Geophysical Journal International

Geophys. J. Int. (2019) **217,** 346–360
Advance Access publication 2019 January 12
GJI Seismology

doi: 10.1093/gji/ggz005

Wave equation dispersion inversion of surface waves recorded on irregular topography

Jing Li [1,2] Fan-Chi Lin,[3] Amir Allam,[3] Yehuda Ben-Zion,[4] Zhaolun Liu [2] and Gerard Schuster[2]

[1] *College of Geo-Exploration Science and Technology, Jilin University, Changchun 130021, China. E-mail: inter.lijing@gmail.com*
[2] *Department of Earth Science and Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal* 23955-6900*, Kingdom of Saudi Arabia*
[3] *Department of Geology and Geophysics, University of Utah, Salt Lake City, UT 84112, USA*
[4] *Department of Earth Sciences, University of Southern CA, Los Angeles, CA 90089-0740, USA.*

Accepted 2019 January 4. Received 2018 November 13; in original form 2017 May 16

SUMMARY

Significant topographic variations can strongly influence the amplitudes and phases of propagating surface waves. Such effects should be taken into account, otherwise the *S*-velocity model inverted from the Rayleigh or Love dispersion curves will contain significant inaccuracies. Here, we show that the recently developed wave equation dispersion inversion (WD) method naturally takes into account the effects of topography to give accurate *S*-velocity tomograms. Application of topographic WD (TWD) to synthetic data demonstrates that TWD can accurately invert for the *S*-velocity model from dispersion curves computed from data recorded over variable topography. This method is applied to ambient noise seismic field data recorded along the Clark strand of the San Jacinto fault zone in Southern California. The resulting TWD *S*-velocity tomogram appears to be more consistent with the geology than the standard WD tomogram.

Key words: surface wave and free oscillations; Tomography; Waveform inversion.

1 INTRODUCTION

There are a number of studies that demonstrate surface waves can strongly scatter from topographic variations along the recording surface (Davies & Heathershaw 1984; Snieder 1986; Spetzler *et al.* 2002; Borisov *et al.* 2016; Nuber *et al.* 2016). Unless these topographic effects are taken into account, the *S*-velocity model inverted from the surface waves will contain significant inaccuracies (Pageot *et al.* 2016; Borisov *et al.* 2017). As an example, Figs 1(a) and (b) depict a velocity model with a sloped topography surface and a horizontal surface. The model size is $40 \times 120\,\mathrm{m}^2$ where the grid spacing and time sampling intervals are 1 m and 0.03 ms, respectively, and the centre frequency of the source wavelet is 30 Hz. We used a 2-D elastic finite-difference algorithm with a topography-free surface condition (Robertsson 1996; Wang *et al.* 2015) to compute the vertical-component shot gather shown in Fig. 1(c) for sources and receivers on the sloped surface. This compares to the traces in Fig. 1(d) recorded along a horizontal free surface denoted by the black dots and red stars in Fig. 1(b). It is obvious that there are notable differences between the shot gather in Fig. 1(c) compared to the one recorded over a horizontal surface in Fig. 1(d). In addition, Nuber *et al.* (2016) and Pan *et al.* (2018) demonstrated through synthetic tests that even minor topographic variations of the free surface will have a significant effect in full waveform inversion (FWI). They found that neglecting topography with an amplitude fluctuation greater than half the minimum seismic wavelength leads to appreciable artefacts in the inversion image (Nuber *et al.* 2016).

Li & Schuster (2016) developed a wave equation dispersion inversion (WD) method for inverting dispersion curves associated with surface waves. Liu *et al.* (2018) extended it to 3-D case. This method has the benefit of robust convergence compared to the tendency of FWI (Masoni *et al.* 2014; Solano *et al.* 2014; Yuan *et al.* 2015; Kohn *et al.* 2016) to getting stuck in local minima. It has the advantage over the traditional inversion of dispersion curves (Xia *et al.* 1999; Maraschini *et al.* 2010; Socco *et al.* 2010) in that it does not assume a layered model and is valid for arbitrary 2-D or 3-D media. The standard WD method can also account for topographic effects by incorporating the free-surface topography into the finite-difference solution to the elastic wave equation. We will denote the WD method as the topographic WD (TWD) if it takes topography into account by solving the elastic wave equation for sources and receivers on the actual topography of the free surface.

In this paper, we demonstrate the effectiveness of the TWD method for surface waves recorded on the free surface with strong variations in topography. First, we briefly review the theory of WD with mathematical details provided in Li *et al.* (2017a). This section also provides

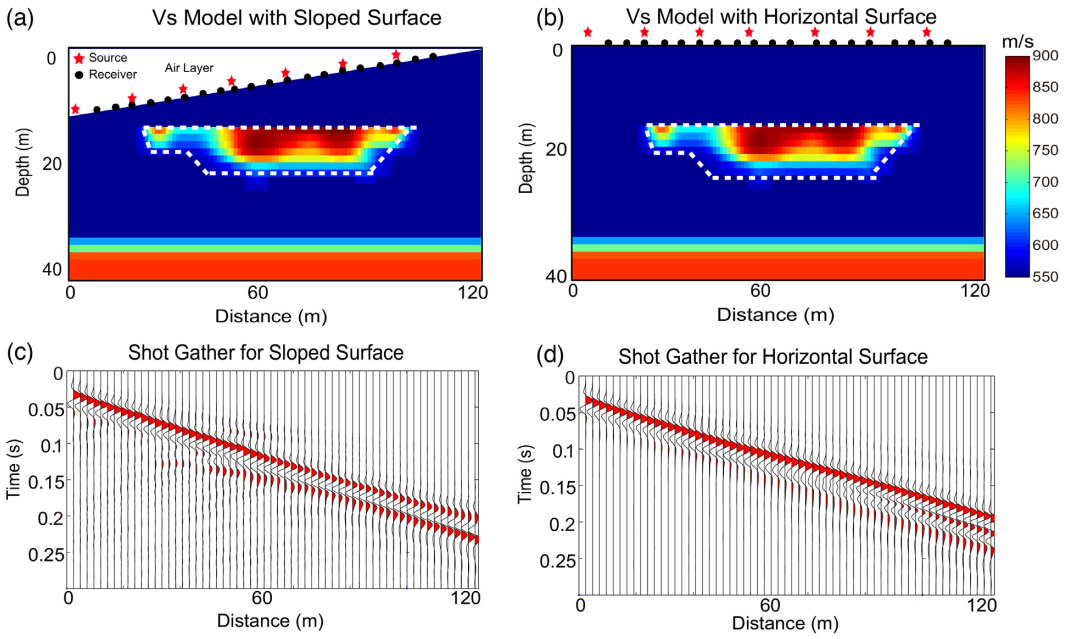

Figure 1. (a) and (b) are *S*-velocity models with sloped and horizontal free surfaces, respectively. (c) and (d) are the corresponding shot gathers for vertical-component records computed by finite-difference solutions to the 2-D elastic wave equation.

the workflow for implementing the TWD method for traces recorded on irregular topography. The next section presents numerical results that demonstrate the efficacy of the TWD method for synthetic data and field data recorded along a line with significant topographic variations in Southern California. The topography of the surface is similar in both the synthetic and field data examples with an elevation change of over 300 m along a recording line with length 2.5 km. The final section presents the conclusions.

2 THEORY OF TWD

The basic theory of TWD is same as that for the WD method (Li *et al.* 2017a), which inverts Rayleigh dispersion curves for the *S*-velocity model that minimizes the dispersion-curve misfit function ε. The difference is that the finite-difference solution to the 2-D elastic wave equation of TWD needs to consider the topography-free surface condition. The dispersion-curve misfit function is defined as

$$\epsilon = \frac{1}{2}\sum_{\omega}(\overbrace{\kappa(\omega) - \kappa(\omega)^{\text{obs}}}^{\text{residual}=\Delta\kappa(\omega)})^2, \tag{1}$$

where $\kappa(\omega)$ represents the predicted dispersion curves picked from the simulated spectrum and $\kappa(\omega)^{\text{obs}}$ denotes the dispersion curve obtained from the recorded spectrum. In practice, these spectra are computed by applying a linear Radon transform (LRT; Luo *et al.* 2009; Li & Hanafy 2016) to the common shot gather in the frequency domain. Any order of dispersion curve can, in principle, be picked and inverted, but for the examples in this paper, we only use the fundamental mode of Rayleigh waves.

The slowness gradient $\gamma(\mathbf{x})$ of eq. (1) is given by

$$\gamma(\mathbf{x}) = \frac{\partial\epsilon}{\partial s(\mathbf{x})} = \sum_{\omega}\Delta\kappa(\omega)\frac{\partial\kappa(\omega)}{\partial s(\mathbf{x})}, \tag{2}$$

where $s(\mathbf{x})$ represents the shear-slowness model. The optimal shear-slowness model $s(\mathbf{x})$ is obtained from the steepest-descent formula:

$$s(\mathbf{x})^{(k+1)} = s(\mathbf{x})^{(k)} - \alpha\sum_{\omega}\Delta\kappa(\omega)\frac{\partial\kappa(\omega)}{\partial s(\mathbf{x})}, \tag{3}$$

where α is the step length by any backtracking line-search method (Nocedal & Wright 1999) and the superscript (k) denotes the kth iteration. For pedagogical simplicity, we assume a single shot gather $\hat{D}(\mathbf{g}, \omega)$ for a source at $\mathbf{s}$ and geophones at $\mathbf{g}$, and the notation for the source location is silent. The misfit function will include an additional summation over different shot gathers if more than one shot gather is used.

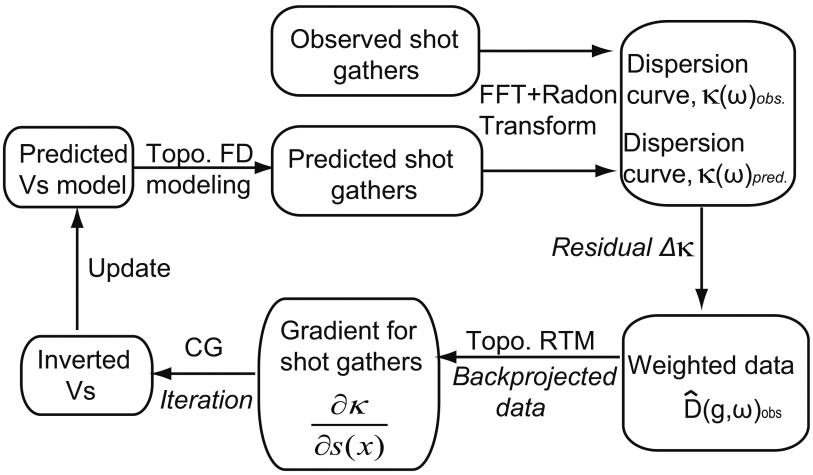

Figure 2. The workflow for implementing the TWD (Li *et al.* 2017b).

Mathematical details for deriving the Fréchet derivative $\frac{\partial\kappa(\omega)}{\partial s(\mathbf{x})}$ are given in Li *et al.* (2017a), where

$$\frac{\partial\kappa}{\partial s(\mathbf{x})} = \frac{4s(\mathbf{x})\rho(\mathbf{x})}{A}\text{Real}\left\{ \overbrace{\frac{\partial D(\mathbf{x},\omega)_x}{\partial x}}^{\text{source}=f(\mathbf{x},\mathbf{s},\omega)_1} \int \overbrace{\hat{D}(\mathbf{g},\omega)^*_{\text{obs}}\left(\frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial z}\right)}^{\text{backprojected data}=B(\mathbf{x},\mathbf{s},\omega)^*_1} dx_g + \overbrace{\frac{\partial D(\mathbf{x},\omega)_z}{\partial z}}^{\text{source}=f(\mathbf{x},\mathbf{s},\omega)_2} \int \overbrace{\hat{D}(\mathbf{g},\omega)^*_{\text{obs}}\frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial x}}^{\text{backprojected data}=B(\mathbf{x},\mathbf{s},\omega)^*_2} dx_g \right.$$
$$\left. -\frac{1}{2}\overbrace{\left(\frac{\partial D(\mathbf{x},\omega)_z}{\partial x}+\frac{\partial D(\mathbf{x},\omega)_x}{\partial z}\right)}^{\text{source}=f(\mathbf{x},\mathbf{s},\omega)_3} \int \overbrace{\hat{D}(\mathbf{g},\omega)^*_{\text{obs}}\left(\frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial z}+\frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial x}\right)}^{\text{backprojected data}=B(\mathbf{x},\mathbf{s},\omega)^*_3} dx_g \right\}. \quad (4)$$

Here, $\hat{D}(\mathbf{g},\omega)^*_{\text{obs}}$ is the harmonic weighted conjugated data for the vertical-component displacement recorded at $\mathbf{g}$ (Appendix A), $G(\mathbf{g}|\mathbf{x})_{\mathbf{ij}}$ is the Green's tensor for the ith component of the displacement field at (x, z) for the jth-component of the virtual source at $(x_g, 0)$, $s(\mathbf{x})$ is the shear-slowness model for the background medium and A is the normalization term. Here, $f(\mathbf{x}, \mathbf{s}, \omega)_i$ for $i \in (1, 2, 3)$ is the ith-component of the downgoing source field at $\mathbf{x}$ that originates at $\mathbf{s}$, and $B(\mathbf{x}, \mathbf{s}, \omega)_i$ represents the ith-component of the backprojected scattered field at $\mathbf{x}$ for a source at $\mathbf{s}$. The interpretation of eq. (3) is that the traces are weighted by terms proportional to the wavenumber residual $\Delta\kappa(\omega)$ and backpropagated into the medium along wavepaths to update the slowness of the *S*-velocity model.

2.1 Workflow for WD with topography

The steps for implementing the TWD algorithm are the following (Fig. 2).

(1) Mute the body waves and scattered waves in the observed and predicted shot gathers by adaptive windowing. Then apply a 1-D Fourier transform (FFT) along the time axis of the shot gather to get the frequency spectrum of each trace in the shot gather. The predicted shot gather is computed by a 2-D finite-difference solution to the elastic wave equation with a free-surface condition along the topographic boundary (Wang *et al.* 2012).

(2) Apply an LRT (Appendix B) to the frequency spectra of the predicted and observed shot gathers to get the phase-velocity curve $C(\omega)$ of the fundamental Rayleigh mode. Here, the finite-difference modelling of the elastic wave equation is for sources and receivers on an irregular free surface. The fundamental dispersion curves are automatically picked according to the maximum amplitudes of the magnitude spectrum that are nearest to the slowness axis. Details for picking the dispersion curves are described in Li *et al.* (2017a).

(3) Select the linear gradient velocity model as the initial model. Use eq. (A3) (Appendix A) to calculate the weighted data $\hat{D}(\mathbf{g}, \omega)_{\text{obs}}$, which is then used for constructing the backprojected data $B(\mathbf{x}, \mathbf{s}, \omega)$ in eq. (4). The source field $f(\mathbf{x}, \mathbf{s}, \omega)$ in eq. (4) is also computed by a finite-difference solution to the elastic wave equation.

(4) Estimate the step-length α by any backtracking line-search method (Nocedal & Wright 1999).

(5) The gradients for each migrated shot gather are added together to get the *S*-velocity update. The background *S*-velocity model is updated and the above steps are repeated until the RMS residual falls below a specified value.

The multi-offset WD strategy is used to update the *S*-velocity model. We first invert for higher frequencies and shorter offsets to invert for the shallow part of the model. Then we subsequently increase the offsets and lower the frequencies to invert for the deeper areas of the model. The details are decribed in Appendix C.

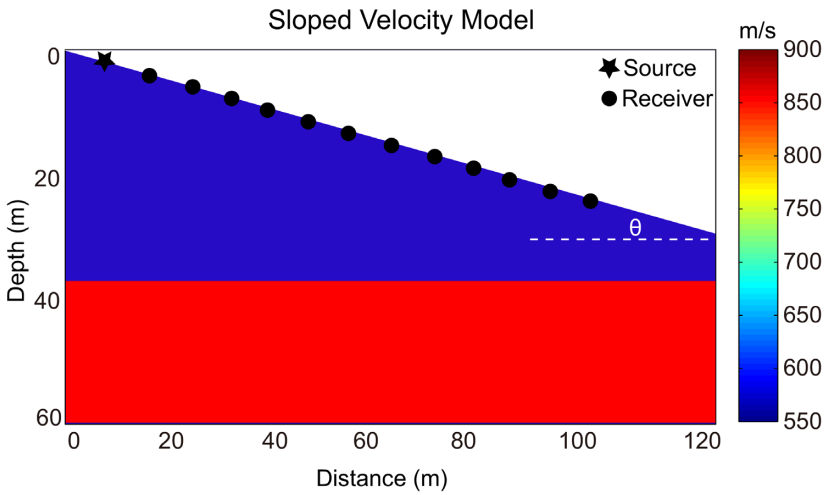

Figure 3. *S*-velocity model with a sloped free surface. The star and circles represent the source and receivers and θ is the angle of the sloped free surface.

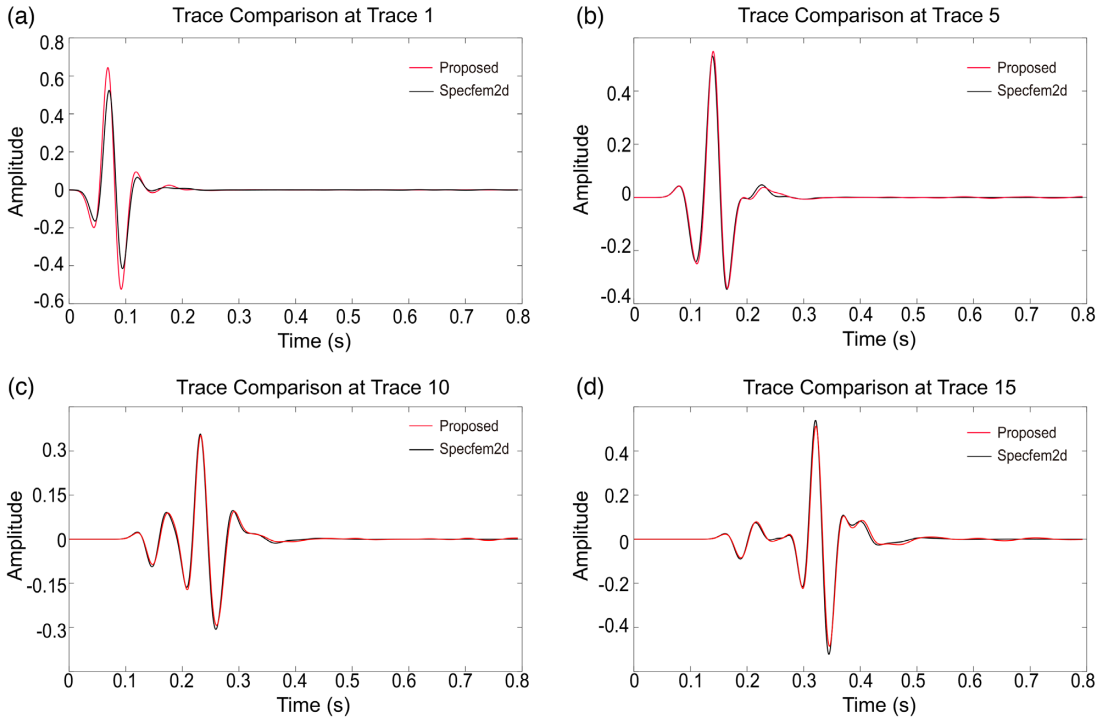

Figure 4. Vertical component traces recorded in the velocity model with a sloped free surface in Fig. 3 for the source–receiver distances of (a) 0 m, (b) 20 m, (c) 40 m and (d) 60 m. The red lines are the records generated with our method (Wang *et al.* 2012) and the black lines are the traces computed by the code SPECFEM2D (Komatitsch *et al.* 2012).

2.2 Accuracy test

First, we compare the accuracy of the elastic finite-difference solutions to those computed by a spectral element method. The spectral element code is the SPECFEM2D software (Komatitsch & Tromp 1999) and the test model is a two-layered *S*-velocity model with a sloped free surface (Fig. 3). The model size is $60 \times 120\,\mathrm{m}^2$ where the grid-spacing and time-sampling intervals for the 2-D elastic finite-difference algorithm are 1 m and 0.025 ms, respectively. The centre frequency of the source wavelet is 30 Hz. The star and circles denote the source and receivers, respectively, and θ is the angle of the sloped free surface. The observed data are generated by one shot at the beginning of the survey line and the traces are recorded by 60 receivers evenly distributed every 2 m on the surface. The single-trace vertical-component records are displayed in Fig. 4 at source–receiver offsets of 0, 20, 40, and 60 m, respectively. The modelling results from SPECFEM2D are also displayed and demonstrate that our finite-difference solutions are nearly the same as the spectral element method solutions.

Then, we test the accuracy of the modelling algorithm by comparing the phase velocities computed for two different models. The first model shown in Fig. 5(a) is a two-layer model with a flat free-surface boundary as depicted by the dashed white line AB. Then, this model is rotated counterclockwise by 10 degrees to obtain the second model in Fig. 5(b). The model size is $50 \times 120\,\mathrm{m}^2$ where the grid spacing

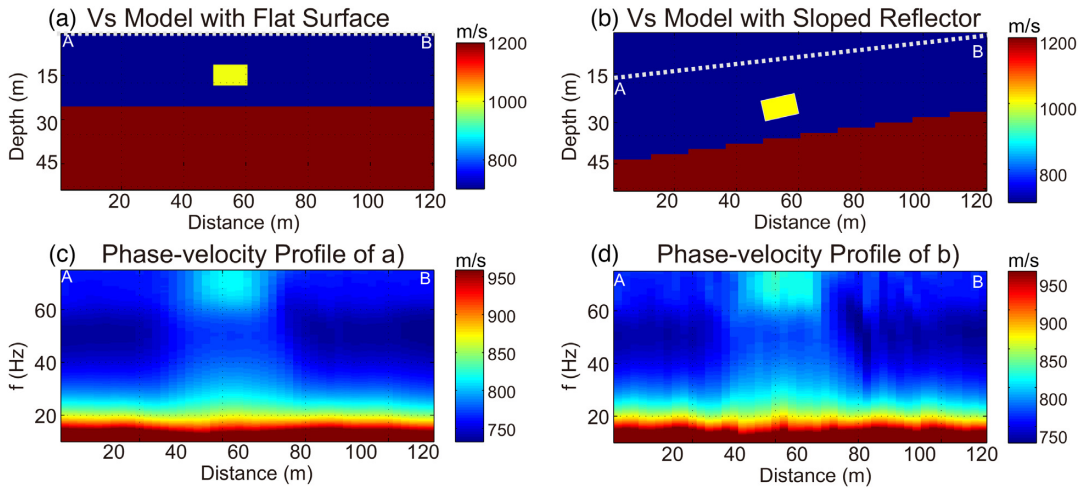

Figure 5. (a) and (b) are *S*-velocity models with flat and sloped free surfaces, (c) phase-velocity profiles for model (a), and (d) phase-velocity profiles for the sloped free-surface model (b). The white dashed line AB represents the free surface. The (d) image is displayed with respect to the Cartesian coordinate system where a coordinate axis is aligned along the sloped free surface.

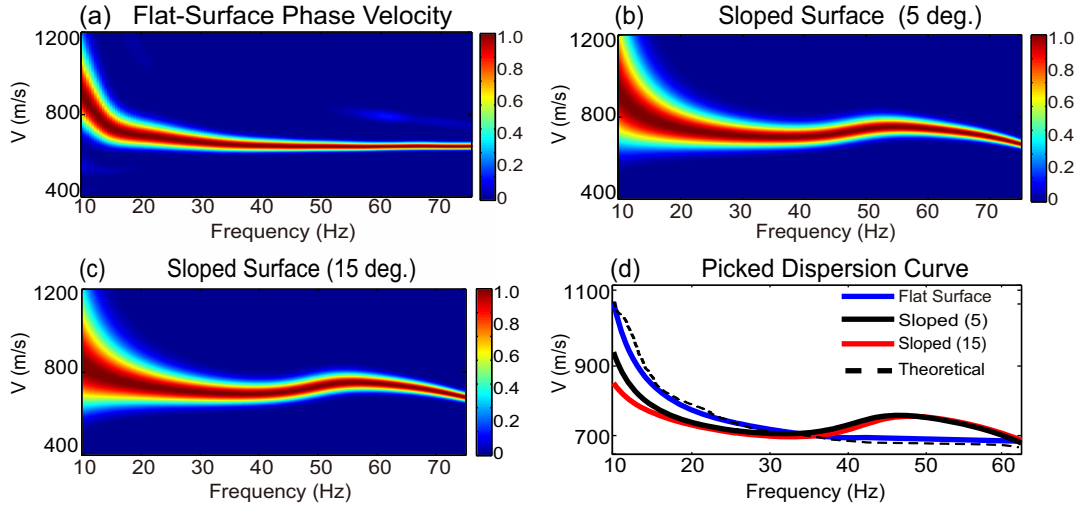

Figure 6. The dispersion curves for free surfaces with slope angles of (a) 0, (b) 5 and (c) 15 degrees. The thickness of the slow layer beneath the free surface varies with slope angle, which naturally leads to variations in the (d) dispersion curves.

and time sampling intervals for the 2-D elastic finite-difference algorithm are 1 m and 0.025 ms, respectively, and the centre frequency of the source wavelet is 30 Hz. The observed data are generated from 30 shots and the traces for each shot are recorded by 60 receivers evenly distributed every 2 m on the surface. We select 15 near-offset traces for each shot to compute the dispersion curve. Then, 30 shot gathers can be used to generate 30 distinct dispersion curves. These dispersion curves can be concatenated together to get a pseudosection of the phase-velocity distribution as a function of frequency. If the modelling algorithm is accurate, then the phase-velocity profiles should be nearly the same for each model. The phase-velocity profiles shown in Figs 5(c) and (d) mostly resemble one another, where the high-wavenumber discrepancies are due to stair-step diffractions along the dipping interface in Fig. 5(d).

2.3 The effect of surface topography on dispersion curves

The influence of topography on Rayleigh waves dispersion curves is analysed for the model in Fig. 3 with different slope angles θ. We now apply the FFT and the LRT to a selected part of the data to estimate the phase-velocity curves $C(\omega)$, where the minimum and maximum source–receiver offsets are 10 and 60 m, respectively. Figs 6(a)–(c) show the phase-velocity profiles as a function of frequency for free surfaces with different slope angles. Fig. 6(d) depicts the picked dispersion curves for the flat and dipping-free surfaces and indicates that the dispersion curves are notably influenced by surface topography (Pageot *et al.* 2016). In addition, the dispersion curves for the flat free surface (blue line) closely resemble the black theoretical dispersion curves (Haskell 1953).

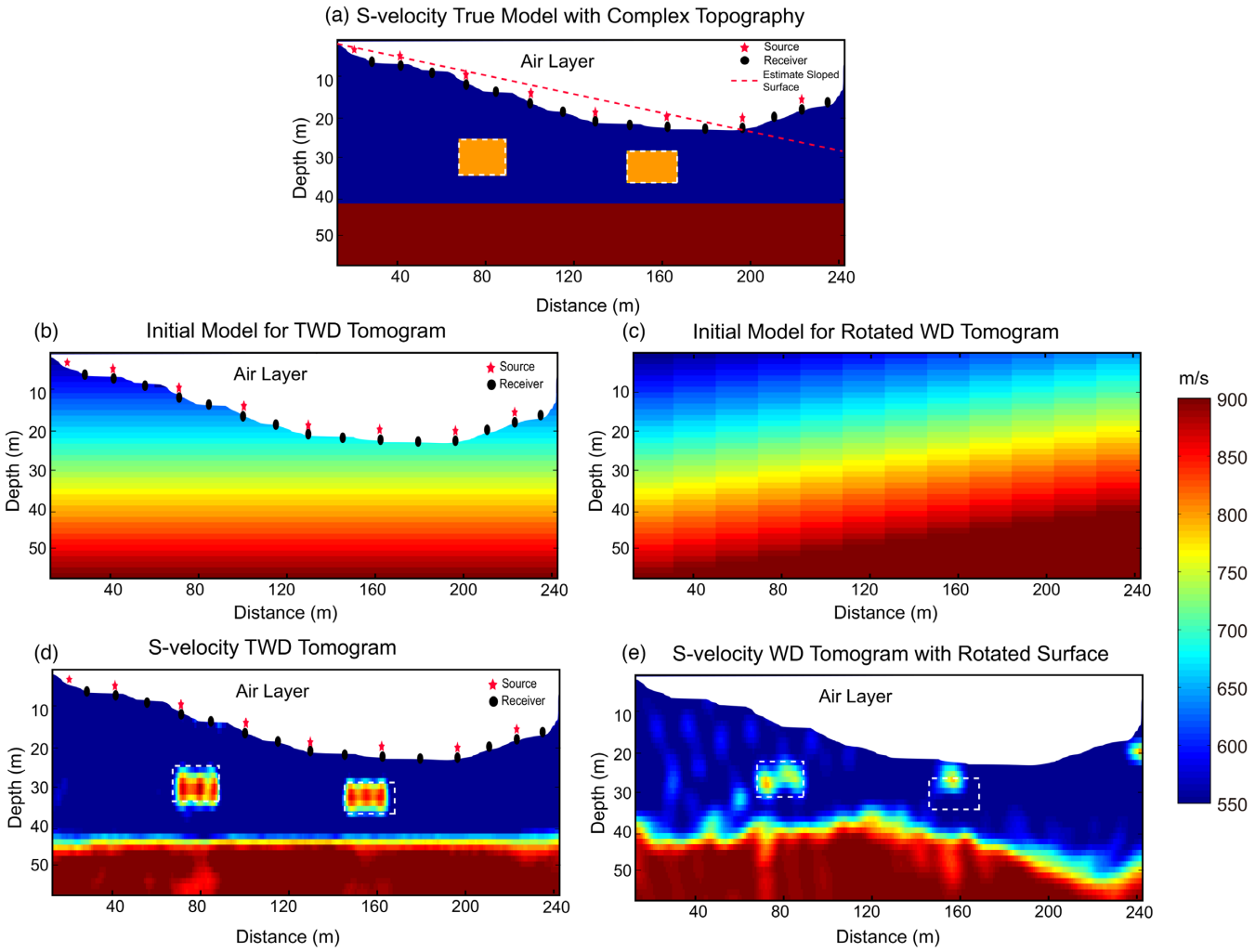

Figure 7. (a) *S*-velocity model with an irregular free surface, (b) and (c) are the initial models for the TWD and rotated WD tomograms. The predicted data for the WD method are computed along a straight sloped surface that rotates the initial model to obtain a horizontal surface. (d) and (e) are the TWD and rotated WD tomograms, respectively.

3 NUMERICAL TESTS

The effectiveness of the TWD method is now demonstrated with synthetic and field data examples, where the data are recorded on surfaces with significant variations in topography. The synthetic example is a complex topographic model with both canyons and horst-like features beneath the free surface. The initial *S*-velocity model consists of a linear velocity gradient $v_s(z) = v_0 + \alpha z$, where v_0 is the *S*-velocity at the surface and $\alpha = (v_{\max} - v_{\min})/Z$ is the velocity gradient with respect to depth. Z is the depth and $v_{\max}$ and $v_{\min}$ are the maximum and minimum *S*-velocity values in the model, respectively. In these examples, the *P*-wave velocity is updated by assigning $v_p = \sqrt{3}v_s$ and the density is taken to be a constant value of $1.5\,\mathrm{g\,cm^{-3}}$. The field-data example is computed from ambient noise recorded in Southern California. In the synthetic examples, the observed data are generated by a staggered-grid solution of the 2-D elastic wave equation for a free surface with strong topographic variations. The WD tomograms that neglect topography are computed from predicted data with sources and receivers on a horizontal free surface. In this case, the predicted data do not take into account the phase-velocity variations in the observed surface waves recorded on the irregular recording topography.

3.1 Synthetic model

For the synthetic topographic model test, the recording surface along the irregular free surface approximates that for the ambient seismic survey carried out in Southern California (Fig. 7a), where the black dots and red stars denote the positions of the receivers and sources, respectively. The model size is $50 \times 240\,\mathrm{m^2}$ where the grid spacing and the time sampling intervals for the 2-D elastic finite-difference algorithm are 1 m and 0.03 ms, respectively, and the central frequency of the source wavelet is 25 Hz. The observed data are generated by 60 shots evenly distributed every 4 m and the data are recorded by 120 receivers evenly distributed every 2 m spaced across the irregular free surface. The minimum and maximum source–receiver offsets are 10 and 236 m, respectively.

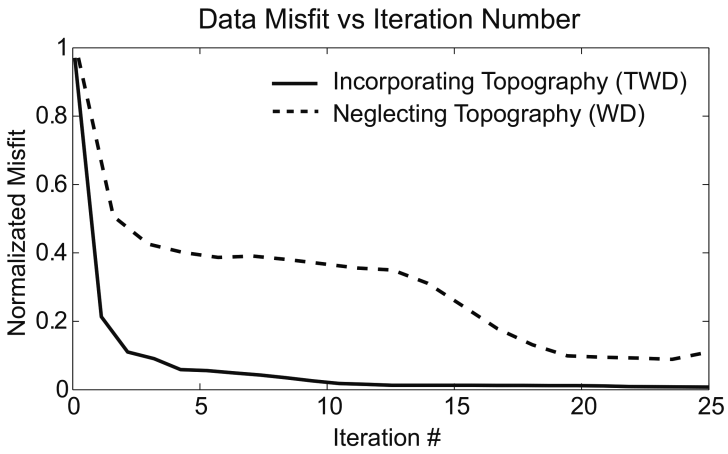

Figure 8. Plot of normalized RMS residual versus iteration number for the synthetic examples. The *y*-axis represents the normalized frequency-shift data residual, and the solid and dashed lines represent the WD results that, respectively, take into account and neglect the effects of an irregular free surface.

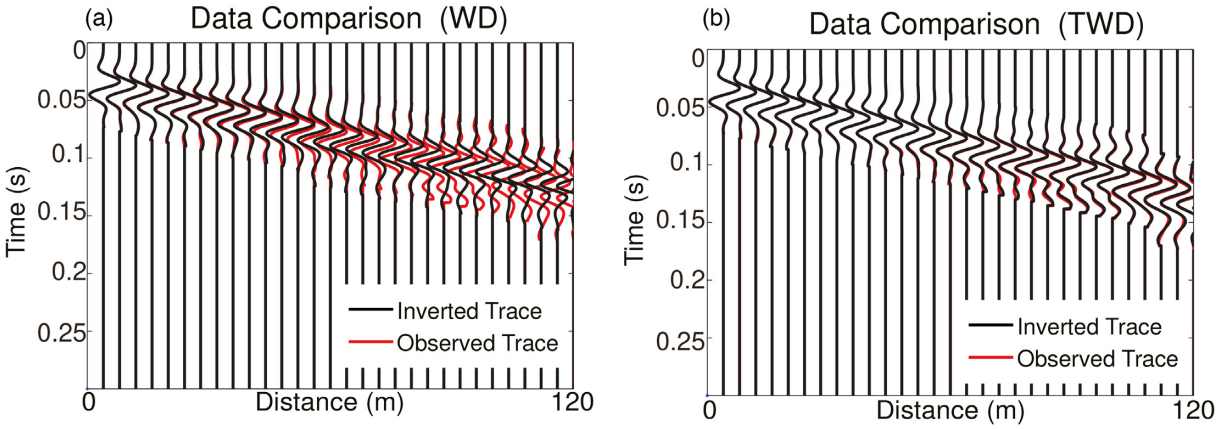

Figure 9. Shot gather comparison between the observed red traces and the inverted black traces that (a) neglect the irregular topography and (b) take into account the irregular topography. In both (a) and (b), a matched filter was used to correct for the source wavelet at the nearest offset trace, and then this same filter was applied to the other traces.

The initial model for the TWD is shown in Fig. 7(b) and the *S*-velocity tomogram is displayed in Fig. 7(d). For WD, we interpolate the topography by a straight line shown as the red dashed line in Fig. 7(a). All the sources and receivers are projected to the straight line. We use the initial model shown in Fig. 7(c) for WD, which is modified from the initial model of TWD by interpolating the true topography by the red dashed line and then rotating the model to the horizontal surface. Here, the straight sloped surface (red dashed line) does not match with the true irregular surface, especially in the right-hand side. The result is that the WD tomogram inverted from these data does not accurately represent the velocity anomalies (Fig. 7e). In contrast, the TWD tomogram is much more accurate than the WD tomogram that neglects topography. In addition, Fig. 8 shows that the normalized TWD residual (solid line) is much less than the WD residual (dashed line) that neglects topographic effects. After 25 iterations, the TWD residual is lower than 0.1 but the standard WD residual is more than 0.2. Figs 9(a) and (b) show the waveform comparison for both TWD and WD, where the TWD traces more closely resemble the observed ones for each receiver.

3.2 Field data tests

The TWD and WD methods are now tested on ambient noise data recorded along the Clark strand of the San Jacinto fault zone (SJFZ) in Southern California, USA (Fig. 10a). The location of the field experiment is shown in Fig. 10(a) (red square) and the white dashed lines denote the SJFZ along with the Elsinore and San Andreas Faults. Fig. 10(b) shows the location of each recording station across the fault in a month-long deployment of a linear array of 134 Fairfield three-component nodal geophones (5 Hz corner frequency) that have a variable sampling interval (up to 0.5 ms). Each of the receivers recorded autonomously for up to 36 days. With a recording aperture of 2.4 km and a mean station spacing of 20 m, the array locally spans the fault zone located over a low-velocity crustal block on the South-West (SW) next to the damaged structure of the fault that is adjacent to a high-velocity crustal block on the North-East (NE) (Allam & Ben-Zion 2012; Ben-Zion *et al*. 2015; Zigone *et al.* 2015; Share *et al.* 2017). The data were continuously recorded for 36 days at a 1000 Hz sampling rate. Fig. 10(c)

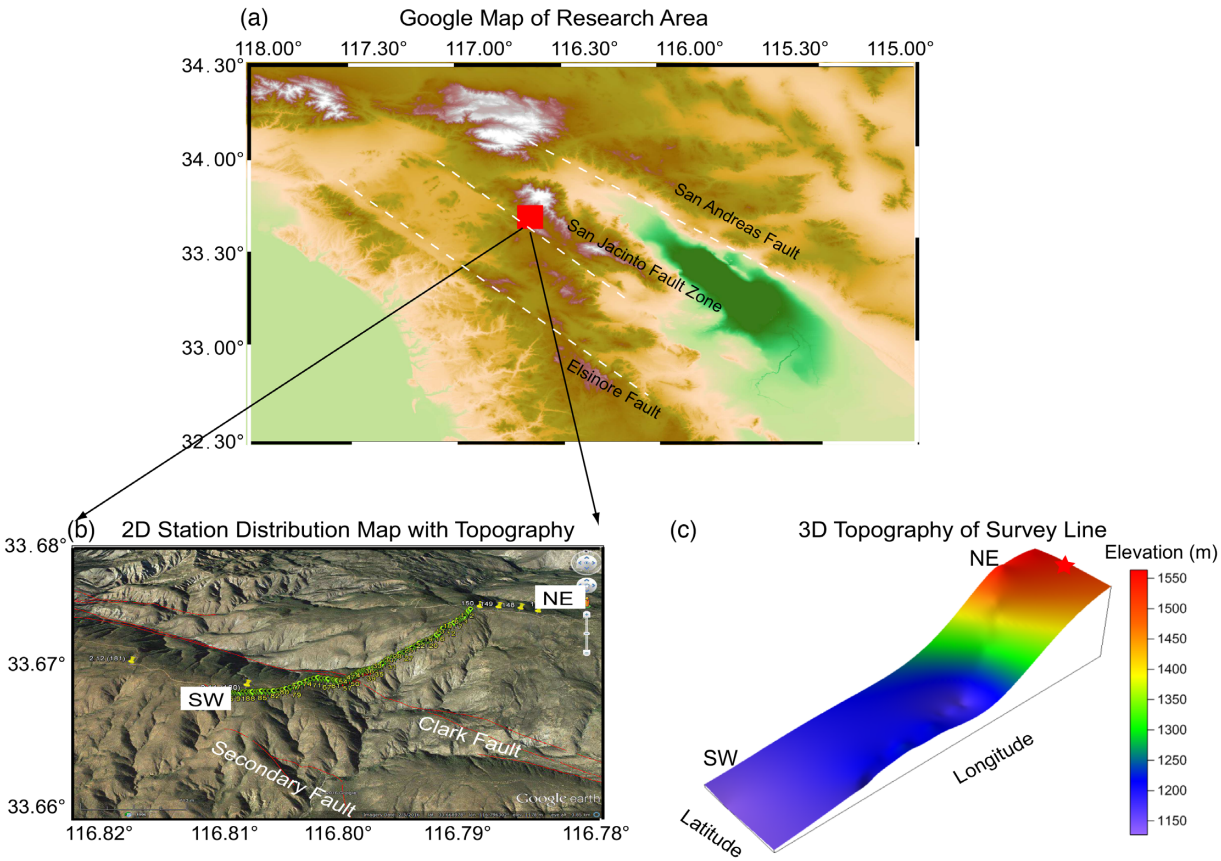

Figure 10. The survey line where the ambient noise data are recorded along the crest of a mountain in Southern California. Illustration (a) is the study region centred on the Clark fault in the San Jacinto Fault Zone (SJFZ), (b) is the 2-D station distribution map and (c) depicts the 3-D topographic map associated with the survey line.

shows the irregular topography along the survey line, where the largest elevation difference is more than 300 m. Before TWD inversion, the following processing steps are applied to the raw ambient-noise data (Bensen *et al.* 2007; Lin *et al.* 2013)

(1) Remove instrument response, remove mean, remove trend, band-pass filter and segment the entire trace to a sequence of shorter traces .
(2) Apply time domain normalization and spectral whitening.
(3) Cross-correlation and temporal stacking of the traces (Lin *et al.* 2013).

A virtual shot gather for a virtual shot at an endline recording station is shown in Fig. 11(a). Adding the time-reversed acausal portion of the trace with the causal portion gives the result shown in Fig. 11(b). There is strong energy near the zero time which is likely caused by body waves from local earthquakes. The shot gathers are muted so only the strongest surface waves are retained, as shown in Fig. 11(c). The dispersion curve associated with a shot gather and the source at station 2 is shown in Fig. 11(d). According to the available geological and seismological information (Sharp 1967; Rempe *et al.* 2013; Share *et al.* 2017), the Clark strand of the SJFZ in this area is at station 55 (The red line in Fig. 10b).

We computed 125 virtual common shot gathers with virtual sources space at about 20 m intervals along the recording line (Fig. 10b). Each virtual shot is recorded by 125 receivers with an average spacing of 20 m. The Radon transform and Fourier transform are applied to each shot gather, and the fundamental dispersion curves are picked. Both the standard WD and TWD methods are then applied to these dispersion curves using the starting *S*-velocity model. This starting model was estimated from the 1-D inversion of these dispersion curves (Li & Hanafy 2016). The initial model for the TWD tomogram includes the actual topographic free surface shown in Fig. 12(a). For the initial model in the WD tomogram, we estimate a straight sloped surface (the black dashed line in Fig. 12a) according to the actual topography in Fig. 11 and then rotate the model to the horizontal free surface (Fig. 12b). Fig. 12(d) shows the standard WD *S*-velocity tomogram and indicates that there is a low-velocity zone in the offset range 900 m$<$x$<$1400 m, which is generally consistent with the local damaged structure of the SJFZ and other major faults (Rempe *et al.* 2013; Zigone *et al.* 2015; Share *et al.* 2017). The black dots and red stars are the receivers and sources position. The TWD *S*-velocity tomogram in Fig. 12(d) shows an obvious low-velocity anomaly that coincides with the low-velocity zone observed in the WD tomogram. In addition, the TWD tomogram reveals a sharper velocity contrast across the fault that is more consistent with the expected geology.

As a final sanity check, Fig. 12(e) shows a common offset gather (COG) with the source–receiver offset of 200 m. The white dashed line in this figure indicates the location of the main fault in the offset range 900 m$<x<$1000 m. Compared to the WD tomogram, the locations

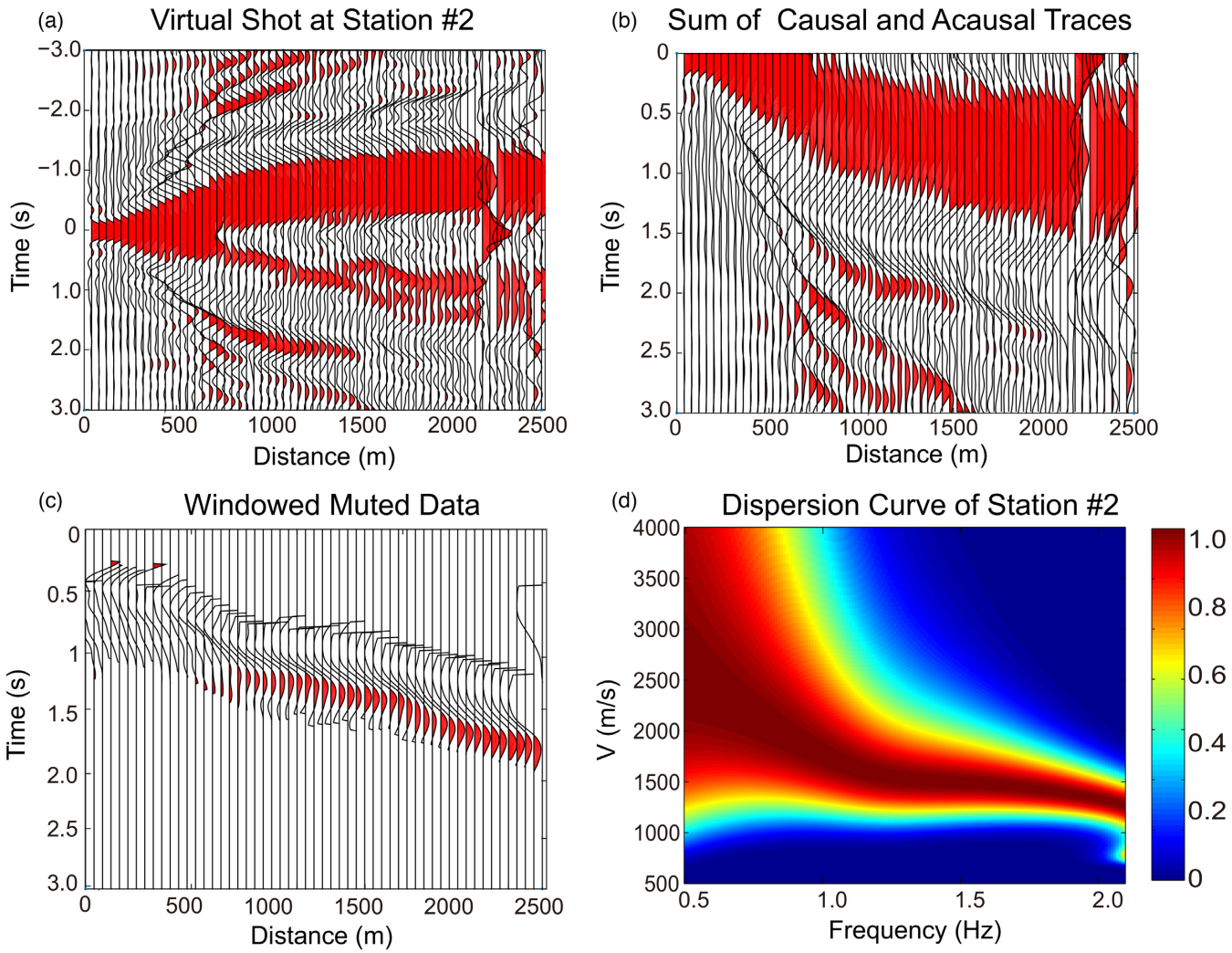

Figure 11. Data processing results for the Southern California data. (a) Virtual shot gather at station No. 2, (b) the sum of the causal and time-reversed non-causal parts of the shot gather, (c) the sum of the shot gather after windowing, (d) and the dispersion curve for station No. 2.

of the low-velocity anomalies in the TWD tomogram are more consistent with the interpreted faults (white lines) in the COG profile. For example, there are strong indications of antithetic faulting seen in the TWD tomogram and COG not seen in the WD tomogram. Figs 13(a) and (b) depict the WD and TWD phase-velocity curves $C(f)$ at stations 2 and 65, respectively. The red line is the observed phase velocity and the blue and black lines are the TWD and standard WD results, where the TWD curves more closely resemble the observed ones. Station 2 is at the top of the mountain, therefore the effect of topography is very strong and thus results in significant discrepancies between the observed and WD phase-velocity curves (Fig. 13a, black line). Fig. 14 shows the RMS misfit values plotted versus iteration number. As expected, after 12 iterations, the last normalized TWD residual (black line) is about 0.73, but that of WD (red line) is more than 0.8.

We also apply TWD to data generated from three checkerboard velocity models (see Figs 15a, c and e) to assess its resolution capabilities. The average background *S*-wave velocity is 800 m s^{-1} in all models. To construct the checkerboard perturbations, the *S*-wave velocity is changed by ±6 per cent and we use a 1.5 Hz Ricker as the source wavelet with a dominant *S*-wavelength of about 500 m. The first checkerboard model has seven rectangular velocity anomalies where each one has an area of 330 × 220 m^2 (Fig. 15a), while the second and third models have two and three layers with checkerboard areas of 240 × 220 m^2 and 240 × 150 m^2, respectively (Figs 15c and e). There are 60 vertical-component shot gathers computed by solving the 2-D elastic wave equation with 60 geophones with a spacing of 40 m. Here, the shots are located every 40 m along the irregular surface. The topographic variations, aperture width of the geophone array and the frequency band of the source are similar to those for the San Jacinto experiment. The reconstructed *S*-wave velocity TWD tomograms are shown in Figs 15(b), (d) and (f) for the different checkerboard models. These tomograms suggest that the TWD method should be capable of estimating the *S*-wave velocity variations with a resolution of about half the *S*-velocity wavelength to a depth of about 200 m below the surface.

The low-velocity zones narrow with depth and are interpreted as the damage zones produced by intense coseismic fracturing. The core damage zone is asymmetric relative to the surface trace of the fault, with the most intense damage occurring to the NE extending from 900 to 1300 m. The broader damage zone, which is shallower and with higher velocity, extends from 800 to 1900 m. A secondary fault structure is visible at 2250 m. Failure to account for topography can lead to the erroneous interpretation of a much wider damage zone, especially to

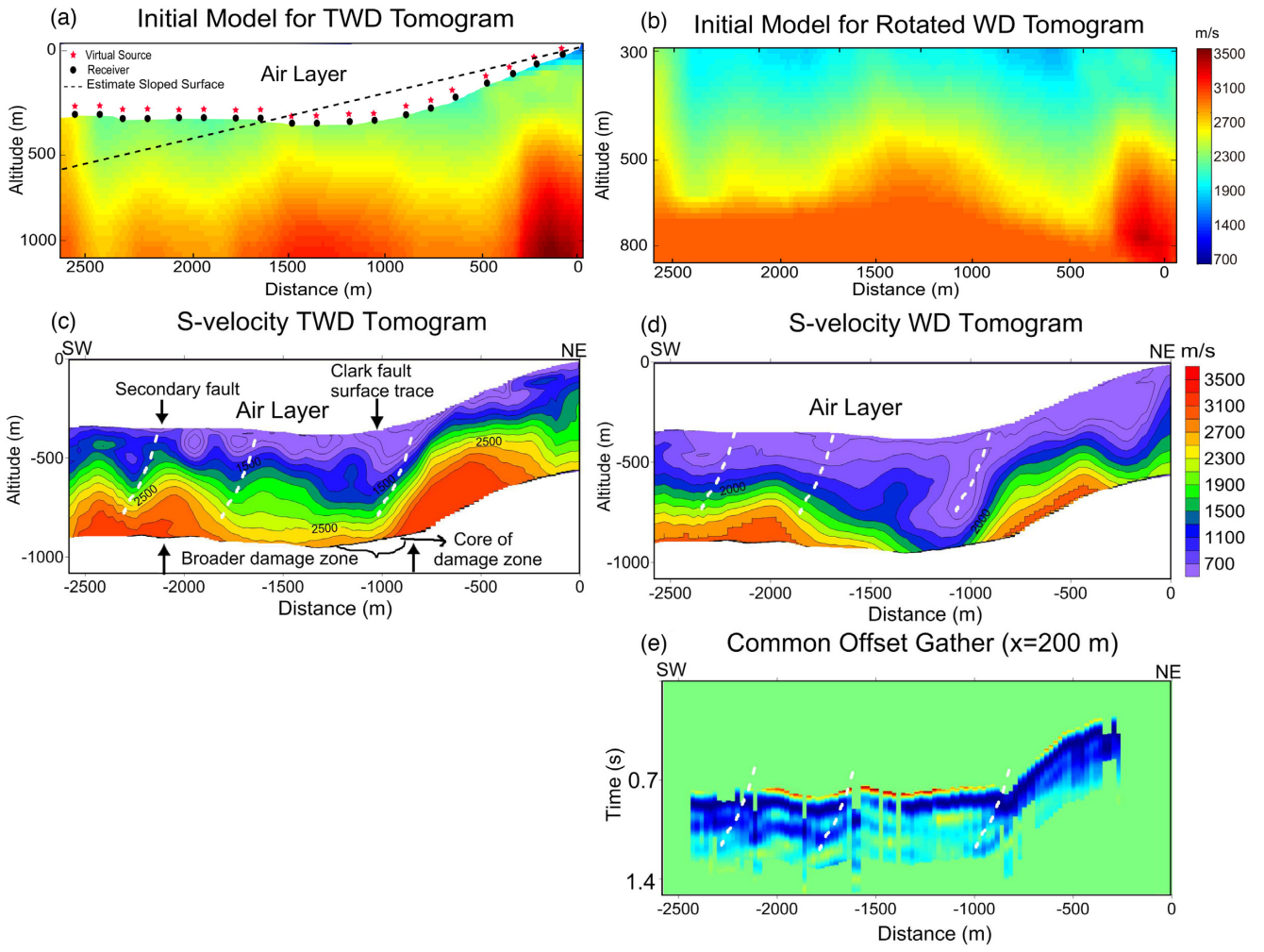

Figure 12. (a) and (b) are the initial models for the TWD and rotated WD tomograms. (c) and (d) are the TWD and WD *S*-velocity tomograms inverted from the Southern California dispersion curves, (e) recorded virtual COGs at the source–receiver offset of 200 m. The dashed white faults are the interpreted faults, where the dashed white line at the far left is the main fault seen at the surface.

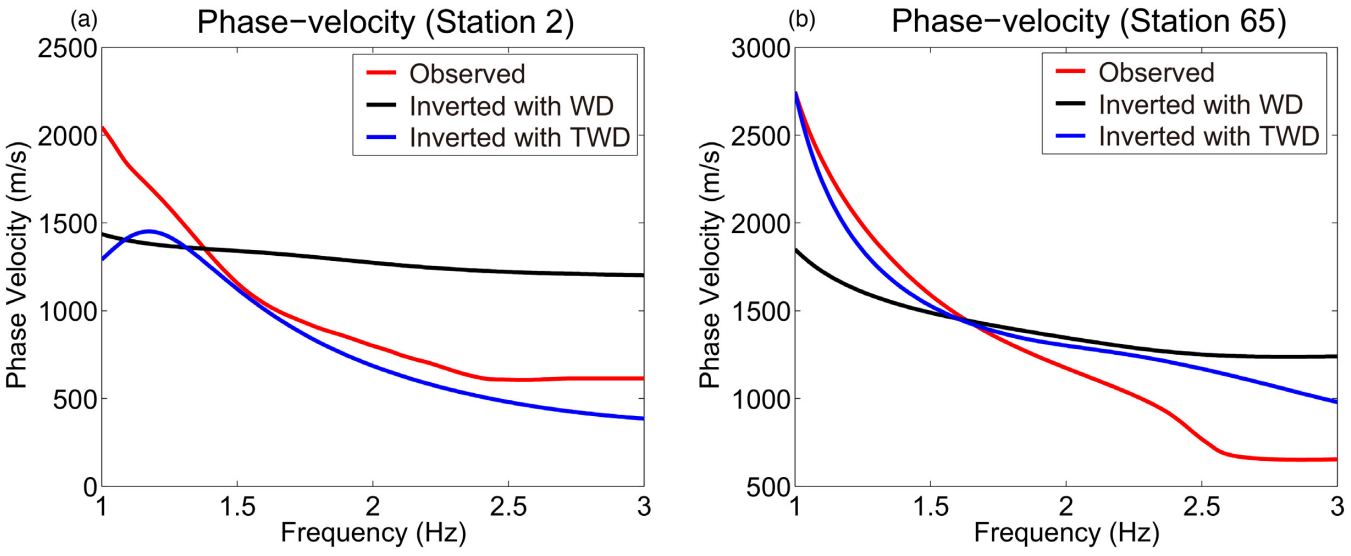

Figure 13. Dispersion curves for standard WD (black line), TWD (blue line) and observed (red line) at station 2 (a) and station 65 (b) virtual shot gather, respectively.

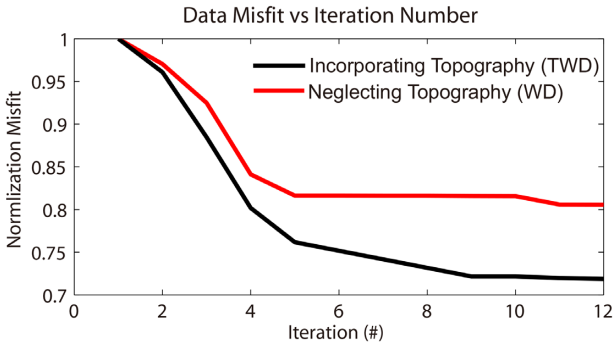

Figure 14. Normalized RMS residual versus iteration number for the two WD tomograms, where the red and black lines are for the WD and TWD residuals, respectively.

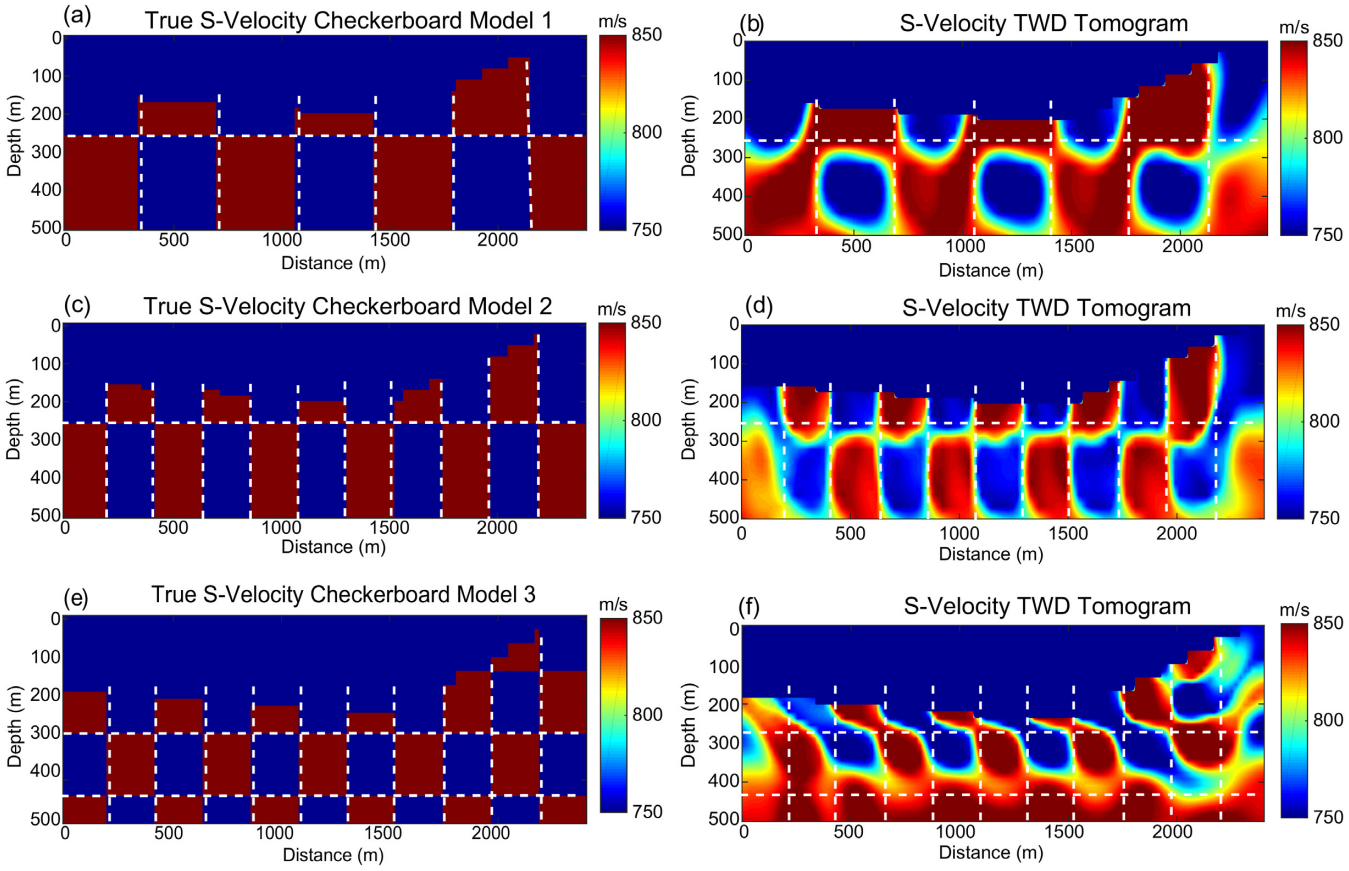

Figure 15. Checkerboard test models on the left and the inverted *S*-velocity tomograms on the right. The checkerboard areas are 330 × 220 m, 240 × 220 m, 240 × 150 m in (a), (c) and (e), respectively.

the NE where the topography is the steepest. In addition, the tomogram reveals a velocity contrast across the fault, with the NE side having a higher velocity on average. These images are in agreement with the previous observational work on the local damage structure and velocity contrast (Share *et al.* 2017).

To improve the accuracy of the TWD method we suggest two extra processing steps.

(1) The Q model should be obtained and then used in the forward and backward modelling for estimating the velocity model. At the near surface, the Q values can be very small and therefore strongly affect the character of the dispersion curves.

(2) A multiscale method should be used to estimate the final velocity model (Liu *et al.* 208; Li *et al.* 2018). The multiscale method should start with the highest significant frequency in the dispersion curves to invert for the shallowest velocities. Shorter offset traces are initially inverted and then longer offset traces are inverted. Once the shallow velocity model is reconstructed then the deeper ones can be reconstructed by inverting the dispersion curves at the lower frequencies. Details are given in Appendix C.

4 DISCUSSION AND CONCLUSIONS

We present the TWD method that accounts for an irregular recording surface. The effectiveness of this method is numerically demonstrated with synthetic data and field traces recorded on an irregular free surface. Results with synthetic data suggest that there will be significant errors in the tomogram if topographic phase delays larger than a quarter of a cycle are not accounted for in the inversion. The TWD method is also applied to field data recorded over mountainous terrain in Southern California. The *S*-velocity TWD tomogram shows more fault details than the WD tomogram, and some fault-like structures appear to be more consistent with those seen in the COG image.

In addition, we propose that surface wave analysis should strive to replace the 1-D inversion method with 2-D or 3-D dispersion inversion methods. In addition all tomograms should be cross-checked for accuracy by direct comparison of the predicted COGs with the recorded COGs at an appropriate source–receiver offset. As shown in our results, this provides an accuracy estimate of the tomogram as well as an indicator of the location of sharp jumps in velocity, which might be faults.

The limitations of TWD are that it is an order-of-magnitude more expensive compared to standard 1-D inversion of surface waves, and TWD only provides a moderate resolution of the *S*-velocity model compared to the theoretically higher resolution from FWI. FWI is much less robust than TWD because TWD only inverts the skeletonized dispersion curves and does not have to account for the amplitudes of every wiggle in the trace. Better resolution with TWD can, in principle, be obtained by picking and inverting higher-order modes. The TWD tomogram can also be used as an excellent starting model for FWI.

ACKNOWLEDGEMENTS

The research reported in this publication was supported by the King Abdullah University of Science and Technology (KAUST), Saudi Arabia. We are grateful to the sponsors of the Center for Subsurface Imaging and Modeling (CSIM) Consortium for their financial support. We also thank KAUST for providing funding by the CRG grant OCRF-14 2014-CRG3-2300. Finally, we would like to express thanks to Mrinal Sinha, anonymous journal referees and Editor Dr. Wolfgang Friederich for valuable comments.

REFERENCES

Allam, A. & Ben-Zion, Y., 2012. Seismic velocity structures in the Southern California plate-boundary environment from double-difference tomography, *Geophys. J. Int.,* **190,** 1181–1196.

Ben-Zion, Y., Vernon, F.L., Ozakin, Y., Zigone, D., Ross, Z.E., Meng, H. & Barklage, M., 2015. Basic data features and results from a spatially dense seismic array on the San Jacinto fault zone, *Geophys. J. Int.,* **202**(1), 370–380.

Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A., Lin, F., Moschetti, M., Shapiro, N. & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.,* **169,** 1239–1260.

Borisov, D., Modrak, R., Gao, F.C. & Tromp, J., 2017. 3D elastic full-waveform inversion of surface waves in the presence of irregular topography using an envelope-based misfit function, *Geophysics,* **81,** R1–R11.

Borisov, D., Modrak, R., Rusmanugroho, H., Yuan, Y., Simons, F., Tromp, J. & Gao, F., 2016. Spectral-element based 3D elastic full-waveform inversion of surface waves in the presence of complex topography using an envelope-based misfit function, in *SEG Technical Program Expanded Abstracts 2016,* Society of Exploration Geophysicists, Tulsa, pp. 1211–1215.

Davies, A. & Heathershaw, A., 1984. Surface-wave propagation over sinusoidally varying topography, *J. Fluid Mech.,* **144,** 419–443.

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

Kohn, D., Meier, T., Fehr, M., De Nil, D. & Auras, M., 2016. Application of 2D elastic Rayleigh waveform inversion to ultrasonic laboratory and field data, *Near Surf. Geophys.,* **14,** 461–476.

Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophys. J. Int.,* **139,** 806–822.

Komatitsch, D. *et al.*, 2012. *SPECFEM2D v7.0.0, Computational Infrastructure for Geodynamics*. Available at: https://geodynamics.org/cig/software/specfem2d/.

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

Li, J. & Hanafy, S., 2016. Skeletonized inversion of surface wave: Active source versus controlled noise comparison, *Interpretation,* **4**(3), SH11–SH19.

Li, J., Hanafy, S. & Schuster, G.T., 2018. Wave-equation dispersion inversion of guided P waves in a waveguide of arbitrary geometry, *Journal of Geophysical Research: Solid Earth,* **123**(9), 7760–7774.

Li, J. & Schuster, G., 2016. Skeletonized wave equation of surface wave dispersion inversion, in *SEG Technical Program Expanded Abstracts 2016,* Society of Exploration Geophysicists, Tulsa, pp. 3630–3635.

Li, J., Schuster, G.T., Lin, F.C. & Alam, A., 2017b. Wave-equation dispersion inversion of surface waves recorded on irregular topography, in *SEG Technical Program Expanded Abstracts 2016,* Society of Exploration Geophysicists, Tulsa, pp. 2621–2626.

Lin, F.-C., Li, D., Clayton, R.W. & Hollis, D., 2013. High-resolution 3D shallow crustal structure in Long Beach, California: Application of ambient noise tomography on a dense seismic array, *Geophysics,* **78**(4), Q45–Q56.

Liu, Z.L. & Huang, J., 2018. Multiscale and layer-stripping wave-equation dispersion inversion of Rayleigh waves, in *SEG Technical Program Expanded Abstracts 2018,* Society of Exploration Geophysicists, Tulsa, pp. 2536–2540.

Liu, Z.L., LI, J. & Schuster, G., 2018. 3D wave-equation dispersion inversion of surface waves, in *SEG Technical Program Expanded Abstracts 2018,* Society of Exploration Geophysicists, Tulsa, pp. 4733–4737.

Luo, Y., Xia, J., Miller, R.D., Xu, Y., Liu, J. & Liu, Q., 2009. Rayleigh-wave mode separation by high-resolution linear Radon transform, *Geophys. J. Int.,* **179,** 254–264.

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

Masoni, I., Brossier, R., Boelle, J. & Virieux, J., 2014. Generic gradient expression for robust FWI of surface waves, in *76th EAGE Conference and Exhibition 2014,* EAGE, The Netherlands, pp. 1–5.

Masoni, I., Brossier, R., Boelle, J.L. & Virieux, J., 2016. Layer stripping FWI for surface waves, in *SEG Technical Program Expanded Abstracts 2016,* Society of Exploration Geophysicists, Tulsa, pp. 369–1373.

Mi, B.B., Xia, J.H., Shen, C., Wang, L.M., Hu, Y. & Cheng, F., 2017. Horizontal resolution of multichannel analysis of surface waves, *Geophysics,* **82,** EN51–EN66.

Nocedal, J. & Wright, S., 1999. *Numerical Optimization, Volume 35, Springer Series in Operations Research,* Springer Company.

Nuber, A., Manukyan, E. & Maurer, H., 2016. Ground topography effects on near-surface elastic full waveform inversion, *Geophys. J. Int.,* **207,** 67–71.

Pageot, D., Le Feuvre, M., Leparoux, D., Cte, P. & Capdeville, Y., 2016. Importance of a 3D forward modeling tool for surface wave analysis methods, in *EGU General Assembly 2016,* Vienna, pp. 2016–11812-1.

Pan, Y.D., Gao, L. & Bohlen, T., 2018. Time-domain full-waveform inversion of Rayleigh and Love waves in presence of free-surface topography, *J. appl. Geophys.,* **152,** 77–85.

Rempe, M., Mitchell, T., Renner, J., Nippress, S., Ben-Zion, Y. & Rockwell, T., 2013. Damage and seismic velocity structure of pulverized rocks near the San Andreas Fault, influences by the topographic free surface, *J. geophys. Res.,* **118**(6), 2813–2831.

Robertsson, J.O., 1996. A numerical free-surface condition for elastic/viscoelastic finite-difference modeling in the presence of topography, *Geophysics,* **61,** 1921–1934.

Roy, N. & Jakkab, R., 2017. Near-field effects on site characterization using MASW technique, *Soil Dyn. Earthq. Eng.,* **97,** 289–303.

Share, P.-E., Ben-Zion, Y., Ross, Z.E., Qiu, H. & Vernon, F., 2017. Internal structure of the San Jacinto fault zone at Blackburn Saddle from seismic data of a dense linear array, *Geophys. J. Int.,* **210**(2), 819–832.

Sharp, R.V., 1967. San Jacinto fault zone in the Peninsular ranges of southern California, *Bull. geol. Soc. Am.,* **78,** 705–730.

Snieder, R., 1986. The influence of topography on the propagation and scattering of surface waves, *Phys. Earth planet. Inter.,* **44,** 226–241.

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

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

Spetzler, J., Trampert, J. & Snieder, R., 2002. The effect of scattering in surface wave tomography, *Geophys. J. Int.,* **149,** 755–767.

Wang, L., Luo, Y. & Xu, Y., 2012. Numerical investigation of Rayleigh-wave propagation on topography surface, *J. appl. Geophys.,* **86,** 88–97.

Wang, L., Xu, Y., Xia, J. & Luo, Y., 2015. Effect of near-surface topography on high frequency Rayleigh-wave propagation, *J. appl. Geophys.,* **116,** 93–103.

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

Xia, J., Xu, Y., Chen, C., Kaufmann, R.D. & Luo, Y., 2006. Simple equations guide high-frequency surface-wave investigation techniques, *Soil Dyn. Earthq. Eng.,* **26**(5), 395–403.

Xu, Y., Xia, J. & Miller, R.D., 2006. Quantitative estimation of minimum offset for multichannel surface-wave survey with actively exciting source, *J. appl. Geophys.,* **59**(2), 117–125.

Yuan, Y.O., Simons, F.J. & Bozda, E., 2015. Multiscale adjoint waveform tomography for surface and body waves, *Geophysics,* **80,** R281–R302.

Zigone, D., Ben-Zion, Y., Campillo, M. & Roux, P., 2015. Seismic tomography of the southern California plate boundary region from noise-based Rayleigh and Love waves, *Pure appl. Geophys.,* **172,** 1007–1032.

APPENDIX A: CORRELATION IDENTITY

In eq. (4), the functions in the integrand can be replaced by their Fourier transforms

$$\dot{\tilde{D}}(k+\Delta\kappa,\omega)_{\rm obs} = \frac{1}{2\pi}\int [iy' D(y',\omega)_{\rm obs} e^{i(k+\Delta\kappa)y'}]\mathrm{d}y',$$
$$\tilde{D}(k,\omega) = \frac{1}{2\pi}\int D(y,\omega)e^{iky}\mathrm{d}y \tag{A1}$$

to give

$$\begin{aligned}\frac{\partial\dot{\Phi}(\Delta\kappa,s(\mathbf{x}))}{\partial s(\mathbf{x})} &= \frac{1}{4\pi^2}\left\{\int \mathrm{d}x_g\left\{\int \mathrm{d}x_g'\left[\int e^{ik(x_g-x_g')}\mathrm{d}k\right](-ix_g')D(x_g',\omega)^*_{\rm obs}e^{-i\Delta\kappa x_g'}\right\}\frac{\partial D(x_g,\omega)}{\partial s(\mathbf{x})}\right\},\\ &= \frac{1}{4\pi^2}\left\{\int \mathrm{d}x_g\left\{\int \mathrm{d}x_g'[2\pi\delta(x_g-x_g')](-ix_g')D(x_g',\omega)^*_{\rm obs}e^{-i\Delta\kappa x_g'}\right\}\frac{\partial D(x_g,\omega)}{\partial s(\mathbf{x})}\right\},\\ &= \frac{1}{2\pi}\left\{\int \mathrm{d}x_g[-ix_g D(x_g,\omega)^*_{\rm obs}e^{-i\Delta\kappa x_g}]\frac{\partial D(x_g,\omega)}{\partial s(\mathbf{x})}\right\}, = \left\{\int \mathrm{d}x_g\overbrace{[\frac{-ix_g e^{-i\Delta\kappa x_g}}{2\pi}D(x_g,\omega)^*_{\rm obs}]}^{\hat{D}(x_g,\omega)^*_{\rm obs}}\frac{\partial D(x_g,\omega)}{\partial s(\mathbf{x})}\right\},\end{aligned} \tag{A2}$$

where the weighted conjugated data function is

$$\hat{D}(x_g,\omega)^*_{\rm obs} = \frac{-ix_g e^{-i\Delta\kappa x_g}}{2\pi}D(x_g,\omega)^*_{\rm obs}. \tag{A3}$$

APPENDIX B: LINEAR RADON TRANSFORM OF DATA RECORDED ON AN IRREGULAR SURFACE

Assume a vertical-component point source and traces recorded at the coordinates (x_i, z_i), where the geophones are evenly spaced by the distance Δx along the free surface. The vertical component of the recorded data is denoted as $\mathrm{d}(x_g, t)$, where

$$x_g = \sum_{j=2}^{g}\sqrt{(x_j-x_{j-1})^2+(z_j-z_{j-1})^2}. \tag{B1}$$

Here, g is the integer index of the gth geophone, t is the listening time and x_g is the offset distance along the free surface of the gth geophone measured from the endline geophone at (x_1, z_1). Here, we assume interpolation is used so that the geophones on the surface are separated by the same distance Δx.

The discrete Radon transform of the shot gather $\mathrm{d}(x_g, t)$ is

$$p_{\min} \le p \le p_{\max}; \qquad m(p,\tau) = \sum_{x_g=x_{\min}}^{x_{\max}} d(x_g, t = \tau + p x_g), \tag{B2}$$

where $\Delta x = 1$, p is the slowness parameter, τ is the intercept time of the line summation at $p = 0$, $p_{\min}$ and $p_{\max}$ represent the range of discrete slowness values $p = i\Delta p$ for integer values of i, and $x_{\min}$ and $x_{\max}$ denote the near- and far-offset values between the source and receivers. Applying a Fourier transform to eq. (B2) gives

$$\begin{aligned}\tilde{m}(p,\omega) &= \sum_{x_g=x_{\min}}^{x_{\max}} \int_{-\infty}^{\infty} d(x_g, \tau + p x_g) e^{i\omega\tau}\,\mathrm{d}\omega,\\ &= \sum_{x_g=x_{\min}}^{x_{\max}} \left[\int_{-\infty}^{\infty} \mathrm{d}(x_g,\tau) e^{i\omega\tau}\,\mathrm{d}\omega\right] e^{-i\omega p x_g},\\ &= \sum_{x_g=x_{\min}}^{x_{\max}} \tilde{d}(x_g,\omega) e^{-i\omega p x_g},\end{aligned} \tag{B3}$$

where $\tilde{d}(x_g, \omega)$ is the Fourier spectrum of the shot gather $\mathrm{d}(x_g, t)$, and $\tilde{m}(p, \omega)$ is the Fourier spectrum of the Radon-transformed data $m(p, \tau)$. The fundamental-mode dispersion curve $C(\omega)$, also known as the fundamental phase-velocity curve, is picked from the magnitude spectrum of $\tilde{m}(p, \omega)$.

The workflow for estimating the fundamental phase-velocity curve $C(\omega)$ from each shot gather in the $p - \omega$ domain consists of three steps. First, apply an FFT to the interpolated shot gather $\mathrm{d}(x_g, t)$ to get $\tilde{d}(x_g, \omega)$. Then, $\tilde{d}(x_g, \omega)$ is weighted by $e^{-i\omega p x_g}$ and summed over a range of offset values to get $\tilde{m}(p, \omega)$ in eq. (B3). The fundamental dispersion curve $C(\omega)$ is identified and picked.

APPENDIX C: MULTI-OFFSET WAVE EQUATION DISPERSION CURVES

The source–receiver offsets have a crucial effect on the horizontal resolution and accuracy of dispersion curves in the multichannel analysis of surface waves (MASW) surveys. According to empirical tests, longer receiver spreads typically lead to a more accurate estimate of the dispersion curve but can lead to a lower horizontal resolution in the velocity model (Xia *et al.* 2006; Xu *et al.* 2006; Mi *et al.* 2017), while shorter receiver spreads give more accurate dispersion curves but with errors in the low-frequency range (Roy & Jakkab 2017). These characteristics are consistent with the behaviour of the misfit gradient associated with a three-layer velocity model in Figs C1(a) to (c). The long-offset data provide a deeper but lower-resolution update of the velocity model compared with the short-offset data. In order to update the deeper part of the velocity model without sacrificing its resolution, we devise a multi-offset strategy to compute a multiscale WD tomogram. We first invert the traces with higher frequencies and shorter source–receiver offsets to update for the shallow velocities in the model. Then we iteratively increase the offsets and lower the frequencies to invert for the deeper areas of the model. A trial and error procedure can be used to determine whether a layer-stripping strategy (Masoni *et al.* 2016; Liu & Huang 2018) is employed where the shallow layers are first inverted from the short-offset data. These inverted velocities are fixed and then the deeper layer velocities are inverted from the long-offset data.

We use the synthetic model shown in Fig. C2(a) to test and evaluate the feasibility of the multi-offset strategy. The *S*-wave and *P*-wave velocity models are related by a constant Poisson's ratio ($Vp/Vs = 2.3$), with a homogeneous density taken to be a constant value of 1.5 g cm^{-3}. We aim to invert for the high-velocity anomalies located in the shallow and deeper parts of the model. The observed data are generated using finite-difference solutions to the 2-D elastic wave equation. The synthetic data are recorded by 60 vertical-component receivers on the surface. Thirty vertical-component sources on the surface are excited with a 30 Hz Ricker wavelet. The initial *S*-wave velocity model consists of a linear gradient. We select data with maximum source–receiver offsets of 50 and 20 m to compute the dispersion curves. The inverted tomograms for 50 and 20 m offset data are shown in Figs C2(b) and (c), respectively. The results demonstrate that the long-offset (50 m) data (Fig. C2b) updates the deeper areas of the *S*-velocity model, whereas the short-offset data (20 m; Fig. C2c) are mostly sensitive to the shallow part.

According to the multi-offset workflow, the model is updated in a top-to-bottom fashion, using the short-to-long offset strategy. Fig. C2(d) shows the inverted multi-offset *S*-velocity tomogram, where the location of the velocity anomalies and layer interface mostly agree with that of the actual model shown in Fig. C2(a).

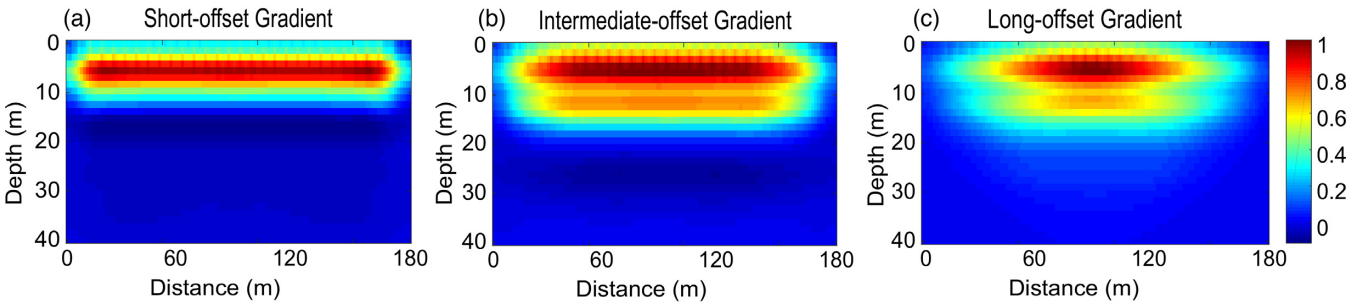

Figure C1. Velocity gradients for different classes of data: (a) Short-offset data, (b) intermediate-offset data and (c) long-offset data [same colour scale used for (a) and (b)]. Surface wave dispersion curves with higher frequency and shorter wavelength have a shallow penetration depth and thus are sensitive to shallow parts of the medium, whereas surface waves with lower frequencies and longer wavelengths can penetrate to much deeper areas.

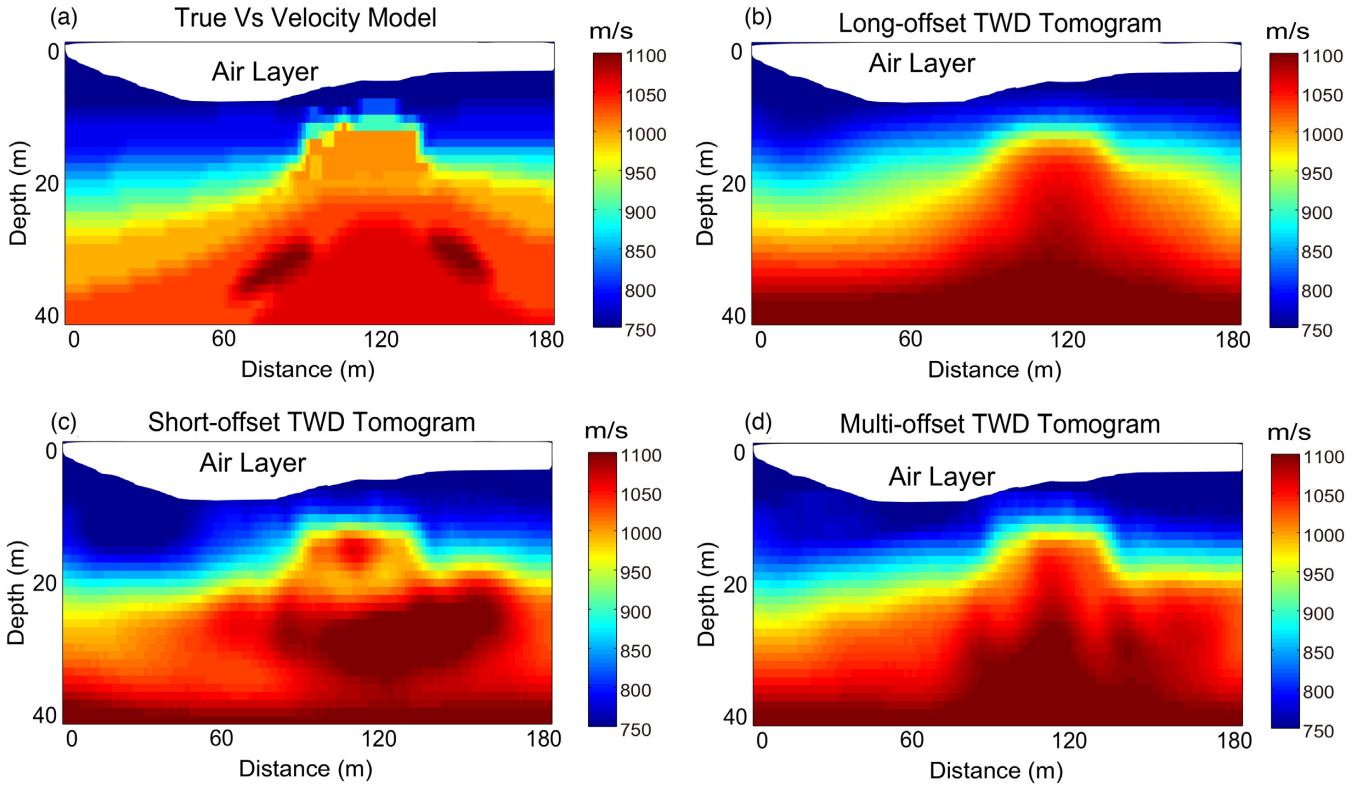

Figure C2. Inverted TWD tomogram for different offset classes. (a) True *Vs* velocity model. TWD tomograms inverted from the (b) long-offset data, (c) short-offset data and (d) multi-offset data.