S- and P-wave velocity model estimation from seismic surface-waves.

Farbod Khosro Anjom[1], Frank Adler[2], Laura Valentina Socco[1]

[1]Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Turin, 10128, Italy
[2] CSTJF, TotalEnergies, Pau, 64000, France

5   *Correspondence to*: Farbod Khosro Anjom (Khosro-Anjom.Farbod@polito.it)

Abstract. The acquisition of seismic exploration data in remote locations presents several logistic and economic criticalities. Carpet layouts, with irregular distribution of sources and/or receivers, facilitate seismic acquisition operations in hard-to-access areas. A very convenient approach is to deploy~, nodal receivers on a regular grid and to locate sources only in accessible locations, creating irregular source-receiver layout. Processing workflows require then to be updated to properly

10   handle irregular acquisition geometries. Surface-wave analysis, T~that is standardly applied to retrieve near-surface velocity models, requires fine-tuning to be successfully applied to the data sets from carpet recording. Here, we applied three surface-wave techniques, namely wavelength/depth method, laterally constrained inversion, and surface-wave tomography to a 3-D large-scale test data set, which was acquired in a hard-rock site using carpet recording approach. The wavelength-depth is a data transformation method that is based on a relationship between skin-depth and surface-wave wavelength and provides

15   both S- and P-wave velocity (V_s and V_p) models. We ~obtainedthen~ used the Poisson's ratios estimated through the wavelength-depth method to constrain the laterally constrained inversion and surface-wave tomography, and retrieve both V_s and V_p also from these methods. The pseudo 3-D V_s and V_p ~models were obtained down to 140 m depth over approximately 900×1500 m^2 area. ~Recently, the sensitivity of surface-wave skin depth to Poisson ratio was applied to also estimate P-wave velocity (VP) models from surface wave records. We use this technique within the framework of three surface-wave

20   methods, the wavelength/depth data transform, the laterally constrained inversion, and surface-wave tomography to estimate both VS and VP models. We apply these methods to a 3-D test data set from a mining site that is characterized by stiff material and by significant elevation contrast. The data were recorded using a regular grid of receivers and an irregular source layout. Pseudo 3D VS and VP models were obtained down to 140 m depth over approximately 900×1500 m^2 area.~ The estimated models from the methods ~well~ match the geological information available for the site. Less than 6%

25   difference ~is~ was observed between the estimated V_s ~VS~ models from the three methods, whereas this value ~is~ was 7.1% for the retrieved V_p ~VP~ models. The ~different~ methods ~are~ were critically compared in terms of resolution and efficiency.

1 Introduction

To address the challenges of seismic data acquisition in remote areas, such as foothills and forests, we can apply the so-called carpet recording, in which the nodal receivers are deployed on a regular grid and the source locations are only limited

30 to reachable locations such as the access roads (Lys et al., 2018). The carpet recording approach creates an irregular source-receiver outline that raises the necessity to evaluate, verify and test the seismic processing workflow. Here, we focus on the application of surface-wave methods to the data recorded through a carpet recording scheme with the purpose of near-surface velocity model estimations. These velocity models can be used for engineering purposes or as input in the exploration processing workflow to improve static corrections and groundroll removal.

35 Surface-wave methods are powerful tools for subsurface characterization. Most of these methods process the data to extract the surface-wave dispersion curve (DC) from seismic records and invert these DCs to estimate the velocity models. Since the energy decay of surface-wave~~s~~ wavefield in depth depends on their wavelength, the investigation depth of ~~the~~ surface-wave methods is related to the maximum recovered wavelength and can be considerably variable, ranging from a few ~~metres~~meters (e.g., Xia et al., 2002; Feng et al., 2004; Comina et al., 2011; Pan et al., 2018) to several tens of ~~metres~~meters

40 (e.g., Mordret et al., 2014; Da col et al., 2019) or even to a few ~~kilometres~~kilometers (e.g., Ritzwoller and Levshin, 1998; Kennet and Yoshizawa, 2002; Fang et al., 2015). The estimated models from surface-wave techniques can be used in many applications such as near-surface site characterization (Lai, 1998;~~;~~ Xia, 2014;Foti et al., 2015~~; Garofalo et al., 2016~~), static corrections (Mari, 1984; Roy et al., 2010~~; Papadopoulou et al., 2020~~), and ground roll prediction and damping (Blonk and Herman, 1994; Ernst et al., 2002; Halliday et al., 2010).

45 Different methods can be adopted for extracting DCs from the seismic records and inverting them (see for instance ~~and invert them.~~ Papadopoulou (2021) ~~provided~~ for a thorough review of the different processing techniques and characteristics of the estimated DCs). The retrieval of a velocity model from the DC can be based on simple data transformations or on model optimization approaches with different inversion strategies. According to the chosen workflow the computational cost and model resolution may vary and identifying the optimal approach for the analysis is an important task. Here, we compare

50 three different methods (wavelength/depth data transform, laterally constrained inversion, and surface-wave tomography) and apply them ~~on~~to a ~~challenging~~ large-scale test data set acquired with carpet recording method.

Regardless of the type of the surface-wave technique, since DCs are known to have lower sensitivity to P-wave velocity (V_p ~~VP~~) compared to S-wave velocity (V_s ~~VS~~), most surface-wave methods focus on V_s ~~VS~~ estimation, and they require *a priori* Poisson's ratio or V_p ~~VP~~ for the inversion stage. Recently, a data transformation method based on the so-called wavelength-

55 depth (W/D) relationship was developed to estimate both V_s ~~VS~~ and V_p ~~VP~~ models (Khosro Anjom et al., 2019). The W/D relationship is obtained by computing wavelength - depth couples corresponding to equal phase-velocity of surface-waves and time-average V_s ~~VS~~ and represents the skin depth of ~~SW~~ surface-waves (Socco et al., (2017). Once estimated, the W/D relationship can be used to directly estimate the time-average V_s ~~VS velocity~~ from DCs. Socco and Comina (2017) showed with synthetic tests and tests on real data~~showed with synthetic and real tests~~ that the W/D relationship is highly sensitive to

60 Poisson's ratio, and it can be used to estimate time-average V_p ~~VP~~. Khosro Anjom et al. (2019) developed a data-driven W/D workflow that directly estimates interval V_s ~~VS~~ and V_p ~~VP~~ models from the DCs and is valid for sites with significant lateral

variations. The method also provides the Poisson's ratio, which can be ~~considered~~ used *a priori* in other surface-wave methods, such as laterally constrained inversion (LCI) and surface-wave tomography (SWT).

The earliest applications of LCI were on resistivity data (Auken and Christiensen, 2004; Wisén et al., 2005; Auken et al., 2005). The first successful application of the LCI to surface-waves was performed by Wisén and Christiansen (2005). In LCI, several multi-channel DCs available along a line or over an area are associated with local relevant 1D models and inverted simultaneously. The parameters of the 1D models are connected laterally and vertically through a set of constraints, whose strength controls the variations between model parameters in adjacent model points (Boiero and Socco 2010). As a result, consistent and smooth estimated pseudo 2-D or 3-D models are usually obtained from the LCI applications.

In the context of earthquake seismology, SWT is a well-stablished method for V_s ~~VS~~ reconstruction of the crust and upper ~~mantel~~ mantle (Wespestad et al., 2019; Bao et al., 2015; Boiero, 2009; Yao et al., 2008; Shapiro et al., 2005). Recently, a few authors showed the application of SWT for ~~the~~ near-surface characterization, using active (Da Col et al., 2019; Socco et al., 2014; Khosro Anjom et al., 2021) and passive data (Badal et al., 2013; Picozzi et al., 2009; Colombero et al., 2022).

In the literature, DC estimation methods are usually categorized into multi-channel and two-station methods, even though there are no theoretical and significant technical differences between the two approaches (Papadopoulou, 2021). The multi-channel technique is the most common approach, in which the recordings from an array of receivers (in 2-D scheme) go through a wavefield transform (e.g., *f-k*, *f-v* and *τ-p*) and the DC is picked on the spectrum as the local maxima within the frequency band of surface-waves. The multi-channel processing stage is repeated to estimate DCs at different locations, which are then inverted individually (e.g., fast simulated annealing of Beaty et al., 2002) or simultaneously (e.g., LCI of Socco et al., 2009) to estimate the S wave velocity (V_s ~~VS)~~ model. For SWT, DC are estimated using the two-station processing method, in which ~~the~~ ~~the recordings of~~ receiver couples aligned with sources are considered to estimate many path-averaged DCs that are later inverted using a tomographic scheme to estimate directly (Boschi and Ekstrom, 2002; Fang et al., 2015; Boiero, 2009; Socco et al., 2014; Karimpour et al., 2022) or indirectly (Yoshizawa and Kennett, 2004; Shapiro and Ritzwoller, 2002; Yao et al., 2008) the V_s ~~VS~~ model.

Here, we show the application of the three surface-wave methods, (W/D, LCI, and SWT), to estimate both V_s ~~VS~~ and V_p ~~VP~~ models. The first two methods are based on multichannel DCs, whereas the latter relies on the DCs from two-station ~~ones~~ method. We apply the methods to a challenging test data set that was acquired in a hard-rock ~~environment~~ site using carpet recording technique. ~~The data were recorded using regular grids of receivers and sources located along roads and accessible locations.~~ ~~Th~~ ~~is acquisition scheme, also known as~~ e irregular source-receiver outline of carpet recording (~~Lys et al., 2018), creates irregular source receiver layout that~~ limits the use of conventional multi-channel processing methods. Conversely, the irregular combination of source-receiver ~~geometry allows~~ is favorable for the estimation of many two-station DCs with different azimuthal angles, providing high data coverage and uniform azimuthal distribution of the paths that is ~~in~~ to SWT's advantage ~~of SWT~~.

In this paper, we first introduce the site and ~~outline~~ describe the acquired data ~~acquisition~~. Then, we explain the multi-channel and two-station DC estimation processing techniques. Then, we briefly describe the W/D, LCI and SWT ~~techniques~~ velocity

model estimation methods and show their application to the data set. We use the W/D method to estimate the *a priori* Poisson's ratio required by LCI and SWT methods, which we then employ to transform their V_s VS results into V_p VP. Finally, we compare the estimated models and the obtained resolution from the application of each method and compare the methods from an efficiency point of view.

2 Site description and field data set

The data were acquiredlocation is in a limestone mining sitequarry in the province of Aurignac in the south of France (Fig. 1a). In Fig. 1b, we show the satellite view of the site superimposed with the elevation map of the area. From north-west to south-east, a significant natural (outside the pits) and human-made elevation (inside the pits) contrast is present, which can cause highly scattered surface-waves. In Fig. 1c, we show the geological map of the area from the website of French geological survey (BRGM). The eastern central, eastern, and northern part of the site are characterized by stiff formations belonging to formationsThanetian and Sparnecian stages, mainly primarily composed of of stiff limestone and marl. In the western zone, recent loose deposits are present (Ypresian)t, creating a significant lateral variation between the east and west portions of the site. The very dense limestone with dolomite layers from Danian stage is outcropping in the north, outside of the investigated area, whichand is expected to be reached in shallow subsurface in the investigated zone.

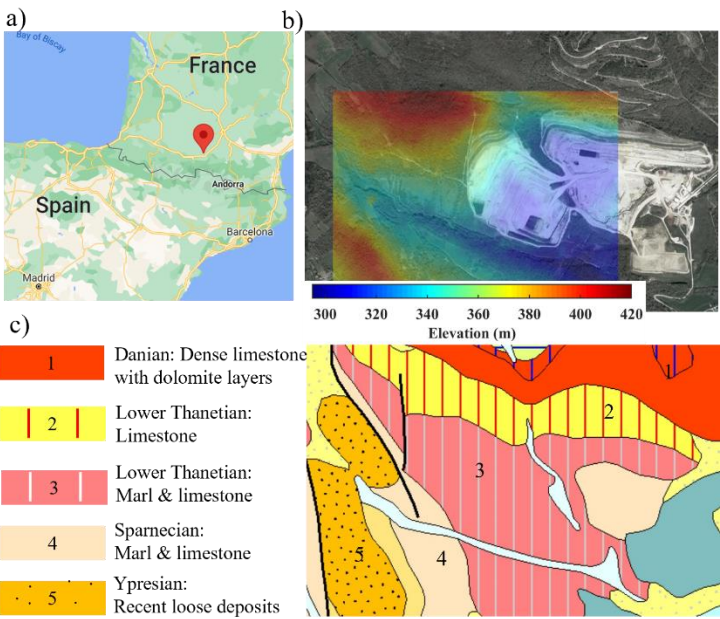



Figure 1: (a) Google map showing the location of the site (© Google Map). (b) Satellite view of the site obtained superimposed with the elevation map of the area (© Google Earth). (c) The geological map of the area obtained from French geological Survey (© BRGM - www.infoterre.brgm.fr).

3 Experimental data

The seismic campaign acquisition was performedconducted inside and outside the two active open mining pits to test the carpet recording technique in a hard rock site and provide an exploration data set to be used for testing different processing approaches. 918 receivers were deployed within on a regular grids (area of 1.7 km × 1 km), and several source points (1077) were considered along the access roads, resulting in a 3-D large-scale acquisition layout. Two Birdwagen Mark IV off-road equipped with 24-ton vibrator were used as the source. The vibration included a sweep of 24 s (3 to 110 Hz) with 5 second listening time. The data were collected in real-time using RT2the RT2 wireless system. In this study, we considered a portion of the data that were collected outside the mining pits. The full description of the acquisition parameters corresponding to this portion of the data is given in Table 1.

Table 1: Acquisition parameters of the data set for outside the mining pits.

Receivers	Sources	Number of receivers	Number of shots	Sampling rate (ms)	Recording time window (s)
5 Hz vertical geophones	Vibroseis truck	581 (spacing 25 and 50 m)	533 (irregular layout)	2	5

To minimize the effect of elevation contrasts (Fig. 1b), we split the data into two sub data sets (north and south), each corresponding to an area with relatively flat topography. In Fig. 2, we show the acquisition layout, where different colourscolors are used for each sub data set.

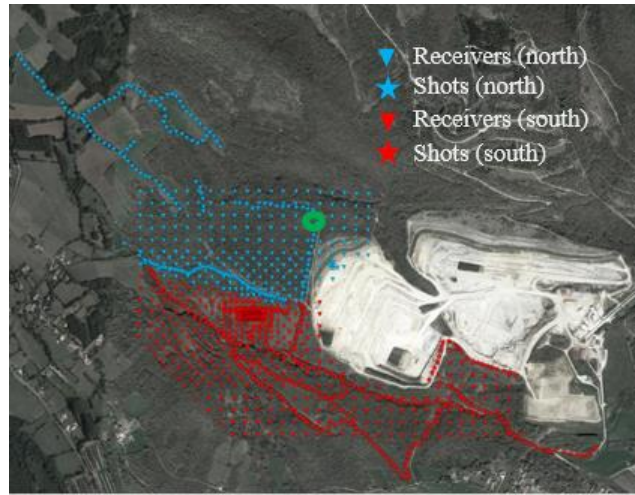



Figure 2: The satellite view of the Aurignac site (© Google Earth) superimposed with the acquisition layout. The data divided into two sub data sets shown with different ~~colours~~colors, each within relatively flat area. The recordings from the highlighted shot (green circle) are plotted in Fig. 3.

In Fig. 3, we show the first 2 s of the recordings from the highlighted source in Fig. 2 in offset domain; only 20 % of the traces are shown for better visualization. The data exhibits ~~low~~a low signal-to-noise ratio as expected for hard rock sites.

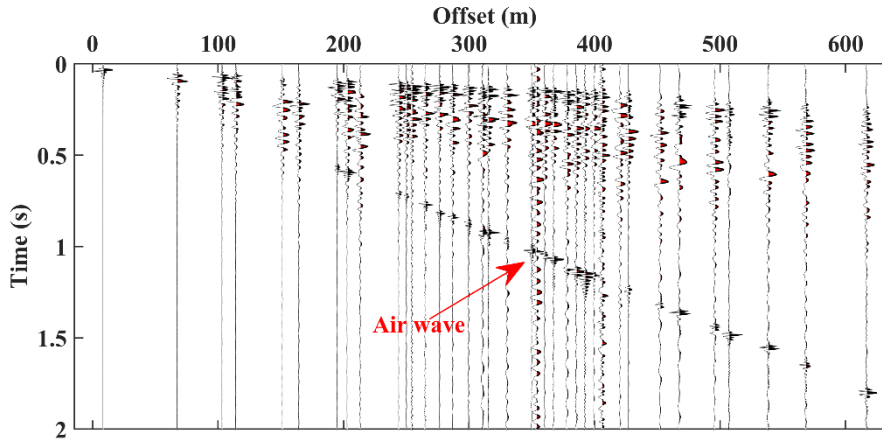



Figure 3: An example seismogram from the north zone of the Aurignac site. The shot location is highlighted with a green circle in Fig. 2.

~~34 W/D data transfo~~Surface-wave processing of 3-D data~~rm~~

3.1 Multi-channel DCs

Multi-channel dispersion analysis is usually performed by selecting recordings from multiple inline receivers with a source and performing domain transform (i.e., $f - v$, $f - k$, $\tau - p$, etc.). Nevertheless, this approach can lead to inconsistencies in the estimated DCs when the technique is applied to 3-D data sets with irregular source-receiver geometry at sites with significant lateral variations. In Fig. 4, we show two examples of DC estimation from the field data set for the same location

using recordings from two linear receiver arrays. The estimated DCs for the location shown as green × in Fig. 4b are on average more than 15% different. The main reason for this inconsistency is the impact of lateral variations and the entirely
150   different surface-wave propagation path along the two arrays.

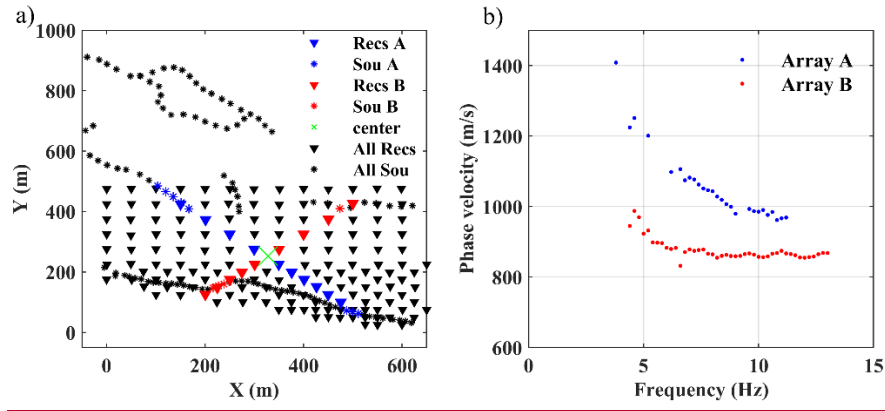



Figure 4: Two example DC estimations for the same location using two different receiver arrays. (a) The two receiver arrays and aligned shots. The middle of both arrays is coincident and shown with the green ×. (b) The estimated DCs for the location of the green × in a using the two arrays' records.~~4.1 Multi-channel processing of 3-D data~~

155

To minimize the impact of lateral variability on the DC estimation of the 3-D data, we consider~~ed~~ the recordings from receiver spread over an area (Wang et al., 2015; Xia et al., 2009; Park, 2019). For each DC estimation, we select~~ed~~ the receivers inside a square area (window) ~~of 100 × 100 m²~~ and consider~~ed~~ the ~~recordings from~~ sources within a ~~250 m~~certain
160   distance from the ~~centre~~center of the square ~~(Fig. 4a)~~. We us~~ed~~e the phase-shift method (Park et al., 1998) to estimate the *f-v* spectrum. We stack~~ed~~ the spectra corresponding to the recordings of the same receivers but from different source locations to increase the signal-to-noise ratio (Neducza, 2007). Then, the fundamental mode ~~was~~ is picked as the maxima on the spectrum and ~~was~~ is assigned to the ~~centre~~center of the receiver spread ~~(Fig. 4b)~~. We ~~shifted~~slide the window by one inter receiver spacing ~~(50 m in the northern side and 25 m in the southern part)~~ to estimate DCs corresponding to different
165   locations of the site.

For the field data, we considered a window of 100×100 m² to select the receiver and sources within 250 m of the center of the square. In Fig. 5~~4~~, we show an example ~~of the~~ DC picking for the north zone. In Fig. 5~~4~~a, the 9 selected receivers inside the square are shown in blue, and the selected shots are ~~presented~~ plotted in green asterisks. In Fig. 5~~4~~b, we show the computed spectrum and the picked DC. Overall, we estimated 545 DCs for northern and southern zones shown in Fig. 5c.

170

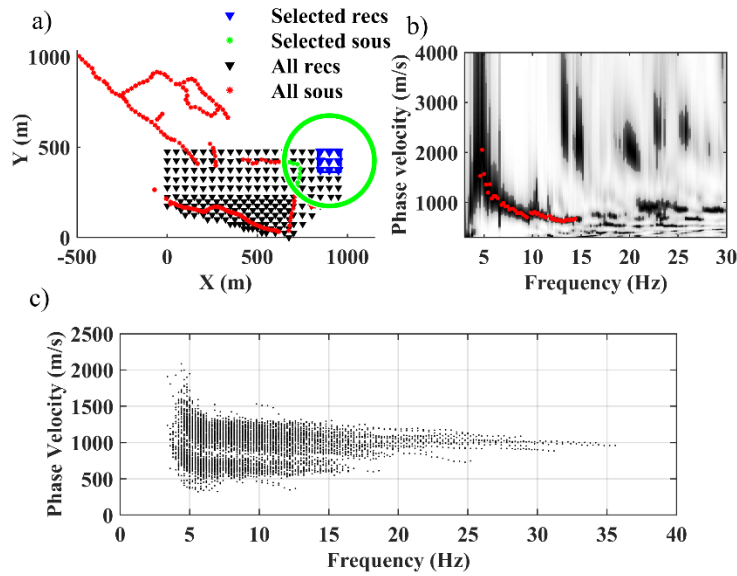



Figure 54: Multi-channel DC estimation from the field data. (a) ~~Example DC picking from the north zone of the site by applying phase-shift method to the recordings of receiver over the square window. (a)~~ An example geometry of the selected receivers and sources for DC estimation in the northern zone. (b) The computed spectrum and estimated DC from the recordings of the receivers shown in (a). (c) All the estimated DCs for the north and south zones. ~~The geometry of the selected receivers inside the blue rectangle. The green circle shows the distance limit of the shot location (250 m) from the centre of the selected receivers. (b) Computed spectrum and picked DC in red.~~

3.2 Two-station DCs

The two-station DCs are estimated ~~using~~applying interferometry ~~between~~to the recordings ~~from receiver~~from receiver couples aligned with a source, assuming straight ray approximation ~~using interferometry~~. We use the algorithm developed by Da Col (2019) and modified by Khosro Anjom et al. (2021). First, an automatic search is performed to find the receiver couples aligned with the source at each azimuth angle, considering 1° tolerance for the deviation from a ~~linear~~straight path. Given the scale of the site, ~~the aligned receivers are along~~we consider the propagation path as occurring over a plain area, and we neglect ~~the great-circle approximation. Then, the traces are narrow band filtered at various frequencies, using zero-phase Gaussian filters and ~~. Next,~~ the filtered traces of the receiver couples are cross-correlated and assembled to form the cross-multiplication matrix. We use a 3rd order spline interpolator to convert the cross-multiplication matrix to the frequency-velocity domain. We stack the cross-multiplication matrices computed from the records of the same two stations but different sources, to increase the signal-to-noise ratio. Finally, at each frequency, the phase-velocity is picked as the maximum of the cross-multiplication matrix. To avoid cycle skipping, we use, as reference, the closest multichannel DC, and we automatically pick maxima closest to the reference DCs. To minimize the contamination of the fundamental mode by higher surface-wave modes, we damp the higher mode data using the muting strategy of Khosro Anjom et al. (2019).

We applied the ~~SWT~~two-station DC estimation method ~~only~~to the data from the north of the site only (blue markers in Fig. 2). We performed an automatic search of the receiver couples aligned with sources within 250 m offset, which resulted in

4710 possible receiver couples and source settings. We used the local DCs from the multi-channel analysis (Fig. 5ca) as references to locate the correct trend of the path-averaged fundamental mode DCs. We discarded noisy or in inconsistent cross-multiplication matrices and in total, 1301 path-averaged DCs were estimated. In Fig. 126a and b, we show the estimated path-averaged DCs and the observed azimuthal illumination. The data show uniform coverage with most paths showing angles between 0° to 40° and 140° to 180° angles. The uniform coverage mitigates the directionality of the tomographic inversion toward dominant directions (Khosro Anjom et al., 2021).

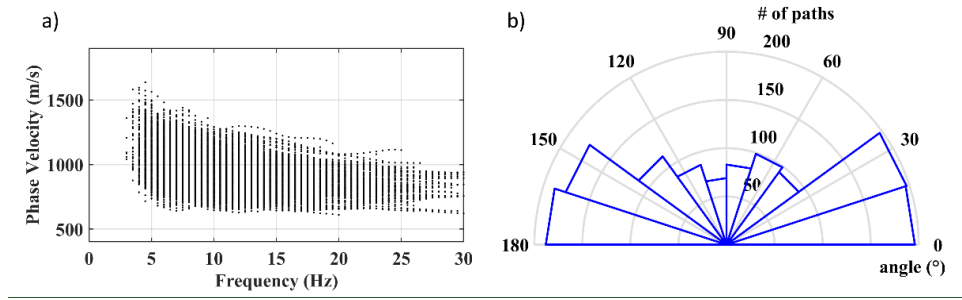



Figure 126: (a) The estimated path-averaged DCs corresponding to northern part of the data setthe site. (b) The obtained azimuthal illumination with the numbers around the great circle showing the angles and other circles representing the obtained coverage.

In Fig. 137a to f, we show the data coverage within different wavelength ranges, where the colourcolor scale shows the path-average phase-velocity. The data exhibit very high coverage for wavelengths between 40 to 220 m, beyond which it decreases substantially.

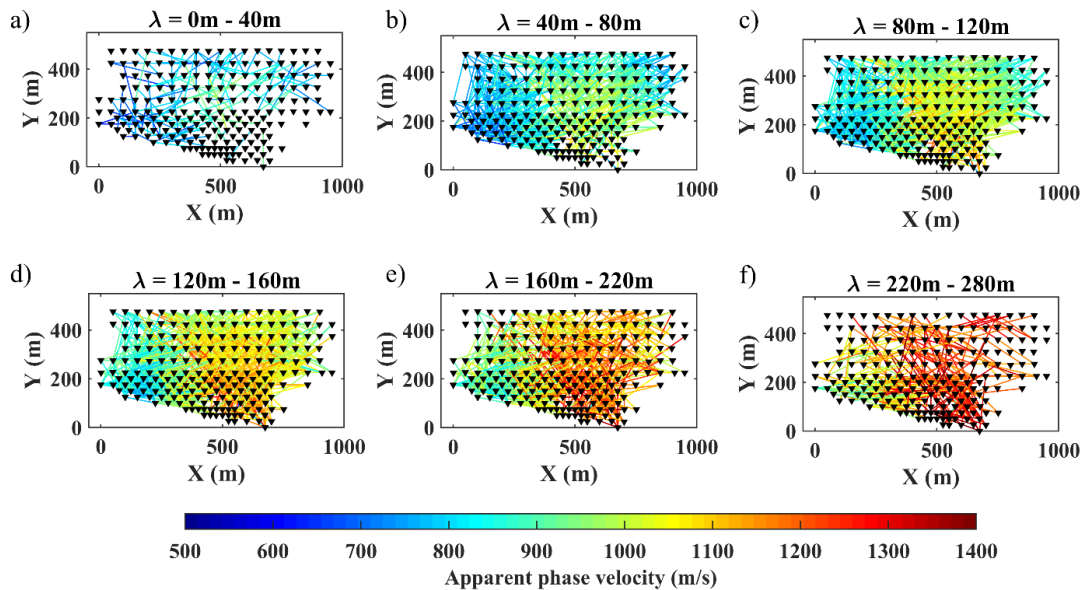



Figure 137: Pseudo-slices of the estimated path-averaged DCs from the north of the site shown within the wavelength ranges: (a) 0 to 40 m. (b) 40 to 80 m. (c) 80 to 120 m. (d) 120 to 160 m. (e) 160 to 220 m. (f) 220 to 280 m.

210 **4.~~1~~4.2 W/D ~~data transformation~~data transform**

The only inputs of the W/D method ~~is~~are the estimated multi-channel DCs. The ~~proposed~~ method, as described in Khosro Anjom et al. (2021), is composed of four main steps: (i) The clustering of DCs, (ii) the selection of a reference DC for each cluster and the estimation of the corresponding time-average V_s ~~VS~~ velocity model, (iii) the W/D relationship and apparent Poisson's ratio estimation for the reference DC of each cluster, and (iv) the direct transformation of the DCs into time-

215 average and then interval V_s ~~VS~~ and V_p ~~VP~~ model~~s~~.

Since the same W/D relationship is applied to different DCs to transform them into velocity models, to apply the method at sites with significant lateral variations, the DCs must be clustered into more homogenous sets, and one W/D relationship should be estimated and applied separately to each ~~group~~ cluster of DCs. We use the hierarchical clustering algorithm developed by Khosro Anjom et al (2017) to cluster the DCs.

220 For each cluster, a reference DC and its corresponding time-average V_s ~~VS~~ model is needed to estimate the W/D relationship. The reference DC is selected based on the quality control proposed by Karimpour (2018). The ~~t~~Time-average velocity at a ~~considered~~given depth z is the weighted average velocity from the surface to the considered depth, for which the one-way time is equal to the one-way time of the layered velocity model to the same depth, and can be directly computed from the layered velocity model as:

$$V_{sz}(z) = \frac{\sum_n h_i}{\sum_n \dfrac{h_i}{V_{s,i}}}, \quad \text{~~} VS_z = \frac{\sum_n h_i}{\sum_n \frac{h_i}{VS_i}} \text{,w~~} \tag{1}$$

where h and n are the layers thickness and the ~~total~~ number of layers down to the depth of z. The ~~t~~Time-average velocity is ~~directly~~ used in many applications, ranging from static correction of the reflection data to seismic hazard analysis ~~using~~where the time average velocity down to 30m of depth, the so called V_{s30}, is used as a proxy for seismic response classification. ~~The reference DC is selected based on the quality control proposed by Karimpour (2018).~~ To estimate the

230 required time-average V_s ~~VS~~ model, we invert the reference DC using an optimized Monte Carlo inversion (Socco and Boiero, 2008). The density has a minor impact on the surface-waves velocity (Xia et al., 1999; Foti and ~~s~~Strobbia, 2002) and is ~~Among the model parameters, density is considered~~ defined *a priori* based on the geological information ~~from~~about the site.~~, whereas~~ The V_s ~~VS~~, the Poisson's ratio, and the thicknesses are ~~uniformly~~ randomly sampled within the boundaries of a wide uniform model space. Then, the synthetic DCs corresponding to each model ~~is~~are computed and compared with the

235 ~~real~~experimental DC. The algorithm ~~is~~uses ~~of~~ the scaling properties of DCs to create an optimal model space based on the initial model space: before computing the misfit with the experimental DC, the synthetic DCs of the random models are shifted as close as possible to the experimental DC; then, the scaling factor is obtained from the DC shift and is used to scale the models. These scaling steps, which are performed in a fully automatic manner, highly optimize the model space sampling

that is focused on low misfit regions and reduce the number of required simulations ~~to find the inversion results~~. Unlike deterministic inversions that result in a single output, the Monte Carlo inversion leads to a set of possible solutions. ~~After the inversion, the~~ The best~~accepted~~ fitting models are selected according to a ~~statistical~~ one-tailed Fisher test~~,~~ ~~with~~imposing a certain level of confidence. Then~~, Unlike deterministic inversions that result in a single output, the Monte Carlo inversion leads to a set of possible solutions. We first compute~~, the selected V_s models are transformed to time-average V_s ~~VS~~ models ~~from the VS layered models as~~ $VS_z = \frac{\sum_m h_i}{\sum_m \frac{h_i}{VS_i}}$, ~~where h and n are the layers thickness and total number of layers~~. ~~Then, we average t~~The values of the selected time-average V_s models are averaged ~~of the accepted model population~~ at each depth to obtain a unique time-average V_s ~~VS~~ model corresponding to the reference DC. Next, we ~~collect~~ estimate the ~~wavelength-depth relationship (~~W/D relationship~~)~~ that consists of the pairs of wavelength and depth values for which the phase velocity of the DC and the time-average V_s ~~VS~~ of the ~~MC~~ Monte Carlo solution have the same value. ~~The experimental W/D relationship is significantly sensitive to Poisson's ratio.~~ Comprehensive synthetic and real data analyses have been performed by Socco et al. (2017), Khosro Anjom et al. (2019), and Khosro Anjom (2021) to show the high sensitivity of the W/D relationship to Poisson's ratio. We use the method of Socco and Comina (2017) to estimate from the experimental W/D relationship an apparent Poisson's ratio that relates the time-average V_s ~~VS~~ and time-average V_p ~~VP~~: First, we generate synthetic DCs corresponding to the estimated V_s ~~VS~~ model from the Monte Carlo inversion and different Poisson's ratios. Then, we consider these DCs and the time-average V_s ~~VS~~ to retrieve synthetic W/D relationships that are each corresponding to a specific Poisson's ratio value. Next, we deduce an apparent Poisson's ratio at each depth by comparing the experimental W/D relationship with the synthetic ones.

We use the estimated experimental W/D relationship to directly transform all DCs of the cluster into time-average V_s ~~VS~~ models. Then, we transform the estimated time-average V_s ~~VS~~ models into time-average V_p ~~VP~~ through the estimated apparent Poisson's ratio. The time-average velocities can be transformed to interval velocities using a ~~DIX~~Dix-type equation. To mitigate the effect of noise in the data, we use the regularized ~~DIX~~Dix-type formulation proposed by Khosro Anjom (2019) to transform the time-average V_s ~~VS~~ and V_p ~~VP~~ models to interval V_s ~~VS~~ and v_p ~~VP~~. Finally, we assemble all the estimated models from the clusters to build a pseudo 2-D/3-D model. ~~Given the W/D method provides both VS and VP models, we use the estimated models at the reference location to retrieve a Poisson's ratio corresponding to the area covered by the cluster. The estimated Poisson's ratios will then be used *a priori* in the reference model of the LCI and SWT.~~

4.2 LCI

The method's inputs are the multi-channel DCs and the initial models at the location of the local DCs ~~obtained in section 4.1~~. ~~The~~An initial model defined as the thickness, density, Poisson's ratio, and V_s, is set for each cluster of DCs. The thickness and V_s are ~~is composed of the thickness, density, Poisson's ratio, and V_s VS~~of the 1D models. ~~The initial thickness and V_s are defined~~ based on the Monte Carlo inversions of reference multi-channel DCs in section 4.1. The Poisson's ratio is

270 selected based on the W/D analysis. ~~Given the W/D method provides both V_s and V_p models, we use the estimated models at the reference location to retrieve a Poisson's ratio corresponding to the area covered by each cluster.~~ We use information from the site to define the densities of the model. ~~We defined the model composed of 9 layers overlying half-space with constant thicknesses of 15 m, except for the first layer which was set 20 m, giving an investigation depth of about 140 m. Estimated Poisson's ratios retrieved from W/D method (blue lines in Fig. 8c and d) were averaged and extrapolated to match~~

275 ~~the layers of the LCI and SWT. In Fig. 8c and d, in red, we show the adopted Poisson's ratio used for the clusters A and B, respectively. We defined a constant density of 2200 kg/m3, except for the first layer (2000 kg/m3) based on the geological formations in the area.~~

The inversion method is a deterministic least-square inversion based on Auken and Christiansen (2004), which was developed by Boiero (2009) and ~~was~~ modified by Khosro Anjom (2021) to support parallel computing. At each iteration the

280 V_s ~~VS~~ and thicknesses are updated, and the ~~VP~~ Poisson's ratio and density are fixed *a priori*. All the DCs are inverted ~~contemporarily~~ simultaneously for a set of 1D models ~~which~~ that are tied by lateral constraints between parameters of ~~neighbouring~~ neighboring models. The constraints act as spatial regularization, and their strength is defined to avoid both overfitting and over smoothing. Week lateral constraints or lack of lateral constraints can create unrealistic lateral changes in the final model, whereas too strong constraints can result in an oversmoothed model, masking the sharp lateral variations in

285 the site. We use the data misfit as an indicator for choosing the level of constraints: t~~T~~he inversion with the highest level of constraint that does not impact the DC misfit compared to unconstrained inversion is selected. A thorough description of the method is available in Boiero (2009) and Socco et al. (2009) and the strategy for constraints selection is provided in Boiero and Socco (2010).

290 ~~In Fig. 6 and 7, we show the steps of estimating the reference W/D relationship and apparent Poisson's ratio for the reference DCs of cluster A (Fig. 6a) and B (Fig. 7a). We considered variable Poisson's ratios between 0.1 and 0.45 for the Monte Carlo inversion. We used density of 2000 kg/m3 for the first layer and constant density of 2200 kg/m3 for the other layers. We imposed a 0.05 level of confidence for the Fisher test to accept the best-fitting models among 1,000,000 sampled models. In Fig. 6b and c, as well as Fig. 6b and c, we show the estimated VS and time-average VS models for cluster A and~~

295 ~~B, respectively. The estimated W/D relationship for cluster A and B are given in Fig. 6d and 7d, while the obtained apparent Poisson's ratios are provided in Fig. 6e and 7e. In Fig. 6d and e, as well as Fig. 7d and e, we also show the uncertainty associated with the reference W/D and apparent Poisson's ratio of the clusters, which was obtained based on the method of Khosro Anjom et al. (2019).~~

~~For both clusters, the W/D relationship and apparent Poisson's ratio were not available for the first 20 m due to the lack of~~

300 ~~short wavelength data in the experimental DCs. The investigation depth of 128 m was reached for cluster A, whereas this value was increased to 140 m for cluster B.~~

4.3 SWT

The inputs of the tomographic inversion are the path-averaged DCs from two-station method and the initial model. The parameters of the initial model are the thickness, V_s, Poisson's ratio, and density. The model points are defined with equal distances in X and Y directions. The distance of the model points depends on the required resolution and also the data coverage (i.e., path-averaged DCs). The parameters of the initial model are selected the same as the initial model of the LCI (section 4.2).

We use the tomographic inversion algorithm developed by Boiero (2009) and modified by Khosro Anjom et al. (2021). An essential part of the ~~SWT~~tomographic inversion is the computation of synthetic path-averaged DCs corresponding to the observed ones. We compute the path-averaged DCs ~~from the local DCs of the model points~~, assuming a ~~linear ray~~straight ray path approximation between the two receivers, and as reciprocal of the average slowness along the paths discretized over the model grid. The phase velocities at the location of the discretized paths are computed by bi-linear interpolation of the phase velocities from local DCs corresponding to the adjacent model points (Boiero, 2009)~~, using the algorithm developed by Boiero (2009).~~

Similar to the LCI algorithm, a damped least-least square method is used to iteratively update the model until the minimum misfit between synthetic and observed DCs is reached. The only parameter that updates in the inversion is V_s~~VS~~, while the

others are fixed *a priori*. The method allows the implementation of ~~horizontal (lateral) and vertical~~lateral constraints. We consider the same criteria explained in section 4~~5~~.2 to select the optimal constraint level. In contrast with the other two methods, the SWT method is applied to the northern data set only due to computational capacity restrictions that will be explained in the discussion in section 7.

5.~~4.3~~ Results

5.1. W/D data transform

The clustering of all the estimated DCs ~~from~~generated two clusters ~~the northern and southern zones were clustered into two zones~~ ~~.~~ In Fig. 8a, we show the estimated DCs with the color scale based on the clustering of the DCs in Fig. 8b. The DCs of the western cluster (cluster A, shown in blue in Fig. 8a) present lower phase-velocities compared to the eastern DCs (cluster B, shown in green in Fig. 8a).

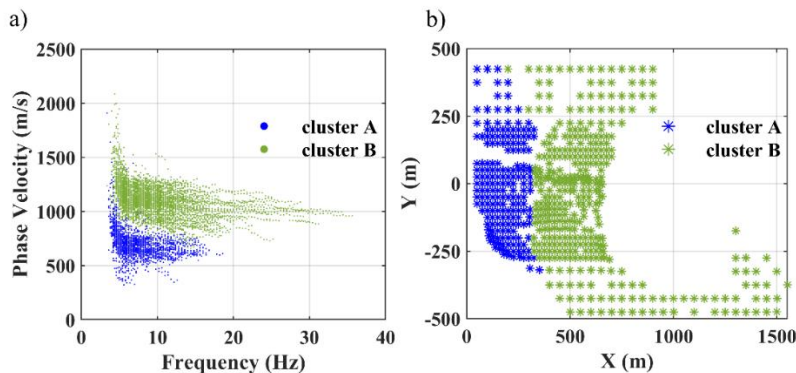



Figure 8: Clustering of the multi-channel DCs. (a) The estimated DCs. (b) The spatial view of the estimated DCs and obtained clusters.

In Fig. 9~~6~~ and 10~~7~~, we show the steps of estimating the reference W/D relationship and apparent Poisson's ratio for the reference DCs of cluster A ~~(Fig. 6a)~~ and B. ~~(Fig. 7a).~~ We considered variable Poisson's ratios between 0.1 and 0.45 for the Monte Carlo inversion. Based on the information from the site, ~~W~~we ~~used~~considered density of 2000 kg/m^3 for the first layer and constant density of 2200 kg/m^3 for the other layers. We imposed a 0.~~05~~001 level of confidence for the Fisher test to accept the best-fitting models among 1,000,000 sampled models. In Fig. 9~~6~~b and c, as well as Fig. 10~~6~~b and c, we show the estimated V_s ~~VS~~ and time-average V_s ~~VS~~ models for cluster A and B, respectively. In Fig. 9b and Fig. 10b, we also show the boundaries of the V_s model space for the Monte Carlo inversion. Due to the application of the scale properties, the selected models can be ~~moved~~scaled to outside of the original boundaries of the model space ~~The selected models are sometimes beyond these boundaries that are caused by the scaling of the models~~ (Socco and Boiero, 2008). ~~Basically, the algorithm scaled the models to a new model space that fits the DC better than the initial model space (the boundaries).~~ The estimated

W/D relationship for cluster A and B are ~~given~~shown in Fig. ~~9~~6d and ~~10~~7d, ~~while~~whereas the obtained apparent Poisson's ratios are provided in Fig. ~~9~~6e and ~~10~~7e. In Fig. ~~9~~6d and e, as well as Fig. ~~10~~7d and e, we also show the uncertainty associated ~~with~~to the reference W/D and to the apparent Poisson's ratio of the clusters, which was obtained based on the method of Khosro Anjom et al. (2019).

For both clusters, the W/D relationship and apparent Poisson's ratio were not available for the first 20 m due to the lack of short wavelength data in the experimental DCs. The investigation depth of 128 m was reached for cluster A, whereas this value was increased to 140 m for cluster B.

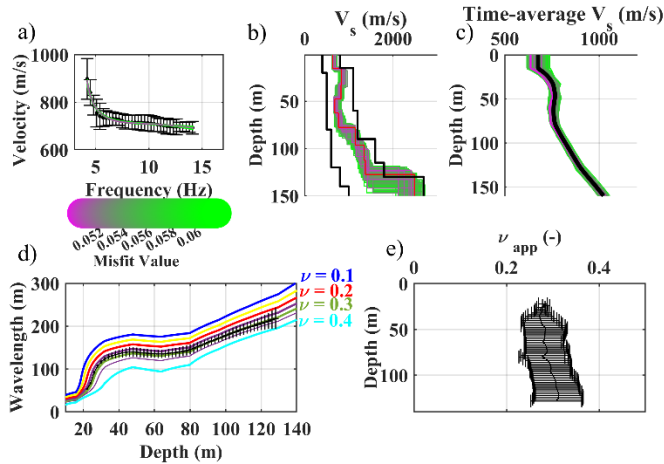



Figure ~~9~~6: The steps of estimating the reference W/D relationship and apparent Poisson's ratio for cluster A. (a) The reference DC with the phase-velocity and uncertainty and the synthetic accepted DCs from the Monte Carlo inversion. (b) The accepted V_s ~~VS~~ models from the Monte Carlo inversion, where the black lines show the initial boundaries of the model space. (c) The accepted time-average V_s ~~VSZ~~ models from the Monte Carlo inversion. In black, the reference time-average V_s. ~~VS~~ (d) Estimated reference W/D relationship. The ~~coloured~~colored W/D relationships are the synthetic ones, each with constant Poisson's ratio, used for apparent Poisson's ratio estimation. (e) The estimated reference apparent Poisson's ratio.

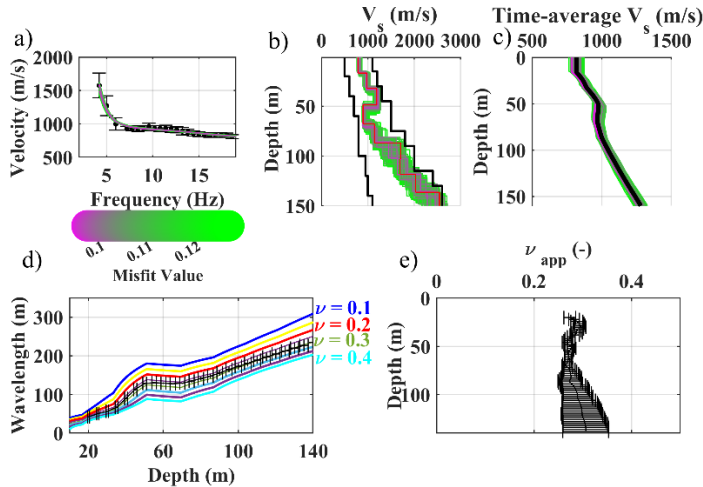
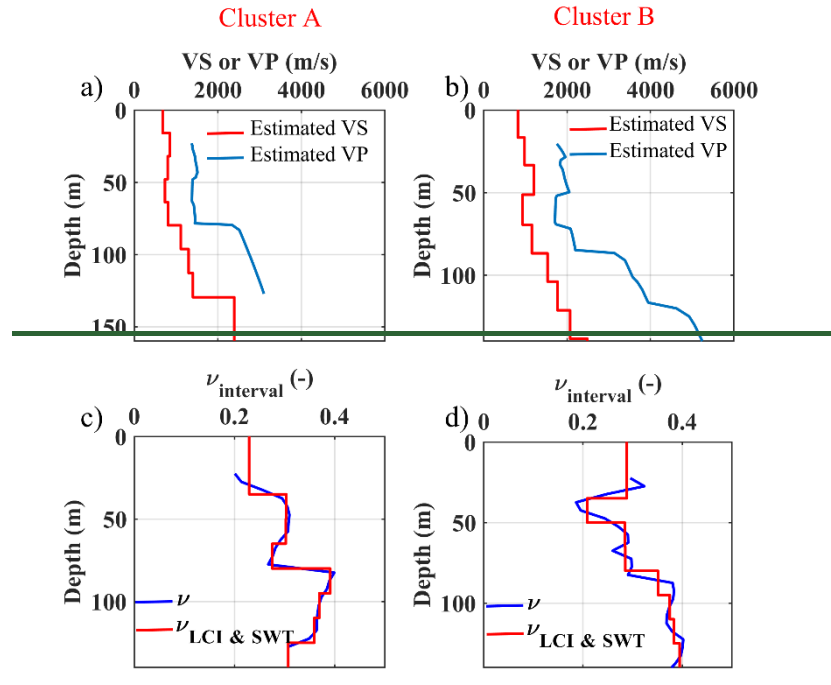



Figure 107: The steps of estimating the reference W/D relationship and apparent Poisson's ratio for cluster B. (a) The reference DC with the phase-velocity and uncertainty and the synthetic accepted DCs from the Monte Carlo inversion. (b) The accepted V_s ~~VS~~ models from the Monte Carlo inversion, where the black lines show the initial boundaries of the model space. (c) The accepted time-average V_s ~~VSZ~~ models from the Monte Carlo inversion. In black, the reference time-average V_s. ~~VS~~ (d) Estimated reference W/D relationship. The ~~coloured~~colored W/D relationships are the synthetic ones, each with constant Poisson's ratio, used for apparent Poisson's ratio estimation. (e) The estimated reference apparent Poisson's ratio.

~~Using the estimated VS and VP models at the reference location of each cluster (Fig. 8a and b), we obtained Poisson's ratios required by LCI and SWT methods. In Fig. 8c and d, in blue, we show the estimated Poisson's ratio for cluster A and cluster B, respectively.~~

~~**Figure 8:** Poisson's ratio estimation for the cluster A and B. (a and b) The estimated VS and VP models corresponding to the reference DC of Clusters A and B. (c and d) In blue, the estimated Poisson's ratio for the cluster A and B. In red, the averaged and extrapolated Poisson's ratio corresponding to the layers of LCI and SWT.~~

The estimated DCs from the two clusters were transformed to interval V_s ~~and~~ and V_p ~~—~~models using the reference W/D relationships and apparent Poisson's ratios (Fig 9d and e, and Fig 10d and e). In Fig. 119, we show several horizontal slices of the estimated V_s ~~VS~~ model averaged over the depth intervals indicated on top of ~~the~~ each plots. Similarly, in Fig. 120, we show the horizontal slices of the estimated V_p ~~VP~~ averaged over the same depth intervals. Both models show a sharp ~~significant~~ velocity ~~contrast~~ transition between the east and west side of the area in the shallow depths above 65 m (Fig. 11a, b and c and Fig. 12a, b and c)~~, which reduces in the deeper portion of the model~~. This contrast is created by the transition from the high-velocity limestone and marl formations in the east to loose materials in the west characterized by lower

velocity. Below, 110 m (Fig. 11e and f, and Fig. 12e and f) the contrast disappears, reaching the high-velocity formation probably from the Danian stage.

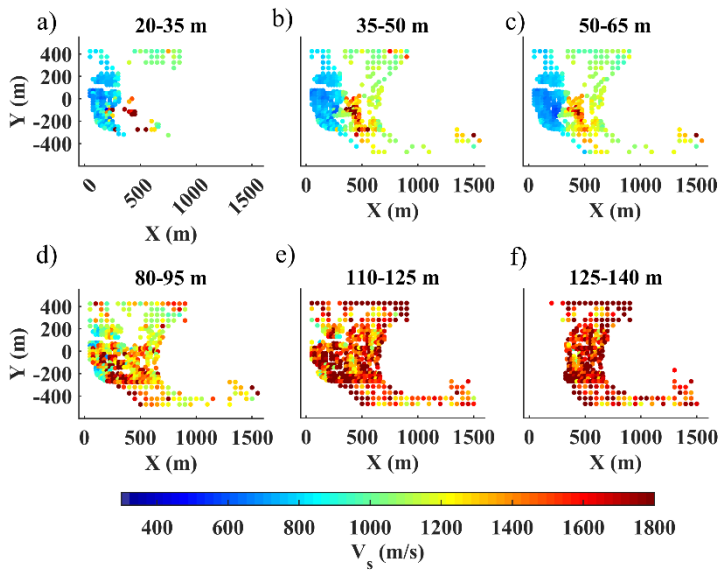



Figure 911: The estimated V_s VS model using the W/D method. (a to f) The horizontal slices at different depth intervals indicated on top of each plot.

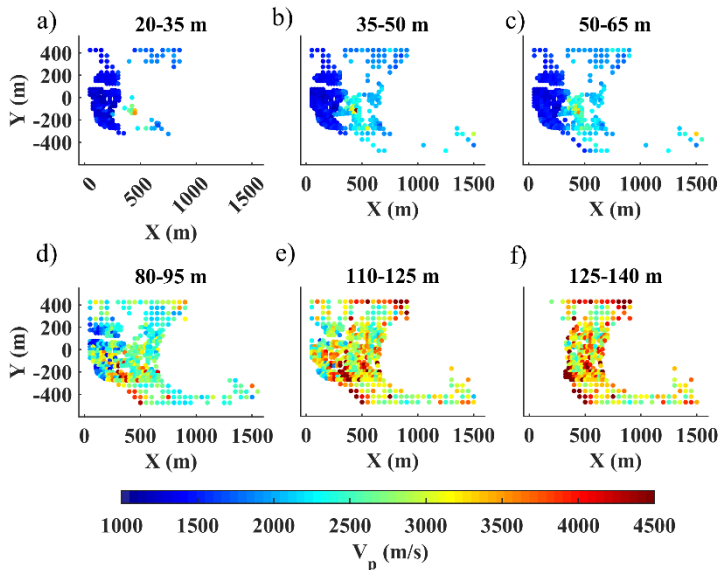



Figure 1012: The estimated V_p VP model using the W/D method. (a to f) The horizontal slices at different depth intervals indicated on top of each plot.

5.2. LCI

We defined ~~the~~an initial model composed of 9 layers overlying half-space with constant thicknesses of 15 m, except for the first layer which was set at 20 m, giving an investigation depth of about 140 m. We assigned the Poisson's ratios of the model based on the results of the W/D analysis: using the estimated V_s and V_p models at the reference location of each cluster (Fig. 13a and b), we obtained the Poisson's ratios, ~~we obtained the estimated Poisson's ratio in~~shown in Fig. 13c and d (in blue), corresponding to cluster A and cluster B, respectively. Since the Poisson's ratios are assumed invariant within each cluster, we used all estimated V_s and V_p models of the clusters from the W/D method to obtain an uncertainty for the estimated Poisson's ratios at each depth. The horizontal error bars in Fig. 13c and d show the standard deviation of the estimated Poisson's ratios. We averaged and extrapolated the values of the Poisson's ratio to match them with the layers of the LCI model (red lines in Fig. 13c and d). Based on the clustering analysis of the W/D method (Fig. 8b) and location of the LCI model points, we assigned the appropriate Poisson's ratio to each 1D model. We defined a constant density of 2200 kg/m^3, except for the first layer (2000 kg/m^3) based on the geological formations in the area.

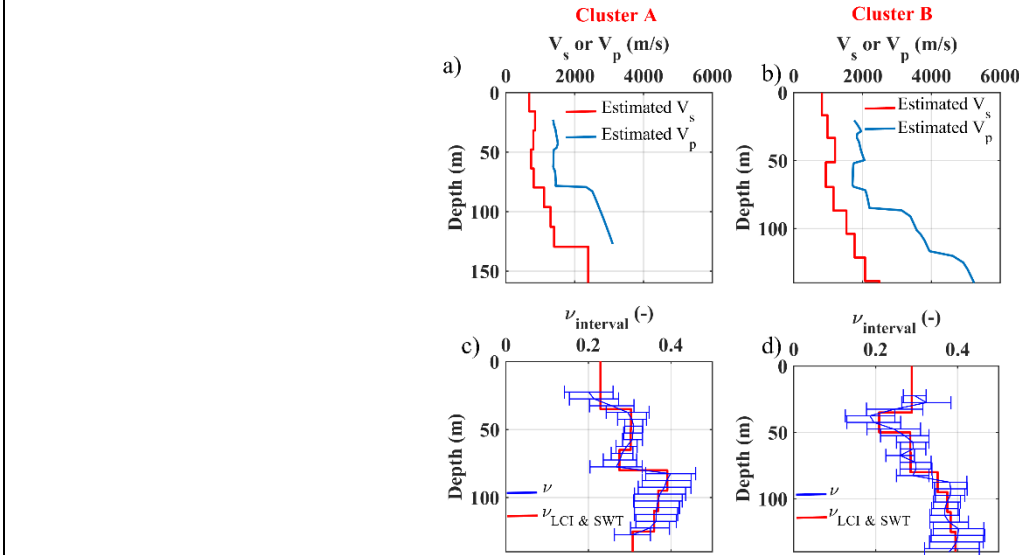



Figure 13: Poisson's ratio estimation for the cluster A and B. (a and b) The estimated V_s and V_p models corresponding to the reference DC of Clusters A and B. (c and d) In blue, the estimated Poisson's ratio for the cluster A and B. In red, the averaged and extrapolated Poisson's ratio corresponding to the layers of LCI and SWT.

5 LCI

5.1 The inputs: multi-channel DCs and initial model

The method's inputs are the multi-channel DCs and the initial models at the location of the local DCs obtained in section 4.1. The initial model is composed of the thickness, density, Poisson's ratio, and VS. We defined the model composed of 9 layers overlying half-space with constant thicknesses of 15 m, except for the first layer which was set 20 m, giving an investigation depth of about 140 m. Estimated Poisson's ratios retrieved from W/D method (blue lines in Fig. 8c and d) were averaged and extrapolated to match the layers of the LCI and SWT. In Fig. 8c and d, in red, we show the adopted Poisson's ratio used for the clusters A and B, respectively. We defined a constant density of 2200 kg/m3, except for the first layer (2000 kg/m3) based on the geological formations in the area.

5.2 Inversion algorithm

The method is a deterministic least-square inversion based on Auken and Christiansen (2004) developed by Boiero (2009) and modified by Khosro Anjom (2021) to support parallel computing. At each iteration the VS and thicknesses are updated, and the VP and density are fixed a priori. All the DCs are inverted contemporarily for a set of 1D models which are tied by lateral constraints between parameters of neighbouring models. The constraints act as spatial regularization, and their strength is defined to avoid both overfitting and over-smoothing. We use the data misfit as an indicator for choosing the level of constraints: The inversion with the highest level of constraint that does not impact the DC misfit compared to unconstrained inversion is selected. A thorough description of the method is available in Boiero (2009) and Socco et al. (2009) and the strategy for constraints selection is provided in Boiero and Socco (2010).

5.3 Results

We performed an unconstrained and several laterally constrained inversions to find the optimal level of constraints according to the strategy described in Boiero and Socco (2010). We chose a lateral constraint on V_s VS equal to 50 m/s, that was the highest level of constraints that did not significantly impact the inversion's residual misfit. In Fig. 14+, we show the horizontal slices of the estimated V_s VS model at various depths corresponding. Even though the inversion is laterally constrained, the algorithm was able to depict a sharp transition between the east and west (Fig. 14a to c), which is in line with the results of the W/D method (Fig 11a to c). The LCI model below 87.5 m shows high velocities with insignificant variations between east and west.

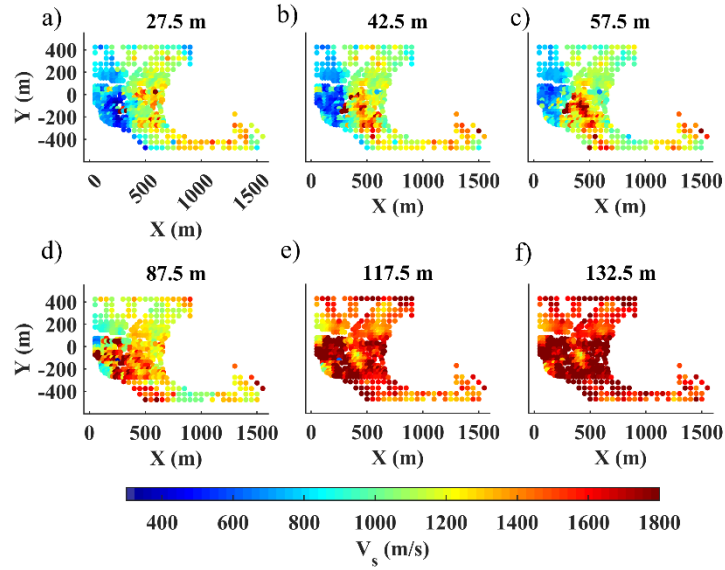



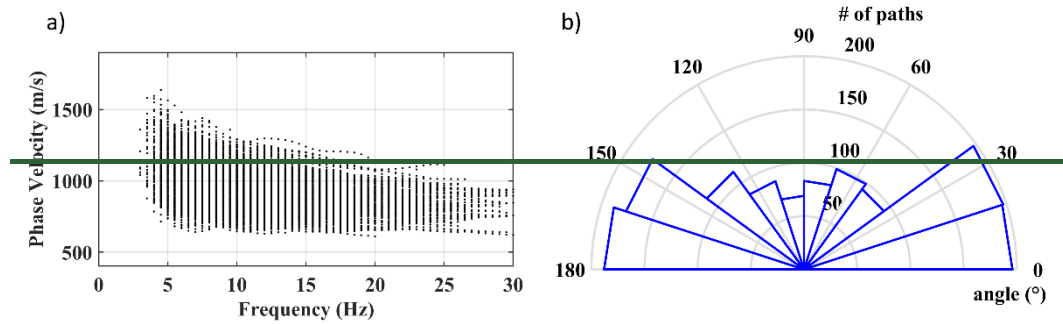
Figure 1114: The estimated V_s VS model using the LCI method. (a to f) The horizontal slices at different depths indicated on top of each figure.

5.36 SWT

Given the high data coverage from the estimated path-averaged DCs (Fig. 7), we defined a dense model grid on the considered northern zone, composed of 300 1D models, aiming at obtaining high resolution model. For the same zone (north), 174 multi-channel DCs and model points were obtained for the W/D and LCI methods, which is almost half of the considered model points for SWT. We used the same initial models defined considered parameters (V_s, Poisson's ratio, density, and thicknesses) of the 1D models equal to the ones defined for the LCI (section 5.2).

6.1 The inputs: two-station DCs and initial model

The two-station DCs are estimated using the recordings from receiver couples aligned with a source using interferometry. We use the algorithm developed by Da Col (2019) and modified by Khosro Anjom et al. (2021). First, an automatic search is performed to find the receiver couples aligned with the source at each azimuth angle, considering 1° tolerance for the deviation from a linear path. Then, the traces are narrow-band filtered at various frequencies, using zero-phase Gaussian filters. Next, the filtered traces of the receiver couples are cross-correlated and assembled to form the cross-multiplication matrix. We use a 3rd-order spline interpolator to convert the cross-multiplication matrix to the frequency-velocity domain. Finally, at each frequency, the phase-velocity is picked as the maximum of the cross-multiplication matrix. To avoid cycle skipping we use as reference the closest multichannel DC, and we automatically pick maxima closest to the reference DCs. To minimize the contamination of the fundamental mode by higher surface wave modes, we damp the higher-mode data using the muting strategy of Khosro Anjom et al. (2019).

470 We applied the SWT to the data from the north of the site only (blue markers in Fig. 2). We performed an automatic search of the receiver couples aligned with sources within 250 m offset, which resulted in 4710 possible receiver couples and source settings. We used the local DCs from the multi-channel analysis (Fig. 5a) as references to locate the correct trend of the path-averaged fundamental mode DCs. We discarded noisy or in inconsistent cross-multiplication matrices and in total, 1301 path-averaged DCs were estimated. In Fig. 12a and b, we show the estimated path-averaged DCs and the observed azimuthal illumination. The data show uniform coverage with most paths showing angles between 0° to 40° and 140° to 180° angles.

475 Figure 12: (a) The estimated path averaged DCs corresponding to north of the data set. (b) The obtained azimuthal illumination with the numbers around the great circle showing the angles and other circles representing the obtained coverage.

In Fig. 13a to f, we show the data coverage within different wavelength ranges, where the colour scale shows the path average phase velocity. The data exhibit very high coverage for wavelengths between 40 to 220 m, beyond which it decreases substantially.

480

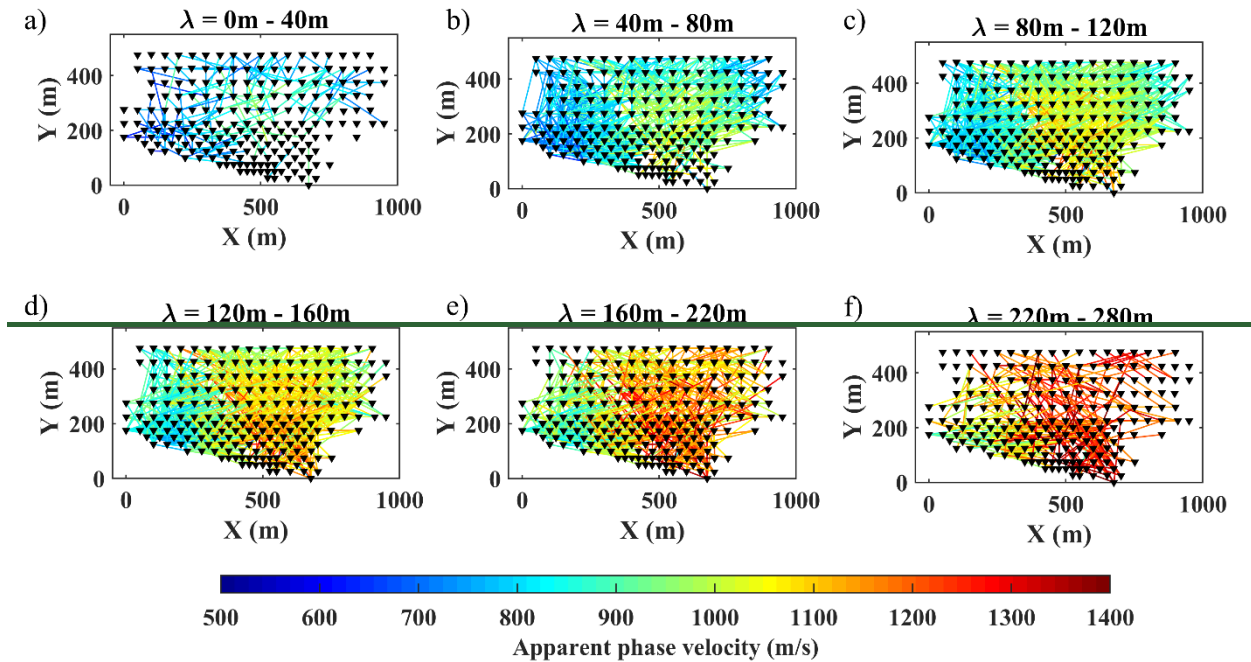



Figure 13: Pseudo-slices of the estimated path-averaged DCs from the north of the site shown within the wavelength ranges: (a) 0 to 40 m. (b) 40 to 80 m. (c) 80 to 120 m. (d) 120 to 160 m. (e) 160 to 220 m. (f) 220 to 280 m.

The parameters of the initial model are the thickness, VS, Poisson's ratio, and density. Given the high data coverage, we defined dense model grid on the considered northern zone, composed of 300 aiming at obtaining high resolution model. For the same zone (north), 174 multi-channel DCs and model points were obtained for the W/D and LCI methods, which is almost half of the considered model points for SWT. We considered parameters (VS, VP, density and thicknesses) of the 1D models equal to the ones defined for the LCI (section5.1).

6.2 Inversion algorithm

We use the algorithm developed by Boiero (2009) and modified by Khosro Anjom et al. (2021). An essential part of the SWT is the computation of synthetic path-averaged DCs corresponding to the observed ones. We compute the path-averaged DC from the local DCs of the model points, assuming linear ray path between the two receivers, using the algorithm developed by Boiero (2009).

Similar to the LCI algorithm, a damped least least square method is used to iteratively update the model until the minimum misfit between synthetic and observed DCs is reached. The only parameter that updates in the inversion is VS, while the others are fixed a priori. The method allows the implementation of horizontal (lateral) and vertical constraints. We consider the same criteria explained in section 5.2 to select the optimal constraint level.

In Fig. 15~~4~~, we show the estimated V_s ~~VS~~ at the different layers, which is in correspondence of the depth intervals defined for the estimated V_s ~~VS~~ of W/D method (Fig. 11~~8~~). We chose a lateral constraint of V_s equal to 50 m/s, that was the highest level of constraints that did not significantly impact the inversion's residual misfit. Similar to the estimated model from W/D and LCI, the estimate model from SWT shows a significant velocity contrast between the east and west. Nevertheless, this contrast is smoother than the other two models. The model shows high velocities with no significant lateral variation below 110 m of depth, an indication of reaching the high-velocity limestone/marl formation from the Danian stage (Fig. 1c).

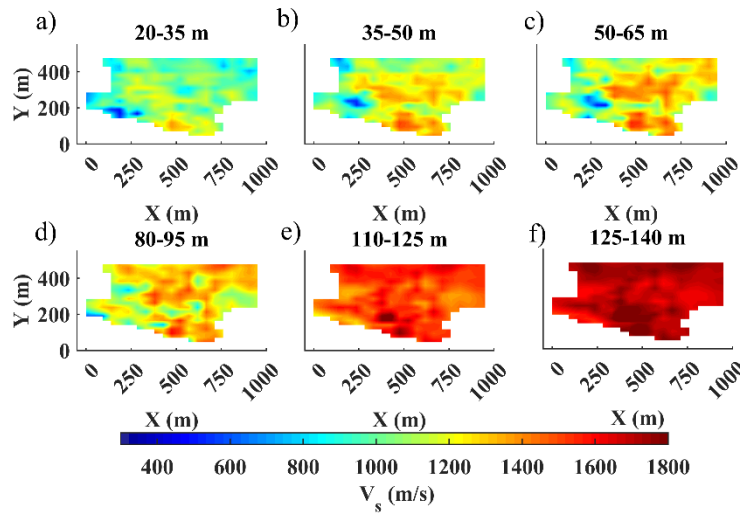



Figure 15~~14~~: The estimated V_s ~~VS~~ model for the north of the site using the SWT method. (a to f) The horizontal slices at different layers indicated on top of each figure.

7 Discussion

We showed the application of three surface-wave methods for V_s ~~VS~~ model estimation, out of which, the W/D and LCI methods provided the velocity models at the location of DCs, both in the northern zone (174 locations) and southern zone (371 locations). The W/D method provided the V_p model and the reference Poisson's ratios used in LCI and SWT. The SWT was applied only to the northern zone, which provided the V_s ~~VS~~ models at 300 defined model points. ~~The W/D method provided continuous velocities in depth, while the SWT and LCI provided layered models. Although we defined similar reference models for the LCI and SWT methods, the LCI was set to also update the thicknesses at each iteration, leading to different parameterization in depth compared to SWT.~~ Here, we further ~~analyse~~evaluate the results of the three methods in terms of vertical resolution, spatial resolution, differences of the estimated models and computational efficiency of each method.

In Fig. 1~~5~~6, we show the wavelength distribution of the estimated multi-channel DCs in blue that shows dense data sampling up to 300 m of wavelength, suggesting good vertical resolution also in deeper portions. These DCs were obtained from the recordings of receivers spread over a square window of 100×100 m^2. The receiver ~~selection square~~window was ~~shifted~~ shifted by one receiver spacing (50 m in the north and 25 m in the south). Neglecting the smoothing effect of superimposed receiver windows, the shifting distance ~~, which~~ can be considered as the spatial resolution of the multi-channel DC data.

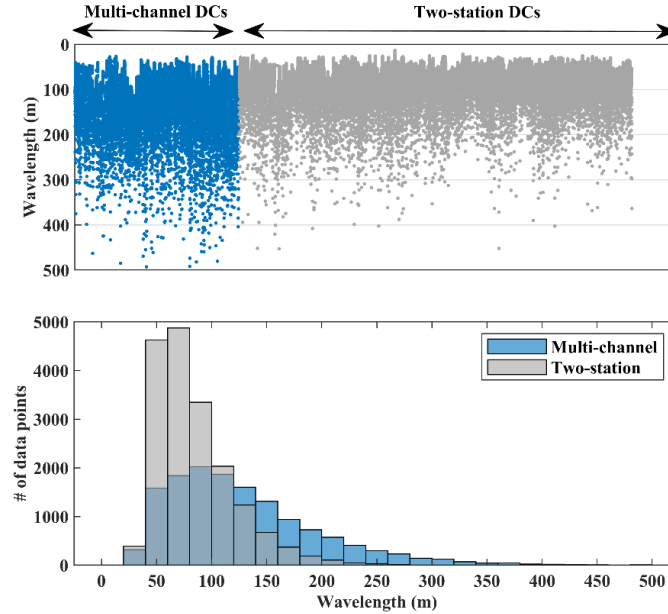



Figure ~~15~~16: Comparison between the wavelength distributions of the multichannel and two-station dispersion data. (Top) The distribution of the wavelength shown separately for each estimated DC. (Bottom) The histogram showing the wavelength distribution of all DCs within 20 m wavelength bins.

In Fig. 1~~5~~6, in grey, we also show the wavelength distribution of the estimated two-station DCs. Even though the total number of DCs from two-station analysis (1301) is far more than multi-channel ones (545), the large wavelength datapoints (>120 m) are sparser than the ones obtained with the multi-channel ~~analysis~~ method (top plot in Fig. 1~~5~~6), mainly due to low signal-to-noise ratio of the cross-multiplication matrices ~~in~~ at low frequencies. This shows that multi-channel DC analysis provides ~~better resolution in depth~~greater investigation depth compared to the two-station method. To mitigate ~~this issue that prevents the estimation of the deepest portion of the model in scheme of SWT~~the lack of investigation depth of SWT, we employed the wavelength-based weighting ~~system~~ developed by Khosro Anjom et al. (2021), to increase the score of large-wavelength data points in the tomographic inversion, aiming at enhancing the resolution ~~in~~ at depth. We performed a checkerboard test to evaluate horizontal and vertical resolution of the SWT. We perturbed the estimated V_s ~~VS~~model from SWT method (Fig. 1~~5~~4) by 8% negatively and positively which alternated every two layers (Fig. 1~~7~~6a and b). In Fig. 1~~7~~6c to f, we show the results of the inversion at various layers. The 50×50 m^2 perturbations were well-recovered up to depth of 50 m (Fig. 1~~7~~6c and d), providing similar spatial resolution compared to LCI and W/D methods. The resolution slightly

decreases towards the deepest portion of the model (Fig. 1~~6~~7e and f), especially in the northern part where long-wavelength data are lacking (Fig. 13f).

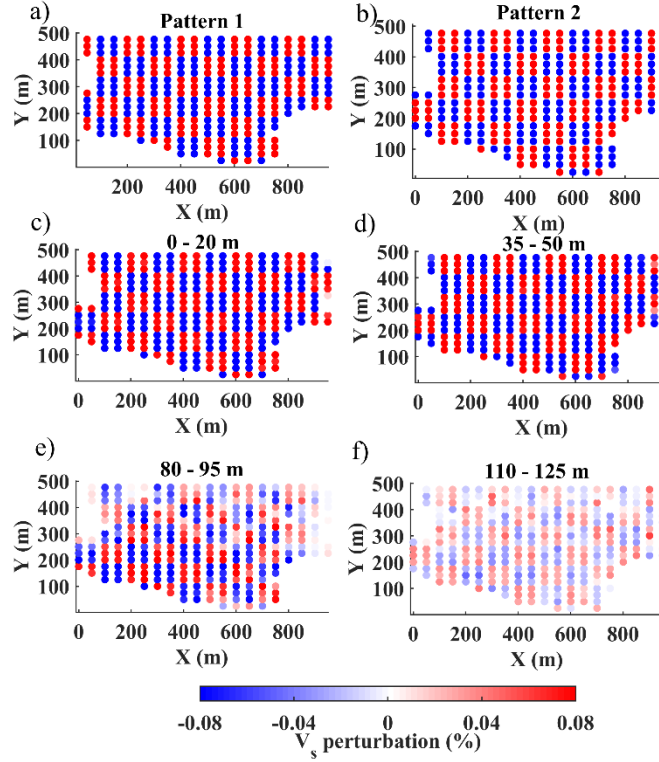



Figure ~~16~~17: Checkerboard test. (a) The pattern 1 used to perturb V_s ~~VS~~ of layer 1, 2, 5, 6, and 9. (b) The pattern used to perturb V_s ~~VS~~ of layer 3, 4, 7, and 8. (c) 0 to 20 m. (d) 35 to 50 m. (e) 80 to 95 m. (f) 125 to 140 m.

The application of the W/D method to the data set provided both V_s ~~VS~~ and V_p ~~VP~~ models. We considered the estimated Poisson's ratio of the two clusters as prior information in the reference model of the LCI and SWT methods. Now, we use the same Poisson's ratios to transform the V_s ~~VS~~ results of these two methods to V_p ~~VP~~ models. We also linearly interpolate the V_s ~~VS~~ and V_p ~~VP~~ results from all three methods to obtain the velocity models at common voxels of $10 \times 10 \times 0.1$ m^3 within x, y, and z (depth) directions, respectively. In Fig. 1~~8~~7, we compare the retrieved pseudo 3-D V_s ~~VS~~ and V_p ~~VP~~ models from the three methods at various iso-surfaces. Very similar trend of variations for V_s ~~VS~~ (left panels) and V_p ~~VP~~ (right panels) are obtained from the application of the methods, and they all depict a significant variation between the east and west side of the site.

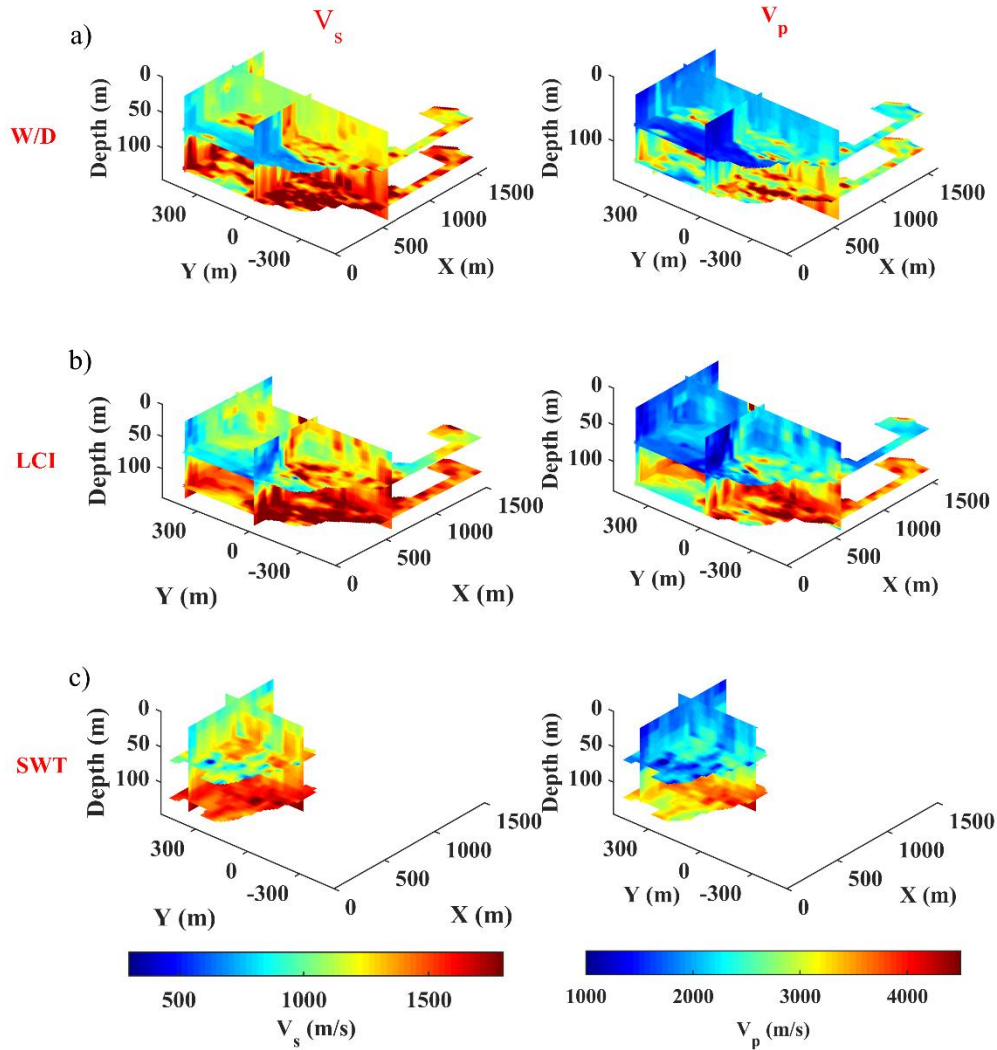



Figure ~~17~~18: Iso-surfaces of the estimated V_s ~~VS~~ models (left panel) and V_p ~~VP~~ models (right panels) using: (a) W/D, (b) LCI, and (c) SWT methods. The sections are at plains x= 600 m, y =0, and 400 m, and depth = 70, and 125 m.

555   To compare the estimated models from each method quantitatively, we compute the difference between estimated V_s ~~VS~~ and V_p ~~VP~~ of every two methods separately, as:

$$\varepsilon(i,j,k) = 2 \times \left[\frac{V(i,j,k)_{method.1} - V(i,j,k)_{method.2}}{V(i,j,k)_{method.1} + V(i,j,k)_{method.2}} \right], \qquad\qquad (\underline{2}\text{\sout{1}})$$

where i, j, and k are the indices of the voxels in x, y, and z (depth) directions, respectively, and V is the velocity. In Fig. 19~~8~~, we show the boxplots of the differences compartmentalized within different depth intervals. ~~The box plot is defined by~~

560   ~~three lines showing the 25th percentile, median and 75th percentile of the residual's distribution, and whisker lines extending~~

~~from the box's edges up to 1.5 times the distance between the edges of the box. The rest of the data are considered as outliers and are shown with "+".~~ The ~~boxplots~~ differences in Fig. 18 are computed for depths between 20 and 140 m, except for the V_s ~~VS~~ comparison of the LCI and SWT (Fig. 1~~9~~8e), which also includes the first 20 m. This is due to the lack of short wavelength data for the W/D method. The significant registered differences ~~registered of the outliers~~ (red "+") are mainly caused by the methods' different parameterization in depth~~:~~. the W/D method provided continuous velocities in depth (every 10 cm), while the SWT and LCI provided layered models; although we defined similar reference models for the LCI and SWT methods, the LCI was set to also update the thicknesses at each iteration, leading to different parameterization in depth compared to SWT; in addition, the W/D and LCI took as input multi-channel DCs and provided velocity models at their locations, whereas the SWT considered path-averaged DCs and resulted in velocity models in defined model point locations.

~~The difference between the estimated V_s VS and V_p VP from W/D and LCI methods are small and uniform within different depth ranges (Fig. 1~~9~~8a). Nevertheless, for the deepest layers, a slight over estimation of the V_p VP from the LCI method is registered compared to the V_p VP from W/D technique (right panel in Fig. 1~~9~~8a). The differences are increased in depth when the V_s VS and V_p VP of the W/D and SWT methods are compared (Fig. 1~~9~~8b). The differences obtained for the estimated V_s VS from the LCI and SWT methods (Fig. 1~~9~~8c) are very similar to the ones observed for the V_p VP, since the V_p VP models of both methods were obtained from the estimated V_s VS and same a priori Poisson's ratios; the differences are mainly less than 5%.~~

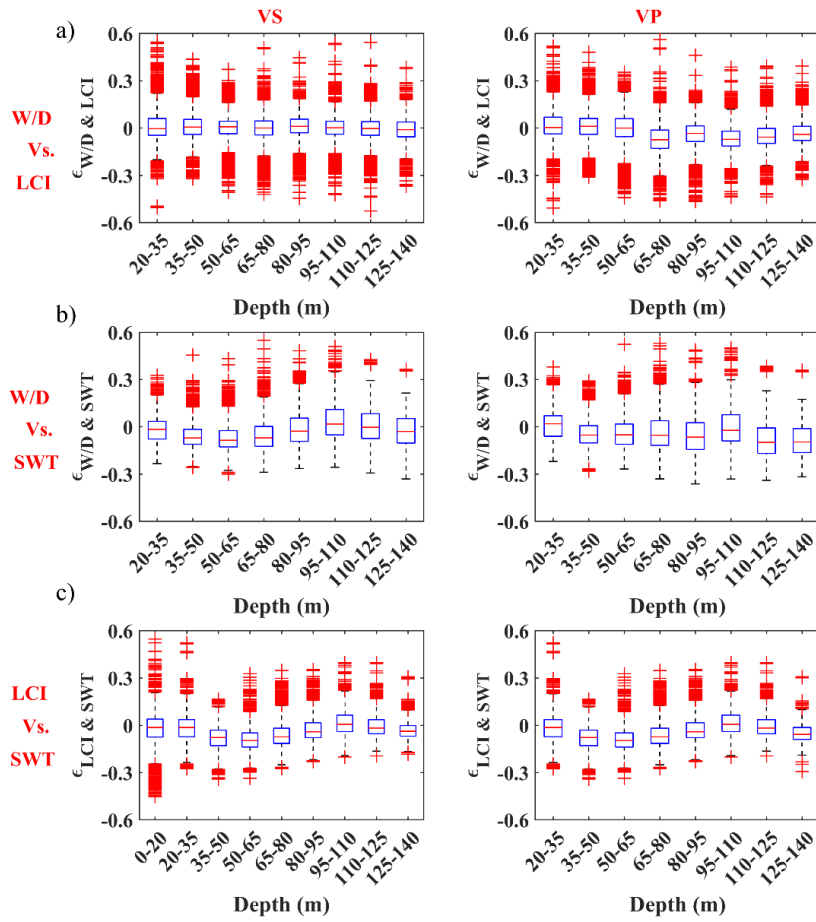



Figure ~~18~~19: The box plot showing the difference between the V_s ~~VS~~ (left panel) and V_p ~~VP~~ (right panel) models obtained from the three methods computed using equation (~~2~~1). The box plot is defined by three lines showing the 25th percentile, median and 75th percentile of the residual's distribution, and whisker lines extending from the box's edges up to 1.5 times the distance between the edges of the box. The rest of the data are considered as outliers and are shown with "+". The box plots show the differences between the models obtained from: (a) W/D and LCI, (b) W/D and SWT, and (c) LCI and SWT.

The difference between the estimated V_s and V_p from W/D and LCI methods are small and uniform within different depth ranges (Fig. 19a). Nevertheless, for the deepest layers, over-estimation of the V_p from the LCI method is registered compared to the V_p from W/D technique (right panel in Fig. 19a). The differences are increased in depth when the V_s and V_p of the W/D and SWT methods are compared (Fig. 19b). The differences obtained for the estimated V_s from the LCI and SWT methods (Fig. 19c) are very similar to the ones observed for the V_p, since the V_p models of both methods were obtained from the estimated V_s and same *a priori* Poisson's ratios; the differences are mainly less than 5%.

We compute the total differences between the estimated models of every two methods as:

$$\varepsilon_{tot} = \left[\frac{1}{n.q}\sum_{j=1}^{q}\sum_{i=1}^{n}\left|\frac{1}{m}\sum_{k=1}^{m}\varepsilon(i,j,k)\right|\right] \times 100, \qquad (32)$$

where m, n, and q are the overall number of voxels in z (depth), x, and y directions, respectively. In Table 2, we report the values of the total differences obtained by comparing the models from the three methods.

Table 2: The total difference between the estimated V_s VS and V_p VP models obtained from the application of W/D, LCI, and SWT methods.

Total difference	ε_{tot} (V_s VS)	ε_{tot} (V_p VP)
W/D vs. LCI	3.3%	4.67%
W/D vs. SWT	6%	7.06%
LCI vs. SWT	4.74%	4.52%

In Table 3, we provide the approximated computational costs for each part of the three methods. The most time-consuming step of all methods is the DC estimation, which also involves expert user intervention. Compared to W/D and LCI, the SWT usually requires more DCs to reach adequate data coverage for the tomographic inversion. We estimated 1301 DCs for SWT applied to the north of the site, whereas only 174 DCs were estimated for the application of the W/D and LCI methods to the same zone. The W/D relationship and Poisson's ratio estimation is a common stage for all three methods. The inversion running times (for LCI and SWT) given in Table 3 are for a single inversion trial using 10 CPU cores. Usually, in addition to an unconstrained inversion, several constrained inversions are performed to reach a satisfactory model in scheme of SWT and LCI methods, whereas the W/D method can be applied faster and is efficient for processing large-scale data sets. It is noteworthy to mention that the SWT was limited to the northern zone due to computational limitations. The tomographic inversion with 1301 DCs and 300 model points was performed by a workstation equipped with 128 GB of memory and a 10-core CPU. The simultaneous inversion of both zones with at least twice the number of DCs and model points would have required exponentially higher memory and computational capabilities that our workstations could not provide. We could have under sampled the DCs and reduced the number of model points in order to invert both zones together, but we decided to maintain the resolution of the tomographic inversion by focusing to the northern zone only.

Table 3: The approximated computational costs for each method.

	Processing (DC estimation)	W/D relationship and Poisson's ratio estimation	Model estimation
W/D	1 min/DC	24 hrs	5 s/1D model
LCI	1 min/DC	24 hrs	5 hrs
SWT	1 min/DC	24 hrs	48 hrs

In Fig. ~~19~~20, we show the geological map superimposed with the satellite view of the area and with the horizontal slice of the estimated V_p ~~VP~~ model from W/D corresponding to the depth between 35-50 m. The two diagonal and vertical faults at the north-west of the investigated area separate the east from the west ~~sides~~. In the region between the two faults, a gap within the estimated model from the W/D method is observed: ~~t~~The scattering and complex propagation of surface-waves th~~at~~e passes through these discontinuities resulted in inconsistencies in the spectrum ~~that~~ and prevented the estimation of reliable DCs. The west of the area is characterized by loose ~~materials~~ formation from recent deposits (outcrop 5 in Fig. ~~19~~20). The rest of the region is known for stiffer materials, composed of limestone and marl. The estimated V_p ~~VP~~ also shows a higher velocity in the eastern region. The fastest V_p ~~VP~~ is registered in the correspondence of the Sparnecian formation (outcrop 4 in Fig. ~~19~~20). The high-velocity formation from the Danian stage that is outcropping outside of the investigated area, is probably reached below the depth of 110 m as all V_s models from the three methods show ver~~r~~y high velocity with minor lateral variations below this depth (Fig. 11e and f, Fig. 14e and f, and Fig. 15e and f).

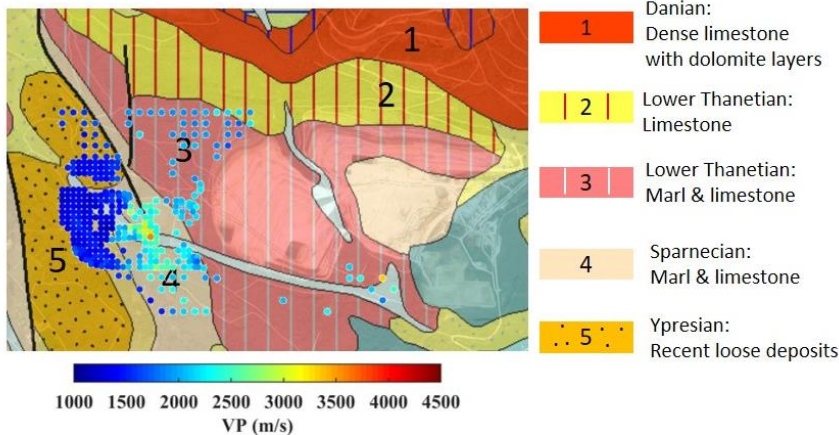



Figure ~~19~~20: The geological map of the site, obtained from French Geological Survey (© BRGM - www.infoterre.brgm.fr), superimposed with the area's satellite view (© Google Earth) and the estimated V_p ~~VP~~ corresponding to depth 35-50 m using the W/D method.

The application of the calibrated multi-channel (W/D and LCI) and two-station (SWT) methods showed promising results for the processing of the carpet recorded data. All three methods depicted a contrast between the limestone rich area in the

635   east~~west~~ and loose material in the ~~east~~west, which are corroborated by the geological information. Given the site is an open-cast limestone mining site, the estimated models can offer valuable insights for strategizing the expansion and excavation of the quarry in the investigated zone to support the nearby cement production facility.

8 Conclusions

We showed the application of three surface-wave methods, W/D, LCI, and SWT, to a large-scale test data set acquired by

640   carpet recording approach and from a stiff site ~~to~~ for the purpose of estimating~~e both~~ V_s ~~VS~~ and V_p ~~VP~~ models. The fine-tuned multi-channel W/D and LCI methods showed potential in the processing of the seismic data from carpet recording technique. Also, the irregular source-receiver outline in scheme of carpet recording provided high DC coverage within two-station method, facilitating high-resolution tomographic inversion. The W/D and LCI methods were applied to both zones outside the mining pits, ~~while~~ whereas SWT application was limited to the north of the site. We used the W/D method to

645   estimate *a priori* Poisson's ratios required for the LCI and SWT methods. The estimated V_s ~~VS~~ and V_p ~~VP~~ models from the three methods were less than 6% and 7.1% different, respectively. The retrieved lateral variation by the methods showed good similarity with the geological information available for the site. The most time-consuming part of the methods are the ~~dispersion curve~~ DC pickings, especially for the SWT method which requires more DC to reach an adequate data coverage compared to the other two methods. As a result, the automation of DC picking can be viewed as an important milestone in

650   the industrialization of the surface-wave methods, facilitating their swift application to data sets even larger than the one used in this study.~~As a result, the automation of the DC picking can be considered a great breakthrough in industrialization of these methods, which enables their fast applications to even larger data sets than the one used in this paper.~~

655   *Code and data availability.* The data are licensed to Politecnico di Torino that allows research activities conducted only by Politecnico di Torino. As a result, the data cannot be made publicly available. The code~~s~~ for the three surface-wave methods were developed in other projects. Nevertheless, the code for W/D data transformation method may be available by contacting the corresponding author.

660   *Author contributions.* FKA worked on the application of the methods to the data set with supervision of FA and LVS. FKH wrote the original paper draft, with the contribution and revisions of FA and LVS.

Competing interests. The contact author declares that neither him nor his co-authors have any competing interests.

665

Acknowledgements

We would like to thank Gallego Technic Geophysics for licensing the field data. The first author would like to thank TotalEnergies for supporting his PhD during which this study was carried out. We also thank the topical Editor Caroline

670   Beghein and the anonymous reviewer for their thorough reviews and useful suggestions.

References

Auken, E. and Christiansen, A. V.: Layered and laterally constrained 2D inversion of resistivity data, Geophysics, 69, 752–761, https://doi.org/10.1190/1.1759461, 2004.

Auken, E., Christiansen, A. V., Jacobsen, B. H., Foged, N., and Sorensen, K. I.: Piecewise 1D laterally constrained inversion

675   of resistivity data, Geophys. Prospect., 53, 497–506, https://doi.org/10.1111/j.1365-2478.2005.00486.x, 2005.

Badal, J., Chen, Y., Chourak, M., and Stankiewicz, J.: S-wave velocity images of the Dead Sea Basin provided by ambient seismic noise, J. Asian Earth Sci., 75, 26–35, https://doi.org/10.1016/j.jseaes.2013.06.017, 2013.

Bao, X., Song, X., and Li, J.: High-resolution lithospheric structure beneath Mainland China from ambient noise and earthquake surface-wave tomography, Earth Planet. Sci. Lett., 417, 132–141, https://doi.org/10.1016/j.epsl.2015.02.024,

680   2015.

Beaty, K. S., Schmitt, D. R., and Sacchi, M.: Simulated annealing inversion of multimode Rayleigh wave dispersion curves for geological structure, Geophys. J. Int., 151, 622–631, https://doi.org/10.1046/j.1365-246X.2002.01809.x, 2002.

Blonk, B. and Herman, G. C.: Inverse scattering of surface waves: A new look at surface Consistency, Geophysics, 59, 963–972, https://doi.org/10.1190/1.1443656, 1994.

685   Boiero, D. and Socco, L.V.: Retrieving lateral variations from surface wave dispersion curves, Geophysical Prospecting, 58, 977-996, https://doi.org/10.1111/j.1365-2478.2010.00877.x, 2010.

Boiero, D.: Surface wave analysis for building shear wave velocity models, Ph.D. thesis, Politecnico di Torino, 233 pp., 2009.

Boschi, L., and Ekström, G.: New images of the Earth's upper mantle from measurements of surface wave phase-velocity

690   anomalies: J. Geophys. Res. Solid Earth, 107, 1-14, https://doi.org/10.1029/2000JB000059, 2002.

Comina, C., Foti, S., Boiero, D., and Socco, L. V.: Reliability of VS,30 evaluation from surface-wave tests, J. Geotech. Geoenviron. Eng., 137, 579–586, https://doi.org/10.1061/(asce)gt.1943-5606.0000452, 2011.

Da Col, F., Papadopoulou, M., Koivisto, E., Sito, Ł., Savolainen, M., and Socco, L. V.: Application of surface-wave tomography to mineral exploration: a case study from Siilinjärvi, Finland, Geophys. Prospect., 68, 254–269,

695   https://doi.org/10.1111/1365-2478.12903, 2020.

Ernst, F. E., Herman, G. C., and Ditzel, A.: Removal of scattered guided waves from seismic data, Geophysics, 67, 1240–1248, https://doi.org/10.1190/1.1500386, 2002.

Fang, H., Yao, H., Zhang, H., Huang, Y.-C., and van der Hilst, R. D.: Direct inversion of surface wave dispersion for three-dimensional shallow crustal structure based on ray tracing: methodology and application, Geophys. J. Int., 201, 1251–1263, https://doi.org/10.1093/gji/ggv080, 2015.

Feng, S., Sugiyama, T., and Yamanaka, H.: Effectiveness of multi-mode surface wave inversion in shallow engineering site investigations, Explor. geophys., 36, 26–33, https://doi.org/10.1071/eg05026, 2005.

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

Foti, S., Lai, C. G., Rix, G. J., and Strobbia, C.: Surface wave methods for near-surface site characterization, 1st ed., CRC Press, London, England, 2017.

Garofalo, F., Foti, S., Hollender, F., Bard, P. Y., Cornou, C., Cox, B. R., Dechamp, A., Ohrnberger, M., Perron, V., Sicilia, D., Teague, D., and Vergniault, C.: InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part II: Inter-comparison between surface-wave and borehole methods, Soil Dyn. Earthq. Eng., 82, 241–254, https://doi.org/10.1016/j.soildyn.2015.12.009, 2016.

Halliday, D. F., Curtis, A., Vermeer, P., Strobbia, C., Glushchenko, A., van Manen, D.-J., and Robertsson, J. O.: Interferometric ground-roll removal: Attenuation of scattered surface waves in single-sensor data, Geophysics, 75, SA15–SA25, https://doi.org/10.1190/1.3360948, 2010.

Karimpour, M.: processing workflow for estimation of dispersion curves from seismic data and QC = Extraction of Dispersion Curves from Field Data, MSc thesis, Politecnico di Torino, 71 pp., 2018.

Karimpour, M., Slob, E., Socco, L.V.: Comparison of straight-ray and curved-ray surface wave tomography approaches in near-surface studies, Solid Earth, 13, 1569-1583, https://doi.org/10.5194/se-13-1569-2022, 2022.

Kennett, B. L. N. and Yoshizawa, K.: A reappraisal of regional surface wave tomography, Geophys. J. Int., 150, 37–44, https://doi.org/10.1046/j.1365-246x.2002.01682.x, 2002.

Khosro Anjom, F.: S-wave and P-wave velocity model estimation from surface waves, Ph.D. thesis, Politecnico di Torino, 165 pp., 2021.

Khosro Anjom, F., Teodor, D., Comina, C., Brossier, R., Virieux, J., and Socco, L. V.: Full-waveform matching of VP and VS models from surface waves, Geophys. J. Int., 218, 1873–1891, https://doi.org/10.1093/gji/ggz279, 2019.

Khosro Anjom, F., Browaeys, T. J., and Socco, L. V.: Multimodal surface-wave tomography to obtain S- and P-wave velocities applied to the recordings of unmanned aerial vehicle deployed sensors, Geophysics, 86, R399–R412, https://doi.org/10.1190/geo2020-0703.1, 2021.

Khosro Anjom, F., Arabi, A., Socco, L.V. and Comina, C.: Application of a method to determine S and P wave velocities from surface waves data analysis in presence of sharp lateral variations, 36th GNGTS national convention, 632-635, https://hdl.handle.net/11583/2740539, 2017.

730   Lai, C.G.: Simultaneous inversion of Rayleigh phase-velocity and attenuation for near-surface site, Ph.D. thesis, Georgia Institute of Technology, 1998

Lys, P.-O., Elder, K., Archer, J., and the METIS Team: METIS, a disruptive R&D project to revolutionize land seismic acquisition, in: RDPETRO 2018: Research and Development Petroleum Conference and Exhibition, Abu Dhabi, UAE, 9-10 May 2018, 2018.

735   Mari, J. L.: Estimation of static corrections for shear-wave profiling using the dispersion properties of Love waves, Geophysics, 49, 1169–1179, https://doi.org/10.1190/1.1441746, 1984.

Mordret, A., Landès, M., Shapiro, N. M., Singh, S. C., and Roux, P.: Ambient noise surface wave tomography to determine the shallow shear velocity structure at Valhall: depth inversion with a Neighbourhood Algorithm, Geophys. J. Int., 198, 1514–1525, https://doi.org/10.1093/gji/ggu217, 2014.

740   Neducza, B.: Stacking of surface waves, Geophysics, 72, 51-58, https://doi.org/10.1190/1.2431635, 2007.

Papadopoulou, M: Surface-wave methods for mineral exploration, Ph.D. thesis, Politecnico di Torino, 2021.

~~Papadopoulou, M., Da Col, F., Mi, B., Bäckström, E., Marsden, P., Brodic, B., Malehmir, A., and Socco, L. V.: Surface wave analysis for static corrections in mineral exploration: A case study from central Sweden, Geophys. Prospect., 68, 214-231, https://doi.org/10.1111/1365-2478.12895, 2020.~~

745   Park, C. B.: ParkSEIS-3D for 3D MASW Surveys, Fast Times, 24, 2019.

Park, C. B., Miller, R. D., and Xia, J.: Imaging dispersion curves of surface waves on multi-channel record, in: SEG Technical Program Expanded Abstracts 1998, 1998.

Picozzi, M., Parolai, S., Bindi, D., and Strollo, A.: Characterization of shallow geology by high-frequency seismic noise tomography, Geophys. J. Int., 176, 164–174, https://doi.org/10.1111/j.1365-246x.2008.03966.x, 2009.

750   Ritzwoller, M. H., and Levshin, A. L.: Eurasian surface wave tomography: Group velocities, J. Geophys. Res. Solid Earth, 103, 4839-4878, https://doi.org/10.1029/97JB02622, 1998.

Roy, S., Stewart, R. R., and Al Dulaijan, K.: S-wave velocity and statics from ground-roll inversion, Lead. Edge, 29, 1250–1257, https://doi.org/10.1190/1.3496915, 2010.

Shapiro, N. M., Campillo, M., Stehly, L., and Ritzwoller, M. H.: High-resolution surface-wave tomography from ambient
755   seismic noise, Science, 307, 1615–1618, https://doi.org/10.1126/science.1108339, 2005.

Socco, L. V. and Boiero, D.: Improved Monte Carlo inversion of surface wave data, Geophys. Prospect., 56, 357–371, https://doi.org/10.1111/j.1365-2478.2007.00678.x, 2008.

Socco, L. V. and Comina, C.: Time-average velocity estimation through surface-wave analysis: Part 2 — P-wave velocity, Geophysics, 82, U61–U73, https://doi.org/10.1190/geo2016-0368.1, 2017.

760   Socco, L. V., Boiero, D., Foti, S., and Wisén, R.: Laterally constrained inversion of ground roll from seismic reflection records, Geophysics, 74, G35–G45, https://doi.org/10.1190/1.3223636, 2009.

Socco, L. V., Boiero, D., Bergamo, P., Garofalo, F., Yao, H., Hilst, R. D. van der, and Da Col, F.: Surface wave tomography to retrieve near surface velocity models, in: SEG Technical Program Expanded Abstracts 2014, 2014.

Socco, L. V., Comina, C., and Khosro Anjom, F.: Time-average velocity estimation through surface-wave analysis: Part 1 — S-wave velocity, Geophysics, 82, U49–U59, https://doi.org/10.1190/geo2016-0367.1, 2017.

Wang, L., Xu, Y., and Luo, Y.: Numerical Investigation of 3D multichannel analysis of surface wave method, J. Appl. Geophy., 119, 156–169, https://doi.org/10.1016/j.jappgeo.2015.05.018, 2015.

Wespestad, C. E., Thurber, C. H., Andersen, N. L., Singer, B. S., Cardona, C., Zeng, X., Bennington, N. L., Keranen, K., Peterson, D. E., Cordell, D., Unsworth, M., Miller, C., and Williams-Jones, G.: Magma reservoir below Laguna del Maule volcanic field, Chile, imaged with surface-wave tomography, J. Geophys. Res. Solid Earth, 124, 2858–2872, https://doi.org/10.1029/2018jb016485, 2019.

Wisén, R., and Christiansen, V.: Laterally and Mutually Constrained Inversion of Surface Wave Seismic Data and Resistivity Data, J. Env. and Eng. Geophy., 10, 251-262, https://doi.org/10.2113/JEEG10.3.251, 2005.

Wisén, R., Auken, E., and Dahlin, T.: Combination of 1D laterally constrained inversion and 2D smooth inversion of resistivity data with a priori data from boreholes, Near Surf. Geophys., 3, 71–79, https://doi.org/10.3997/1873-0604.2005002, 2005.

Xia, J., Miller, R.D., and Park, C.B., Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves, Geophysics, 64, 691-700, https://doi.org/10.1190/1.1444578, 1999.

Xia, J.: Estimation of near-surface shear-wave velocities and quality factors using multichannel analysis of surface-wave methods, Journ of App. Geophy., 103, 140-151, https://doi.org/10.1016/j.jappgeo.2014.01.016, 2014.

Xia, J., Miller, R. D., Park, C. B., Hunter, J. A., Harris, J. B., and Ivanov, J.: Comparing shear-wave velocity profiles inverted from multichannel surface wave with borehole measurements, Soil Dyn. Earthq. Eng., 22, 181–190, https://doi.org/10.1016/s0267-7261(02)00008-8, 2002.

Xia, J., Miller, R. D., Xu, Y., Luo, Y., Chen, C., Liu, J., Ivanov, J., and Zeng, C.: High-frequency Rayleigh-wave method, J. Earth Sci., 20, 563–579, https://doi.org/10.1007/s12583-009-0047-7, 2009.

Yao, H., Beghein, C., and Van Der Hilst, R.: Surface wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - II. Crustal and upper-mantle structure, Geophys. J. Int., 173, 205-219, https://doi.org/10.1111/j.1365-246X.2007.03696.x, 2008.