Retrieving 2-D laterally varying structures from multistation surface wave dispersion curves using multiscale window analysis

Geophysical Journal International

Geophys. J. Int. (2021) **227,** 1418–1438
Advance Access publication 2021 July 20
GJI Seismology

https://doi.org/10.1093/gji/ggab282

Retrieving 2-D laterally varying structures from multistation surface wave dispersion curves using multiscale window analysis

Shufan Hu [1], Yonghui Zhao [1], Laura Valentina Socco[2] and Shuangcheng Ge[3]

[1] *School of Ocean and Earth Science, Tongji University, Shanghai* 200092, *China. E-mail: zhaoyh@tongji.edu.cn*
[2] *Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Torino* 10129, *Italy*
[3] *Zhejiang University of Water Resources and Electric Power, Hangzhou* 310018, *China*

Accepted 2021 July 13. Received 2021 April 21; in original form 2020 October 27

SUMMARY

The analysis of multistation surface wave records is of increasing popularity in imaging the structure of the Earth due to its robustness on dispersion measurement. Since the representation of multistation surface wave dispersion curves (DCs) is uncertain in laterally varying media, average information beneath the receiver array is assumed to be obtained by inverting the dispersion curves with a horizontally layered model. To retrieve a more realistic 2-D laterally varying structure, we present a multiscale window analysis of surface waves (MWASW) method for analysing 2-D active-source surface wave data. The MWASW method is based on the use of a forward algorithm for calculating the theoretical DCs over 2-D models and multisize spatial windows for estimating the dispersion data. The forward algorithm calculates the theoretical dispersion considering the lateral variation beneath the receiver array; hence, the estimated DC is not treated as representative of the average properties but as data containing the lateral variation information. By inverting the dispersion data extracted from different spatial windows, the subsurface information at different depth ranges and lateral extensions are integrated to produce a shear wave velocity model. The dispersion curves analysed from smaller spatial windows retrieve the shallow structure with a higher lateral resolution, whereas the phase velocity data from larger spatial windows provide average information with a greater depth. We test the effectiveness of the MWASW method using three synthetic examples and two field data sets. Both results show the improved lateral resolution of the *S*-wave velocity structure retrieved with the MWASW method compared to the traditional multistation method in which the local horizontally layered model is adopted.

Key words: Structure of the Earth; Inverse theory; Surface waves and free oscillations.

1 INTRODUCTION

Surface waves represent the interference phenomenon of body waves and usually are the dominant events in the seismogram. As surface waves are dispersive (commonly represented by a dispersion curve) in vertically heterogeneous media, they can be analysed to infer the relevant medium properties, mainly the *S*-wave velocity (*Vs*) near the surface. Currently, surface wave analysis has become a foundational tool for building *Vs* models of the subsurface in fields of near-surface geophysics (Socco *et al.* 2011; Garofalo *et al.* 2016; Adly *et al.* 2017), exploration geophysics (Mordret *et al.* 2013; Mi *et al.* 2020; Papadopoulou *et al.* 2020), as well as regional and global seismology (Fang *et al.* 2019; Zhai *et al.* 2019; Kaviani *et al.* 2020). For 2-D active-source seismic records, there are generally two kinds of methods, the two-station method and the multistation method, to analyse surface waves. Although both methods can determine the phase velocity of surface waves from seismic records, they retrieve dispersion curves with different qualities in complex environments and diverse meanings in laterally inhomogeneous media.

The two-station method, also known as the spectral analysis of surface waves (SASW) method (Stokoe *et al.* 1994) in near-surface applications, evaluates dispersion curves by estimating, for each frequency component, the phase difference between the signals observed at two receivers. Since the evaluated dispersion curve represents path averaged phase dispersion, it is possible to retrieve subsurface lateral variations using the tomography scheme (Yin *et al.* 2016; Ikeda & Tsuji 2020). However, for active-source seismic records, the dispersion curves estimated by the two-station method are very sensitive to individual geophone coupling, as well as random and coherent noise such as body waves and higher modes (Neducza 2007; Hashemi Jokar *et al.* 2019). These difficulties in extracting the dispersion curve may be the reason why surface wave tomography is seldom used for active-source seismic data (Ikeda & Tsuji 2020).

Rather than using only the signals observed at two stations, the multistation method, identified by the multichannel analysis of surface waves (MASW) method (Park *et al.* 1999; Xia *et al.* 1999), transforms the time-space domain data acquired at several receivers into another domain (e.g. the frequency–wavenumber domain) to reconstruct a dispersion image. This dispersion image allows the identification of dispersion trends based on spectral maxima and supplies more robust and accurate dispersion curves (Socco & Strobbia 2004). On the other hand, the dispersion curve evaluated by the multistation method cannot be representative of any specific velocity of the subsurface if strongly lateral variations are presented (Boiero & Socco 2011). In addition, the extracted dispersion curve is usually assigned at the mid-point of the receiver array and inverted with a horizontally layered model, finally introduces errors into spatially 2-D *Vs* imaging. To mitigate the adverse effects of the horizontally layered model assumption, several authors propose approaches to reconstruct a local dispersion curve by applying the common-midpoint cross-correlation and data windowing techniques (Bohlen *et al.* 2004; Hayashi & Suzuki 2004; Grandjean & Bitri 2006; Neducza 2007; Bergamo *et al.* 2012; Ikeda *et al.* 2013; Vignoli *et al.* 2016). These methods tend to improve the lateral resolution thus allowing lateral variations to be retrieved but still neglect the effect of lateral variations on each dispersion curve extracted from a receiver window. Xia *et al.* (2005) present a lateral unblurring processing technique by generalized inversion to reduce acquisition geometry-dependence in pseudo-2-D sections. They also state that the data kernel used in generalized inversion is critical for obtaining meaningful unblurred results. Alternatively, full-waveform inversion is able to exploit the whole information recorded at several stations and theoretically take into account any lateral heterogeneity; but it can easily fall into a local minimum due to the cycle-skipping problem (Virieux & Operto 2009). To avoid this issue, several strategies, such as windowed-amplitude waveform inversion (Pérez Solano *et al.* 2014) and wave equation dispersion inversion (Li *et al.* 2017, 2019), are introduced to obtain a compromise between resolution and stability. Nevertheless, wave equation-based seismic modeling is associated with a high computational cost, which has led researchers to look for more practical and time-efficient solutions (Pasquet & Bodet 2017).

As we can see, despite the limited effectiveness in laterally varying media, the MASW method is popular by its robustness in determining phase velocity and high efficiency (Xia 2014). Moreover, it is clear that dispersion curves measured from a shorter spread can provide higher lateral resolution information concerning the shallow stratum, while the phase velocity data evaluated from a longer spread supply subsurface average information with greater depth (Boiero & Socco 2011; Vignoli *et al.* 2016; Pasquet *et al.* 2021). Therefore, if the meaning of the phase velocity determined by the multistation method is understood in laterally varying sites, using and inverting the dispersion data obtained by multisize spatial windows may help to achieve a more comprehensive understanding of the subsurface.

The goal of this paper is to develop a new method, which is both robust in dispersion measurement and applicable for laterally varying models, for analysing 2-D active-source surface wave data. First, by using the beamforming, we analyse what the multistation surface wave dispersion curves (DCs) represent. Based on this analysis, we calculate the theoretical DCs over 2-D models by searching the local maxima on the beamformer output. Then, we simultaneously invert the dispersion data extracted from multisize spatial windows using the Gauss–Newton optimization algorithm. In the inverse problem, we use the total variation (TV) regularization (Rudin *et al.* 1992) to stabilize the inversion process and preserve sharp edges in the retrieved *Vs* model. At the same time, we utilize the multifrequency data weighing (Hu *et al.* 2009) to prevent the high-frequency data from dominating the inversion process. Finally, we use the synthetic and field data sets to test the performance of the proposed algorithm.

2 METHOD

In this section, we analyse the meaning of the multistation DCs and present the corresponding forward algorithm. Then, we introduce the multiscale window analysis of surface waves (MWASW) method and the algorithm for inverting dispersion data extracted from multisize spatial windows. Finally, we describe the workflow for implementing the proposed MWASW method.

2.1 The meaning of multistation DCs

For 2-D active-source seismic records, the frequency–wavenumber method (Rix *et al.* 2002; Zywicki & Rix 2005) is a common technique to estimate surface wave dispersion. If $\hat{S}_l(\omega)$ represents the Fourier spectrum of the signals observed at the lth receiver with an offset of x_l, the output $Z(\omega, k)$ of the conventional frequency-domain beamforming for L receivers in a linear array is calculated by (Johnson & Dugeon,1993):

$$Z(\omega, k) = \sum_{l=1}^{L} w_l \hat{S}_l(\omega) e^{jkx_l} = \mathbf{e}^{\mathrm{H}}(k)\,\mathbf{W}\mathbf{S}(\omega), \tag{1}$$

where, ω is the angular frequency (rad s^{-1}), k is a trial wavenumber (rad m^{-1}), H denotes the Hermitian transpose, $\mathbf{W}$ is a diagonal matrix containing shading weights w_l for each receiver, $\mathbf{S}(\omega) = [\hat{S}_1(\omega), \hat{S}_2(\omega), \cdots, \hat{S}_L(\omega)]^{\mathrm{T}}$ is a column vector composed of the Fourier spectral of recorded signals (the superscript T indicates the transpose of a vector), and $\mathbf{e}(k) = [e^{-jkx_1}, e^{-jkx_2}, \cdots, e^{-jkx_L}]^{\mathrm{T}}$ is a steering vector.

According to eq. (1), the output of the conventional frequency-domain beamforming is equal to the one produced by the 2-D fast Fourier transform (Nolet & Panza 1976) and the p–ω transform (McMechan & Yedlin 1981) if the shading weights of each receiver are equal to $w_l = 1$. In addition, the conventional frequency-domain beamforming is equivalent to the phase shift method (Park *et al.* 1998) if shading weights are set to $w_l = 1/|\hat{S}_l(\omega)|$ to compensate for the effects of attenuation and geometrical spreading.

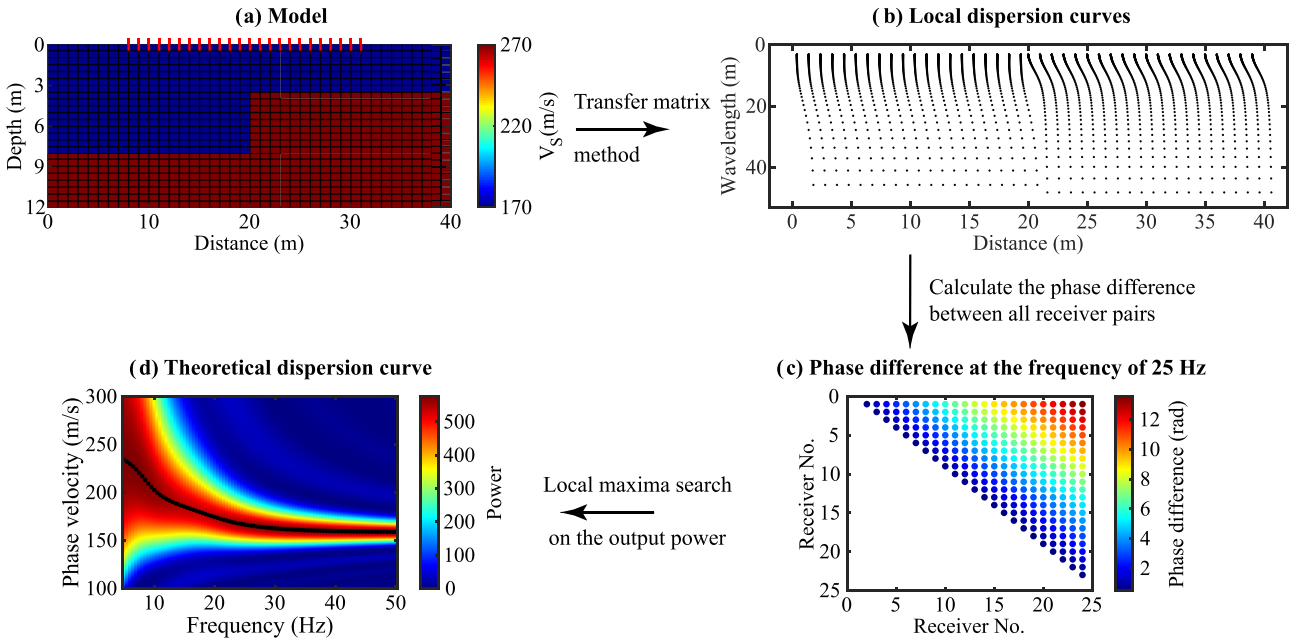

Figure 1. Workflow for the forward algorithm. (a) A cell-based model. The red lines indicate the location of the receivers. (b) Local dispersion curves. (c) Phase difference between all receiver pairs at the frequency of 25 Hz. (d) Representation of the output power calculated by eq. (6). The black dotted line is the theoretical DC obtained from the local maxima search.

The power in the beamformer output is given by the steered response power spectrum:

$$\begin{aligned} P(\omega, k) &= Z(\omega, k)\, Z(\omega, k)^{\mathrm{H}} \\ &= \mathbf{e}^{\mathrm{H}}(k)\, \mathbf{WS}(\omega)\, \mathbf{S}^{\mathrm{H}}(\omega)\, \mathbf{W}^{\mathrm{H}} \mathbf{e}(k) \\ &= \mathbf{e}^{\mathrm{H}}(k)\, \mathbf{WR}(\omega)\, \mathbf{W}^{\mathrm{H}} \mathbf{e}(k), \end{aligned} \tag{2}$$

where, $\mathbf{R}(\omega)$ is the spatio-spectral correlation matrix that is assembled by the cross-power spectra of all receiver pair combinations.

As mentioned, we can normalize the amplitude of each receiver to compensate for the effects of attenuation and geometric spreading, that is $\mathbf{W} = \mathbf{diag}[1/|\hat{S}_l(\omega)|]$. In this way, eq. (2) is simplified to:

$$\begin{aligned} P(\omega, k) &= \sum_{m=1}^{L} \sum_{n=1}^{L} w_m w_n \left|\hat{S}_m(\omega)\right| \left|\hat{S}_n(\omega)\right| e^{j[\Delta\varphi_{mn}(\omega) - k(x_n - x_m)]} \\ &= L + 2 \sum_{m=1}^{L-1} \sum_{n=m+1}^{L} \cos[\Delta\varphi_{mn}(\omega) - k x_{mn}], \end{aligned} \tag{3}$$

where, $\Delta\varphi_{mn}(\omega)$ is the phase difference between receiver m and receiver n, and x_{mn} is the distance between receiver m and receiver n. From this point of view, the spectral maxima in the power spectrum calculated by the amplitude-normalized beamforming is a function of the phase difference between all receiver pairs, therefore, the relevant medium properties under the receiver array. Accordingly, if we can calculate the phase difference between each receiver pair in 2-D models, it is possible to determine the theoretical wavenumber (or phase velocity) corresponding to the spectral maxima in the beamformer output.

2.2 Forward algorithm of multistation DCs

If the seismic data consist of a single noise-free surface wave mode and $s_m(t)$ and $s_n(t)$ represent two signals recorded by two receivers of a linear acquisition system, then the Fourier transform of the second signal after geometric spreading correction $\hat{S}_n(\omega)$ can be expressed in terms of the Fourier transform of the first signal $\hat{S}_m(\omega)$ (Kulesh *et al.* 2008; Askari & Ferguson 2012):

$$\hat{S}_n(\omega) = e^{-jk_{mn}(\omega)x_{mn}}\, e^{-A_{mn}(\omega)x_{mn}}\, \hat{S}_m(\omega), \tag{4}$$

in which $k_{mn}(\omega)$ represents the spatial wavenumber that controls wave propagation from receiver one to receiver two, and $A_{mn}(\omega)$ is the frequency-dependent attenuation function.

For the forward problem, we parametrize the 2-D subsurface model into several rectangular cells whose widths Δx and thicknesses Δy are fixed, as shown in Fig. 1(a). Each cell is defined in terms of *Vs*, *P*-wave velocity (*Vp*) and density. According to eq. (4), the phase

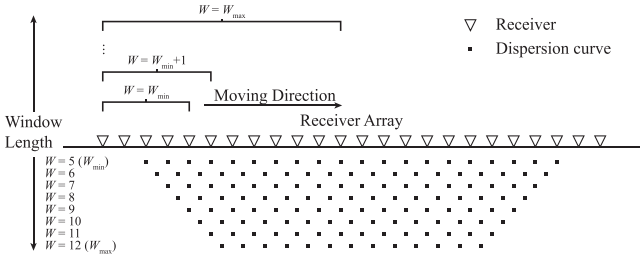

Figure 2. Sequence of dispersion analysis. Note that each black dot under the receiver array indicates one dispersion curve.

difference $\Delta\varphi_{mn}(\omega)$ can be expressed as:

$$\Delta\varphi_{mn}(\omega) = k_{mn}(\omega)\, x_{mn} = \sum_{q=1}^{Q} k_q(\omega)\, \Delta x, \tag{5}$$

where $k_q(\omega)$ is the local spatial wavenumber determined by the model cells of the qth column between receivers m and n.

Substituting eq. (5) into eq. (3) and using $v = \omega/k$, the relationship between the output power calculated by the amplitude-normalized beamforming and the local phase velocity is expressed by:

$$\begin{aligned} P(\omega, v) &= L + 2\sum_{m=1}^{L-1}\sum_{n=m+1}^{L} \cos\left(\Delta\varphi_{mn}(\omega) - \omega\frac{x_{mn}}{v}\right) \\ &= L + 2\sum_{m=1}^{L-1}\sum_{n=m+1}^{L} \cos\left[\omega\left(\sum_{q=1}^{Q}\frac{\Delta x}{v_q(\omega)} - \frac{x_{mn}}{v}\right)\right], \end{aligned} \tag{6}$$

where v and $v_q(\omega)$ are the trial and local phase velocities, respectively.

Since the experimental dispersion curves are always identified by the spectral maxima of the dispersion image, we solve the forward problem using the following three steps, as shown in Fig. 1:

1. Calculating the local phase velocity $v_q(\omega)$ for each column of model cells using the transfer matrix method (Haskell 1953; Thomson 1950).
2. Converting the local phase velocity $v_q(\omega)$ to the local spatial wavenumber $k_q(\omega)$, and calculating the phase difference between all receiver pairs according to eq. (5).
3. Substituting the calculated phase difference into eq. (6), and searching the spectral maximum $v(\omega)$ of the output power using the Golden Section search algorithm.

As a model with a complex structure, eq. (6) may have more than one local maxima, as shown in Fig. 1(d). To obtain a theoretical dispersion curve that can be compared with the experimental one, we carefully design the local maximum search in step 3 that considers the continuity of the dispersion curve. A validation of the proposed forward algorithm is given in Appendix A. We should note that the proposed algorithm is able to calculate the single mode of multistation DCs, including fundamental and high modes, as long as we compute the corresponding local phase velocity in step 1.

2.3 Multiscale window analysis

As mentioned, the length of the receiver spread affects the available wavelength-band (investigation depth) of the estimated phase velocity data and can lead to different lateral resolutions in the retrieved *Vs* model. Therefore, we implement a strategy of multiscale window analysis to achieve a considerable lateral resolution and investigation depth at the same time.

The multiscale window analysis is based on the work of Socco *et al.* (2009), that is using the data windowing (Bohlen *et al.* 2004) to exploit the data redundancy and the dispersion stacking (Grandjean & Bitri 2006; Neducza 2007) to increase the signal-to-noise ratio (SNR) of the dispersion image, except that we use spatial windows with multiple sizes to extract the local wavefield as well as the amplitude-normalized beamforming presented in eq. (3) to generate the dispersion image. The processing parameters are the minimum window length $W_{\min}$, the maximum window length $W_{\max}$, the shift between two successive windows ΔW, and the maximum offset d (source – furthest receiver distance). We determine these parameters by tests according to the workflow described in Socco *et al.* (2009). The principle of data windowing and dispersion stacking and the processing parameter selection criteria are given in Appendix B. Once the processing parameters are defined, dispersion curves with uncertainties are estimated at all potential positions of the spatial window, and several dispersion curves extracted from different windows are possible at the same midpoint, as shown in Fig. 2.

The truncation in the space domain causes leakage in the wavenumber spectral analysis (Lin & Chang 2004), which leads to inaccuracy for estimating phase velocities with wavelengths longer than the receiver window (Park and Carnevale 2010). Therefore, we limit each

estimated dispersion curve according to the criterion of $\lambda < (N_{\rm rec}-1)\Delta r$, where λ is the wavelength, $N_{\rm rec}$ is the number of receivers in the spatial window and Δr is the receiver spacing, that is only the dispersion data with a wavelength shorter than the window length is acceptable.

2.4 Inversion algorithm

To invert the dispersion curves we use a local optimization approach, the Gauss–Newton method, for iteratively updating the model parameters from a given initial value.

To minimize non-linearity and impose positivity, we apply logarithmic data and parameters, as in Auken & Christiansen (2004). The data vector, which consists of surface wave phase velocity $d^{\rm obs}(W_i, \omega_j)$ within the spatial window W_i and corresponding frequency ω_j, is represented by:

$$\mathbf{d}^{\rm obs} = \left[\log\left(d^{\rm obs}\left(W_i, \omega_j\right)\right), i = 1, 2, \ldots, I; j = 1, 2, \ldots J\right]^{\rm T}, \tag{7}$$

where I is the number of spatial windows and J is the number of frequencies.

Since the dispersion curve is more sensitive to *Vs* than *P*-wave velocity (*Vp*) or density (Gabriels *et al.* 1987; Xia *et al.* 1999), we assign *Vp* and density *a priori* value in the inversion. The model vector is a collection of *Vs* in each cell:

$$\mathbf{m} = \left[\log\left(V_s\left(x_p, y_q\right)\right), p = 1, 2, \ldots, P; q = 1, 2, \ldots, Q\right]^{\rm T}, \tag{8}$$

where x_p and y_q denote the center of the 2-D discretization cell.

Under the definition of data vector (eq. 7) and model vector (eq. 8), the non-linear inverse problem is described by the following operator equation:

$$\mathbf{d}^{obs} = \mathbf{F}\left[\mathbf{m}\right] + \varepsilon, \tag{9}$$

where $\mathbf{F}$ is the forward operator and ε denotes the measurement error.

Then, we define the objective function $\phi(\mathbf{m})$ based on the data misfit $\phi_d(\mathbf{m})$ and the TV regularization $\phi_m(\mathbf{m})$:

$$\begin{aligned}\phi\left(\mathbf{m}\right) &= \phi_d\left(\mathbf{m}\right) + \lambda\phi_m\left(\mathbf{m}\right)\\ &= \frac{1}{2}\|\mathbf{W}_d\left(\mathbf{F}\left[\mathbf{m}\right] - \mathbf{d}^{obs}\right)\|_2^2 + \frac{1}{2}\lambda\|\mathbf{L}\mathbf{m}\|_1,\end{aligned} \tag{10}$$

where $\mathbf{W}_d$ is the data weighting matrix; λ represents the regularization parameter, and we determine it using the adaptive regularization (Zhdanov 2009); $\mathbf{L}$ is the roughness matrix, a combination of the first-order differential operators in both the x and y directions:

$$\mathbf{L} = \begin{bmatrix} a_x\mathbf{L}_x \\ a_y\mathbf{L}_y \end{bmatrix}, \tag{11}$$

where a_* can be applied as scalars or diagonal matrices with varying weights for each parameter (Oldenburg & Li 2005). Here, a_x and a_y are scalars and equal to 1.

We choose the data weighting matrix $\mathbf{W}_d$ as the product of the reciprocal of uncertainties estimated during data processing $\mathbf{W}_\sigma$ as well as a multifrequency data weighting matrix $\mathbf{W}_\omega$, that is $\mathbf{W}_d = \mathbf{W}_\sigma\ \mathbf{W}_\omega$. The latter is used to balance the contributions from different frequency measurements and prevent high-frequency data from dominating the inversion process (Hu *et al.* 2009). The expression of the multifrequency data weighting is given by:

$$W_{\omega,j} = \frac{\omega_j^{-2}}{\sum_{j=1}^{J}\omega_j^{-2}}, \tag{12}$$

The TV regularization used in eq. (10) is an L1 regularized problem, which can be solved using the iterative method (Aster *et al.* 2005). At the kth iteration, the TV regularization can be written as $\frac{1}{2}\|\mathbf{L}\mathbf{m}^k\|_1 = \frac{1}{2}\sum_{m=1}^{M}|\gamma_m^k|$, where M is the number of rows of $\mathbf{L}\mathbf{m}^k$ and γ_m^k is the element of $\mathbf{L}\mathbf{m}^k$, then at the $k+1$th iteration:

$$\phi_m^{k+1}\left(\mathbf{m}\right) = \frac{1}{2}\mathbf{m}^{\rm T}\mathbf{L}^{\rm T}\mathbf{W}\mathbf{L}\mathbf{m}, \tag{13}$$

in which $\mathbf{W}$ is a diagonal matrix with elements

$$W_{m,m} = \begin{cases} 1/\left|\gamma_m^k\right| & \left|\gamma_m^k\right| \geq \delta \\ 1/\delta & \left|\gamma_m^k\right| < \delta \end{cases}, \tag{14}$$

where δ is a small positive number.

Finally, for optimization posed in eq. (10), the unknown model parameters are updated using the Gauss–Newton minimization as follows:

$$\mathbf{m}^{k+1} = \mathbf{m}^k + \alpha\Delta\mathbf{m}, \tag{15}$$

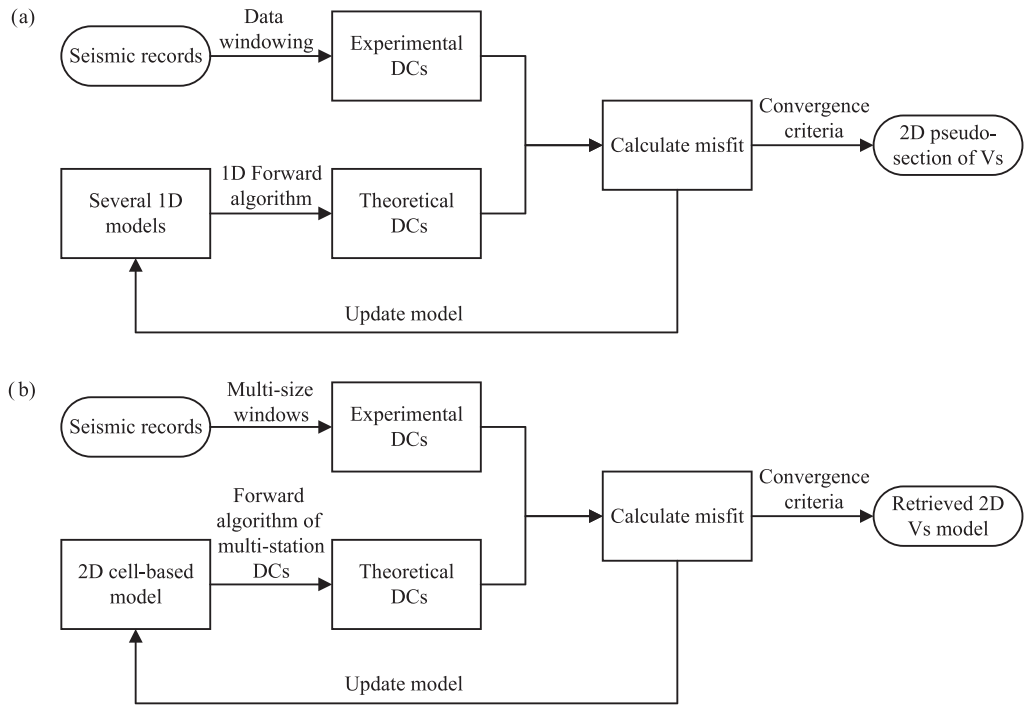

Figure 3. Flowchart of (a) the MASW and (b) the MWASW method.

where, α is a line-search parameter, which is determined by backtracking under a sufficient decrease condition, the Armijo rule (Armijo 1966); $\Delta\mathbf{m}$ is the model update and obtained by solving the following equation using the conjugate-gradient least-squares (CGLS) method (Golub & van Loan 1989):

$$\left(\mathbf{J}(\mathbf{m})^{\mathrm{T}}\mathbf{W}_d^{\mathrm{T}}\mathbf{W}_d\mathbf{J}(\mathbf{m}) + \lambda\mathbf{L}^{\mathrm{T}}\mathbf{W}\mathbf{L}\right)\Delta\mathbf{m} = \mathbf{J}(\mathbf{m})^{\mathrm{T}}\mathbf{W}_d^{\mathrm{T}}\mathbf{W}_d\left(\mathbf{d}^{obs} - \mathbf{F}(\mathbf{m})\right) - \lambda\mathbf{L}^{\mathrm{T}}\mathbf{W}\mathbf{L}\mathbf{m}, \tag{16}$$

where, $\mathbf{J}(\mathbf{m})$ is the Jacobian matrix. The explicit formula for the Jacobian matrix is given in Appendix C.

2.5 Workflow for MWASW

Fig. 3(b) depicts the workflow for the MWASW method, which can be presented in the following five steps:

1. Defining the processing parameters in the multiscale window analysis according to the tests on some representative records, and using multisize spatial windows to extract the experimental dispersion curves from 2-D active-source seismic records.
2. Parametrizing the 2-D subsurface model into several rectangular cells, and assigning an initial value to the model parameter of each rectangular cell.
3. Using the proposed forward algorithm to calculate the theoretical DCs based on the 2-D cell-based model.
4. Calculating the residual between the theoretical dispersion data and experimental dispersion data. Then, using the Gauss–Newton iteration (eq. 15) to update the model parameters.
5. Repeating steps 3 and 4 until the convergence criteria are reached. The final updated model represents the retrieved 2-D *Vs* model.

Compared to the traditional MASW method (Fig. 3a), the main improvements of the proposed MWASW are:

1. The MWASW method uses the forward algorithm of multistation DCs that takes into account the lateral variations in the 2-D subsurface model. Therefore, it is applicable in laterally varying media.
2. The MWASW method uses multisize spatial windows to extract dispersion curves. In this way, it is able to infer the deep medium properties without losing lateral resolution for the shallow stratum.

3 SYNTHETIC DATA TEST

We use three synthetic models to demonstrate the improvements and limitations of the proposed MWASW method. The models are elastic isotropic and with a size of 60 × 12 m, a constant *Vp*/*Vs* ratio of 2, and a homogenous density of 1900 kg m^{-3}. The first model is a three-layer model and we use the dispersion curve calculated by the transfer matrix method to generate the synthetic data set, which is utilized to test the performance of the proposed method in a noise-free situation. The latter two models simulate different laterally varying conditions and the dispersion curves is extracted from the waveform data generated using a 2-D finite-difference time-domain (FDTD) code (used in Qin *et al.*

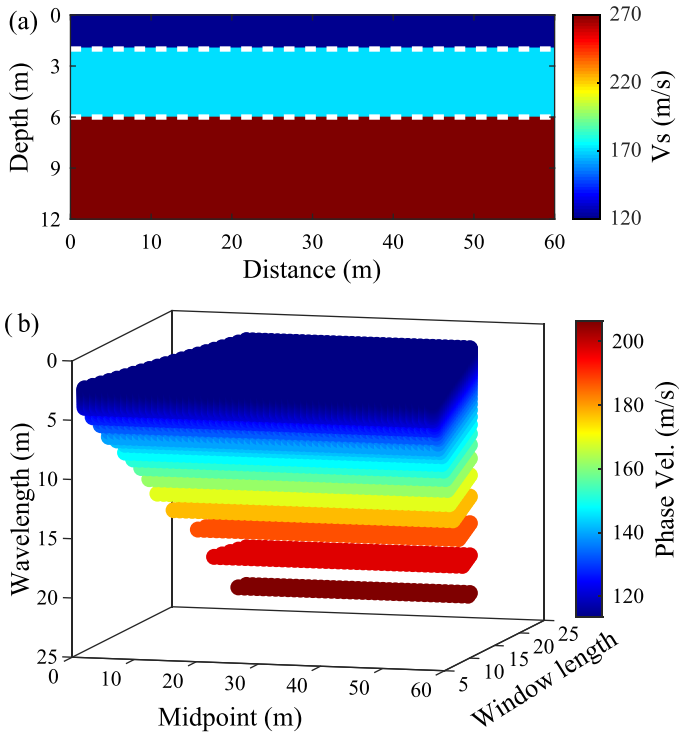

Figure 4. (a) Three-layer model. The white dashed lines indicate the depth of each layer. (b) Dispersion data of all spatial windows.

2020). In FDTD simulation, the model size is extended to 100 m × 50 m and the grid size is 0.1 m × 0.1 m. The sources are evenly distributed every 3 m on the free surface, finally simulates 21 shot gathers. A linear acquisition array, consisting of 61 receivers that are evenly spaced every 1 m, is used to record the vertical component of the particle velocity on the free surface. The exciting source is a Ricker wavelet with a center frequency of 20 Hz; the recording time window is 1 s with a sampling interval of 0.25 ms.

For all cases, we set the minimum and maximum lengths of spatial windows, $W_{\min}$ and $W_{\max}$, to 5 and 25 receivers, respectively. The length of other windows successively increases from $W_{\min}$ to $W_{\max}$. For the latter two cases, a maximum offset of $d = 40$ m is used in dispersion stacking. Moreover, we simulate or extract dispersion data in a frequency band of 5–50 Hz. As mentioned above, we cut each stacked dispersion curve using the criterion of $\lambda < (N_{\text{rec}}-1)\Delta r$. After obtaining the dispersion data of all spatial windows, we use both the proposed MWASW and the traditional MASW method to retrieve the *Vs* model. For the MASW method, only the dispersion data of the maximum window is used. In both these methods, we discretize the subsurface using 23 layers and a half-space. The thickness of each layer is 0.5 m. For the MWASW method, we set the width of the model cells to 1 m since the lateral resolution can theoretically be as high as the receiver spacing; while, the MASW only retrieves the 1-D *Vs* model at the midpoint. For comparison, we also use the TV regularization in the MASW method. The positive number δ in eq. (14) can influence the stability of inversion. We test this value in a wide range from 10^{-1} to 10^{-5} and find that a value of 10^{-3} can always achieve a stable result. Therefore, here we fix δ to be 10^{-3}. The other parameters used in the inversion of the MASW and the MWASW are also the same.

3.1 Three-layer model

Fig. 4(a) shows a simple three-layer model in which the *Vs* increases with depth. It consists of the surface layer ($Vs = 120$ m s^{-1}), the second layer ($Vs = 170$ m s^{-1}), and a half-space ($Vs = 270$ m s^{-1}). Fig. 4(b) represents the all dispersion data that are generated using the transfer matrix method.

The initial model we used in the inversion, shown in Fig. 5(a), is homogenous with a constant velocity of 206 m s^{-1}, which is the maximum phase velocity of the dispersion data. Figs 5(b) and (c) show the *Vs* model retrieved by the MASW and MWASW methods, respectively. Both inversion results suggest that the shallow velocity distribution at the depth range of 0–5 m is accurately retrieved, which validates the reliability of the MWASW method. Moreover, according to the result of the MASW, the horizontally layered model assumption can be safely applied to a situation without any lateral variation. In both retrieved *Vs* models, the velocity below a depth of 5 m presents a smooth increase, which can be seen clearly in the 1-D *Vs* profile at the position of 30 m, as shown in Fig. 5(d). These phenomenon occurred because the surface wave dispersions are sensitive to absolute *S*-wave velocities but are poor at locating interfaces (Julià *et al.* 2000). According to the result obtained from the MASW, the RMS error between the retrieved and true model is 12.8 m s^{-1}; the RMS error is 12.1 m s^{-1}, with regards to the MWASW method. This similar RMS error indicates that both methods retrieve the *Vs* model at the same level. Figs 5(e) and (f) show the phase velocity residual of the maximum window corresponds to the MASW method and the MWASW method, respectively. The residual

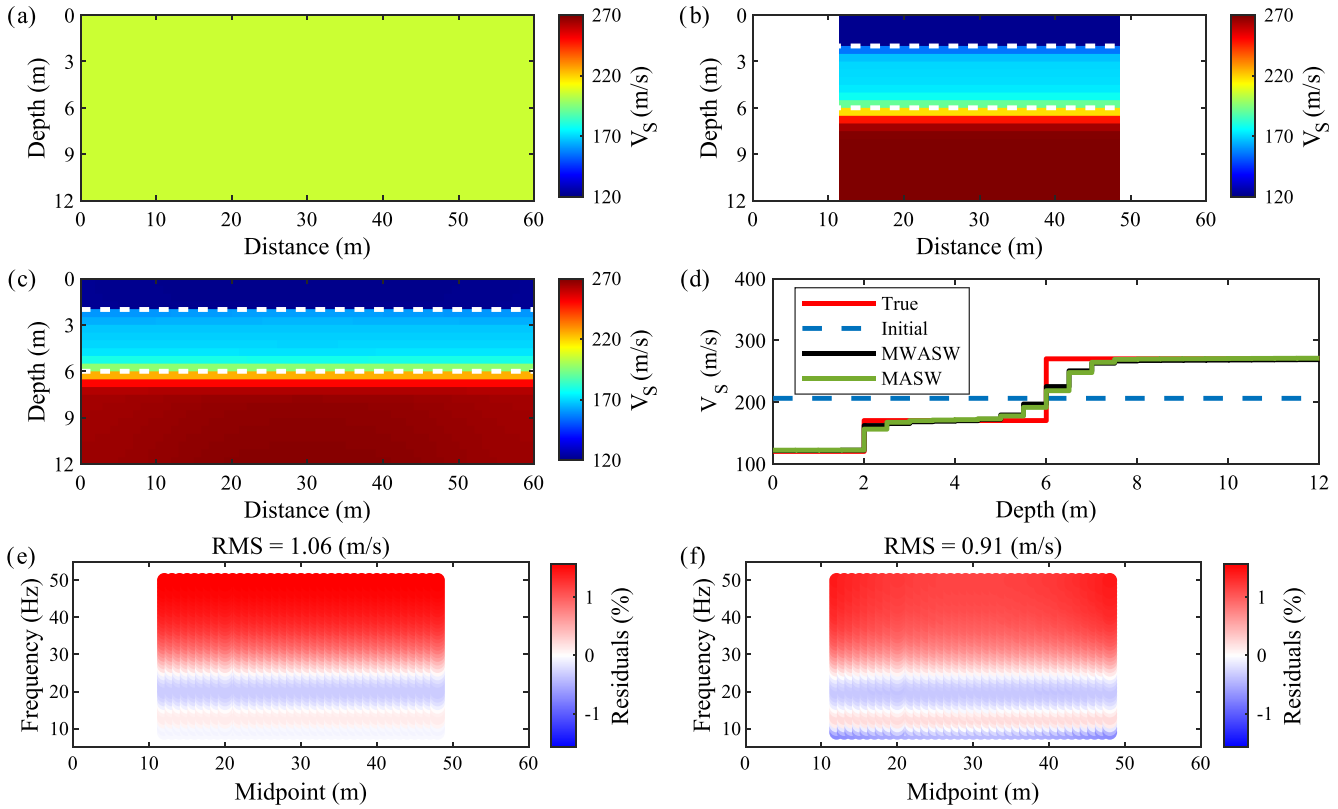

Figure 5. Inversion results of the three-layer model. (a) Initial model. (b and c) 2-D *Vs* sections retrieved via the MASW and MWASW methods, respectively. The white dashed lines depict the true depth of each layer. (d) *Vs* profile at the position of 30 m. (e and f) Phase velocity residuals for the maximum window correspond to the MASW method and the MWASW method, respectively.

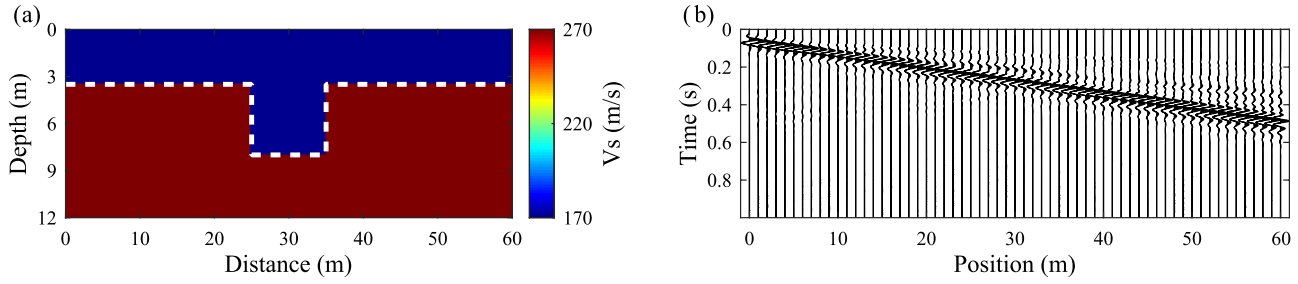

Figure 6. (a) Graben model. The white dashed line indicates the outline of the graben. (b) Simulated first common-shot gather.

plots indicate that both of them fit the dispersion data of the maximum window well. From this synthetic test, we can know that both the MASW and MWASW methods are suitable in the situation without lateral variations.

3.2 Graben model

The second model presents a graben with a length of 10 m, as shown in Fig. 6(a). This model consists of a surface layer ($Vs = 170$ m s^{-1}) and a half-space ($Vs = 270$ m s^{-1}). The simulated first common-shot gather after normalization is shown in Fig. 6(b).

Figs 7(a) and (b) show the stacked dispersion images generated using the first 5 and the first 25 receivers, respectively. The white dotted line indicates the theoretical dispersion curve of the layered model at distances between 0 and 25 m calculated by the transfer matrix method; the blue dotted line with the blue bar depicts the extracted dispersion curve with uncertainties. We can see that the smallest window has an adequate resolution to estimate the phase velocities with a wavelength shorter than the window length. Moreover, the maximum window underestimates phase velocities at the frequency band lower than 20 Hz, which might be due to the near-field effect (the stacking can mitigate the near-field effect rather than eliminate it). Fig. 7(c) represents estimated dispersion data of all spatial windows.

We build the homogenous initial model (see Fig. 8a) based on the maximum phase velocity of 218.5 m s^{-1}. The *Vs* models produced using the MASW and MWASW methods are shown in Figs 8(b) and (c), respectively. The white dashed line depicts the outline of the graben model. Although we use the TV regularization to penalize large gradient magnitudes, the model retrieved by the MASW method presents

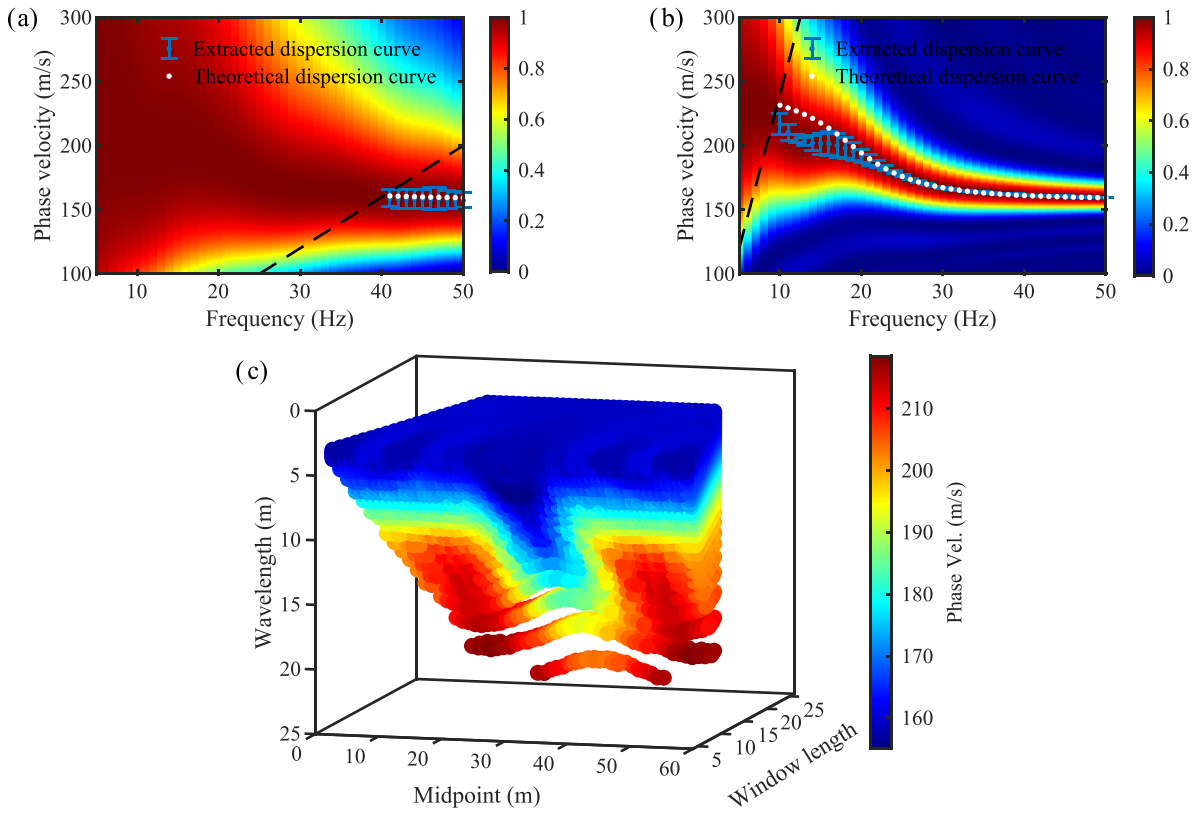

Figure 7. Panels (a) and (b) are the stacked dispersion images corresponding to the first 5 and the first 25 receivers, respectively. The white dotted line indicates the theoretical dispersion curve calculated by the transfer matrix method, the blue dotted line with the blue bar depicts the extracted dispersion curve with uncertainties, and the black dashed line represents the wavelength of $\lambda = (N_{\mathrm{rec}}-1)\Delta r$. (c) All extracted dispersion data.

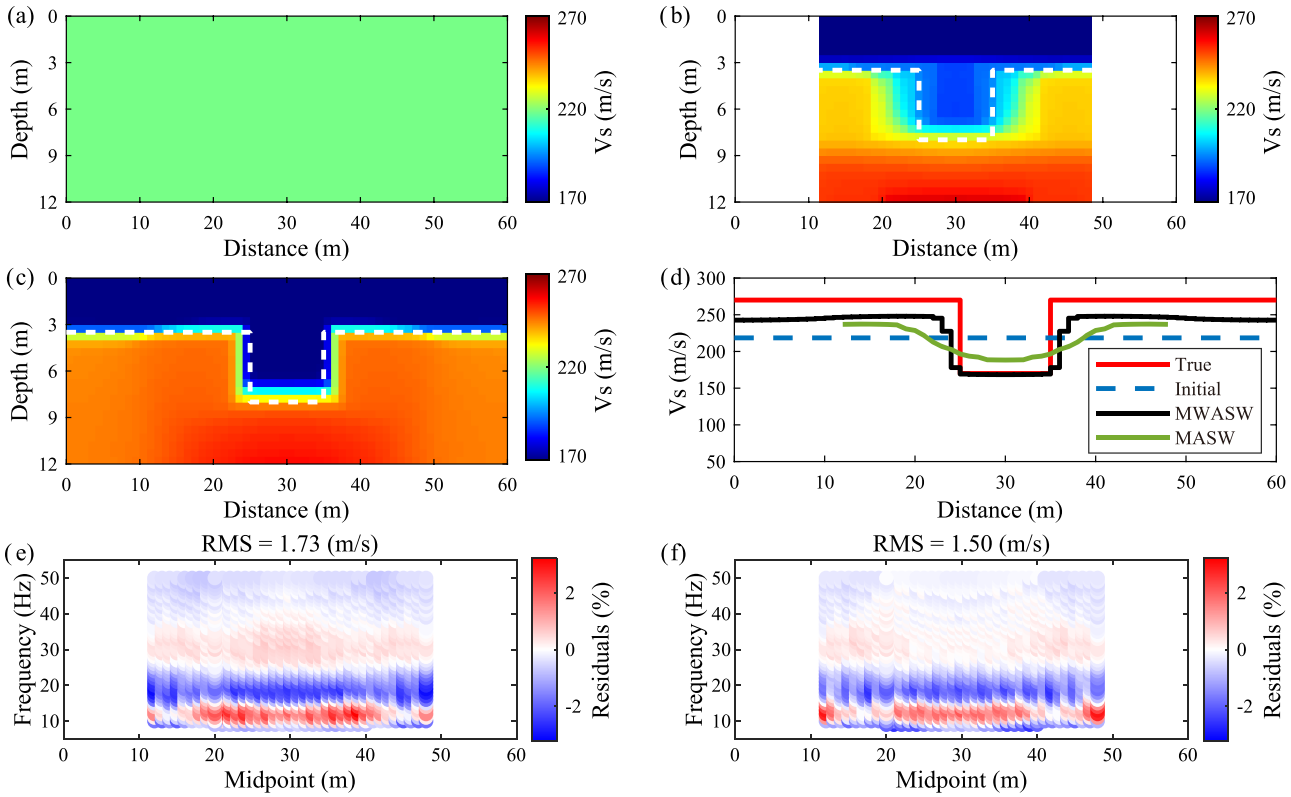

Figure 8. Inversion results of the graben model. (a) Initial model. (b and c) 2-D *Vs* sections retrieved via the MASW and MWASW methods, respectively. The white dashed lines depict the outline of the graben. (d) The plot of the velocity value for the row of model cells between the depths of 5.5 and 6 m. (e and f) Phase velocity residuals for the maximum window correspond to the MASW method and the MWASW method, respectively.

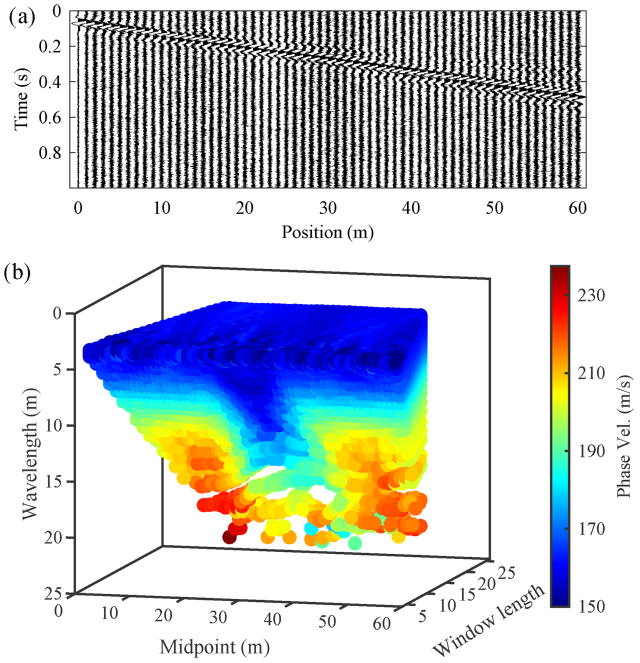

Figure 9. (a) First common-shot gather of the graben model with 5 per cent Gaussian noise added. (b) Corresponding dispersion data of all spatial windows.

a blurry image with a gradational boundary at the center of the graben. This is due to the average effect of the MASW method. The model produced by the MWASW method presents a much more uniform distribution of the *Vs* values in both regions and a significantly sharper boundary that is closer to the true model. In addition, both results underestimate the velocity of the half-space due to the underestimation of phase velocity in the low-frequency band (The RMS error between the model retrieved by the MASW method and the true model is 28.0 m s^{-1}, and the error between the model recovered by the MWASW method and the true model is 22.9 m s^{-1}). Nevertheless, the MWASW method is able to retrieve the laterally varying structure well. Fig. 8(d) shows the plots of the velocity value for the row of model cells between the depths of 5.5 and 6 m. The model retrieved by the MASW yields a smooth lateral variation from 18 to 42 m; while, the MWASW produces a model that presents an abrupt change of *Vs* at the positions of 24 and 36 m, which is closer to the actual scale of the graben model. Figs 8(e) and (f) show the phase velocity residual of the maximum window corresponds to the MASW method and the MWASW method, respectively. The RMS error between the experimental and theoretical dispersion curves of the maximum window for the MWASW method is slightly lower than that for the MASW method, which might be because the MWASW utilizes a forward algorithm that can better consider the dispersion behaviour of the 2-D model.

To study the effects of noise on the inversion results, we add Gaussian random noise at a magnitude of 5 per cent maximum amplitude to the simulated shot gather. The corresponding first common-shot gather after normalization and dispersion data are shown in Figs 9(a) and (b). Although stacking improves the SNR of the dispersion image, the phase velocity with a wavelength longer than 15 m is significantly distorted.

Fig. 10 shows the results of noisy data. The model retrieved by the MASW method is similar to the inversion result (Fig. 8b) without noise, except for the outliers in the half-space. The RMS error between the recovered and true model is 31.0 m s^{-1}. In comparison, the MWASW method produces a more compact model that is closer to the true model. The RMS error between the retrieved and true model is 21.7 m s^{-1}. The difference between the models retrieved by the MASW and MWASW can also be seen clearly in the plots of the velocity value for the row of model cells between the depths of 5.5 and 6 m (Fig. 10c). Figs 10(d) and (e) represent the phase velocity residual of the maximum window corresponds to the MASW and the MWASW methods, respectively. In the case of noise, the RMS error between the experimental and theoretical dispersion curves of the maximum window for the MWASW method is slightly higher than that for the MASW method. This phenomenon is due to the MWASW method uses more DCs and hence less dependent on the dispersion data of the maximum window than the MASW method.

3.3 Low-velocity block model

The third model consists of a low-velocity block embedded in the second layer and a normal fault (Fig. 11a). The *S*-wave velocity of the first layer is 120 m s^{-1}, the second layer is 170 m s^{-1}, the half-space is 270 m s^{-1} and the low-velocity block is 80 m s^{-1}. Fig. 11(b) shows the corresponding first common-shot gather after normalization.

Figs 12(a) and (b) represent the stacked dispersion images generated using the first 5 and the first 25 receivers, respectively. The blue dotted line with the blue bar depicts the extracted dispersion curve with uncertainties. Both dispersion images have an adequate resolution to

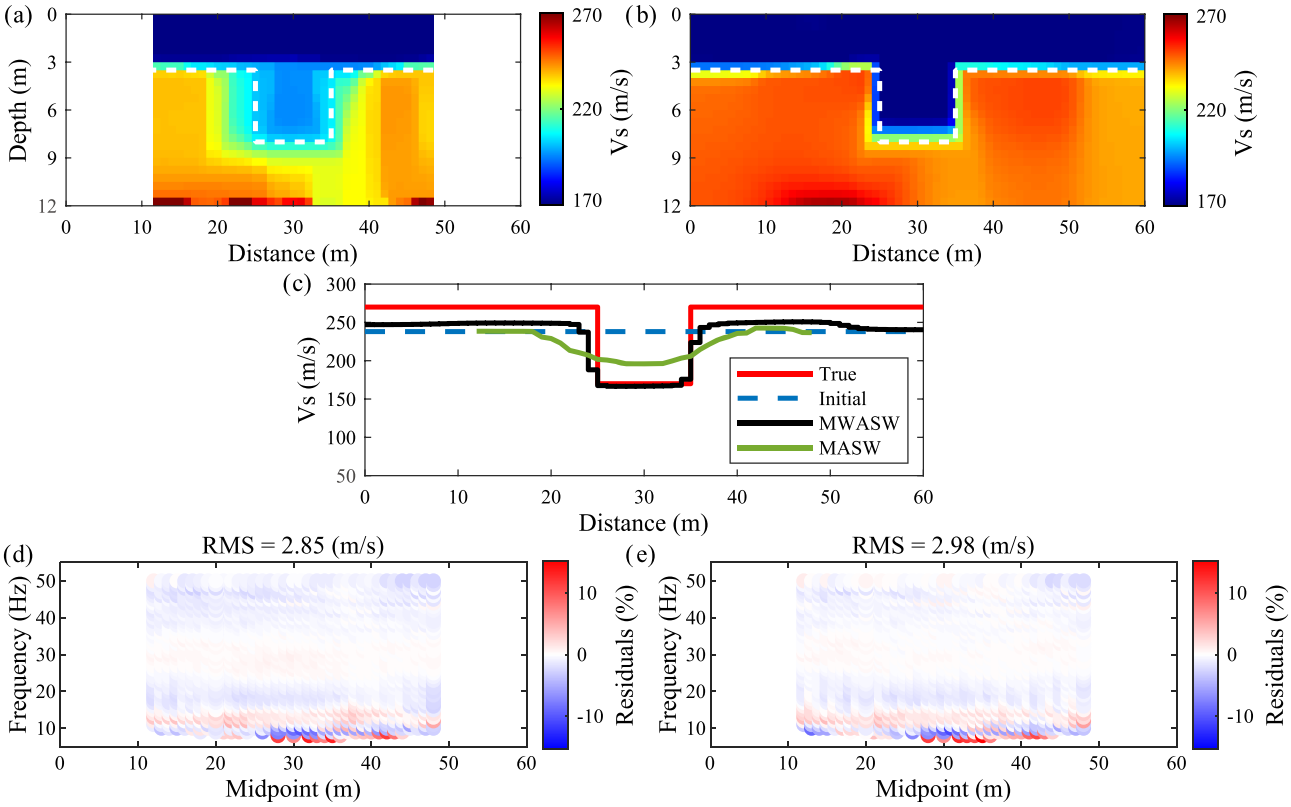

Figure 10. Inversion results of the graben model with 5 per cent Gaussian noise added. (a and b) 2-D *Vs* sections retrieved by the MASW and MWASW methods, respectively. The white dashed lines depict the outline of the graben. (c) The plot of the velocity value for the row of model cells between the depths of 5.5 and 6 m. (d and e) Phase velocity residuals for the maximum window correspond to the MASW method and the MWASW method, respectively.

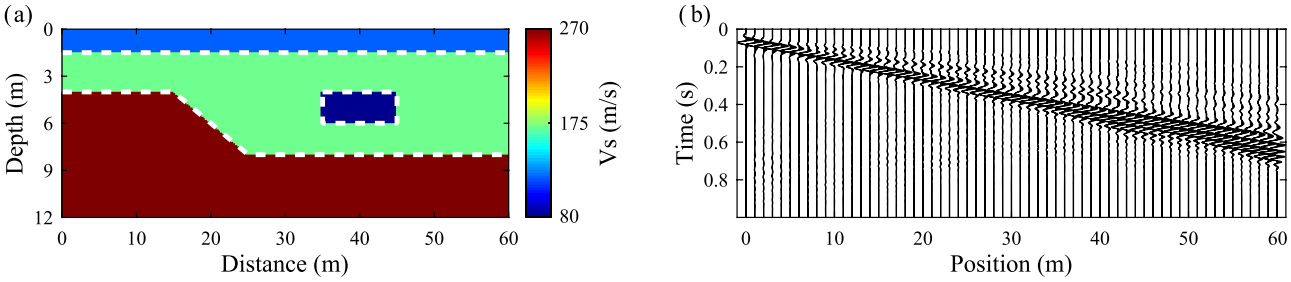

Figure 11. (a) Low-velocity block model. The white dashed line indicates the outlines of the stratum and low-velocity block. (b) Simulated first common-shot gather.

identify the dispersion trend at the wavelength band shorter than the window length. Fig. 12(c) shows estimated dispersion data of all spatial windows. According to the phase velocity plot, there is inverse dispersion at the region corresponding to the low-velocity block.

The initial model used in the inversion (Fig. 13a) is homogenous and built by the maximum phase velocity of 213.5 m s^{-1}. The *Vs* models reconstructed by the MASW and MWASW methods are shown in Figs 13(b) and (c), respectively. The white dashed line indicates the true outlines of the stratum and the low-velocity block. In Fig. 13(b), the *Vs* model retrieved by the MASW shows a weak low-velocity zone that spreads out over a wider region compared to the actual shape of the low-velocity block. The RMS error between the retrieved and true model is 27.7 m s^{-1}. In the *Vs* model retrieved by the MWASW (Fig. 13c), the low-velocity zone and the interfaces between each layer are closer to the true model. In addition, the RMS error between the recovered and true model decreases to 17.8 m s^{-1}. The difference between the MASW and MWASW is also shown clearly in the plot of the velocity values for the row of model cells between the depths of 5.5 and 6 m (Fig. 13d). As we can see, the MASW method could only produce a *Vs* model with smooth variations in the lateral direction. The minimum velocity in the low-velocity zone is approximately 133.6 m s^{-1}. However, the MWASW method more effectively reveals the sharp lateral variations of the subsurface, the velocity reaches 113.0 m s^{-1} in the low-velocity zone of the retrieved model. Figs 13(e) and (f) show the phase velocity residual of the maximum window corresponds to the MASW method and the MWASW method, respectively. We can see that the RMS error between the experimental and theoretical dispersion curves of the maximum window for the MWASW method is slightly lower than that for the MASW method.

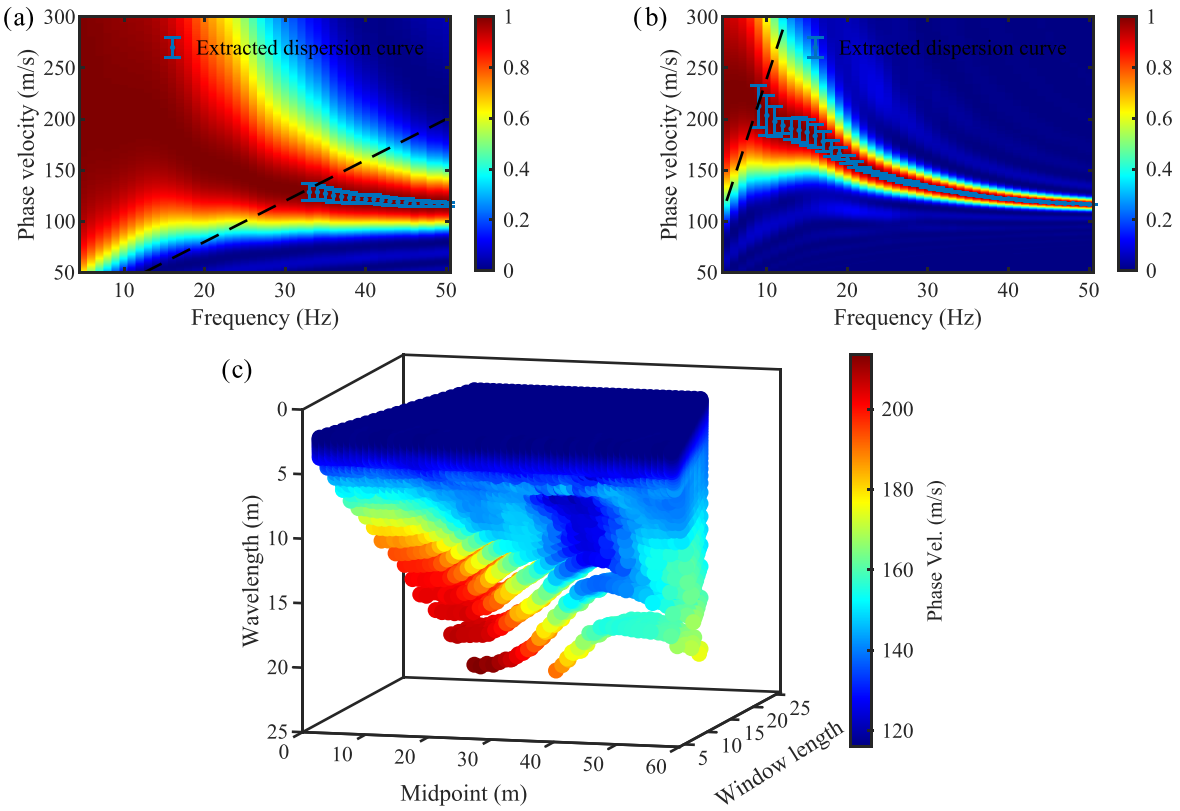

Figure 12. Panels (a) and (b) are the stacked dispersion images corresponding to the first 5 and the first 25 receivers, respectively. The blue dotted line with the blue bar depicts the extracted dispersion curve with uncertainties, the black dashed line represents the wavelength of $\lambda = (N_{rec}-1)\Delta r$. (c) All extracted dispersion data.

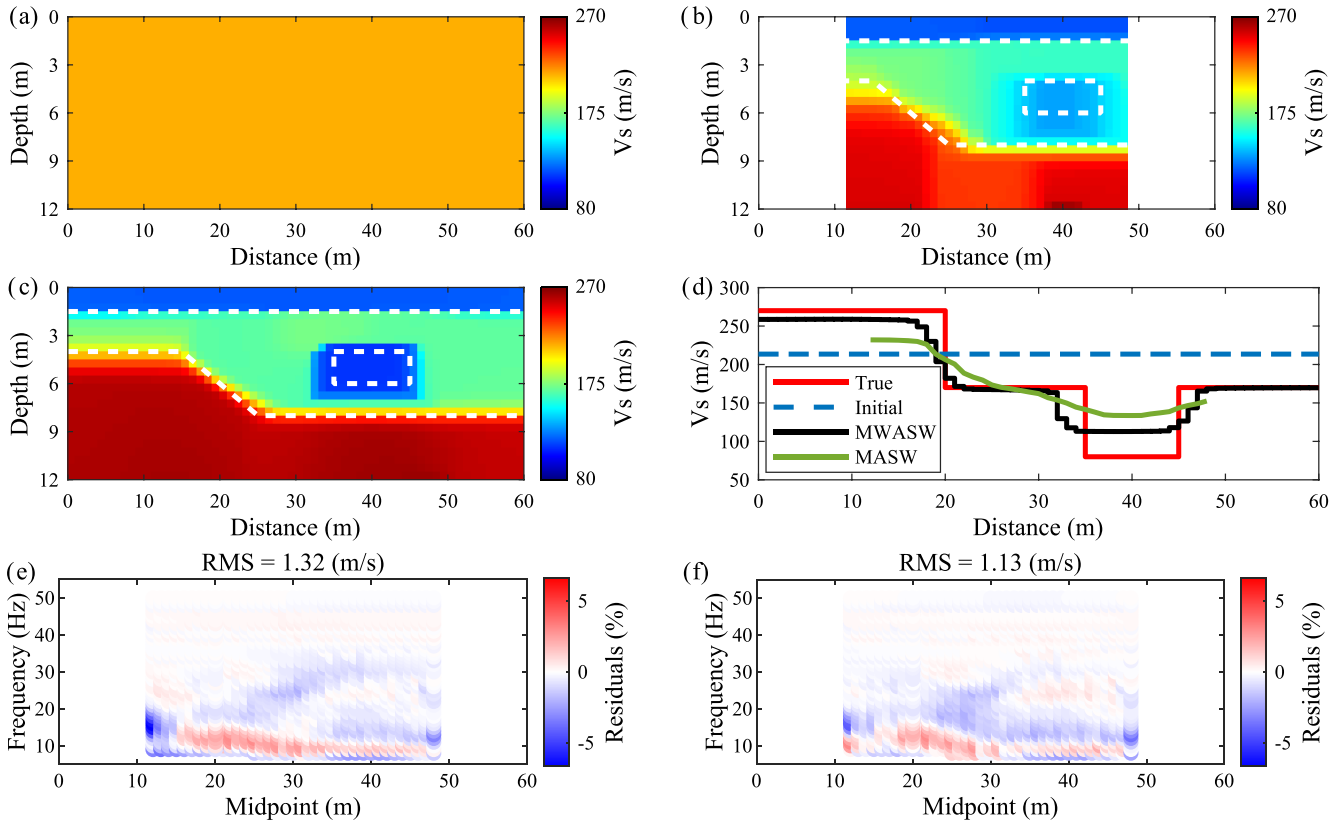

Figure 13. Inversion results of the low-velocity block model. (a) Initial model. (b and c) 2-D *Vs* sections retrieved via the MASW and MWASW methods, respectively. The white dashed lines depict the true outlines of the stratum and low-velocity block. (d) The plot of the velocity value for the row of model cells between the depths of 5.5 and 6 m. (e and f) Phase velocity residuals for the maximum window correspond to the MASW method and the MWASW method, respectively.

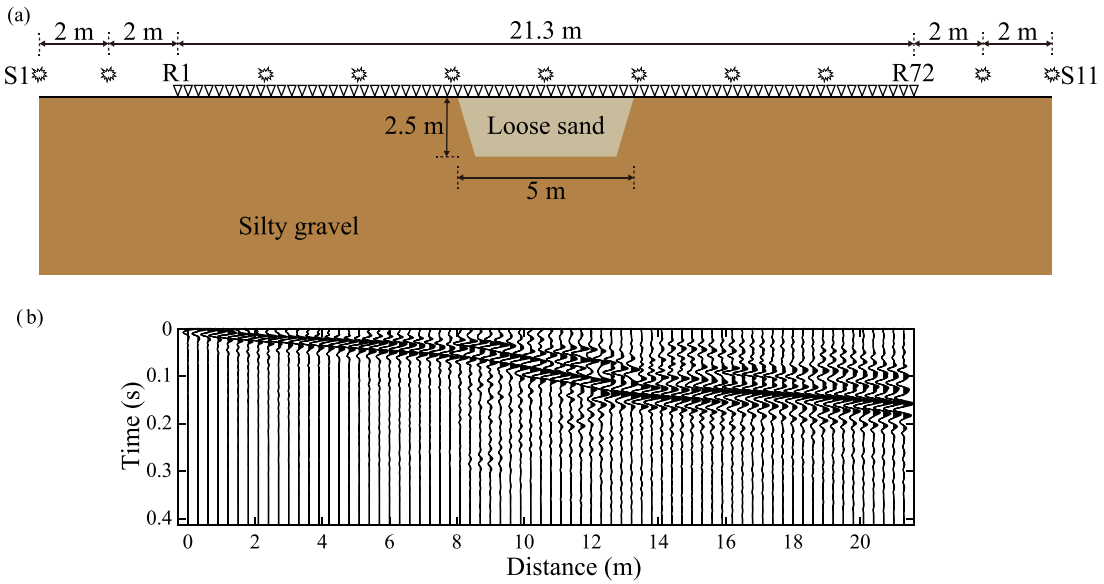

Figure 14. (a) The geometry of the CNR test site and the acquisition layout. R1–R72 and S1–S11 indicate the position of receivers and sources, respectively. (b) The first virtual source gather.

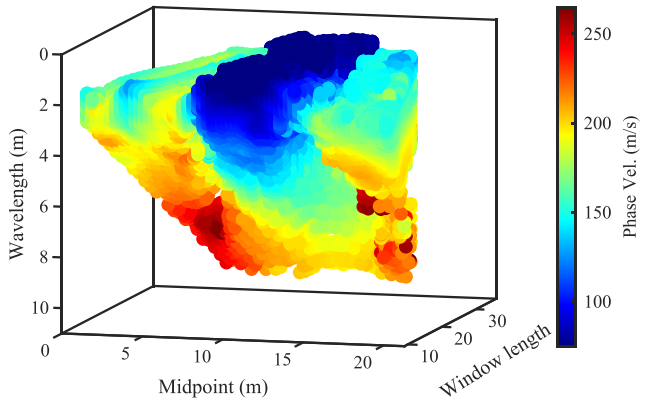

Figure 15. Extracted dispersion data of the CNR test site.

4 FIELD DATA TEST

4.1 CNR test site

This field data set is acquired at the CNR (National Research Council) headquarter in Turin, Italy. In this site, an artificial loose sand body is embedded in the local soil formation, which is mainly composed of thick silty gravel overlain by a shallow soil layer. The shape of the loose sand body on the surface is a square with a size of 5 m × 5 m; the depth of this target is approximately 2.5 m. A seismic survey line, consisting of 72 vertical 4.5 Hz geophones evenly spaced every 1 m, is deployed across the sand body. Eleven shot records are acquired along the survey line by moving a hammer source. Fig. 14(a) shows the geometry of this test site and the acquisition layout. More details about the acquisition parameters can be found in Anjom *et al.* (2019).

To improve the stacking numbers in the dispersion analysis, we generate 72 virtual source gathers by cross-correlating the wavefields recorded at each receiver and summing over all sources. Fig. 14(b) shows the first virtual source gather. After that, we apply the multiscale window analysis to extract dispersion curves. The minimum and maximum length of the spatial windows, $W_{\min}$ and $W_{\max}$, are 10 and 35 receivers, respectively. The maximum offset d is selected to 15 m. Fig. 15 shows the estimated dispersion data of all spatial windows.

We use the MASW and MWASW methods to retrieve the *Vs* model of the subsurface. In both these methods, we discretize the model using 20 layers and a half-space. The thickness of each layer is 0.25 m. We set the width of the model cells to the receiver spacing of 0.3 m for the MWASW method; while, the MASW only retrieves the 1-D model at the midpoint of the maximum spatial window. Fig. 16(a) reports the *Vs* model obtained by the MASW method. Influenced by the horizontally layered model assumption, the MASW method overestimates the scale of the sand body along the horizontal direction (7.5–13.8 m compared to the real value of 8.5–13.5 m). The *Vs* model retrieved by the MWASW method is shown in Fig. 16(b). The range of the low-velocity anomaly is closer to the real situation in the horizontal direction,

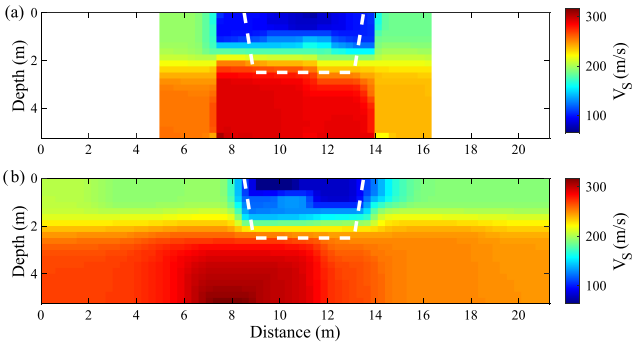

Figure 16. *Vs* models of the CNR test site retrieved by the (a) MASW and (b) MWASW methods, respectively. The white dashed line indicates the outline of the sand body.

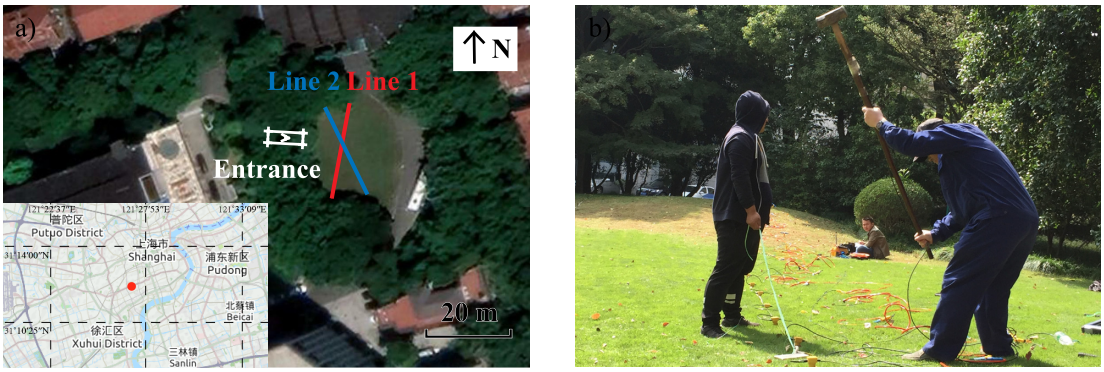

Figure 17. ShanghaiTech University. (a) Seismic survey lines. (b) Implementation of data acquisition along Line 1.

which indicates the improved lateral resolution of the MWASW method. Moreover, both of these two methods underestimate the thickness of the sand body due to the low sensitivity of surface wave dispersion to interfaces.

4.2 ShanghaiTech University

This field data set is collected in an abandoned air-raid shelter located at ShanghaiTech University (Yueyang Campus), Shanghai, China (Fig. 17a). To determine its underground layout, we perform the surface wave survey along the two crossed survey lines, shown in Fig. 17(a). The seismic signal is generated by a hammer source and recorded by a Geode seismograph with 24 vertical 4-Hz geophones at 1 m intervals. We use the common-receiver-spread recording geometry on each survey line. Twenty-five shot records are acquired along this receiver spread at 1 m intervals. Fig. 17(b) shows the performance of the field acquisition along Line 1.

Based on the background information that the depth of this air-raid shelter is about 4–5 m, we apply a series of spatial windows, successively changed from a minimum length W_{min} of 6 to a maximum length W_{max} of 13 receivers, to extract the dispersion data. The maximum offset d is selected to 15 m. Figs 18(a) and (b) show the estimated dispersion data of Line 1 and Line 2, respectively. We can see that both phase velocity plots clearly present characteristics of inverse dispersion at each survey line.

After data processing, we implement the MASW and MWASW methods to map the *Vs* structure of the subsurface. The model discretization is the same as the synthetic test, except that the thickness of each layer is 0.25 m. Fig. 19 shows the 2-D *Vs* sections that are reconstructed using the MASW and MWASW methods. For interpretation, the retrieved *Vs* models are presented by contour maps. Despite the smoothness of the contouring, the low-velocity features corresponding to the air-raid shelter that are retrieved along two survey lines should be very similar to each other; the *Vs* models should be the same at the cross point without considering the 3-D effect. Due to the average effect of the MASW method, the structural features of the two sections retrieved along Line 1 and Line 2 are discrepant, as shown in Figs 19(a) and (c). More details at the cross point are shown in Fig. 19(e). The velocity between these two *Vs* profiles differs at a depth of between 1.25 and 2.5 m and below 5.25 m. The RMS error between these two velocity profiles is 7.3 m s^{-1}. On the contrary, the velocity features retrieved using the MWASW method at two survey lines, shown in Figs 19(b) and (d), are more consistent, which is much more beneficial for understanding the layout of the air-raid shelter. As we have observed from the synthetic tests, the MWASW method overestimates the velocity in the low-velocity zone; however, it has the ability to reconstruct the shape of the anomaly. The 105 m s^{-1} contour line may indicate the underground layout of the air-raid shelter. In addition, the *Vs* profiles retrieved by the MWASW (Fig. 19f) at the cross point are similar, and the RMS error between these two velocity profiles is 4.0 m s^{-1}. This similarity further indicates the improved lateral resolution of the proposed MWASW method.

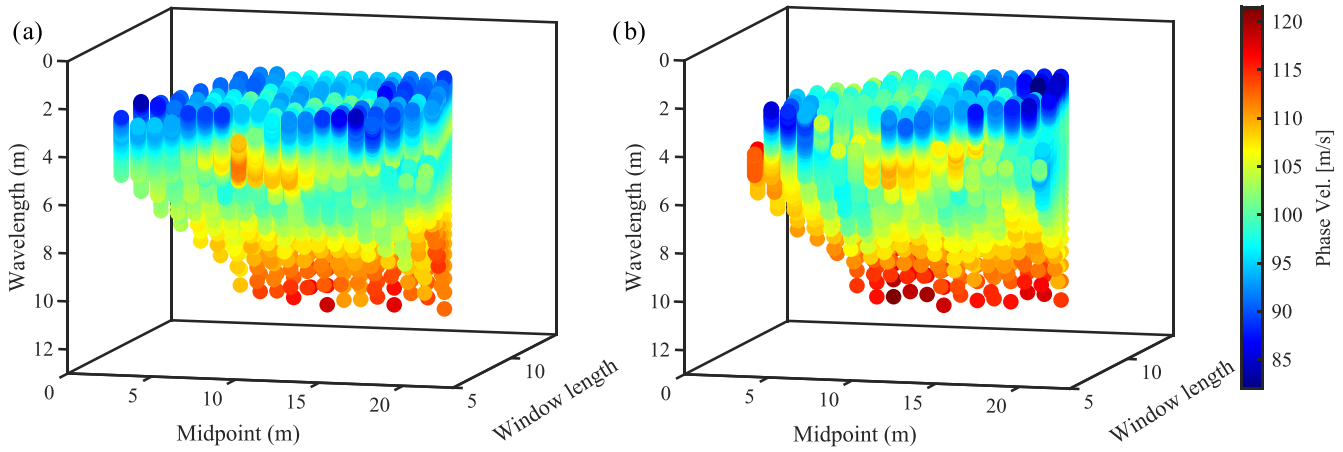

Figure 18. Retrieved dispersion curves of ShanghaiTech University along (a) Line 1 and (b) Line 2, respectively.

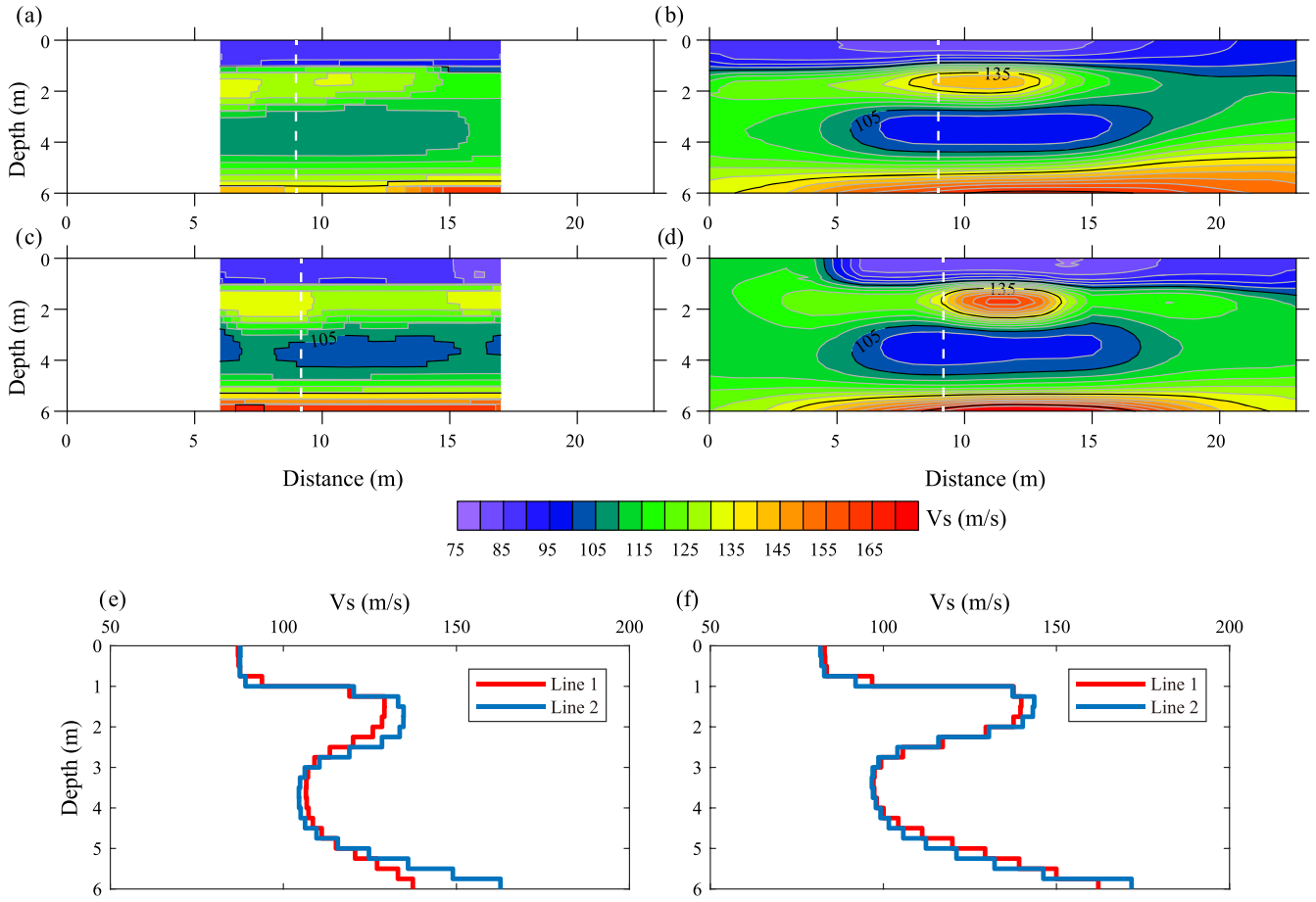

Figure 19. Retrieved *Vs* models of ShanghaiTech University. (a and b) 2-D *Vs* models along Line 1 retrieved via the MASW and MWASW methods, respectively. (c and d) 2-D *Vs* models along Line 2 retrieved via the MASW and MWASW methods, respectively. The white dashed line indicates the position of the cross point. (e) *Vs* profile at the cross point retrieved via the MASW method. (f) *Vs* profile at the cross point retrieved via the MWASW method.

5 DISCUSSION

The MWASW method retains the advantage of the MASW method, that is the robustness in determining the phase velocity of surface waves (Socco *et al.* 2010; Foti *et al.* 2011, 2018). The novelty of the MWASW method is in the integration of two essential aspects: (1) the used forward algorithm of multistation DCs takes into account the lateral variations in the 2-D subsurface model and (2) the strategy of multiscale window analysis exploits the subsurface information of different scales contained in 2-D active-source seismic records. As a result, the MWASW improves the ability of surface wave analysis to retrieve a more realistic 2-D *Vs* model. Synthetic and field data tests presented in the previous section demonstrate the improved lateral resolution of the MWASW method in various heterogeneous conditions.

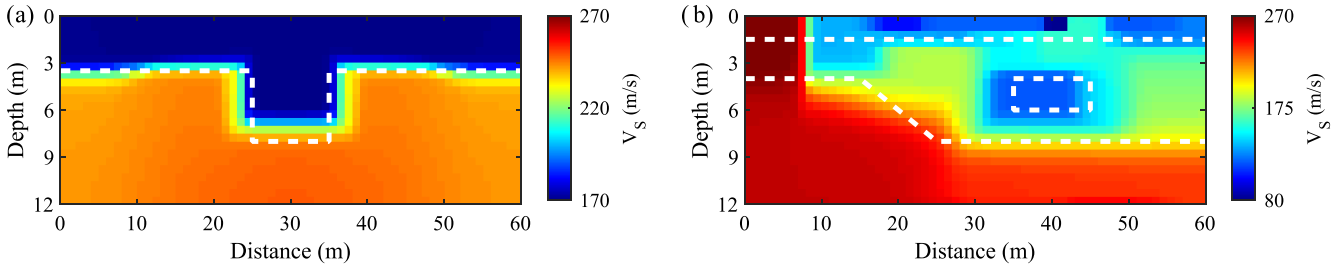

Figure 20. Retrieved 2-D *Vs* section of (a) the graben model and (b) the low-velocity block model using only the dispersion data of the maximum window.

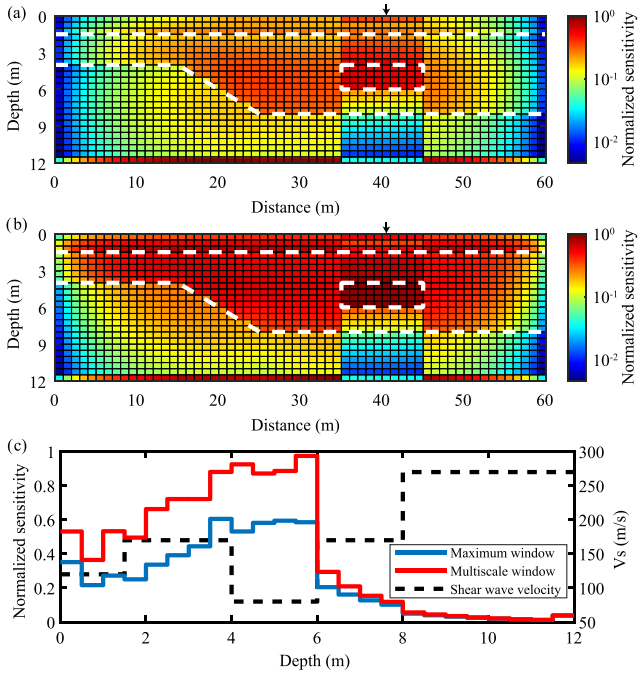

Figure 21. Sensitivity analysis of the low-velocity block model. Panels (a) and (b) are the sensitivity distributions produced by the data of the maximum window and multiscale window, respectively. (c) The sensitivity plot at the position of 40.5 m. The black dashed line indicates the layered *Vs* model at the position of 40.5 m.

The incorporation of more dispersion curves slightly increases the time in dispersion analysis and the computational cost in inversion. Thus, using only the dispersion curve extracted from the maximum spatial window may be preferred. Figs 20(a) and (b) show the retrieved 2-D *Vs* section of the graben model and the low-velocity block model using only the dispersion data of the maximum spatial window. For the simple case (graben model), inverting only the dispersion curves extracted from the maximum window is able to retrieve the lateral variations of the subsurface. However, for the complex case (low-velocity block model), the inversion process becomes unstable and falls into the local minimum (Fig. 20b), despite that the inversion of dispersion data of all spatial windows reveals the sharp lateral variations of the subsurface (Fig. 13c). This is understandable because more data used in the inversion produce a better constraint on the model parameter and thus alleviate the ill-posedness of the inversion problem. To examine the improvement in the use of multiscale window analysis, we implement a sensitivity analysis for the low-velocity block model by taking advantage of the local optimization. We adopt the notion of integrated sensitivity (Zhdanov 2002) for simplifying the presentation of the 2-D sensitivity. Since the MWASW always use the multifrequency data weighting, the weighted integrated sensitivity is more straightforward and takes the form:

$$\mathbf{S} = \mathbf{diag}\left(\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\omega}^{\mathrm{T}}\mathbf{W}_{\omega}\mathbf{J}\right)^{1/2}. \tag{17}$$

Fig. 21 shows the weighted integrated sensitivity for the low-velocity block model. The frequency band and applied spatial windows used in sensitivity analysis are the same as the experimental dispersion curves analysed from the wavefields. We normalize the respective values calculated from eq. (17) by the maximum sensitivity, so the sensitivity distributions produced from different data sets can be comparable.

A general decrease of sensitivity with depth is observed in the sensitivity distributions generated by the dispersion data of the maximum window (Fig. 21a) and multisize windows (Fig. 21b), except in the region above the low-velocity block and the half-space. The high sensitivity of the half-space is because all the dispersion data with a wavelength longer than 11.5 m contribute to those model parameters in the half-space. Compared to the sensitivity distribution produced by the data of the maximum window, the incorporation of dispersion data extracted from multisize spatial windows significantly improves the sensitivity to the shallow formations as well as the low-velocity block. The sensitivity plot

at the position of 40.5 m (Fig. 21c) also presents this improvement. The sensitivity is higher for the multisize windows than for the maximum window at the depth above 8 m. For the MASW method, caused by the lower sensitivity for the maximum window and the horizontally layered model assumption, it overestimates the velocity of the low-velocity block as well as fails to reveal the thin layer below the low-velocity zone (Fig. 13b). Moreover, although both sensitivity distributions show a low sensitivity at the depth below 8 m, the MASW and the MWASW method retrieve the *Vs* model well in this region (Figs 13b and c) thanks to the TV regularization, which implies a lateral constraint and therefore alleviates the non-uniqueness of the inverse problem.

For active-source seismic records, surface wave analysis usually suffers from the near-field effect since the assumption of plane surface waves is not valid and body waves are significant in the near-offset region (Heisey *et al.* 1982; Gucunski & Woods 1992; Zywicki 1999; O'Neil 2003; Zywicki & Rix 2005). The adverse influence of the near-field effect is that the phase velocity in the low-frequency band can be underestimated (Bodet *et al.* 2005, 2009; Yoon & Rix 2009), which result in a slightly lower *S*-wave velocity in the deep formations of the retrieved model (see, for example, Fig. 8c). Although stacking can generally enhance quality and mitigate the near-field effect (Pasquet & Bodet 2017), the underestimation of the phase velocity in the low-frequency band can occur in some situations, as shown in Fig. 7(b). The degree of underestimation seems to be highly site dependent (Park & Carnevale 2010) and is sometimes negligible, such as the result in Fig. 13 (c). Nevertheless, the MWASW can reconstruct the laterally varying structure well. Namely, it is applicable if the absolute *Vs* value of the deeper stratum is not that necessary in the investigation. An option to mitigate the near-field effects is to limit the maximum wavelength to 40–50 per cent of the spread length (O'Neil 2003; Bodet *et al.* 2005; Bodet *et al.* 2009), despite it inevitably leads to a loss of information (Bodet *et al.* 2009). On the other hand, to produce more detailed information about the subsurface, the *Vs* model retrieved by the MWASW method could be used as the initial model for full-waveform inversion.

The MWASW method is able to reconstruct the subsurface *Vs* structure from a simple initial model, even when abrupt lateral variations are presented. However, this method may overestimate the velocity in the low-velocity zone and fail to locate the interfaces, as shown in Figs 13(c) and 16(b). These problems are caused by the non-uniqueness of surface wave inversion, as several different models may produce similar goodness of fit with the experimental data (Foti *et al.* 2009, 2011). In addition, the sensitivity of the dispersion curves generally decreases with depth, as shown in Fig. 21. To improve sensitivity to the model parameters and the consequent reliability of the results, a well-parametrized model is necessary for the inversion of surface wave data (Socco *et al.* 2010; Foti *et al.* 2018). For these reasons, the advanced global search optimization algorithm and the constrained surface wave inversion (Bodin *et al.* 2012; Killingbeck *et al.* 2018) could be used to deal with model parametrization and alleviate the non-uniqueness of the solution.

6 CONCLUSION

We have presented a methodology for retrieving 2-D laterally varying structures from the multistation surface wave DCs. Unlike the traditional MASW method, which assigns one dispersion curve extracted from several receivers to an individual point and adopts a horizontally layered model in the inversion, the proposed MWASW method simultaneously inverts the dispersion data analysed from multisize spatial windows using the forward algorithm of multistation DCs. Smaller spatial windows provide short wavelength data with a high lateral resolution, whereas larger windows yield average information for greater investigation depth. In this way, the subsurface information in different scales contained in the surface wave records is exploited. The forward problem is solved based on amplitude-normalized beamforming and a 2-D cell-based model. The calculation of the theoretical DCs, which is similar to the determination of phase velocity using the multistation method, considers the lateral variations beneath the receiver spread. Consequently, the adverse influence of the horizontally layered model assumption is mitigated.

We have tested the reliability and limitations of the proposed MWASW method using synthetic and field data sets. Compared to the results of the MASW, the MWASW method can produce a more compact 2-D *Vs* structure with a higher lateral resolution. Even when a low-velocity zone is presented, the shape of this anomaly can be well revealed. Moreover, it can provide more information beneath the entire receiver array, which helps to obtain a comprehensive understanding of the subsurface. The main drawback of the MWASW method is that it may overestimate the velocity in the low-velocity zone and fail to locate the interfaces. For these problems caused by solution non-uniqueness, a global search optimization algorithm will be implemented to explore the solution space and reduce dependence on the initial model. Since the proposed forward algorithm can also calculate the multistation DCs of higher modes in theory, another improvement is that incorporating the information coming from higher modes when they are observable in the seismic records.

ACKNOWLEDGEMENTS

SH, YZ and SG thank the National Natural Science Foundation of China for financial support (Nos. 41774124 and 42074177). SH would like to thank Prof Jiansheng Wu and Mr. Tan Qin for their valuable suggestions during this work. He also thanks the China Scholarship Council for supporting his visit to Politecnico di Torino. The authors thank Editor Hervé Chauris and two anonymous reviewers for their constructive comments and suggestions which greatly improved this paper.

DATA AVAILABILITY

Waveform data used in this study is available in repository ZENODO (http://doi.org/10.5281/zenodo.5105871). The MWASW program can be downloaded at http://mwasw.cn/msasw/a/download/81.html.

REFERENCES

Adly, A., Poggi, V., Fäh, D., Hassoup, A. & Omran, A., 2017. Combining active and passive seismic methods for the characterization of urban sites in Cairo, Egypt, *Geophys. J. Int.,* **210**(1), 428–442.

Anjom, F.K., Teodor, D., Comina, C., Brossier, R., Virieux, J. & Socco, L.V., 2019. Full-waveform matching of VP and VS models from surface waves, *Geophys. J. Int.,* **218**(3), 1873–1891.

Armijo, L., 1966. Minimization of functions having Lipschitz continuous first partial derivatives, *Pac. J. Math.,* **16**(1), 1–3.

Askari, R. & Ferguson, R.J., 2012. Dispersion and the dissipative characteristics of surface waves in the generalized S-transform domain, *Geophysics,* **77**(1), V11–V20.

Aster, R.C., Borchers, B. & Thurber, C.H., 2005. *Parameter Estimation and Inverse Problems,* Academic Press.

Auken, E. & Christiansen, A.V., 2004. Layered and laterally constrained 2D inversion of resistivity data, *Geophysics,* **69**(3), 752–761.

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

Bodet, L., Abraham, O. & Clorennec, D., 2009. Near-offset effects on Rayleigh-wave dispersion measurements: physical modeling. *J. appl. Geophys.,* **68**(1), 95–103.

Bodet, L., Wijk, K., Bitri, A., Abraham, O., Côte, P., Grandjean, G. & Leparoux, D., 2005. Surface-wave inversion limitations from laser-Doppler physical modeling. *J. Environ. Eng. Geophys.,* **10**(2), 151–162.

Bodin, T., Sambridge, M., Tkalcic, H., Arroucau, P., Gallagher, K. & Rawlinson, N., 2012. Transdimensional inversion of receiver functions and surface wave dispersion, *J. geophys. Res.,* **117**(B2), doi.org/10.1029/2011jb008560.

Bohlen, T., Kugler, S., Klein, G. & Theilen, F., 2004. 1.5 D inversion of lateral variation of Scholte-wave dispersion, *Geophysics,* **69**(2), 330–344.

Boiero, D. & Socco, L.V., 2011. The meaning of surface wave dispersion curves in weakly laterally varying structures, *Near Surf. Geophys.,* **10**(3), 249–261.

Fang, H., Yao, H., Zhang, H., Thurber, C., Ben-Zion, Y. & Hilst, R.D.., 2019. Vp/Vs tomography in the southern California plate boundary region using body and surface wave traveltime data, *Geophys. J. Int.,* **216**(1), 609–620.

Foti, S. *et al.*, 2018. Guidelines for the good practice of surface wave analysis: a product of the InterPACIFIC project. *Bull. Earthq. Eng.,* **16**(6), 2367–2420.

Foti, S., Comina, C., Boiero, D. & Socco, L., 2009. Non-uniqueness in surface-wave inversion and consequences on seismic site response analyses, *Soil Dyn. Earthq. Eng.,* **29**(6), 982–993.

Foti, S., Parolai, S., Albarello, D. & Picozzi, M., 2011. Application of surface-wave methods for seismic site characterization, *Surv. Geophys.,* **32**(6), 777–825.

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

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

Golub, G.H. & Van Loan, C.F., 1989. *Matrix Computations,* The Johns Hopkins Univ. Press.

Grandjean, G. & Bitri, A., 2006. 2M-SASW: multifold multichannel seismic inversion of local dispersion of Rayleigh waves in laterally heterogeneous subsurfaces: application to the Super-Sauze earthflow, France, *Near Surf. Geophys.,* **4**(6), 367–375.

Gucunski, N. & Woods, R.D., 1992. Numerical simulation of the SASW test. *Soil Dyn. Earthq. Eng.,* **11**(4), 213–227.

Hashemi Jokar, M., Rahnema, H., Boaga, J., Cassiani, G. & Strobbia, C., 2019. Application of surface waves for detecting lateral variations: buried inclined plane, *Near Surf. Geophys.,* **17**(5), 501–531.

Haskell, N.A., 1953. The dispersion of surface waves on multilayered media, *Bull. seism. Soc. Am.,* **43**(1), 17–34.

Hayashi, K. & Suzuki, H., 2004. CMP cross-correlation analysis of multichannel surface-wave data, *Explor. Geophys.,* **35**(1), 7–13.

Heisey, J.S., Stokoe, K.H., II, Hudson, W.R. & Meyer, A.H., 1982. Determination of in situ shear wave velocities from spectral analysis of surface waves, Research Report 256-2, Ctr. for Transp. Res., Univ. of Texas, Austin, TX.

Hu, W., Abubakar, A. & Habashy, T.M., 2009. Simultaneous multifrequency inversion of full-waveform seismic data, *Geophysics,* **74**(2), R1–R14.

Ikeda, T. & Tsuji, T., 2020. Two-station continuous wavelet transform cross-coherence analysis for surface-wave tomography using active-source seismic data, *Geophysics,* **85**(1), EN17–EN28.

Ikeda, T., Tsuji, T. & Matsuoka, T., 2013. Window-controlled CMP cross-correlation analysis for surface waves in laterally heterogeneous media, *Geophysics,* **78**(6), EN95–EN105.

Johnson, D.H. & Dudgeon, D.E., 1993. *Array Signal Processing: Concepts and Techniques,* Prentice Hall.

Julià, J., Ammon, C.J., Herrmann, R.B. & Correig, A.M., 2000. Joint inversion of receiver function and surface wave dispersion observations, *Geophys. J. Int.,* **143**(1), 99–112.

Kaviani, A. *et al.*, 2020. Crustal and uppermost mantle shear wave velocity structure beneath the Middle East from surface wave tomography, *Geophys. J. Int.,* **221**(2), 1349–1365.

Killingbeck, S.F., Livermore, P.W., Booth, A.D. & West, L.J., 2018. Multimodal layered transdimensional inversion of seismic dispersion curves with depth constraints, *Geochem., Geophys., Geosyst.*, **19**(12), 4957–4971.

Kulesh, M., Holschneider, M., Ohrnberger, M. & Lück, E., 2008. Modeling of wave dispersion using continuous wavelet transforms II: wavelet-based frequency-velocity analysis, *Pure appl. Geophys.,* **165**(2), 255–270.

Li, J., Feng, Z. & Schuster, G., 2017. Wave-equation dispersion inversion, *Geophys. J. Int.,* **208**(3), 1567–1578.

Li, J., Lin, F.-C., Allam, A., Ben-Zion, Y., Liu, Z. & Schuster, G., 2019. Wave equation dispersion inversion of surface waves recorded on irregular topography, *Geophys. J. Int.,* **217**(1), 346–360.

Lin, C.-P. & Chang, T.-S., 2004. Multi-station analysis of surface wave dispersion. *Soil Dyn. Earthq. Eng.,* **24**(11), 877–886.

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

Mi, B., Xia, J., Bradford, J.H. & Shen, C., 2020. Estimating near-surface shear-wave-velocity structures via multichannel analysis of Rayleigh and Love waves: an experiment at the Boise hydrogeophysical research site, *Surv. Geophys.,* **41**(2), 323–341.

Mordret, A., Landès, M., Shapiro, N.M., Singh, S.C., Roux, P. & Barkved, O.I., 2013. Near-surface study at the Valhall oil field from ambient noise surface wave tomography, *Geophys. J. Int.,* **193**(3), 1627–1643.

Neducza, B., 2007. Stacking of surface waves, *Geophysics,* **72**(2), V51–V58.

Nolet, G. & Panza, G.F., 1976. Array analysis of seismic surface waves: limits and possibilities, *Pure appl. Geophys.,* **114**(5), 775–790.

O'Neill, A., 2003. Full-waveform reflectivity for modeling, inversion and appraisal of seismic surface wave dispersion in shallow site investigations, *PhD thesis*, University of Western Australia, Perth, Australia.

Oldenburg, D.W. & Li, Y., 2005. Inversion for applied geophysics: a tutorial, in *Near-Surface Geophysics, Investigations in Geophysics,* pp 89–150, ed., Butler, D.K.,Society of Exploration Geophysicists.

Papadopoulou, M. *et al.*, 2020. Surface-wave analysis for static corrections in mineral exploration: a case study from central Sweden. *Geophys. Prospect.,* **68**(1), 214–231.

Park, C.B. & Carnevale, M., 2010. Optimum MASW survey—revisit after a decade of use, In *Proceedings of GeoFlorida 2010,* Am. Soc. Civil Eng., Reston, pp. 1303–1312.

Park, C.B., Miller, R.D. & Xia, J., 1998. Imaging dispersion curves of surface waves on multi-channel record, in *SEG Technical Program Expanded Abstracts 1998,* Society of Exploration Geophysicists, pp. 1377–1380.

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

Pasquet, S. & Bodet, L., 2017. SWIP: an integrated workflow for surface-wave dispersion inversion and profiling, *Geophysics,* **82**(6), WB47–WB61.

Pasquet, S., Wang, W., Chen, P. & Flinchum, B.A., 2021. Multi-window weighted stacking of surface-wave dispersion. *Geophysics,* **86**(2), A15–W19.

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

Qin, T. *et al.*, 2020. An interactive integrated interpretation of GPR and Rayleigh wave Data based on the genetic algorithm, *Surv. Geophys.,* **41,** 549–574.

Rix, G.J., Hebeler, G.L. & Orozco, M.C., 2002. Near-surface Vs profiling in the New Madrid seismic zone using surface-wave methods, *Seismol. Res. Lett.,* **73**(3), 380–392.

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

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

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

Socco, L.V., Boiero, D., Maraschini, M., Vanneste, M., Madshus, C., Westerdahl, H., Duffaut, K. & Skomedal, E., 2011. On the use of the Norwegian Geotechnical Institute's prototype seabed-coupled shear wave vibrator for shallow soil characterization - II. Joint inversion of multimodal Love and Scholte surface waves, *Geophys. J. Int.,* **185**(1), 237–252.

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), 75A83–75A102.

Stokoe, K.H., Wright, S.G., Bay, J.A. & Roesset, J.M., 1994. Characterization of geotechnical sites by SASW method, in *Geophysical Characterization of Sites,* pp. 15–25, ed., Woods, R.D.,Oxford Publisher.

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

Vignoli, G., Gervasio, I., Brancatelli, G., Boaga, J., Della Vedova, B. & Cassiani, G., 2016. Frequency-dependent multi-offset phase analysis of surface waves: an example of high-resolution characterization of a riparian aquifer, *Geophys. Prospect.,* **64**(1), 102–111

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

Xia, J., 2014. Estimation of near-surface shear-wave velocities and quality factors using multichannel analysis of surface-wave methods, *J. appl. Geophys.,* **103,** 140–151.

Xia, J., Chen, C., Tian, G., Miller, R.D. & Ivanov, J., 2005. Resolution of high-frequency Rayleigh-wave data, *J. Environ. Eng. Geophys.,* **10**(2), 99–110.

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.

Yin, X., Xu, H., Wang, L., Hu, Y., Shen, C. & Sun, S., 2016. Improving horizontal resolution of high-frequency surface-wave methods using travel-time tomography, *J. appl. Geophys.,* **126,** 42–51.

Yoon, S. & Rix, G.J., 2009. Near-Field effects on array-based surface wave methods with active sources, *J. Geotech. Geoenviron. Eng.,* **135**(3), 399–406.

Zhai, Q., Yao, H. & Peng, Z., 2019. Upper-mantle shear velocity structure beneath the equatorial East Pacific Rise from array-based teleseismic surface wave dispersion analysis, *Geophys. J. Int.,* **219**(1), 607–618.

Zhdanov, M.S., 2002. *Geophysical Inverse Theory and Regularization Problems,* Elsevier.

Zhdanov, M.S., 2009. New advances in regularized inversion of gravity and electromagnetic data, *Geophys. Prospect.,* **57**(4), 463–478.

Zywicki, D.J. & Rix, G.J., 2005. Mitigation of near-field effects for seismic surface wave velocity estimation with cylindrical beamformers, *J. Geotech. Geoenviron. Eng.,* **131**(8), 970–977.

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

APPENDIX A: VALIDATION OF THE FORWARD ALGORITHM

A numerical experiment is carried out to validate the reliability of the forward algorithm for calculating the theoretical multistation surface wave DCs. The synthetic model is a step model in which the thickness of the first layer is abruptly decreased from 8 to 3.5 m at the model center, as shown in Fig. A1(a). The elastic parameters and model discretization used in this model are the same as the Graben model shown in Fig. 6(a). In the FDTD modeling, a receiver array consisting of 24 receivers evenly spaced every 1 m and a source with a minimum offset of 5 m, as shown in Fig. A1(a) by the black lines and the red circle, are used to simulate the common-shot gather (Fig. A1b). The emitted source is a Ricker wavelet with a center frequency of 20 Hz.

Fig. A1(c) shows the corresponding dispersion image generated by the amplitude-normalized beamforming (eq. 3). The black dot line indicates the extracted dispersion curve identified by the spectral maxima on the dispersion image; the white dashed–dotted line represents the theoretical dispersion curve calculated by the proposed forward algorithm; the blue dashed–dotted line depicts the local dispersion curve calculated using the 1-D model at the midpoint of the receiver array (the position of 19.5 m). At the wavelength shorter than the array length, that is in the wavelength band without the influence of leakage error, the theoretical dispersion curve is in good agreement with the extracted one, the RMS error between these two curves is 1.76 m s^{-1}, which demonstrates the reliability of the proposed forward algorithm. In the same frequency band, the RMS error between the extracted and the local dispersion curves is 11.70 m s^{-1}, which indicates that the multistation dispersion cannot be representative of the phase velocity at the midpoint of the spread.

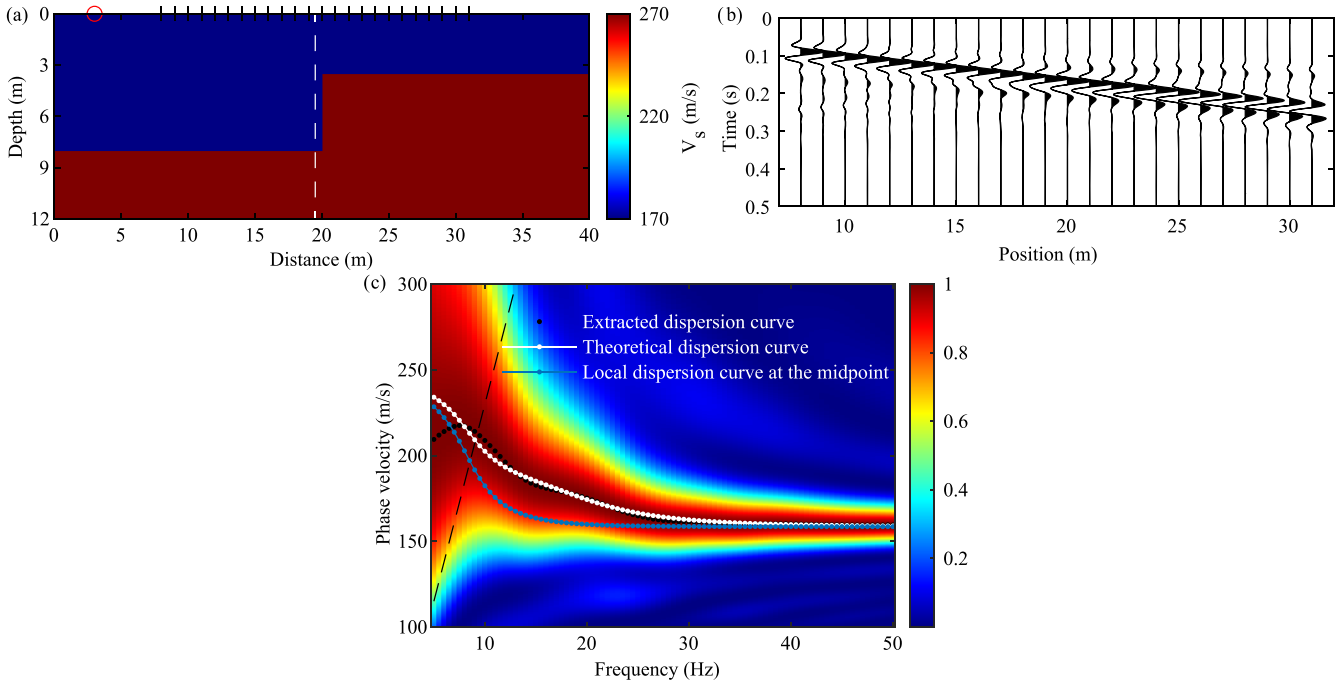

Figure A1. (a) *Vs* model. The red circle and black lines indicate the position of the source and receivers, respectively. The white dash line represents the midpoint of the receiver array. (b) Simulated common-shot gather. (c) The dispersion image corresponding to (b). The black dot line, the white dashed–dotted line, and the blue dashed–dotted line represent the extracted dispersion curve, the theoretical dispersion curve, and the local dispersion curve at the midpoint of the receiver array, respectively. The black dash line indicates the wavelength of $\lambda = (N_{\mathrm{rec}}-1)\Delta r = 23$ m.

APPENDIX B: DATA PROCESSING AND PARAMETER SELECTION

Fig. B1 explains the principle of data windowing and dispersion stacking:

1. Within a specified spatial window, the local wavefield is extracted from the shot gathers with a source offset inside the maximum offset d.
2. The amplitude-normalized beamforming (eq. 3) is used to transform the extracted local wavefields into the frequency-phase velocity domain for producing several dispersion images, which are summed to generate a stacked dispersion image.
3. A stacked dispersion curve with high SNR is picked on the stacked dispersion image; the dispersion curves for each single shot are also measured for estimating the experimental uncertainty σ. The stacked dispersion curve with experimental uncertainty is assigned at the midpoint of the window.
4. The spatial window is moved to the next position according to the shift ΔW, and steps 1–3 are repeated.

The processing parameters used in multiscale window analysis are determined by test. Based on the guideline for selecting the array length (Foti *et al.* 2018), we choose two times the investigation depth as the maximum window length $W_{\max}$. Using the chosen $W_{\max}$, we gradually increase the maximum offset d and check the quality of the stacked dispersion image. Keep in mind that the use of noisy far-offset signals can deteriorate the stacked dispersion images (Socco *et al.* 2009; Pasquet & Bodet 2017). After selecting $W_{\max}$ and d, we progressively decrease the window size until the loss in spectral resolution and SNR becomes too high to fail to identify the dispersion trend. Finally, we choose one receiver spacing Δr as the shift ΔW for all the spatial windows.

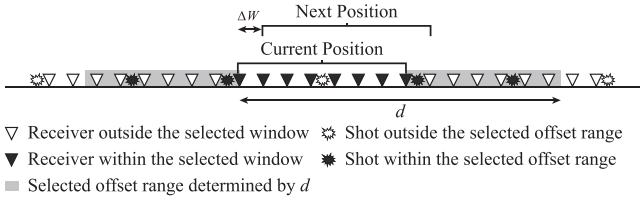

Figure B1. Data windowing and gather selection.

APPENDIX C: JACOBIAN MATRIX

Since the theoretical phase velocity $v(\omega)$ is the local maximum of eq. (6), we obtain

$$\frac{\partial P\left(\omega, v\left(\omega\right)\right)}{\partial v\left(\omega\right)} = -\frac{2\omega}{v(\omega)^2}\sum_{m=1}^{L-1}\sum_{n=m+1}^{L}\left\{x_{mn}\sin\left[\omega\left(\sum_{q=1}^{Q}\frac{\Delta x}{v_q\left(\omega\right)} - \frac{x_{mn}}{v\left(\omega\right)}\right)\right]\right\} = 0. \tag{C1}$$

After taking a derivative operation with respect to the tth S-wave velocity model V_s^t, eq. (C1) becomes

$$\frac{\partial v\left(\omega\right)}{\partial V_s^t} = \frac{v^2\left(\omega\right)\sum_{m=1}^{L-1}\sum_{n=m+1}^{L}\left\{x_{mn}\cos\left[\omega\left(\sum_{q=1}^{Q}\frac{\Delta x}{v_q\left(\omega\right)} - \frac{x_{mn}}{v\left(\omega\right)}\right)\right]\sum_{q=1}^{Q}\frac{\Delta x}{v_q^2\left(\omega\right)}\frac{\partial v_q\left(\omega\right)}{\partial V_s^t}\right\}}{\sum_{m=1}^{L-1}\sum_{n=m+1}^{L}\left\{x_{mn}^2\cos\left[\omega\left(\sum_{q=1}^{Q}\frac{\Delta x}{v_q\left(\omega\right)} - \frac{x_{mn}}{v\left(\omega\right)}\right)\right]\right\}}, \tag{C2}$$

where, $\frac{\partial v_q(\omega)}{\partial V_s^t}$ is the partial derivative of the local phase velocity with respect to the S-wave velocity model V_s^t.

The Jacobian matrix $\mathbf{J}(\mathbf{m})$ contains all the partial derivatives

$$J_{st} = \frac{\partial d_s}{\partial m_t} = \frac{\partial \log\left(v_s\left(\omega\right)\right)}{\partial \log\left(V_s^t\right)} = \frac{V_s^t}{v_s\left(\omega\right)}\frac{\partial v_s\left(\omega\right)}{\partial V_s^t} \tag{C3}$$

for the sth logarithmic phase velocity in the data vector and the tth logarithmic S-wave velocity in the model vector.