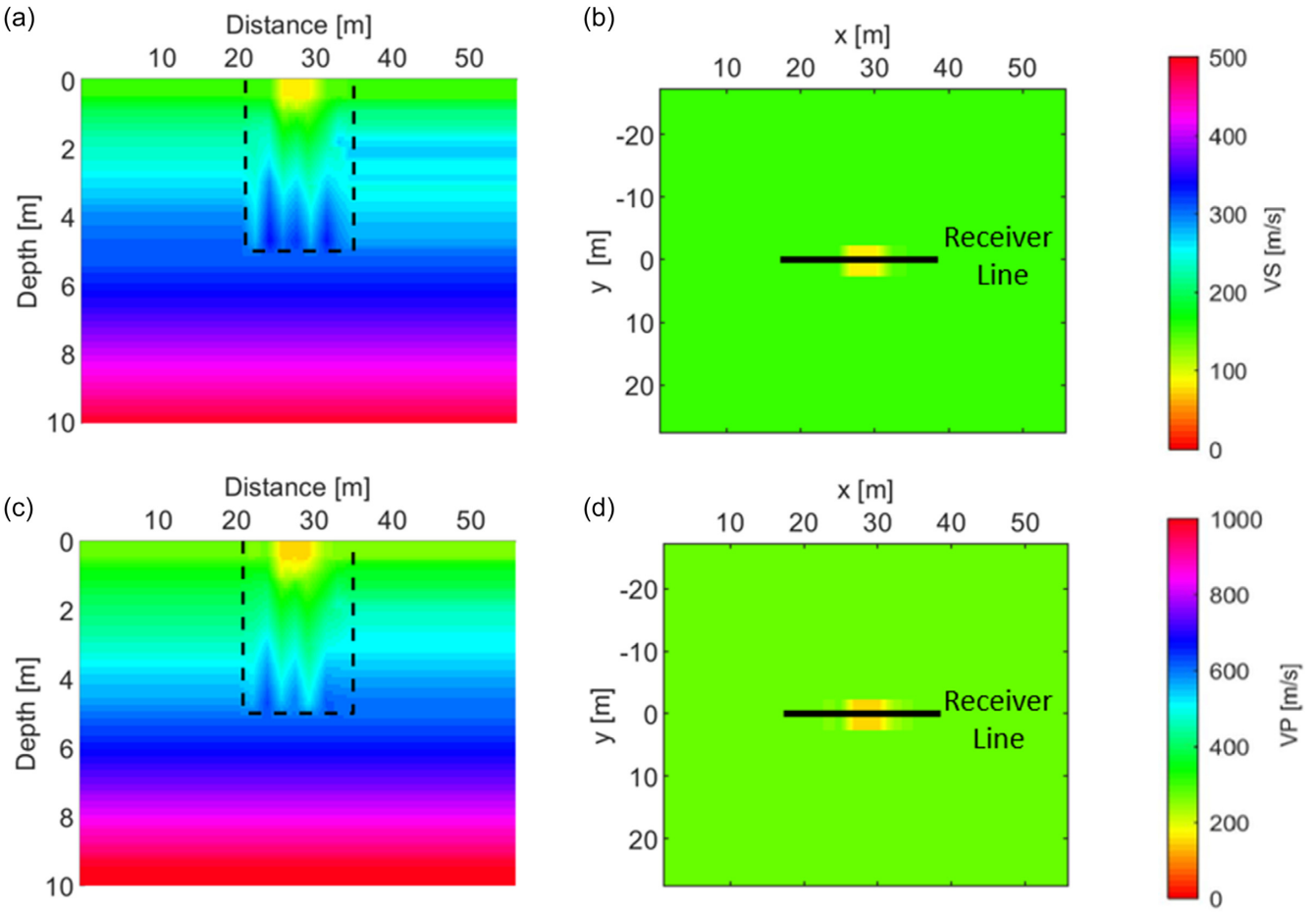

Figure 12. (a) The vertical section view of the extended VS superimposed with the boundaries of the estimated VS before extension. (b) The map view of the extended VS (at $z = 0$). (c) The vertical section view of the extended VP superimposed with the boundaries of the estimated VP before extension. (d) The map view of the extended VP (at $z = 0$).

In addition, the models are extended laterally along the x-direction and in depth, beyond the investigation limits of the method, with the aim of avoiding boundary artefacts during numerical simulations. The vertical extension is performed using a constant gradient, extrapolating the trend of the lower portion of the estimated models. In Figs 12(a) and (c), we show the vertical sections of the extended VS and VP superimposed with the boundaries of the estimated VS and VP before extension (dashed black line). In Figs 12(b) and (d), we show the horizontal section of the 3-D models (at $z = 0$) where the black line shows the position of the receivers along the 2-D seismic line.

The model parameters needed for the simulations are the estimated VP and VS, density and quality factors (QP and QS). The density is assumed constant over the 3-D space (1800 kg m^{-3}), while the values of QP and QS are estimated from the velocity models according to Hauksson & Shearer (2006).

We perform the simulations considering the space and time numerical dispersion criteria. In agreement with the frequency band (10–150 Hz) and the velocity range (80–1050 m s^{-1}), a mesh with constant element size of 0.3 m (in the x-, y- and z-directions) is used for honouring the wavelength sampling to avoid numerical dispersion. This requires a choice for the minimum element size so that at least 5 Gauss–Lobatto–Legendre nodes are present per shortest wavelength to accurately model the elastic waves propagation with the interpolation order equal to 4 implemented in the finite-element code (code SEM46 described in Trinh *et al.* 2019). We set the explicit time integration stepping to 1.4e-5 s for honouring the Courant–Friedrichs–Lewy time stability condition (i.e. the minimum time step, related to the mesh size and the maximum velocity of the model).

The results of the simulation together with the observed (real) records at one shot position (in Fig. 3, the fourth shot from the left) are shown in Fig. 13. The results (in red) of the forward modelling show the same patterns of phases as the real data (in black), except for small diffracted phases. In Figs 13(b) and (c), zoomed sections (shown as black boxes in Fig. 13a) illustrate the in-phase shape of waveforms. This suggests that the VS and VP models could be good candidates as initial models for FWI.

As a further validation, DCs are extracted from the synthetic data and compared with those deduced from real data (Fig. 14). Differences between phase velocities of these DCs are always less than 10 per cent. The two different phase velocity trends, inside and outside the sand body, can be observed both for the real DCs and for the ones deduced from numerical simulations.

Cross-correlation between real and synthetic data can be a fair criterion for the goodness of the full-waveform matching when

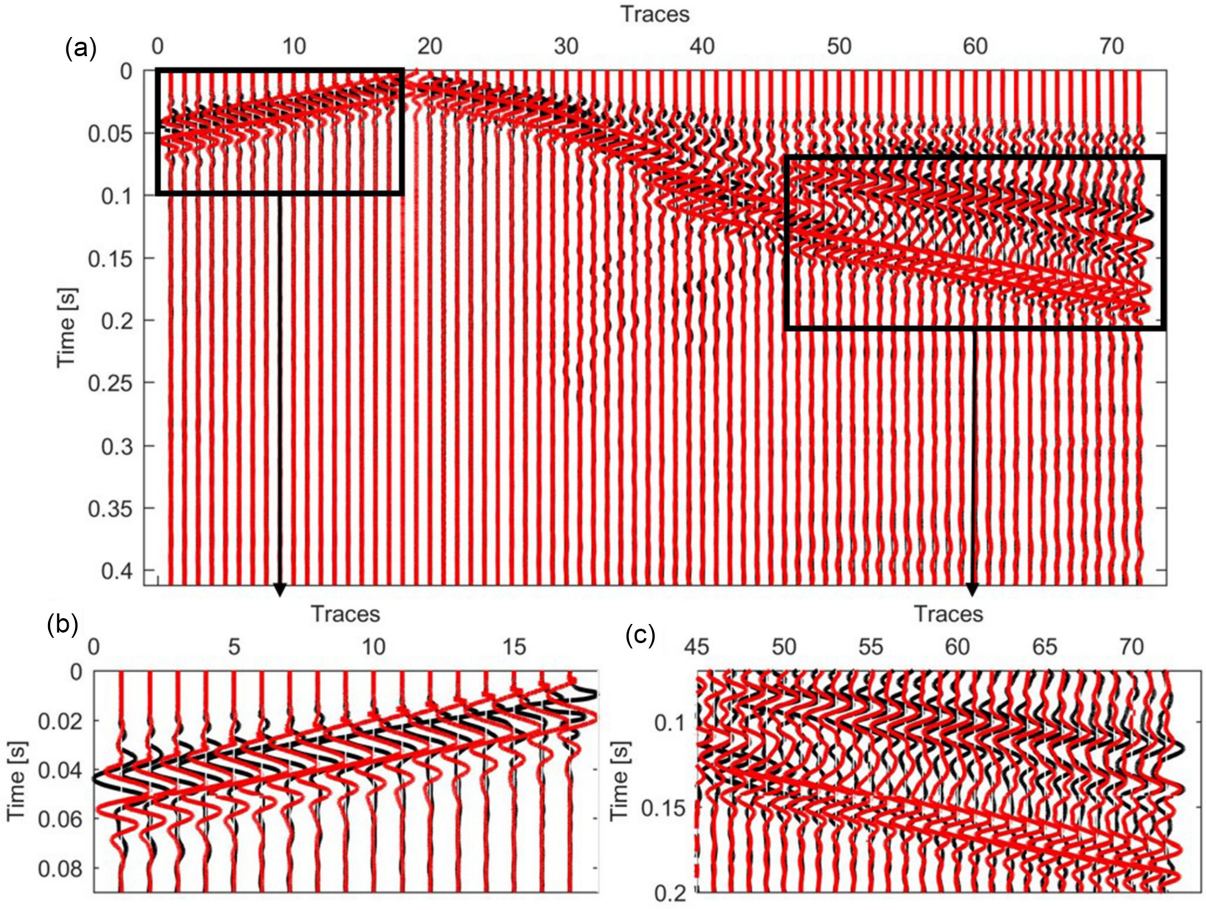

Figure 13. Full-waveform comparison. In black we show the traces of shot number 4 recorded on the field. In red we display the traces, for the same shot position, obtained by running forward simulations on the 3-D estimated VP and VS models. (a) The whole recordings. (b) The zoomed view of the traces 0–17. (c) The zoomed view of the traces 45–72.

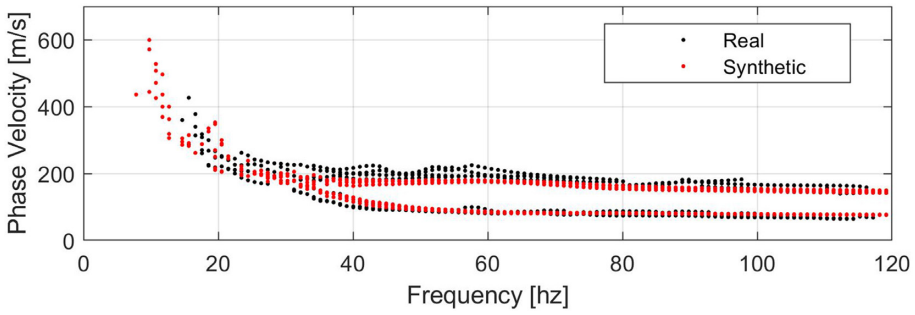

Figure 14. Comparison between the experimental DCs and the DCs from synthetic seismograms of numerical simulations.

considering possible cycle-skipping effects (Fig. 15). In Fig. 15 we show the results of this cross-correlation. The red solid line in Fig. 15 shows the zero-time lag. To avoid cycle-skipping in the FWI process, the error in matching the real and synthetic data should fall below half-the-period of the dominant frequency (Virieux & Operto 2009). For our data, at a dominant frequency of 60 Hz, the half-period corresponds to 8.33 ms. For the shot 4 (the source located between the traces 18 and 19), it can be observed in Fig. 15 that the short-offset traces (0–52) show lower time-lag than half-the-period, while long-offset traces (52–72) show a time lag higher than 8.33 ms, leading to possible cycle-skipping problems, relevant at long offsets, for arrival times greater than 0.12 s (see Fig. 13c). The high-frequency surface waves generated from the source are highly affected by the low velocity of the sand body before reaching the receivers 52–72. Slight errors in the velocity or in the size of the sand body significantly affect the arrival time of the high-frequency phases. These effects, related to the interaction of the wave front with the sand body anomaly, can be mitigated by hierarchical strategies from low-to-high frequencies and from short offsets to large offsets. Dedicated FWI strategies should be designed for overcoming these cycle-skipping issues.

As an illustration of such hierarchical strategy, let us mute this portion of the recordings and consider the muted recordings as inputs of FWI, at least for the first iterations. In Fig. 16(a), the muted real recordings in black and the muted synthetic recording in red are shown. In Figs 16(b) and (c), the zoomed views (again as

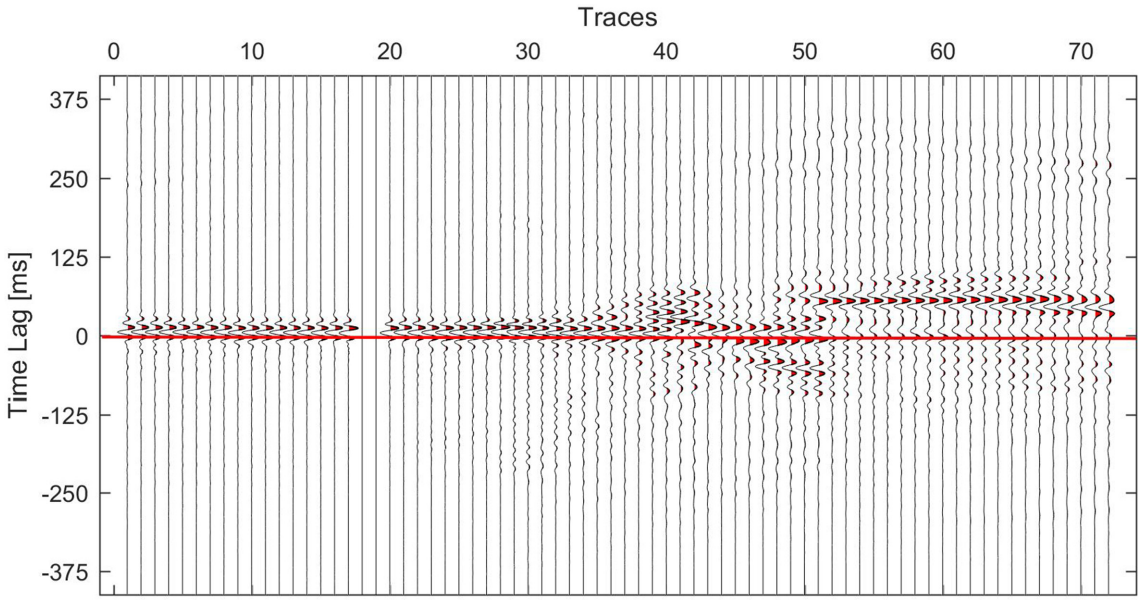

Figure 15. The cross correlationof the experimental data with the synthetic ones.

black boxes in Fig. 16a) show better behaviours of wiggles for far offsets.

The impact of the muting on the cycle-skipping is illustrated by the cross-correlation of real and synthetic recordings in Fig. 16(a) and plotted in Fig. 17. The cross-correlation shows an acceptable time-lag for the whole seismogram considering the main frequency of the recordings as 60 Hz. This will open the road for a more systematic analysis based on frequency, offsets and multisources acquisition for FWI high-resolution reconstruction.

DISCUSSION

The lateral resolution of velocity models corresponds to the lateral resolution of the local DCs (Fig. 8a) and depends on the processing method used for estimating the DCs. The DCs in this work were estimated by spatial Gaussian windowing of the data before computing the *f–k* spectrum and picking the DCs (Bergamo *et al.* 2012). The standard deviation (σ) of the Gaussian window for a fixed number of receivers is related to parameter α (eq. 9). The estimated DCs from this method represent the properties of the subsurface in the position of the Gaussian window's centre and the lateral resolution can be considered as 2σ (one σ at each side of the centre). Specifically, a larger α corresponds to lower standard deviation of the Gaussian window and therefore higher spatial resolution. The value of the parameter α is usually based on the minimum detectable wavenumber required in the *f–k* domain and a trade-off exists between lateral and spectral resolution. In this work a value of α equal to 6 was adopted, with a resulting lateral resolution of 3.6 m. This corresponds to a minimum detectable normalized wavenumber of 0.3 rad. However, these values can be also tuned depending on the desired lateral resolution that is aimed to be obtained within the workflow. According to Bergamo *et al.* (2012), the optimal distance between two neighbouring DCs is twice the standard deviation. This ensures that each DC is affected by independent portion of the subsurface. However, here, we extracted a DC every 1.8 m ($=$ σ) to increase the number of DCs and improve the lateral resolution. More dense set of DC could be extracted by shifting the moving window with smaller steps. Virtually the moving window could be shifted of a step equal to receiver spacing leading to very dense set of DCs. This anyway would not correspond to an improved lateral resolution since the neighbouring windows would be highly overlapped to each other replicating the same information. It is important to stress that the lateral resolution depends on the DC extraction and on the spatial sampling and not on the processing workflow which is applied later. The presented data transform process does not introduce any smoothing along the seismic line.

The vertical resolution of the estimated models is conversely related to the distance between experimental data points of DCs in wavelength (Fig. 4b). To represent the vertical resolution in depth, using the reference W/D relationships, we transformed the wavelengths of DCs into depth (the black dots in Fig 8a). Using a uniform sampling in frequencies in the *f–k* domain and assuming higher phase velocities at lower frequencies, the distance between DC's data points increases with increasing depth (Fig. 8a). This increase leads to a lower resolution in the deeper portion of the investigated medium. Conversely, shallow depth data points are closer to each other which implies higher vertical resolution in the shallow portion of the investigated medium. In the present case, the vertical resolution of the obtained velocity models can be considered in the order of 10 cm.

The velocity models extracted with the proposed approach have also uncertainties that are related to the different steps of the workflow and that are specifically discussed through an uncertainty propagation analysis reported in the Appendix. The results of this uncertainty analysis for the estimated interval VS and VP (Figs 8b and 10d) show minor and uniform uncertainties (less than 10 per cent in most regions). The highest uncertainties (around 15 per cent) for the interval VS and VP were recorded in some portion inside and below the sand body area. The resulting uncertainties are directly related to the DCs' uncertainties and hence depend on the data quality only. Since in this example only four source points were used to estimate the DCs, it was not possible to estimate an

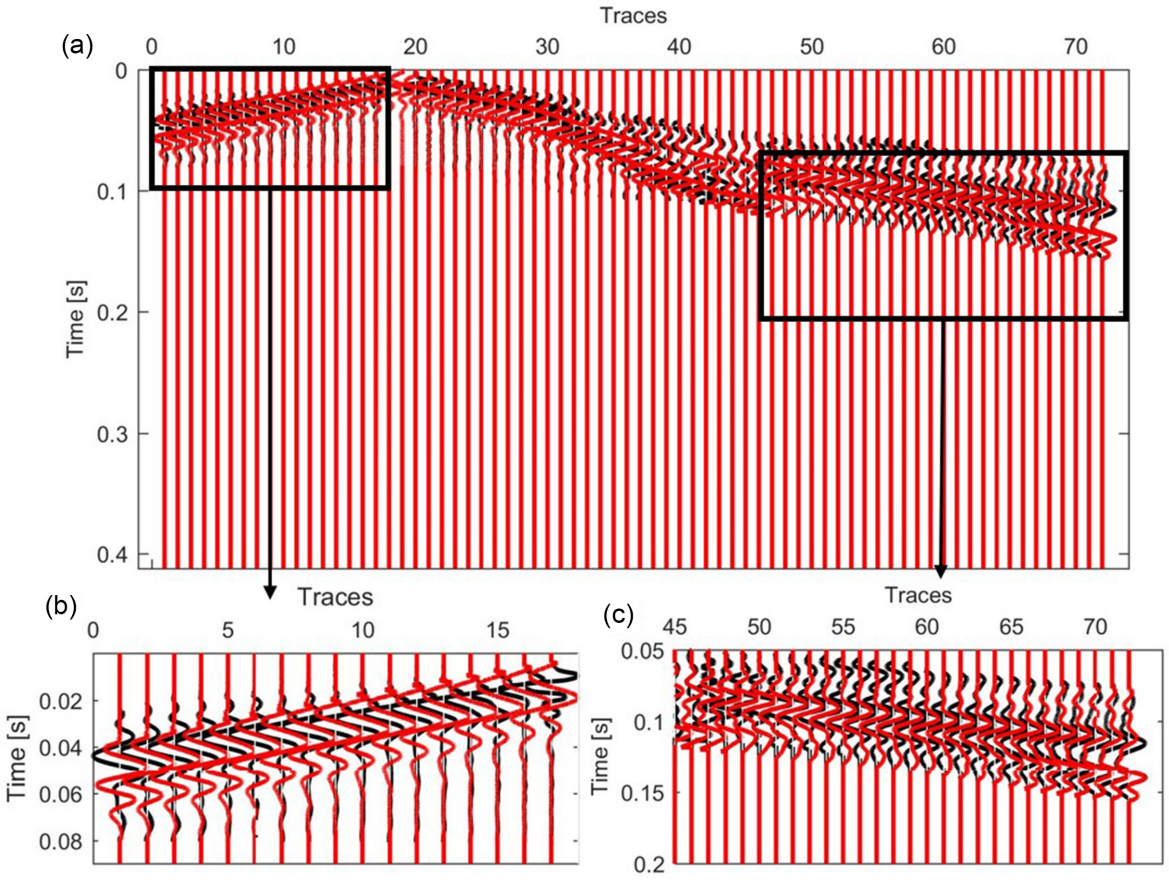

Figure 16. Full-waveform comparisonof the muted data. In black we show the muted traces of shot number 4 recorded on the field. In red we display the traces, for the same shot position, obtained by running forward simulations on the 3-D estimated VP and VS models and muted afterwards. (a) The whole recordings. (b) The zoomed view of the traces 0–17. (c) The zoomed view of the traces 45–72.

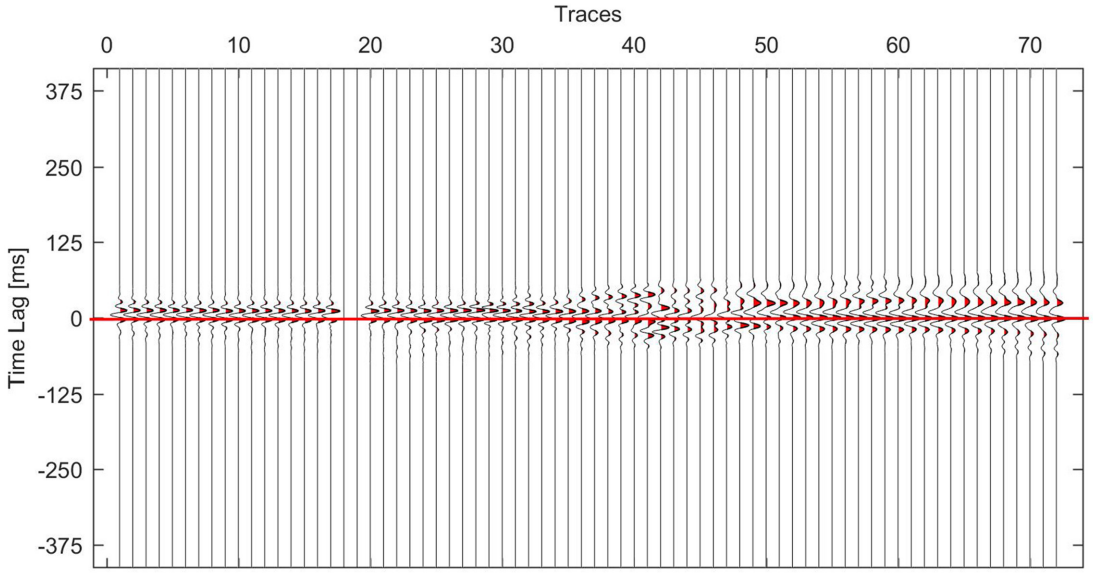

Figure 17. The cross correlation of the muted experimental data with the muted synthetic ones shown in Fig. 16.

experimental uncertainty from the data which is usually associated with the DCs (see e.g. Xia *et al.* 2002; Marosi & Hiltunen 2004; Lai *et al.* 2005; Socco *et al.* 2009). A conventional uncertainty was assumed on the basis of a statistical study on a large number of data sets (Passeri 2019, appendix). If a large number of shots is acquired without stacking the records, the individual shots can be used as element of population of measurement and uncertainties of the DC data points can be estimated from the data (Socco *et al.* 2009). In general, the data quality can be improved using more powerful sources to increase the signal-to-noise ratio or increasing the density of the spatial sampling given the same seismic line length (Socco & Strobbia 2004). Although the Poisson ratio is not a direct output of the method, the uncertainty of the estimated Poisson ratio is calculated to be less than 20 per cent in most parts of the model. The highest uncertainties (more than 40 per cent) for the estimation of the Poisson ratio were registered at shallow portions of the model. This aspect can also be evidenced in apparent Poisson ratio estimation of the reference models. The sensitivity in apparent Poisson ratio conversely increases with depth given that constant Poisson ratio W/D relationships diverge at large wavelengths (Figs 9a and b) and therefore allow for a more consistent apparent Poisson ratio estimation.

The proposed method is based on direct transformation of local DCs. As a result, in the presence of complex topography along the investigated area devoted attention is required in the application of the workflow. In addition, the fundamental modes of DCs are the inputs of the workflow and the method is not considering possible contamination of fundamental mode curve by higher modes. In the present case, the assumption of fundamental mode was strongly supported by the data and by the expected smooth velocity gradient. If higher modes are present in the data, and can be distinguished from the fundamental one, they can be still used within the workflow in a similar approach to the one presented here. An example of the application of the workflow to higher mode DCs is reported in Bamarouf *et al.* (2017).

CONCLUSION

We have described a workflow to transform surface-wave data to VS and VP models without any prior information and without the need for intensive inversions. Direct transformation of DCs gathered by clusters into time-average VS and time-average VP profiles is possible using a W/D relationship specific to each cluster. These profiles are converted into interval VS and VP profiles using total variation regularization of needed derivatives. This specific VP estimation from surface waves shows, in most regions, less than 5 per cent difference when compared to a VP tomographic model. These profiles are interpolated into 2-D velocity models to be used for different analysis. The clustering approach seems to be effective in the presence of sharp lateral variations.

In the perspective of inverting the full content of seismograms for recovering improved velocity structures, the performed full-waveform simulation based on these estimated VS and VP shows good agreement with the real records at short-offset ranges, while far offsets may still require specific strategies for overcoming cycle-skipping issues. Muting and/or filtering traces at least at the first iterations of the FWI should be advisable. Still the global agreement between real and synthetic waveforms makes us confident that FWI will work out when starting from these initial VS and VP models estimated from DCs, opening a door for near-surface high-resolution seismic imaging.

ACKNOWLEDGEMENTS

We would like to thank CNR group for giving us access to do the acquisition. FKA would like to thank TOTAL E&P for supporting his PhD and research.

REFERENCES

Badji, R. *et al.*, 2016. Simultaneous joint inversion for near surface characterization—improving from first break picking to statics, in *78th EAGE Conference and Exhibition,* doi:10.3997/2214-4609.201601151.

Bamarouf, T., Socco, L.V. & Comina, C., 2017. Direct Statics estimation from ground roll data—the role of higher modes, in *79th EAGE Conference and Exhibition,* doi:10.3997/2214-4609.201700958.

Beaty, K.S., Schmitt, D.R. & Sacchi, M., 2002. Simulated annealing inversion of multimode Rayleigh wave dispersion curves for geological structure, *Geophys. J. Int.,* **151**(2), 622–631.

Bergamo, P., Boiero, D. & Socco, L.V., 2012. Retrieving 2D structures from surface-waves data by means of space-varying windowing, *Geophysics,* **77**(4), EN39–EN51.

Bodin, T., Sambridge, M., Gallagher, K. & Rawlinson, N., 2012. Transdimensional inversion of receiver functions and surface wave dispersion, *J. geophys. Res.,* **117**(B2), doi:10.1029/2011JB008560.

Brossier, R., Operto, S. & Virieux, J., 2009. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion, *Geophysics,* **74**(6), WCC105–WCC118.

Chartrand, R., 2011. Numerical differentiation of noisy, nonsmooth data, *ISRN Appl. Math.,* **2011,** 164564, doi:10.5402/2011/164564.

Foti, S. & Strobbia, C., 2002. Some notes on model parameters for surface wave data inversion, in *Symposium on the Application of Geophysics to Engineering and Environmental Problems (SAGEEP)*, doi:10.4133/1.2927179.

Haskell, N., 1953. The dispersion of surface waves on multilayered media, *Bull. seism. Soc. Am.,* **43,** 17–34.

Hauksson, E. & Shearer, P.M., 2006. Attenuation models (*Qp* and *Qs*) in three dimensions of the southern California crust: Inferred fluid saturation at seismogenic depths, *J. geophys. Res.,* **111,** B05302, doi:10.1029/2005JB003947.

Khosro Anjom, F., Arabi, F., Socco, L.V. & Comina, C., 2017. Application of a method to determine S and P wave velocities from surface waves data analysis in presence of sharp lateral variations, in 36th GNGTS national convention, Expanded Abstract.

Knowles, I. & Renka, R.J., 2014. Methods for numerical differentiation of noisy data, *Electon. J. Differ. Equ.,* **21,** pp 235–246.

Lai, C.G., Foti, S. & Rix, G.J., 2005. Propagation of data uncertainty in surface wave inversion, *J. Environ. Eng. Geophys.,* **10**(2), 67–234.

Maraschini, M., 2008. A new approach for the inversion of Rayleigh and Scholte waves in site characterization, *PhD thesis*, Politecnico di Torino.

Marosi, K.T. & Hiltunen, D.R., 2004. Characterization of spectral analysis of surface waves shear wave velocity measurement uncertainty, *J. Geotechnical Geoenviron. Eng.,* **130**(10), 1034–1041

Masoni, I., Brossier, R., Virieux, J. & Boelle, J.L., 2014. Robust full waveform inversion of surface waves, in *SEG Technical Program Expanded Abstracts,* pp. 1126–1130, doi:10.1190/segam2014-1077.1.

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

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

Pérez Solano, C.A., Donno, D. & Chauris, H., 2014. Alternative waveform inversion for surface wave analysis in 2-D media, *Geophys. J. Int.,* **198**(3), 1359–1372.

Rokach, L. & Maimon, O., 2005. *Clustering Methods, Data Mining and Knowledge Discovery Handbook,* Springer, pp. 321–352.

Rudin, L., Osher, S. & Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms, *Phys. D,* **60**(1-4), 259–268.

Socco, L.V. & Boiero, D., 2008. Improved Monte Carlo inversion of surface wave data, *Geophys. Prospect.,* **56,** 357–371.

Socco, L.V. & Comina, C., 2017. Time-average velocity estimation through surface-wave analysis: Part 2—*P*-wave velocity, *Geophysics,* **82**(3), U61–U73.

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

Socco, L.V., Boiero, D., Foti, S. & Wisen, R., 2009. Laterally constrained inversion of ground roll from seismic reflection records, *Geophysics,* **74**(6), G35–G45.

Socco, L.V., Foti, S. & Boiero, D., 2010. Surface-wave analysis for building near-surface velocity models—established approaches and new perspectives, *Geophysics,* **75**(5), A83–A102.

Socco, L.V., Comina, C. & Khosro Anjom, F., 2017. Time-average velocity estimation through surface-wave analysis: Part 1—*S*-wave velocity, *Geophysics,* **82**(3), U49–U59.

Teodor, D., Comina, C., Socco, L.V., Brossier, R., Trinh, P.T. & Virieux, J., 2017. Initial model design for full-waveform inversion—preliminary elastic modeling from surface waves data analysis, Extended Abstract, in 36th GNGTS national convention, Expanded Abstract.

Teodor, D., Comina, C., Khosro Anjom, F., Socco, L.V., Virieux, J., Trinh, P.T. & Brossier, R., 2018. Building initial models for Full-Waveform Inversion of shallow targets by surface waves dispersion curves clustering and data transform, SEG, doi.org/10.1190/segam2018-2997848.1.

Thomson, W. T., 1950. Transmission of elastic waves through a stratified solid medium, *J. Appl. Phys.,* **21,** 89–93.

Trinh, P.T., Brossier, R., Métivier, L., Tavard, L. & Virieux, J., 2019. Efficient time-domain 3D elastic and visco-elastic full-waveform inversion using a spectral-element method on flexible Cartesian-based mesh, *Geophysics,* doi:10.1190/geo2018-0059.1.

Van Leeuwen, T. & Herrmann, F.J., 2013. Mitigating local minima in full-waveform inversion by expanding the search space, *Geophys. J. Int.,* **195**(1), 661–667.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics,* **74**(6), WCC1–WCC26.

Warner, M. & Guasch, L., 2016. Adaptive waveform inversion: theory, *Geophysics,* **81**(6), R429–R445.

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., Hunter, J.A., Harris, J.B. & Ivanov, J., 2002. Comparing Shear-wave velocity profiles inverted from multichannel surface wave with borehole measurements. *Soil Dyn. Earthq. Eng.,* **22**(3), P181–P190.

Yuan, Y.O, Simsons, F.J. & Bozdag, E., 2015. Multiscale adjoint waveform tomography for surface and body waves, *Geophysics,* **80**(5), R281–R302.

APPENDIX: UNCERTAINTY ANALYSIS WITHIN THE PROPOSED WORKFLOW

Starting from the uncertainties of the estimated surface wave DCs we compute the uncertainties of the estimated VS, VP and Poisson ratio profiles obtained with the proposed workflow. First, the uncertainties of the estimated reference W/D relationships and apparent Poisson ratios for the two reference DCs (outside and inside the sand body) are computed. Then, we propagate the uncertainty to the estimated models at the position of other DCs.

The experimental uncertainties of the DCs can be assumed following Passeri (2019) who studied the experimental uncertainties distribution of DCs from a wide data set (52 different sites). He observed a frequency dependence of the phase velocity standard deviation according to the following:

$$\sigma_{\mathrm{VR}} = \left[0.2822\, e^{-0.1819\, f} + 0.0226\, e^{0.0077\, f}\right] * \mathrm{VR}, \tag{A1}$$

where f is the frequency, VR is the phase velocity at f and σ_{VR} is the standard deviation of the phase velocity. Alternatively, in the presence of enough pool of data from the recordings, the standard deviation of the DCs can be estimated directly from the experimental data and the following uncertainty propagation within the workflow can be applied to these last uncertainties.

Considering f as a function of g and h, and assuming no correlation between the error of g and h, the variance of g and h can be propagated to $f(g,h)$ as

$$\sigma_f^2 = \left(\frac{\partial f}{\partial g}.\sigma_g\right)^2 + \left(\frac{\partial f}{\partial h}.\sigma_h\right)^2, \tag{A2}$$

where σ_f, σ_g and σ_h are the standard deviation of f, g and h, respectively. Considering that wavelength is given by phase velocity divided by frequency and assuming negligible error for frequency, using eq. (A2), we compute the uncertainties of surface wave's wavelength as

$$\sigma_\lambda = \frac{\sigma_{\mathrm{VR}}}{f}, \tag{A3}$$

where f is the frequency, σ_λ is the standard deviation of wavelength and σ_{VR} is the standard deviation of phase velocity. This last can be computed from a given frequency using eq. (A1). In Figs A1(a) and (b), in red, we show the reference DC number 11 (outside the sand body) and DC number 8 (inside the sand body), in terms of wavelength. The horizontal error bars of the DCs show the standard deviation of the phase velocity obtained using eq. (A1) and the vertical error-bars show the standard deviation of the wavelength estimated using eq. (A3). The next step is to estimate VSZ of the reference models by applying an MCI to the reference DCs. The result of the MCI is an ensemble of accepted VSZ models according to the implemented statistical test. The estimated reference VSZ is computed by averaging these accepted models at each depth. Evidently, the MCI of the reference DCs includes non-uniqueness which should be considered in the uncertainty analysis. The velocity standard deviation (σ_{VSZ}) of the reference VSZs is estimated based on the velocity distribution of the accepted models at each depth. Considering depth given by VSZ multiplied by one-way traveltime, and assuming negligible error for time measurement, using eq. (A2) the depth standard deviation of the reference VSZ models can be obtained through

$$\sigma_z = \sigma_{\mathrm{VSZ}}.t_{\text{one-way}}, \tag{A4}$$

where σ_z is the depth standard deviation, $t_{\text{one-way}}$ is the one-way arrival time and σ_{VSZ} is the velocity standard deviation of the estimated reference VSZ. The one-way traveltime at each depth is computed by dividing depth by the estimated reference VSZ model. In Figs A1(a) and (b), in black, we show the reference VSZ for outside and inside the sand body, respectively. The horizontal error bars show the velocity standard deviation for VSZ and the vertical error bars present the depth standard deviation.

The computed depth and wavelength uncertainties of the reference models are directly translated to the reference W/D relationships estimated using the Socco *et al.* (2017) method. In Figs A2(a) and (b), in red, we show the reference W/D relationships for outside and inside the sand body, respectively with vertical error bars showing wavelength standard deviations and horizontal error bars showing depth standard deviations.

The next step is to calculate the apparent Poisson ratio of the reference models using the Socco & Comina (2017) method. As a result, we are now considering the uncertainty associated to the reference apparent Poisson ratios. Assuming normal Gaussian distribution of error at each depth-wavelength point and considering the estimated W/D relationships (Fig. A2) as the centre of the distribution, we randomly sample 10 000 points corresponding to each

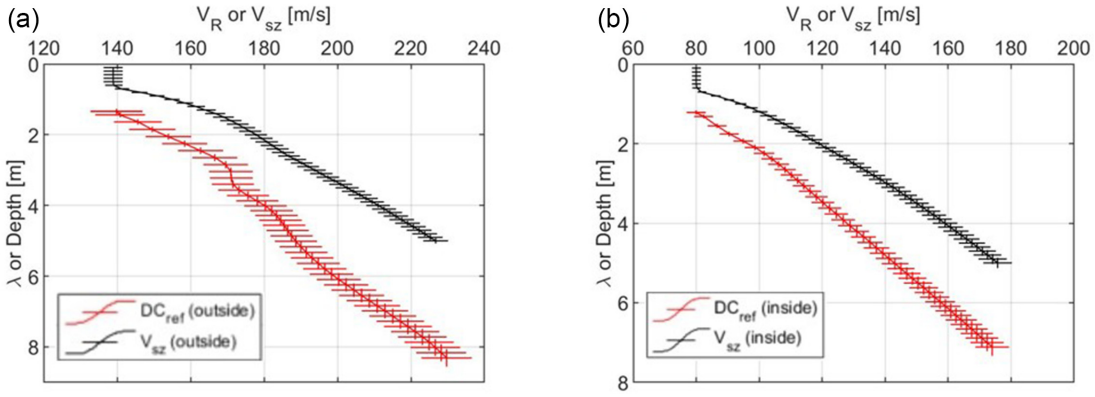

Figure A1. Uncertainty of the reference DCs and VSZs with the standard deviation error bars: (a) the reference DC number 11 and VSZ for outside the sand body and (b) the reference DC number 8 and VSZ for inside the sand body.

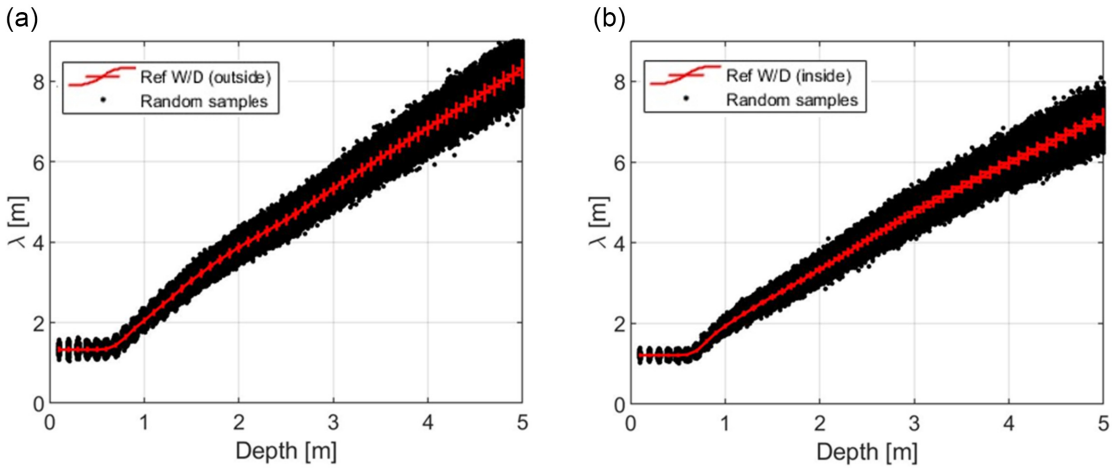

Figure A2. The reference W/D relationships, in red, with the standard deviation error bars for depth and wavelength; in black, the random samples corresponding to the W/D relationships and standard deviations: (a) for outside the sand body and (b) for inside the sand body.

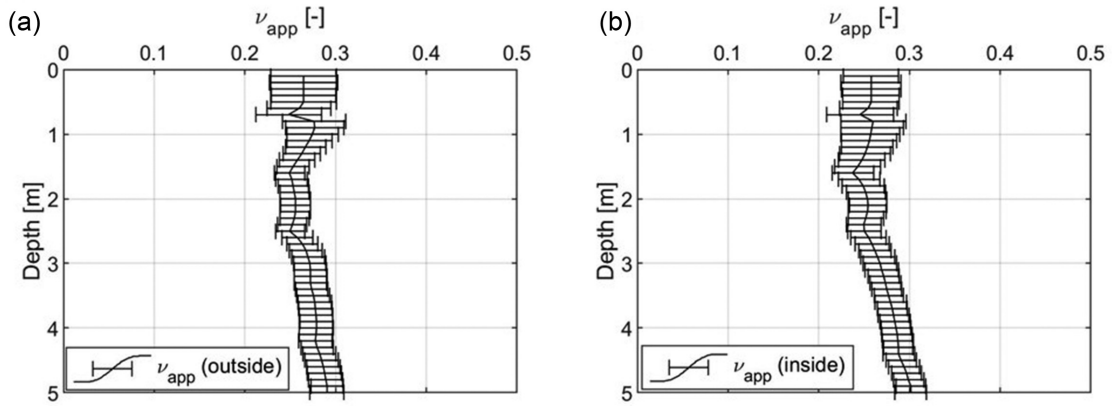

Figure A3. The estimated reference apparent Poisson ratio profiles with the standard deviation error bars: (a) for outside the sand body and (b) for inside the sand body.

depth–wavelength point. In Figs A2(a) and (b), we show the random samples of the W/D relationships for outside and inside the sand body in black. Applying the Socco & Comina (2017) method, from all these randomly sampled points, we obtain a distribution of apparent Poisson ratio values at each depth. This allows us to estimate the standard deviation for the estimated reference apparent Poisson ratio profile at each depth. In Figs A3(a) and (b), we show the estimated reference apparent Poisson ratio profiles with standard deviation error-bars for outside and inside the sand body, respectively.

We now have to propagate the uncertainty estimated at the reference model to the other models of the cluster, which are obtained by direct transform of the DCs through the reference W/D relationship and the reference apparent Poisson ratio. Hence, we have

to consider the uncertainty of the reference W/D relationships and apparent Poisson ratios and also the uncertainty of the local DCs. As an example, in the following, we consider DC number 3 and DC number 6 which belong to the same cluster of the reference DCs in Figs A1(a) located outside the sand body and Fig A1(b) located inside the sand body.

From the same random samples of the reference W/D relationship for outside the sand body (Fig. A2a) and applying the direct transformation method (Socco *et al.* 2017) to DC number 3, we get a distribution of the estimated VSZ at each depth. Based on the distribution of the estimated VSZ, we estimate the standard deviation of VSZ at each depth. In addition, there is the uncertainty of DC number 3 (from eq. A1) that needs to be considered for the uncertainty of the estimated VSZ model. Using the reference W/D relationship for outside the sand body (Fig. A2a), we transform the uncertainty of the velocities in wavelength to uncertainty of velocities in depth. We compute the final standard deviation of the estimated VSZ model at the position of DC number 3 as the summation of the standard deviation from the proposed workflow and the standard deviation from the experimental uncertainty of the DC. In Fig. A4(a), in black, we show the estimated VSZ model and its standard deviation at the position of the DC number 3. Similarly, in Fig. A4(b), we show the estimated VSZ, with its corresponding standard deviation, at the position of DC number 6, using the reference randomly sampled W/D relationship for inside the sand body (Fig. A2b) and the experimental uncertainty of DC number 6.

We now consider the estimation of the VPZ at the position of DCs numbers 3 and 6. Considering the relationship between VPZ, VSZ and reference apparent Poisson ratio and following eq. (A2), we propagate the uncertainty associated to the estimated VSZ and the reference apparent Poisson ratio to the estimated VPZ as

$$\sigma_{\mathrm{vpz}}^2 = \left(\sqrt{\frac{2\upsilon_{\mathrm{app}}-2}{2\upsilon_{\mathrm{app}}-1}}.\sigma_{\mathrm{vsz}}\right)^2 + \left(\mathrm{VSZ}.\ \frac{1}{\left(2\upsilon_{\mathrm{app}}-1\right)^2\sqrt{\frac{2\upsilon_{\mathrm{app}}-2}{2\upsilon_{\mathrm{app}}-1}}}\sigma_{\upsilon_{\mathrm{app}}}\right)^2, \tag{A5}$$

where σ_{vpz}, σ_{vsz} and $\sigma_{\upsilon_{\mathrm{app}}}$ are the standard deviations of VPZ, VSZ and apparent Poisson ratio, respectively. Applying eq. (A5) to the estimated VSZs and reference apparent Poisson ratios (Figs A3 and A4), we estimate the standard deviation of the estimated VPZs. In Figs A4(a) and (b), in blue, we show the estimated VPZs and standard deviations as error-bars at the position of DCs number 3 and 6, respectively.

The interval VS and VP are estimated using the relationship between interval velocity and time-average velocity and performing total variation regularizations. We are now considering the uncertainties of the interval VS and VP. To estimate the uncertainties associated to interval velocities, assuming normal Gaussian distribution for time-average velocity estimates in Fig. (A4), we randomly sample 10 000 points at each depth of time-average velocity models. Using the proposed workflow, we transform time-average velocities into interval velocities. Then, we use the distribution of the estimated interval velocity at each depth to estimate the standard deviations of the interval velocities. In Figs A5(a) and (b), we show the estimated interval velocities (VS and VP) and the associated standard deviations at the position of the DCs numbers 3 and 6, respectively.

Considering the relationship between the interval VS, interval VP and Poisson ratio and using eq. (A2), we propagate the standard deviation of the interval velocities to Poisson ratio as

$$\sigma_{\upsilon}^2 = \frac{V_p^2 V_s^2}{\left(V_p^2 - V_s^2\right)^4}\left[\, V_p^2\sigma_{\mathrm{VS}}^2 + V_s^2\sigma_{\mathrm{VP}}^2\right], \tag{A6}$$

where σ_{VS}, σ_{VP} and σ_{υ} are the standard deviations of VS, VP and Poisson ratio, respectively. In Figs A6(a) and (b), we show the estimated Poisson ratio with the standard deviation error-bars at the position of the DC number 3 and DC number 6, respectively.

The above reported uncertainty propagation approach can be repeated for the estimated models at the position of all the DCs along the investigated line. Then, the obtained uncertainties are linearly interpolated to create a pseudo 2-D view of the uncertainty corresponding to the estimated VS, VP and Poisson ratio. The uncertainties normalization for VS is obtained using

$$\sigma_{\mathrm{norm,\ VS}} = \frac{\sigma_{\mathrm{VS}}}{\mathrm{VS}}, \tag{A7}$$

where, the $\sigma_{\mathrm{norm,\ VS}}$ is the normalized uncertainty of the VS. Following the same normalization approach, we normalize also the standard deviation of the estimated VP and Poisson ratio.

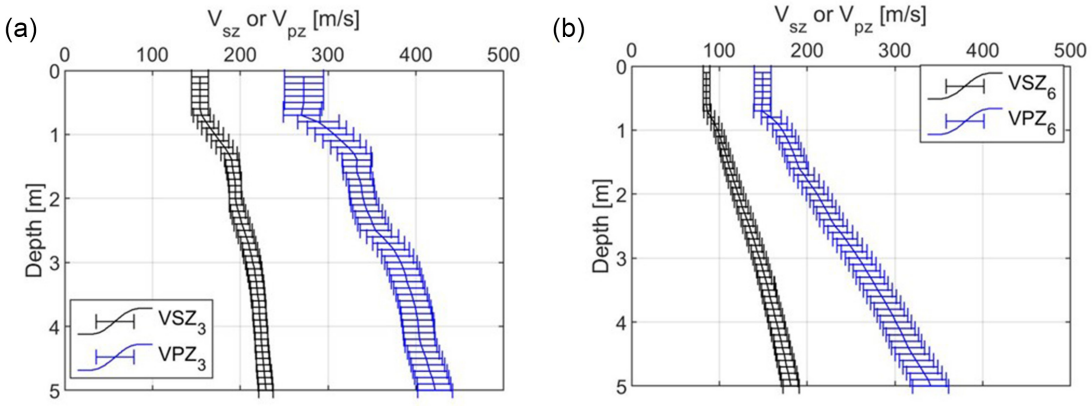

Figure A4. The estimated VSZ and VPZwith the standard deviation error bars: (a) at the position of DC 3 (outside the sand body); (b) at the position of DC 6 (inside the sand body).

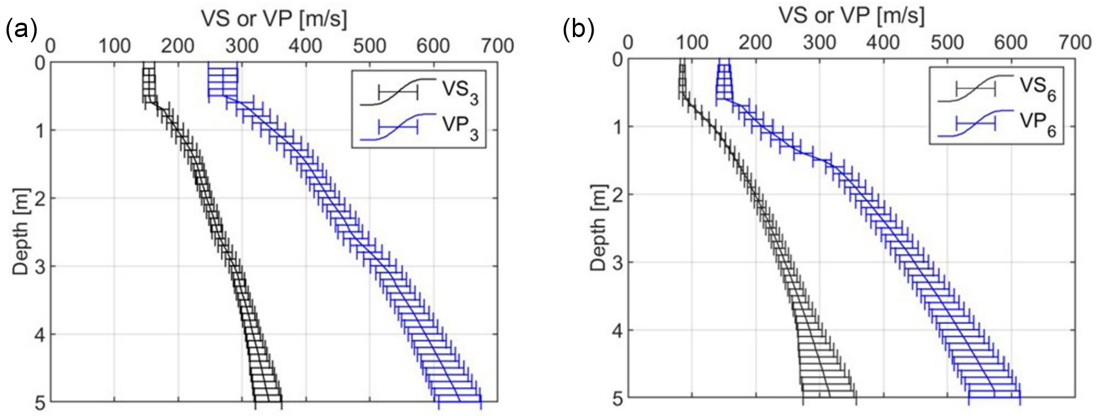

Figure A5. The estimated VP and VS with the standard deviation error bars: (a) at the position of DC number 3 (outside the sand body); (b) at the position of DC number 6 (inside the sand body).

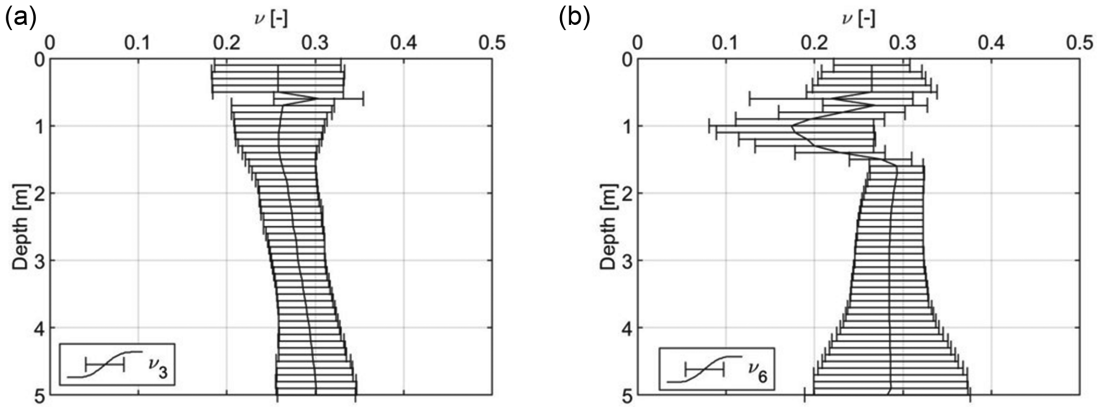

Figure A6. The estimated Poisson ratio with the standard deviation error bars: (a) at the position of DC 3 (outside the sand body); (b) at the position of DC 6 (inside the sand body).