3D Wave-equation Dispersion Inversion of Rayleigh Waves

Zhaolun Liu*, Jing Li*, Sherif M. Hanafy* and Gerard Schuster*

* *Department of Earth Science and Engineering,*

King Abdullah University of Science and Technology,

Thuwal, Saudi Arabia, 23955-6900.

Email: zhaolun.liu@kaust.edu.sa; jing.li@kaust.edu.sa; sherif.mahmoud@kaust.edu.sa;

gerard.schuster@kaust.edu.sa.

(August 14, 2018)

Running head: **3D Wave-equation Dispersion Inversion**

ABSTRACT

We extend the 2D wave-equation dispersion inversion (WD) method to 3D wave-equation inversion of surface waves for the shear-velocity distribution. The objective function of 3D WD is the frequency summation of the squared wavenumber $\kappa(\omega)$ differences along each azimuth angle of the fundamental or higher modes of Rayleigh waves in each shot gather. The S-wave velocity model is updated by the weighted zero-lag cross-correlation between the weighted source-side wavefield and the back-projected receiver-side wavefield for each azimuth angle. The synthetic and field data examples demonstrate that 3D WD can accurately reconstruct the 3D S-wave velocity model of a laterally heterogeneous medium and has much less of a tendency to getting stuck in a local minimum compared to full waveform inversion. It has nearly the same resolution as FWI and does not require the modeling of true amplitudes in the arrivals.

INTRODUCTION

Surface-wave methods are powerful tools for characterizing the near-surface S-velocity distribution. An advantage of surface-wave imaging over body-wave imaging is that the seismic energy of surface waves spreads out as $1/r$ from the source, compared to the $1/r^2$ geometrical spreading of body waves (Aki and Richards, 2002). Here, r is the distance along the horizontal propagation path between the source and receiver on the free surface. Thus, the recorded data are usually dominated by surface waves for a wide range of source-receiver offsets within the time window of surface-wave arrivals. A practical use of surface waves is that they can be inverted to detect shallow drilling hazards down to the depth on the order of the dominant shear wavelength (Ivanov et al., 2013).

The conventional dispersion-inversion method calculates the S-wave velocity model directly from the surface-wave dispersion curves (Haskell, 1953; Xia et al., 1999, 2002; Park et al., 1999) by assuming a 1D velocity beneath the recording data. Unfortunately, this assumption is violated when there are strong lateral gradients in the S-velocity model, such as faults, vugs or gas channels. To partially mitigate this problem, spatial interpolation of 1D velocity models (Pan et al., 2016) and laterally constrained inversion (Socco et al., 2009; Bergamo et al., 2012) can be used to compute an approximation to the 2D S-velocity model.

As an alternative, full-waveform inversion (FWI) (Groos et al., 2014; Pérez Solano et al., 2014; Dou and Ajo-Franklin, 2014; Groos et al., 2017) estimates the S-velocity model that accurately predicts the surface waves recorded in a heterogeneous S-velocity model. But in practice, FWI easily gets stuck in a local minimum due to the strongly dispersive nature of surface waves and an inadequate initial velocity model. To mitigate this problem,

Pérez Solano et al. (2014) changed the misfit function of FWI into the l^2 misfit of magnitude spectra of surface waves, and their synthetic data results showed this to be an effective method for reconstructing the S-wave velocity model at the near surface. Until now there are few studies to assess the full benefits and limitations of this method so its effectiveness on a wider variety of data sets is still to be determined.

To combine the inversion of both surface waves with body waves, Yuan et al. (2015) developed a wavelet multi-scale adjoint method for the joint inversion of both surface and body waves. The efficacy of this method is validated with synthetic data. However, further studies are needed to assess its robustness in convergence. To enhance robustness, layer stripping FWI of surface waves was presented by Masoni et al. (2016) who first invert the high-frequency and near-offset data for the shallow S-velocity model, and gradually incorporates lower-frequency data with longer offsets to estimate the deeper parts of the model. This procedure partly mitigates the local minima problem. All of these methods, however, are still under development and require more tests to fully understand their relative benefits and limitations.

To avoid the assumption of a layered medium and also mitigate FWI's sensitivity to the local minima, Li and Schuster (2016) and Li et al. (2017c) proposed a new surface-wave dispersion inversion method, which is denoted as wave-equation dispersion inversion (WD). The WD method skeletonizes the complicated surface-wave arrivals as simpler data, namely the picked dispersion curves in the wavenumber-angular frequency ($k-\omega$) domain. These curves are obtained by applying a combination of temporal Fourier and spatial Radon transforms to the Rayleigh waves recorded by vertical-component geophones. The sum of the squared differences between the wavenumbers along the predicted and observed dispersion curves is used as the objective function, and solutions to the elastic wave equation and an

iterative optimization method are then used to invert these curves for the S-wave velocity models. Numerical tests on the 2D synthetic and field data show that WD can accurately reconstruct the S-wave velocity distributions in laterally heterogeneous media. The WD method also enjoys robust convergence because the skeletonized data, namely the dispersion curves, are simpler than traces with many dispersive arrivals. The penalty, however, is that the inverted S-velocity model has lower resolution than a model that accurately fits both the waveform and phase information. Recently, Fu et al. (2018a) showed that only inverting the phase information and ignoring the amplitudes gives almost the same resolution as obtained by full waveform inversion.

In this paper, we extend the 2D WD method to invert dispersion curves for the 3D S-wave velocity model that accounts for strong velocity variations in all three dimensions. After the introduction, we describe the theory of 3D WD and its implementation. We also discuss the procedure for estimating a good initial model for 3D WD: first use the 1D dispersion-curve inversion method and then use the 2D WD method. Numerical tests on synthetic and field data are presented in the third section to validate the theory. The limitations of the proposed method are discussed in the fourth section and the summary is given in the last section.

THEORY

Let $d(\boldsymbol{g}, t)$ denote a shot gather of vertical particle-velocity traces recorded by the receiver on the surface at $\boldsymbol{g} = (x_g, y_g, 0)$. The surface waves are excited by a vertical-component force on the surface at $\boldsymbol{s} = (x_s, y_s, 0)$, where the horizontal recording plane is at $z = 0$. We will assume that the effects of attenuation on the dispersion curves are insignificant. But, if important, such effects can be accounted for by using solutions to viscoelastic wave equation

(Li et al., 2017a,b). Assume the data have been filtered so that $d(\boldsymbol{g},t)$ only contains the fundamental mode of Rayleigh waves. A 3D Fourier transform is then used to transform $d(\boldsymbol{g},t)$ into $\tilde{D}(\boldsymbol{k},\omega)$ in the $\boldsymbol{k}-\omega$ domain:

$$\begin{aligned}\tilde{D}(\boldsymbol{k},\omega) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d(\boldsymbol{g},t)e^{-i(\boldsymbol{k}\cdot\boldsymbol{g}+\omega t)}d\boldsymbol{g}dt \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} D(\boldsymbol{g},\omega)e^{-i\boldsymbol{k}\cdot\boldsymbol{g}}d\boldsymbol{g}\end{aligned} \tag{1}$$

where $d\boldsymbol{g} = dx_g dy_g$ and $D(\boldsymbol{g},\omega)$ represents the data in the space-frequency $(\boldsymbol{x}-\omega)$ domain. Here, the $z=0$ notation is silent. The wavenumber vector $\boldsymbol{k}=(k_x,k_y)$ can be represented in polar coordinate as (k,θ), where $\theta = \arctan\frac{k_y}{k_x}$ is the azimuth angle and $k=\sqrt{k_x^2+k_y^2}$ is the radius. Following this notation, the Fourier transformed data $\tilde{D}(\boldsymbol{k},\omega)$ are denoted as $\tilde{D}(k,\theta,\omega)$. We skeletonize the spectrum $\tilde{D}(k,\theta,\omega)$ as the dispersion curves associated with the fundamental mode of the Rayleigh waves, which are the wavenumbers $\kappa(\theta,\omega)$ obtained by picking the (k,θ,ω) coordinates of the fundamental dispersion curve.* This curve is recognized as the maximum magnitude spectrum $\tilde{D}(k,\theta,\omega)$ along the azimuth angle θ and is denoted as $\kappa(\theta,\omega)_{obs}$ for the observed data. In this paper, we assume that the dispersion curves are those for Rayleigh waves recorded by vertical-component geophones, but this approach is also valid for Love waves or guided waves at the near surface (Li et al., 2018b).

Misfit Function

The 3D WD method inverts for the S-wave velocity model that minimizes the objective function J of the dispersion curve:

$$J = \frac{1}{2}\sum_{\omega}\sum_{\theta}\overbrace{[\kappa(\theta,\omega)_{pre}-\kappa(\theta,\omega)_{obs}]}^{residual=\Delta\kappa(\theta,\omega)}{}^2 + \text{penalty term}, \tag{2}$$

*Higher-order modes can also be picked and inverted.

where the penalty term can be any model-based function that penalizes solutions far from an apriori model. Here, $\kappa(\omega,\theta)_{pre}$ represents the predicted dispersion curve picked from the simulated spectrum along the azimuth angle θ and $\kappa(\omega,\theta)_{obs}$ describes the observed dispersion curve obtained from the recorded spectrum along the azimuth θ. For pedagogical clarity, we will ignore the penalty term in further manipulations of the objective function.

The gradient $\gamma(\boldsymbol{x})$ of J with respect to the S-wave velocity $v_s(\boldsymbol{x})$ is given by

$$\gamma(\boldsymbol{x}) = \frac{\partial J}{\partial v_s(\boldsymbol{x})} = \sum_{\omega}\sum_{\theta} \Delta\kappa(\theta,\omega)\frac{\partial \kappa(\theta,\omega)_{pre}}{\partial v_s(\boldsymbol{x})}, \tag{3}$$

so that the optimal S-wave velocity model $v_s(\boldsymbol{x})$ is obtained from the steepest-descent formula (Nocedal and Wright, 2006)

$$v_s(\boldsymbol{x})^{(k+1)} = v_s(\boldsymbol{x})^{(k)} - \alpha \sum_{\theta}\sum_{\omega} \Delta\kappa(\theta,\omega)\frac{\partial \kappa(\theta,\omega)_{pre}}{\partial v_s(\boldsymbol{x})}, \tag{4}$$

where α is the step length and the superscript (k) denotes the k^{th} iteration. In practice, a preconditioned conjugate gradient method is preferred for faster convergence.

Connective Function

The Fréchet derivative $\frac{\partial \kappa(\theta,\omega)_{pre}}{\partial v_s(\boldsymbol{x})}$ in equation 3 is derived by forming a connective function that relates the dispersion curve $\kappa(\theta,\omega)_{pre}$ to the S-wave velocity model $v_s(\boldsymbol{x})$ (Luo and Schuster, 1991a,b; Li et al., 2017d; Lu et al., 2017; Schuster, 2017). This connective function $\Phi(\kappa, v_s(\boldsymbol{x}))$ is defined as the cross-correlation between the predicted $\tilde{D}(k,\theta,\omega)$ and observed $\tilde{D}(k,\theta,\omega)_{obs}$ spectra along a specified azimuth θ at frequency ω in the $\mathbf{k}-\omega$ domain:

$$\Phi(\kappa, v_s(\boldsymbol{x})) = \Re\Big\{ \int \tilde{D}(k+\kappa,\theta,\omega)^*_{obs}\tilde{D}(k,\theta,\omega)dk \Big\}, \tag{5}$$

where $\Re$ denotes the real part and the superscript $*$ stands for the complex conjugation. Here, κ is an arbitrary wavenumber shift between the predicted and observed spectra at

frequency ω and along the azimuth angle θ. We seek the value of κ that shifts the predicted spectrum $\tilde{D}(k,\theta,\omega)$ so that it "best" matches the observed spectrum $\tilde{D}(k,\theta,\omega)_{obs}$. The criterion for "best" match is defined as the wavenumber residual $\Delta\kappa$ that maximizes the cross-correlation function $\Phi(\kappa, v_s(\boldsymbol{x}))$ in equation 5 and the predicted data $\tilde{D}$ is an implicit function of the shear-velocity model $v_s(\boldsymbol{x})$. In this case, the derivative of the cross-correlation function Φ with respect to the wavenumber shift κ should be zero at $\Delta\kappa$:

$$\dot{\Phi}(\kappa, v_s(\boldsymbol{x}))|_{\kappa=\Delta\kappa} = \Re\Big\{ \int \dot{\tilde{D}}(k+\Delta\kappa,\theta,\omega)^*_{obs} \tilde{D}(k,\theta,\omega) dk \Big\} = 0, \tag{6}$$

where $\dot{\tilde{D}}(k+\Delta\kappa,\theta,\omega)_{obs} = \frac{\partial \tilde{D}(k+\kappa,\theta,\omega)_{obs}}{\partial\kappa}|_{\kappa=\Delta\kappa}$. Equation 6 connects the S-wave velocity model with the dispersion curve which will be used to derive the Fréchet derivative $\frac{\partial\kappa}{\partial v_s(\boldsymbol{x})}$.

Fréchet Derivative

For equation 6, the implicit function theorem (Luo and Schuster, 1991a,b; Li et al., 2017d; Lu et al., 2017; Schuster, 2017) implies that $\Delta\kappa$ is an implicit function of $v_s(\boldsymbol{x})$ so that

$$d\dot{\Phi} = \frac{\partial\dot{\Phi}}{\partial v_s(\boldsymbol{x})} dv_s(\boldsymbol{x}) + \frac{\partial\dot{\Phi}}{\partial\Delta\kappa} d\Delta\kappa = 0. \tag{7}$$

Rearranging this equation gives the Fréchet derivative

$$\frac{\partial\Delta\kappa}{\partial v_s(\boldsymbol{x})} = \frac{\partial\kappa_{pre}}{\partial v_s(\boldsymbol{x})} = -\frac{\partial\dot{\Phi}/\partial v_s(\boldsymbol{x})}{\partial\dot{\Phi}/\partial\Delta\kappa}, \tag{8}$$

where the denominator is the normalization term

$$A = \Re\Big\{ \int \ddot{\tilde{D}}(k+\Delta\kappa,\theta,\omega)^*_{obs} \tilde{D}(k,\theta,\omega) dk \Big\}, \tag{9}$$

and the numerator is

$$\frac{\partial\dot{\Phi}(\Delta\kappa, v_s(\boldsymbol{x}))}{\partial v_s(\boldsymbol{x})} = \Re\Big\{ \int \dot{\tilde{D}}(k+\Delta\kappa,\theta,\omega)^*_{obs} \frac{\partial\tilde{D}(k,\theta,\omega)}{\partial v_s(\boldsymbol{x})} dk \Big\}. \tag{10}$$

Plugging equations 9 and 10 into equation 8 gives the Fréchet derivative

$$\frac{\partial \kappa_{pre}}{\partial v_s(\boldsymbol{x})} = -\frac{\partial \dot{\Phi}/\partial v_s(\boldsymbol{x})}{\partial \dot{\Phi}/\partial \Delta\kappa} = -\frac{\Re\left\{\int \dot{\tilde{D}}(k+\Delta\kappa,\theta,\omega)^*_{obs}\frac{\partial \tilde{D}(k,\theta,\omega)}{\partial v_s(\boldsymbol{x})}dk\right\}}{A}. \tag{11}$$

[Figure 1 about here.]

The integral with respect to the wavenumber k in equation 11 can be transformed into an integral with respect to the receiver location $\boldsymbol{g}$ (see Appendix A), so that equation 11 becomes

$$\frac{\partial \kappa_{pre}}{\partial v_s(\boldsymbol{x})} = -\frac{\Re\left\{\int d\boldsymbol{g}\frac{\partial D(\boldsymbol{g},\omega)}{\partial v_s(\boldsymbol{x})}\hat{D}(\boldsymbol{g},\theta,\omega)^*_{obs}\right\}}{A}, \tag{12}$$

where $D(\boldsymbol{g},\omega)$ is the inverse Fourier transform of $\tilde{D}(k,\theta,\omega)$, and $\hat{D}(\boldsymbol{g},\theta,\omega)^*_{obs}$ is the weighted conjugated data function defined in equation A-2:

$$\hat{D}(\boldsymbol{g},\theta,\omega)^*_{obs} = 2\pi i\boldsymbol{g}\cdot\boldsymbol{n}e^{i\boldsymbol{g}\cdot\boldsymbol{n}\Delta\kappa}\int_C D(\boldsymbol{g}',\omega)^*_{obs}d\boldsymbol{g}'. \tag{13}$$

Here, $\boldsymbol{n} = (\cos\theta, \sin\theta)$ and C is the line $(\boldsymbol{g}'-\boldsymbol{g})\cdot\boldsymbol{n} = 0$. Figure 1 depicts the procedure for calculating the weighted conjugated data $\hat{D}(\boldsymbol{g},\theta,\omega)^*_{obs}$.

For equation 12, $\frac{\partial D(\boldsymbol{g},\omega)}{\partial v_s(\boldsymbol{x})}$ can be obtained according to the Born approximation of elastic waves (see Appendix B):

$$\frac{\partial D(\boldsymbol{g},\omega)}{\partial v_s(\boldsymbol{x})} = 4v_{s0}(\boldsymbol{x})\rho_0(\boldsymbol{x})\left\{G_{3k,k}(\boldsymbol{g}|\boldsymbol{x})D_{j,j}(\boldsymbol{x},\omega) - \frac{1}{2}G_{3n,k}(\boldsymbol{g}|\boldsymbol{x})\left[D_{k,n}(\boldsymbol{x},\omega)+D_{n,k}(\boldsymbol{x},\omega)\right]\right\}, \tag{14}$$

where $v_{s0}(\boldsymbol{x})$ and $\rho_0(\boldsymbol{x})$ are the reference S-velocity and density models, respectively, at location $\boldsymbol{x}$. $D_i(\boldsymbol{x},\omega)$ denotes the i^{th} component of the particle velocity recorded at $\boldsymbol{x}$ due to a vertical-component force. Einstein notation is assumed in equation 14 where $D_{i,j} = \frac{\partial D_i}{\partial x_j}$

for $i, j \in \{1, 2, 3\}$. The 3D harmonic Green's tensor $G_{3j}(\boldsymbol{g}|\boldsymbol{x})$ is the particle velocity at location $\boldsymbol{g}$ along the j^{th} direction due to a vertical-component force at $\boldsymbol{x}$ in the reference medium.

Gradient Update

Plugging equations 12 and 14 into equation 3 gives the final expression of the gradient:

$$\gamma(\boldsymbol{x}) = \frac{\partial J}{\partial v_s(\boldsymbol{x})} = -\sum_{\omega} \frac{4v_{s0}(\boldsymbol{x})\rho_0(\boldsymbol{x})}{A} \Re\Bigg\{ \overbrace{\int \sum_{\theta} \Delta\kappa(\theta,\omega)\hat{D}(\boldsymbol{g},\theta,\omega)^*_{obs} G_{3k,k}(\boldsymbol{g}|\boldsymbol{x}) d\boldsymbol{g}}^{backprojected\ data=B_{k,k}(\boldsymbol{x},\boldsymbol{s},\omega)^*} \overbrace{D_{j,j}(\boldsymbol{x},\omega)}^{source=f_{j,j}(\boldsymbol{x},\boldsymbol{s},\omega)}$$
$$\overbrace{-\frac{1}{2}\int \sum_{\theta} \Delta\kappa(\theta,\omega)\hat{D}(\boldsymbol{g},\theta,\omega)^*_{obs} G_{3n,k}(\boldsymbol{g}|\boldsymbol{x}) d\boldsymbol{g}}^{backprojected\ data=B_{n,k}(\boldsymbol{x},\boldsymbol{s},\omega)^*} \overbrace{\left[D_{k,n}(\boldsymbol{x},\omega) + D_{n,k}(\boldsymbol{x},\omega)\right]}^{source=f_{n,k}(\boldsymbol{x},\boldsymbol{s},\omega)} \Bigg\}, \tag{15}$$

where $f_{i,j}(\boldsymbol{x}, \boldsymbol{s}, \omega)$ for i and $j \in \{1, 2, 3\}$ is the downgoing source field at $\boldsymbol{x}$, and $B_{i,j}(\boldsymbol{x}, \boldsymbol{s}, \omega)$ for i and $j \in \{1, 2, 3\}$ is the backprojected scattered field at $\boldsymbol{x}$. The above equation indicates that the gradient is computed by a weighted zero-lag correlation of the source-side and receiver-side wavefield.

From equation 15, we can see that the back-propagated data (adjoint source) for azimuth angle θ in the time domain is

$$\bar{D}(\theta, \boldsymbol{g}, t) = \mathcal{F}^{-1}(\Delta\kappa(\theta,\omega)\hat{D}(\boldsymbol{g},\theta,\omega)_{obs}), \tag{16}$$

where $\mathcal{F}^{-1}$ is the inverse Fourier operator in the frequency domain. Plugging equation 15

into equation 4 gives the steepest-descent formula for updating the S-wave velocity model:

$$v_s(\boldsymbol{x})^{(k+1)} = v_s(\boldsymbol{x})^{(k)} + \alpha \sum_{\omega} \frac{4 v_s(\boldsymbol{x}) \rho(\boldsymbol{x})}{A} \Re \Big\{ B_{k,k}(\boldsymbol{x}, \boldsymbol{s}, \omega)^* f_{j,j}(\boldsymbol{x}, \boldsymbol{s}, \omega) + B_{n,k}(\boldsymbol{x}, \boldsymbol{s}, \omega)^* f_{n,k}(\boldsymbol{x}, \boldsymbol{s}, \omega) \Big\}. \quad (17)$$

WORKFLOW AND IMPLEMENTATION

The workflow for implementing the 3D WD method is summarized in the following six steps.

1. Remove the first-arrival body waves and higher-order modes of the Rayleigh waves in the shot gather (Li et al., 2017c).

2. Determine the range of the azimuth angles θ for each shot gather.

3. Apply a 3D Fourier transform to the predicted and observed common shot gather (CSG) to compute the dispersion curves $\kappa(\theta, \omega)$ and $\kappa(\theta, \omega)_{obs}$ along each azimuth angle θ. Calculate the sum of the squared dispersion residuals in equation 2.

4. Calculate the weighted conjugated data $\hat{D}(\boldsymbol{g}, \omega)^*_{obs}$ according to equation 13, which is then used for constructing the backprojected data $\bar{D}(\theta, \boldsymbol{g}, t)$ in equation 16. The source-side wavefield $f_{i,j}(\boldsymbol{x}, \boldsymbol{s}, \omega)$ in equation 15 is also computed by a finite-difference solution to the 3D elastic wave equation.

5. Calculate and sum the gradients for all the shot gathers. Source illumination is sometimes needed as a preconditioner (Plessix and Mulder, 2004).

6. Calculate the step length and update the S-wave tomogram using the steepest-descent or conjugate gradient methods. In practice we use a conjugate gradient method.

3D Dispersion Curves for 3D Data

[Figure 2 about here.]

[Figure 3 about here.]

This sub-section describes how to compute the 3D dispersion curves of a shot gather. Here, we assume a 3D seismic survey (Boiero et al., 2011), where either shots or receivers are located on a dense areal grid (Figure 2). Figures 3a and 3b depict a CSG recorded by an array of geophones and its spectrum respectively. The frequency slice of the spectrum at 50 Hz is displayed in Figure 3c, which shows that the dispersion curves are distinct only along a range of azimuth angles, denoted as the "dominant azimuth angles". For example, the dispersion curve along θ_1 is much more pronounced than that along θ_2 because there are more seismic geophones along the θ_1 direction. In practice, the dominant angles can be determined by the following two steps:

- Calculate the distances $H(\theta)$ from the source point to all the geophones along the boundary of the acquisition zone. The maximum distance is denoted as H_{max}. Figure 4a shows the lines from the source point to the geophones along the boundary.

- Define the ratio $R(\theta)$ as

$$R(\theta) = \frac{H(\theta)}{H_{max}}. \tag{18}$$

  The azimuth angle θ is the dominant angle when R exceeds a threshold value R_0. For example, Figure 4b shows the dominant azimuth angles for the source in Figure 4a by setting $R_0 = 0.6$. The threshold value R_0 is selected by trial and error.

[Figure 4 about here.]

After determining the dominant azimuth angles, we can pick the maximum values along these angles from the dispersion curves. Figure 3c shows the picked dispersion curves (red circles) for a frequency slice, and the whole dispersion surface is depicted in Figure 3d. An efficient machine learning algorithm can be used to pick the dispersion curves (Li et al., 2018a).

Initial Model for 3D WD

Figure 5 shows the workflow for calculating the initial model used with 3D WD. First, we extract 2D in-line profiles from the 3D data set and retrieve their Rayleigh wave dispersion curves. Then, a pseudo 1D S-velocity model is obtained from the dispersion curves and the result is the initial model for the next iteration of inverting the dispersion curves. The depth z and velocity value v_s of the pseudo 1D S-velocity model are calculated by scaling the wavelength λ and phase velocity c with factors of 0.5 and 1.1, respectively (O'Neill and Matsuoka, 2005),

$$z = 0.5\lambda, \quad v_s = 1.1c.$$

We use SURF96, a dispersion-curve inversion code developed by Herrmann (2013), to invert the dispersion curves for the 1D S-velocity models. By interpolating the 1D S-velocity models, a 2D S-velocity model can be computed, which then serves as the starting model for 2D WD (Li and Schuster, 2016; Li et al., 2017c). Finally, We interpolate the 2D WD tomograms to obtain a starting model for 3D WD.

[Figure 5 about here.]

NUMERICAL EXAMPLES

In this section, the 3D WD inversion method is evaluated with synthetic and field data examples including data associated with 1) a simple checkerboard model, 2) the complex 3D Foothills model, and 3)a surface seismic experiment carried out near the Qademah area north of KAUST.

In the synthetic examples, the observed and predicted data are generated by an O(2,8) time-space-domain solution to the first-order 3D elastic wave equation with a free-surface boundary condition (Graves, 1996). For 3D WD, only the S-wave velocity model is inverted and the true P-wave velocity model is used for modeling the predicted surface waves. The density model is homogeneous with ρ =2000 kg/m^3 for all synthetic models and the source wavelet is a Ricker wavelet, for the synthetic data. For the field data, the source wavelet is estimated from the direct arrivals. The fundamental dispersion curves associated with each shot gather are picked along the dominant azimuth angles, where the dominant azimuth angle is defined in equation 18 and the dispersion curves are picked for amplitudes above a specified threshold value R_0. For each iteration, source-side illumination compensation is used as a preconditioner (Plessix and Mulder, 2004; Feng and Schuster, 2017) for the WD gradient:

$$\gamma_{pre}(\boldsymbol{x}) = \frac{1}{\sum_{t,\boldsymbol{s}} \sqrt{D_1^2(t,\boldsymbol{x},\boldsymbol{s}) + D_2^2(t,\boldsymbol{x},\boldsymbol{s}) + D_3^2(t,\boldsymbol{x},\boldsymbol{s})}} \gamma(\boldsymbol{x}), \tag{19}$$

where $D_i(t,\boldsymbol{x},\boldsymbol{s})$ is the i^{th} component of the wavefield at $\boldsymbol{x}$ generated by the source located at $\boldsymbol{s}$.

Checkerboard Test

The 3D checkerboard model shown in Figure 6a is used to test the 3D WD method. The depth slice at $z = 6$ m of the model is shown in Figure 6b where the values of the high and low S-velocities are 690 m/s and 510 m/s, respectively. The initial S-velocity model is homogeneous with $v_s = 600$ m/s. The P velocity is set to be $v_p = \sqrt{3}v_s$. 100 vertical-component shots are uniformly distributed on a 10×10 source grid with an interval of 12 m along both the x and y directions. Each shot gather is recorded by a 40×40 receiver array with a 3 m spacing. The center frequency of the source wavelet is 30 Hz.

[Figure 6 about here.]

For the source located at $\boldsymbol{s} = (60, 0, 0)$ m, the adjoint-source wavefield $\bar{D}(\theta, \boldsymbol{g}, t)$ with $\theta = 90°$ is shown in Figure 7a, where the length r_1 of the receiver spread is 80 m. This adjoint source can be interpreted as a plane-wave source with the azimuth angle of 90°. The associated gradient at the depth slice $z = 6$ m is shown in Figure 7b. Figures 7c and 7d show the accumulated adjoint-source wavefield $\sum_\theta \bar{D}(\theta, \boldsymbol{g}, t)$ for all the azimuth angles from 0° to 180° with an interval of 5° and its gradient at $z = 6$ m, respectively.

The choice of the maximum source-receiver offset will affect the inversion results. For example, if we change the maximum source-receiver offsets r_1 to the values 40 m and 120 m, the associated depth slices of the gradients are shown in Figures 8a and 8b, respectively. The comparison suggests that the small offset value gives better horizontal resolution than that of the long offset value, yet, the longer offset value will lead to deeper velocity updates (Pérez Solano et al., 2014). As a rule of thumb, surface-wave methods can be sensitive to S-velocities down to a depth of about one-half of the total aperture of the receiver array (Foti

et al., 2014). To make sure that WD has enough penetration depth and lateral resolution, we usually choose r_1 to be about two or three times greater than the interested penetration depth.

Next, we reset the length of the receiver spread to be 40 m and repeat the 3D WD computations. The fundamental dispersion curves for each shot gather are picked along the dominant azimuths from 0° to 360° with an interval of 10° in the $k_x - k_y - f$ domain. For example, Figure 9 shows the observed dispersion curves from the CSGs for the sources located at points A, B, C and D marked in Figure 6b, where the black dashed lines represent the contours of the observed dispersion curves. The cyan dash-dot lines in Figure 9 represent the contours of the initial dispersion curves.

Figure 6c displays the inverted S-wave velocity model after 10 iterations, and one of its depth slices at $z = 6$ m is shown in Figure 6d, which agrees well with the true model. The contours of the predicted dispersion curves for the sources at A, B, C and D are represented by the red lines in Figure 9, which correlate well with the contours of the observed dispersion curves. After 10 iterations, the normalized misfit residual decreases to 2.8% of the starting value at the first iteration.

Figure 10 compares the depth slices of the true and the inverted tomograms at z= 15 m and 24 m. It is evident that the deep part of the velocity model is less updated compared to the shallow part. This indicates that the sensitivity of surface waves to the S-velocity decreases with depth.

[Figure 7 about here.]

[Figure 8 about here.]

[Figure 9 about here.]

[Figure 10 about here.]

[Figure 11 about here.]

Modified Foothills Model

The modified 3D Foothills S-wave velocity model is shown in Figure 11a, and the P-wave velocity is defined as $v_p = \sqrt{3}v_s$. The physical size of the velocity model is 1.2 km in the x and y directions and is 80 m deep in the z direction. An array of geophones is distributed on the surface, where 3600 receivers are arranged in 60 parallel lines along the x direction and each line has 60 receivers. The in-line and cross-line receiver intervals are both 20 m. There are 400 vertical-component shots distributed on a 20×10 grid with source intervals of 60 m and 120 m in the x and y directions, respectively. The peak frequency of the source is 15 Hz and the observed data are recorded for 0.90 seconds with a 0.3 ms sampling rate.

To construct the initial velocity model, we follow the workflow shown in Figure 5. Ten 2D in-line data sets are extracted from the 3D data set. We take the second line as an example, and all the sources on this line are located at $y = 120$ m. The vertical slice of the true S-velocity model at $y = 120$ m is shown in Figure 12a, where the locations of the first and last sources are indicated by the left-hand-side (LHS) and right-hand-side (RHS) black stars, respectively. The black horizontal lines near the stars represent the receiver spreads used to calculate the dispersion curves.

First, an initial model is obtained by applying the 1D inversion method to the dispersion curves. In Figure 13a, the dispersion curve for the first CSG from the 2D line at $y = 120$

m is shown as the blue solid line. The pseudo 1D S-velocity model is displayed as the blue dashed line in Figure 13b. We use SURF96 (Herrmann, 2013) to invert for the 1D velocity model (see the red solid line in Figure 13b). The predicted dispersion curve is shown as the red triangle line in Figure 13a, which matches well with the observed one. The inverted 1D depth profile is assumed at the middle of the receiver spreads, marked by the black solid circles in Figure 12. For comparison with the 3D tomogram by 3D WD, the inverted 1D model is interpolated as the 3D tomogram shown in Figure 11b.

The interpolated 1D profiles (see Figure 12b) are used as the starting model for 2D WD. Figure 12c shows the inverted 2D model at $y = 120$ m. Figures 14a and b show the observed dispersion curves for all of the 2D shot gathers at $y = 120$ m along the azimuth angles of $\theta = 0^\circ$ and $\theta = 180^\circ$, respectively, where the black dashed lines, the cyan lines and the red dash-dot lines represent the contours of the observed, initial and inverted dispersion curves, respectively. All ten inverted 2D models are interpolated to form a 3D tomogram (Figure 11c), which is the starting model for 3D WD.

During the workflow of 3D WD, the fundamental dispersion curves for each shot gather are picked along the dominant azimuths from 0° to 360° with an interval of 3° in the $k_x - k_y - f$ domain. For example, Figure 15 shows the observed dispersion curves calculated from the CSGs for the sources located at points A, B, C and D indicated in Figure 11a, where the black dashed lines represent the contours of the observed dispersion curves. The cyan lines represent the contours of initial dispersion curves.

Figure 11d displays the inverted S-wave velocity model, and its 2D slices at $y = 120$ m are shown in Figure 12d. Figure 16 shows the corresponding depth slices at $z = 20$ m for the models shown in Figure 11. The contours of the predicted dispersion curves for the

sources at A, B, C and D are represented by the red dash-dot lines in Figure 15, which agree better with the contours of the observed dispersion curves.

Figures 17 show the root-mean-square deviation between the true model and the inverted models by the 1D inversion, 2D WD and 3D WD methods. Compared with the 1D inversion and 2D WD methods, the root-mean-square deviation of the 3D WD is less than 35 percent and 22 percent less, respectively.

Figure 18 compares the observed (red) and synthetic (blue) traces at far source-receiver offsets predicted from the initial and inverted models for (a) and (b) with CSG No.1, and (c) and (d) with CSG No.15. It can be seen that the synthetic waveforms computed from the 3D WD tomogram more closely agree with the observed ones compared to those computed from the 2D WD velocity model.

Mitigating Cycle Skipping by WD

We will show that WD is less sensitive to cycle skipping compared to FWI by using the synthetic 2D model shown in Figure 12a. Figure 19a displays the initial model, which is far from the true model. The S-velocity tomogram is shown in Figure 19b. Figure 20 compares the observed (red) and synthetic (blue) seismograms predicted from the initial model (LHS panels) and the inverted model (RHS panels) for CSG No.1. We can see that the synthetic seismograms from the initial model have a time delay greater than half of the period of wavelet, which can lead to cycle skipping when using FWI (Virieux and Operto, 2009). However, the synthetic waveforms computed from the WD tomogram closely agree with the observed ones, which indicates that WD can sometimes mitigate the cycle skipping problem of FWI. This seems reasonable because, similar to wave-equation travel

time inversion (WT) (Luo and Schuster, 1991a,b), WD computes a simple dispersion curve to explain the observed seismogram and does not get confused with neighboring curves in FWI.

[Figure 12 about here.]

[Figure 13 about here.]

[Figure 14 about here.]

[Figure 15 about here.]

[Figure 16 about here.]

[Figure 17 about here.]

[Figure 18 about here.]

[Figure 19 about here.]

[Figure 20 about here.]

[Figure 21 about here.]

Qademah Fault Seismic Data

A 3D land survey was carried out along the Red Sea coast over the Qademah fault system, about 30 km north of the KAUST campus and the location is shown on the Google map

in Figure 21a (Hanafy, 2015). The survey consisted of 288 receivers arranged in 12 parallel lines, and each line has 24 receivers. The inline receiver interval is 5 m and the crossline interval is 10 m, and the source is a 40 kg weight drop striking a metal plate on the ground next to each geophone position. The receiver geometry is shown in Figure 21b, where one shot is fired at each receiver location for a total of 288 shot gathers. The observed data were recorded for 0.7 seconds with a 4 ms sampling rate. The following processing steps are first applied to the data:

- Each trace is normalized to compensate for the effects of attenuation and geometrical spreading. The traces of CSG No. 12 in the first line before and after amplitude compensation are shown in Figures 22a and 22b, respectively.
- All other arrivals but the fundamental-mode Rayleigh waves are masked in the CSG by a muting window, the length of the window is marked by the red dashed lines in Figure 22b.

The dispersion images shown in Figures 22c and 22d are computed by the frequency-sweeping method (Park et al., 1998), where the red asterisks represent the maximum value for each frequency. A false high-mode dispersion curve can be observed clearly on the top-right area, which is caused by the spatial aliasing due to the large receiver interval. The observed dispersion curves (see the blue lines) are picked and used for inversion. Figure 23 shows the observed dispersion curves for all of the 2D CSGs at the first line, where the black dashed lines represent the contours of the observed dispersion curves. At certain frequency ranges, it is difficult to pick the dispersion curves because of the low signal-to-noise ratio of the data so that some dispersion curves are missing in Figure 23.

Evaluation of the accuracy of the picked dispersion curves is important. A reciprocity test is needed to determine if the dispersion curves of a shot gather are the same as those of a receiver gather at the same location. The schematic diagram of the reciprocity test is shown in Figure 24.

We first apply the 1D inversion method to invert the dispersion curves. For example, the dispersion curve for the 12th CSG from the first line is shown as the solid blue line in Figure 25a. The pseudo 1D S-velocity model is displayed as the blue dashed line in Figure 25b. SURF96 (Herrmann, 2013) is used to invert for the 1D velocity model (see the red solid line in Figure 25b). The predicted dispersion curve is shown as the red triangle line in Figure 25a. The inverted 1D depth profile is assumed to be at the middle of the receiver spread. The 1D velocity profiles are interpolated as the starting model for 2D WD, which is shown in Figure 26a. For comparison with the 3D WD tomogram, we interpolated the 1D velocity profiles as the 3D model shown in Figure 27a.

Then, we apply 2D WD to invert for the 2D velocity model along the twelve lines. Figures 26a and 26b show the initial and inverted S-velocity models beneath the first line. The cyan lines in Figure 23 represent the contours of the initial dispersion curves. The contours of the predicted dispersion curves are represented by the red dash-dot lines in Figure 23, which more closely agree with the contours of the observed dispersion curves.

The twelve inverted 2D S-velocity models are then interpolated to obtain an initial velocity model (see Figure 27b) for 3D WD. For each shot gather, only the receivers within the distance $r_1 = 50$ m from the source are used to retrieve the dispersion curves. The frequency range used in WD is from 20 Hz to 60 Hz. Figure 28 displays the fundamental dispersion curves calculated from the CSGs for the sources located at A, B, C and D,

which are indicated in Figure 21. Here the black dashed lines represent the contours of the observed dispersion curves. The contours of the initial dispersion curves are represented by the cyan lines in Figure 28.

The S-wave velocity tomogram is shown in Figure 27c, where the red line labeled with "F1" indicates the location of the interpreted Qademah fault[†] and the red line labeled with "F2" refers to a possible small antithetic fault. The low-velocity anomaly between the two faults is interpreted as a colluvial wedge labeled with "CW". The red dash-dot lines in Figure 28 show the predicted $\kappa(\omega)$ curves calculated from the CSGs with sources located at A, B, C and D indicated in Figure 21. It is evident that the WD tomogram has decreased the differences between the initial and observed dispersion curves.

To further test the accuracy of the 3D tomogram, Figures 29 show the comparison between the observed (blue) and synthetic (red) traces at far offsets predicted from the initial model (LHS panels) and 3D tomogram (RHS panels) for (a) and (b) with CSG No. 9, and (c) and (d) with CSG No. 105. Here, a matched filter is applied to reshape the synthetic waveform. We can see that the predicted fundamental-mode surface waves closely match the observed ones. Figure 30 shows the common offset gathers (COGs) with the offset of 30 m for several 2D CSGs, where the blue and red wiggles represent the observed and predicted COGs, respectively. The predicted COGs are consistent with the raw data.

The slices of the S-wave velocity tomogram are shown in Figure 31a and the dashed lines indicate the locations of the conjectured Qademah fault. The low-velocity zone (LVZ) in Figure 31a next to the conjectured fault is consistent with the downthrown-side of an interpreted normal fault. The LVZ is also consistent with the reflectivities of the migration

[†]The low-velocity zone in this tomogram is interpreted as a fault.

image (Liu et al., 2016, 2017) indicated by the blue zone next to the dashed fault in Figure 31b. Figure 32 compares the 2D zoom view of the tomogram, migration image and COG, in which the LVZ of the tomogram is consistent with the location of the delay in the COG arrivals and the reflectivity of the migration image. This observation improves our confidence in the accuray of the S-wave velocity tomogram.

[Figure 22 about here.]

[Figure 23 about here.]

[Figure 24 about here.]

[Figure 25 about here.]

[Figure 26 about here.]

[Figure 27 about here.]

[Figure 28 about here.]

[Figure 29 about here.]

[Figure 30 about here.]

[Figure 31 about here.]

[Figure 32 about here.]

DISCUSSION

The lateral resolution of the WD tomogram is related to the length of the receiver spread. Different receiver-spread lengths can lead to different lateral-resolution limits of the retrieved dispersion curves (Mi et al., 2017; Bergamo et al., 2012). A wide receiver-spread for a specific azimuth angle can lead to poor lateral resolution along the azimuth angle of the gradient(Figure 8), but can provide a deep penetration depth (Foti et al., 2014). These resolution limits suggest a possible multiscale strategy: use the long-offset and low-frequency data to update the deep areas and use the short-offset and high-frequency data to update the shallow regions.

In our work we assumed that the effects of attenuation on dispersion curves are insignificant. However, if the attenuation is important and the attenuation model is known, the effects can be accounted for by solving the visco-elastic wave equation to compute the theoretical dispersion curves. For example, in the fault zones, attenuation is typically quite significant near the surface and in damaged faulted-zone layers, which can modify considerably the dispersion curves (Peng et al., 2003). Instead of inverting just for velocity, the WD method can be modified to invert for both the velocity and attenuation models (Li et al., 2017a,b). In this case the visco-elastic wave equation and its numerical solutions must be computed to estimate the gradients for the attenuation parameters. However, there is an inherent non-uniqueness problem is inverting for both velocity and attenuation models, so it is likely that both dispersion curves and normalized amplitudes should be used as input data.

A limitation of 3D WD is that the fundamental dispersion curves must be picked for each shot gather. This process can be prone to errors when there is a strong overlap with

higher-order modes (Li et al., 2017c) or there is spatial and temporal aliasing due to large spatial and temporal sampling intervals. A supervised machine learning method (Li et al., 2018a) can be used to expedite the picking of dispersion curves for large data sets. In addition, guided waves that are trapped in near-surface wave guides can be inverted by 3D WD for the near-surface P-velocity model (Li et al., 2018b).

CONCLUSIONS

We extend the 2D WD methodology to 3D, where the objective function is the sum of the squared differences between the wavenumbers along the predicted and observed dispersion curves for each azimuth angle. The Fréchet derivative with respect to the 3D S-wave velocity model is derived by the implicit function theorem. The WD gradient is calculated by correlating the back-propagated wavefield with the forward-propagated source field in the model based on the Born approximation in an isotropic, elastic reference earth model.

We provide a comprehensive approach to build the initial model for 3D WD, which starts from the pseudo 1D S-velocity model, which is then used to get the 2D WD tomogram, which in turn is used as the starting model for 3D WD. Our numerical results from both synthetic and field data show that the 3D WD method can reconstruct the 3D S-wave velocity tomogram for a laterally heterogeneous medium so that the predicted surface waves closely match the observed ones for the fundamental modes. This suggests that the WD result can serve as a good starting model for surface-wave FWI. The 3D WD method can be easily adapted to also invert the higher-order modes for a more detailed velocity model. In addition, guided waves that are trapped in near-surface wave guides can be inverted by 3d WD for the near-surface P-velocity model.

The main limitation of 3D WD is its high computational cost, which is more than an order of magnitude greater than that of 2D WD. However, the improvement in accuracy compared to 2D WD can make this extra cost worthwhile when there are significant near-surface lateral variations in the S-velocity distribution. If the attenuation is important, then its effects can be accounted for by solving the visco-elastic wave equation to compute the theoretical dispersion curves. To expedite the picking of dispersion curves we recommend supervised machine learning methods that adapt to the data recorded at different sites.

REFERENCES

Aki, K., and P. G. Richards, 2002, Quantitative Seismology: University Science Books.

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

Boiero, D., P. Bergamo, R. B. Rege, and L. V. Socco, 2011, Estimating surface-wave dispersion curves from 3D seismic acquisition schemes: Part 1-1D models: Geophysics, **76**, G85–G93.

Dou, S., and J. B. Ajo-Franklin, 2014, Full-wavefield inversion of surface waves for mapping embedded low-velocity zones in permafrost: Geophysics, **79**, EN107–EN124.

Feng, Z., and G. T. Schuster, 2017, Elastic least-squares reverse time migration: GEOPHYSICS, **82**, S143–S157.

Foti, S., C. G. Lai, G. J. Rix, and C. Strobbia, 2014, Surface wave methods for near-surface site characterization: CRC Press.

Fu, L., B. Guo, and G. T. Schuster, 2018a, Multiscale phase inversion of seismic data: GEOPHYSICS, **83**, R159–R171.

Fu, L., Z. Liu, and G. Schuster, 2018b, Superresolution near-field imaging with surface waves: Geophys. J. Int., **212**, 1111–1122.

Graves, R. W., 1996, Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences: Bull. Seismol. Soc. Am., **86**, 1091–1106.

Groos, L., M. Schäfer, T. Forbriger, and T. Bohlen, 2014, The role of attenuation in 2D full-waveform inversion of shallow-seismic body and Rayleigh waves: Geophysics, **79**, R247–R261.

———, 2017, Application of a complete workflow for 2D elastic full-waveform inversion to recorded shallow-seismic Rayleigh waves: Geophysics, **82(2)**, R109–R117.

Hanafy, S. M., 2015, Mapping the Qademah fault with traveltime, surface-wave, and resistivity tomograms: SEG Technical Program Expanded Abstracts 2015, Society of Exploration Geophysicists, Society of Exploration Geophysicists, 3347–3351.

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

Herrmann, R. B., 2013, Computer programs in seismology: An evolving tool for instruction and research: Seismol. Res. Lett., **84**, 1081.

Ivanov, J., B. Leitner, W. Shefchik, J. T. Shwenk, and S. L. Peterie, 2013, Evaluating hazards at salt cavern sites using multichannel analysis of surface waves: The Leading Edge, **32**, 298–305.

Li, J., Y. Chen, and G. Schuster, 2018a, Separation of multi-mode surface waves by supervised machine learning methods: Presented at the 2018 SEG/SPE Beijing Workshop: Maximizing asset value through Artificial Intelligence and Machine Learning.

Li, J., G. Dutta, and G. Schuster, 2017a, Wave-equation Qs inversion of skeletonized surface waves: Geophys. J. Int., **209**, 979–991.

Li, J., G. Dutta, and G. T. Schuster, 2017b, Skeletonized wave-equation Qs tomography using surface waves: SEG Technical Program Expanded Abstracts 2017, Society of Exploration Geophysicists, Society of Exploration Geophysicists, 2726–2731.

Li, J., Z. Feng, and G. Schuster, 2017c, Wave-equation dispersion inversion: Geophys. J. Int., **208**, 1567–1578.

Li, J., S. Hanafy, and G. Schuster, 2018b, Wave-equation dispersion inversion of guided P waves in a waveguide of arbitrary geometry: Journal of Geophysical Research, **Accepted for Publication**.

Li, J., K. Lu, and G. Schuster, 2017d, Two robust imaging methodologies for challenging

environments: Wave equation dispersion inversion of surface waves and supervirtual interferometry+tomography for far-offset refractions: SEG/SINOPEC Foothill Technical Forum, Nanjing, China, 23-24 April 2017, 17–21.

Li, J., and G. Schuster, 2016, Skeletonized wave equation of surface wave dispersion inversion: SEG Technical Program Expanded Abstracts 2016, Society of Exploration Geophysicists, Society of Exploration Geophysicists, 3630–3635.

Liu, Z., A. Altheyab, S. Hanafy, and G. Schuster, 2016, Imaging near-surface heterogeneities by natural migration of surface waves: SEG Technical Program Expanded Abstracts 2016, Society of Exploration Geophysicists, 4946–4950.

Liu, Z., A. AlTheyab, S. M. Hanafy, and G. Schuster, 2017, Imaging near-surface heterogeneities by natural migration of backscattered surface waves: Field data test: Geophysics, **82(3)**, S197–S205.

Lu, K., J. Li, B. Guo, L. Fu, and G. Schuster, 2017, Tutorial for wave-equation inversion of skeletonized data: Interpretation, **5**, SO1–SO10.

Luo, Y., and G. T. Schuster, 1991a, Wave equation inversion of skeletalized geophysical data: Geophys. J. Int., **105**, 289–294.

———, 1991b, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Masoni, I., J.-L. Boelle, R. Brossier, and J. Virieux, 2016, Layer stripping FWI for surface waves: SEG Technical Program Expanded Abstracts 2016, Society of Exploration Geophysicists, 1369–1373.

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

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228.

Nocedal, J., and S. Wright, 2006, Numerical optimization: Springer Science & Business Media.

O'Neill, A., and T. Matsuoka, 2005, Dominant higher surface-wave modes and possible inversion pitfalls: J. Environ. Eng. Geophys., **10**, 185–201.

Pan, Y., J. Xia, Y. Xu, Z. Xu, F. Cheng, H. Xu, and L. Gao, 2016, Delineating shallow S-wave velocity structure using multiple ambient-noise surface-wave methods: An example from western Junggar, China: Bull. Seismol. Soc. Am., **106**, 327–336.

Park, C. B., R. D. Miller, and J. Xia, 1998, Imaging dispersion curves of surface waves on multichannel record: SEG Technical Program Expanded Abstracts 1998, Society of Exploration Geophysicists, 1377–1380.

———, 1999, Multichannel analysis of surface waves: Geophysics, **64**, 800–808.

Peng, Z., Y. Ben-Zion, A. J. Michael, and L. Zhu, 2003, Quantitative analysis of seismic fault zone waves in the rupture zone of the 1992 landers, california, earthquake: evidence for a shallow trapping structure: Geophysical Journal International, **155**, 1021–1041.

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

Plessix, R.-E., and W. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophys. J. Int., **157**, 975–987.

Schuster, G. T., 2017, Seismic Inversion: SEG Society of Exploration Geophysicists.

Snieder, R., 2002a, General theory of elastic wave scattering: Chapter 1.7.1, *in* Scattering and Inverse Scattering in Pure and Applied Science: Academic Press, San Diego, California, 528 – 542.

———, 2002b, Scattering of surface waves: Chapter 1.7.3, *in* Scattering and inverse scattering in pure and applied science: Academic Press, San Diego, California, 562–577.

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

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Xia, J., R. Miller, C. Park, and G. Tian, 2002, Determining Q of near-surface materials from Rayleigh waves: J. Appl. Geophys., **51**, 121–129.

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

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

APPENDIX A

CORRELATION IDENTITY

The integrand in equation 10 can be replaced by its Fourier transform (Li et al., 2017c)

$$\dot{\tilde{D}}(k+\Delta\kappa,\theta,\omega)_{obs} = -\int\int i(x'_g\cos\theta + y'_g\sin\theta)$$
$$D(x'_g,y'_g,\omega)_{obs}e^{-i(k+\Delta\kappa)(\cos\theta x'_g+\sin\theta y'_g)}dx'_g dy'_g,$$
$$\tilde{D}(k,\theta,\omega)_{pre} = \int\int D(x_g,y_g,\omega)_{pre}e^{-ik(\cos\theta x_g+\sin\theta y_g)}dx_g dy_g,$$

to give

$$\frac{\partial\dot{\Phi}(\Delta\kappa, v_s(\boldsymbol{x}))}{\partial v_s(\boldsymbol{x})} = \Re\Big\{ \int\int dx_g dy_g \Big[\int\int dx'_g dy'_g \big[\int e^{ik(\cos\theta(x'_g - x_g) + \sin\theta(y'_g - y_g))} dk \big] i(x'_g \cos\theta + y'_g \sin\theta) D(x'_g, y'_g, \omega)^*_{obs} e^{i\Delta\kappa(\cos\theta x'_g + \sin\theta y'_g)} \Big] \frac{\partial D(x_g, y_g, \omega)_{pre}}{\partial v_s(\boldsymbol{x})} \Big\}$$

$$= \Re\Big\{ \int\int dx_g dy_g \Big[\int\int dx'_g dy'_g \big[2\pi\delta(\cos\theta(x'_g - x_g) + \sin\theta(y'_g - y_g)) \big] i(x'_g \cos\theta + y'_g \sin\theta) D(x'_g, y'_g, \omega)^*_{obs} e^{i\Delta\kappa(\cos\theta x'_g + \sin\theta y'_g)} \Big] \frac{\partial D(x_g, y_g, \omega)_{pre}}{\partial v_s(\boldsymbol{x})} \Big\}$$

$$= \Re\Big\{ \int\int dx_g dy_g \frac{\partial D(x_g, y_g, \omega)_{pre}}{\partial v_s(\boldsymbol{x})} 2\pi \Big[i \frac{(x_g \cos\theta + y_g \sin\theta)}{\cos\theta} e^{i\Delta\kappa(\cos\theta x_g + \sin\theta y_g)} \int dy'_g D(x_g - (y'_g - y_g)\tan\theta, y'_g, \omega)^*_{obs} \Big] \Big\}$$

$$= \Re\Big\{ \int\int dx_g dy_g \frac{\partial D(x_g, y_g, \omega)_{pre}}{\partial v_s(\boldsymbol{x})} \hat{D}(x_g, y_g, \omega)^*_{obs} \Big\}, \tag{A-1}$$

where the weighted conjugated data function is

$$\hat{D}(\boldsymbol{g}, \theta, \omega)^*_{obs} = 2\pi i \frac{\boldsymbol{g}\cdot\boldsymbol{n}}{\cos\theta} e^{i\boldsymbol{g}\cdot\boldsymbol{n}\Delta\kappa} \int dy'_g D(x_g - (y'_g - y_g)\tan\theta, y'_g, \omega)^*_{obs}, \tag{A-2}$$

in which $\boldsymbol{g} = (x_g, y_g)$ and $\boldsymbol{n} = (\cos\theta, \sin\theta)$.

To calculate the weighted conjugated data, we first compute the integration of $D(x_g - (y'_g - y_g)\tan\theta, y'_g, \omega)^*_{obs}$ along the line $x'_g = x_g - (y'_g - y_g)\tan\theta$ (see the schematic diagram in Figure 1). The line is passing through the point $\boldsymbol{x} = (x_g, y_g)$ and is perpendicular to the direction of $\boldsymbol{n}$. We can see that the weighted conjugated data along the line are almost identical, which means it will generate a plane-wave surface wave for the backprojected wavefield.

Next, we will interpret the weighted conjugated data function by the stationary phase

method. The Green's function for the fundamental mode of Rayleigh waves excited by a vertical-component force in the far field can be approximated as (Snieder, 2002b):

$$G(x_g, y_g, \omega) \simeq A' \frac{e^{-ik\sqrt{x_g^2+y_g^2}+i\pi/4}}{\sqrt{0.5\pi k\sqrt{x_g^2+y_g^2}}}. \tag{A-3}$$

where the source location the term A' accounts for the source amplitude and radiation patten for a trace at (x_g, y_g) by a point source at $(0, 0)$.

Replacing $D(x_g, y_g)_{obs}$ in the equation A-2 with the Rayleigh Green's function in the equation A-3 and the source $W(\omega)$, we can get

$$\hat{D}(\boldsymbol{g}, \theta, \omega)^*_{obs} = 2\pi i W(\omega) \frac{\boldsymbol{g}\cdot\boldsymbol{n}}{\cos\theta} e^{i\boldsymbol{g}\cdot\boldsymbol{n}\Delta\kappa} A' e^{i\pi/4} \int dy'_g \frac{e^{ikf(y'_g)}}{\sqrt{0.5\pi k\sqrt{x_g^2+y_g^2}}}, \tag{A-4}$$

where

$$f(y'_g) = \sqrt{{y'_g}^2/\cos\theta - 2(x_g + y_g\tan\theta)\tan\theta y'_g + (x_g + y_g\tan\theta)^2}. \tag{A-5}$$

According to the stationary phase approximation, for $k \gg 1$, the stationary point is located at $y'_g = c_y$ so that $\frac{\partial f(c_y)}{\partial y'_g} = 0$ and $c_y = (x_g\cos\theta + y_g\sin\theta)\sin\theta$. Because (x'_g, y'_g) is located at the line: $(\boldsymbol{g}' - \boldsymbol{g})\cdot\boldsymbol{n} = 0$, the x coordinate of the stationary point is $c_x = (x_g\cos\theta + y_g\sin\theta)\cos\theta$. Then equation A-4 can be approximated as

$$\begin{aligned}\hat{D}(\boldsymbol{g}, \theta, \omega)^*_{obs} &= 2\pi i W(\omega)\frac{\boldsymbol{g}\cdot\boldsymbol{n}}{\cos\theta} e^{i\boldsymbol{g}\cdot\boldsymbol{n}\Delta\kappa} A' e^{i\pi/4} \frac{e^{ik(x_g\cos\theta+y_g\sin\theta)}}{\sqrt{0.5k\sqrt{x_g^2+y_g^2}/\pi}} \\ &\sqrt{\frac{2\pi i(\cos^2\theta(x_g\cos\theta + y_g\sin\theta))}{k}} \\ &= 2\pi i\mu\boldsymbol{g}\cdot\boldsymbol{n} W(\omega) e^{i\boldsymbol{g}\cdot\boldsymbol{n}\Delta\kappa} A' e^{i\pi/4} \frac{e^{ik\boldsymbol{g}\cdot\boldsymbol{n}}}{\sqrt{0.5k|\boldsymbol{g}|/\pi}} \sqrt{\frac{2\pi i\boldsymbol{g}\cdot\boldsymbol{n}}{k}},\end{aligned} \tag{A-6}$$

where $\mu \cos\theta = |\cos\theta|$.

APPENDIX B

ELASTIC GRADIENT

The gradient for the WD method is now derived. For an isotropic heterogeneous medium, the Born approximation in terms of the 3D elastic Green's functions for a harmonic source (Snieder, 2002a) is

$$\begin{aligned}\delta D_i(\boldsymbol{x},\omega) &= \omega^2 \int G_{ij}(\boldsymbol{x}|\boldsymbol{x}')\delta\rho(\boldsymbol{x}')D_i(\boldsymbol{x}',\omega)dx'^3 \\ &- \int G_{ik,k}(\boldsymbol{x}|\boldsymbol{x}')\delta\lambda(\boldsymbol{x}')D_{j,j}(\boldsymbol{x}',\omega)dx'^3 \\ &- \int G_{in,k}(\boldsymbol{x}|\boldsymbol{x}')\delta\mu(\boldsymbol{x}')(D_{k,n}(\boldsymbol{x}',\omega)+D_{n,k}(\boldsymbol{x}',\omega))dx'^3, \end{aligned} \tag{B-1}$$

where $\delta D_i(\boldsymbol{x},\omega)$ denotes the i^{th} component of the perturbed particle velocity recorded at $\boldsymbol{x}$ due to the scattering from the perturbations of density $\delta\rho$ and Lamé parameters $\delta\lambda$ and $\delta\mu$. Einstein notation is assumed in equation B-1. $D_{i,j} = \frac{\partial D_i}{\partial_{x_j}}$ for $i,j \in \{1,2,3\}$. G_{ij} is the 3D harmonic Green's tensor (Snieder, 2002a) for the background medium with the Lamé parameters λ and μ, and density ρ. If we assume density ρ is a constant, equation B-1 yields the derivative of $\delta D_i(\boldsymbol{x},\omega)$ with respect to $\delta\lambda$ and $\delta\mu$ at $\boldsymbol{x}'$

$$\begin{aligned}&\frac{\delta D_i(\boldsymbol{x},\omega)}{\delta\lambda(\boldsymbol{x}')} = -G_{ik,k}(\boldsymbol{x}|\boldsymbol{x}')D_{j,j}(\boldsymbol{x}',\omega), \\ &\text{and} \quad \frac{\delta D_i(\boldsymbol{x},\omega)}{\delta\mu(\boldsymbol{x}')} = -G_{in,k}(\boldsymbol{x}|\boldsymbol{x}')(D_{k,n}(\boldsymbol{x}',\omega)+D_{n,k}(\boldsymbol{x}',\omega)).\end{aligned} \tag{B-2}$$

Our interest is confined to the derivative of the vertical component of the particle velocity at $\boldsymbol{g}$, so equation B-2 for $i=3$, $\{1,2,3\} \to \{x,y,z\}$ and $D_i(\boldsymbol{x},\omega) \to D(\boldsymbol{g},\omega)$ with respect

to λ and μ at $\boldsymbol{x}$ can be written as

$$\frac{\partial D(\boldsymbol{g},\omega)}{\partial \lambda(\boldsymbol{x})} = -\Bigg(\frac{\partial G_{zx}(\boldsymbol{g}|\boldsymbol{x})}{\partial x} + \frac{\partial G_{zy}(\boldsymbol{g}|\boldsymbol{x})}{\partial y} + \frac{\partial G_{zz}(\boldsymbol{g}|\boldsymbol{x})}{\partial z}\Bigg)\Bigg(\frac{\partial D_x(\boldsymbol{x},\omega)}{\partial x} + \frac{\partial D_y(\boldsymbol{x},\omega)}{\partial y} + \frac{\partial D_z(\boldsymbol{x},\omega)}{\partial z}\Bigg), \tag{B-3}$$

and

$$\begin{aligned}\frac{\partial D(\boldsymbol{g},\omega)}{\partial \mu(\boldsymbol{x})} = &-2\Bigg(\frac{\partial G_{zx}(\boldsymbol{g}|\boldsymbol{x})}{\partial x}\frac{\partial D_x(\boldsymbol{x},\omega)}{\partial x} + \\ &\frac{\partial G_{zy}(\boldsymbol{g}|\boldsymbol{x})}{\partial y}\frac{\partial D_y(\boldsymbol{x},\omega)}{\partial y} + \frac{\partial G_{zz}(\boldsymbol{g}|\boldsymbol{x})}{\partial z}\frac{\partial D_z(\boldsymbol{x},\omega)}{\partial z}\Bigg) \\ &-\Bigg(\frac{\partial G_{zx}(\boldsymbol{g}|\boldsymbol{x})}{\partial z} + \frac{\partial G_{zz}(\boldsymbol{g}|\boldsymbol{x})}{\partial x}\Bigg)\Bigg(\frac{\partial D_x(\boldsymbol{x},\omega)}{\partial z} + \frac{\partial D_z(\boldsymbol{x},\omega)}{\partial x}\Bigg) \\ &-\Bigg(\frac{\partial G_{zx}(\boldsymbol{g}|\boldsymbol{x})}{\partial y} + \frac{\partial G_{zy}(\boldsymbol{g}|\boldsymbol{x})}{\partial x}\Bigg)\Bigg(\frac{\partial D_x(\boldsymbol{x},\omega)}{\partial y} + \frac{\partial D_y(\boldsymbol{x},\omega)}{\partial x}\Bigg) \\ &-\Bigg(\frac{\partial G_{zy}(\boldsymbol{g}|\boldsymbol{x})}{\partial z} + \frac{\partial G_{zz}(\boldsymbol{g}|\boldsymbol{x})}{\partial y}\Bigg)\Bigg(\frac{\partial D_y(\boldsymbol{x},\omega)}{\partial z} + \frac{\partial D_z(\boldsymbol{x},\omega)}{\partial y}\Bigg),\end{aligned} \tag{B-4}$$

where $D_x(\boldsymbol{x},\omega)$, $D_y(\boldsymbol{x},\omega)$ and $D_z(\boldsymbol{x},\omega)$ are finite-difference solutions to the 3D elastic wave equation for the background velocity model. From the definitions $v_p = \sqrt{(\lambda+2\mu)/\rho}$ and $v_s = \sqrt{\mu/\rho}$, the Fréchet derivative of $D(\boldsymbol{g},\omega)$ with respect to v_s (Mora, 1987) can be obtained:

$$\begin{aligned}\frac{\partial D(\boldsymbol{g},\omega)}{\partial v_s(\boldsymbol{x})} = &-4v_s(\boldsymbol{x})\rho(\boldsymbol{x})\frac{\partial D(\boldsymbol{g},\omega)_{pre}}{\partial \lambda(\boldsymbol{x})} \\ &+ 2v_s(\boldsymbol{x})\rho(\boldsymbol{x})\frac{\partial D(\boldsymbol{g},\omega)_{pre}}{\partial \mu(\boldsymbol{x})}\end{aligned} \tag{B-5}$$

Inserting equations B-3 and B-4 into equation B-5 gives,

$$\frac{\partial D(\boldsymbol{g},\omega)}{\partial v_s(\boldsymbol{x})} = 4v_s(\boldsymbol{x})\rho(\boldsymbol{x})\Big\{\big(\frac{\partial G_{zy}(\boldsymbol{g}|\boldsymbol{x})}{\partial y} + \frac{\partial G_{zz}(\boldsymbol{g}|\boldsymbol{x})}{\partial z}\big)$$
$$\frac{\partial D_x(\boldsymbol{x},\omega)}{\partial x} + \big(\frac{\partial G_{zx}(\boldsymbol{g}|\boldsymbol{x})}{\partial x} + \frac{\partial G_{zz}(\boldsymbol{g}|\boldsymbol{x})}{\partial z}\big)\frac{\partial D_y(\boldsymbol{x},\omega)}{\partial y} +$$
$$\big(\frac{\partial G_{zx}(\boldsymbol{g}|\boldsymbol{x})}{\partial x} + \frac{\partial G_{zy}(\boldsymbol{g}|\boldsymbol{x})}{\partial y} + \big)\frac{\partial D_z(\boldsymbol{x},\omega)}{\partial z}$$
$$-\frac{1}{2}\big(\frac{\partial G_{zx}(\boldsymbol{g}|\boldsymbol{x})}{\partial z} + \frac{\partial G_{zz}(\boldsymbol{g}|\boldsymbol{x})}{\partial x}\big)\big(\frac{\partial D_x(\boldsymbol{x},\omega)}{\partial z} + \frac{\partial D_z(\boldsymbol{x},\omega)}{\partial x}\big)$$
$$-\frac{1}{2}\big(\frac{\partial G_{zx}(\boldsymbol{g}|\boldsymbol{x})}{\partial y} + \frac{\partial G_{zy}(\boldsymbol{g}|\boldsymbol{x})}{\partial x}\big)\big(\frac{\partial D_x(\boldsymbol{x},\omega)}{\partial y} + \frac{\partial D_y(\boldsymbol{x},\omega)}{\partial x}\big)$$
$$-\frac{1}{2}\big(\frac{\partial G_{zy}(\boldsymbol{g}|\boldsymbol{x})}{\partial z} + \frac{\partial G_{zz}(\boldsymbol{g}|\boldsymbol{x})}{\partial y}\big)\big(\frac{\partial D_y(\boldsymbol{x},\omega)}{\partial z} + \frac{\partial D_z(\boldsymbol{x},\omega)}{\partial y}\big)\Big\}. \tag{B-6}$$

LIST OF FIGURES

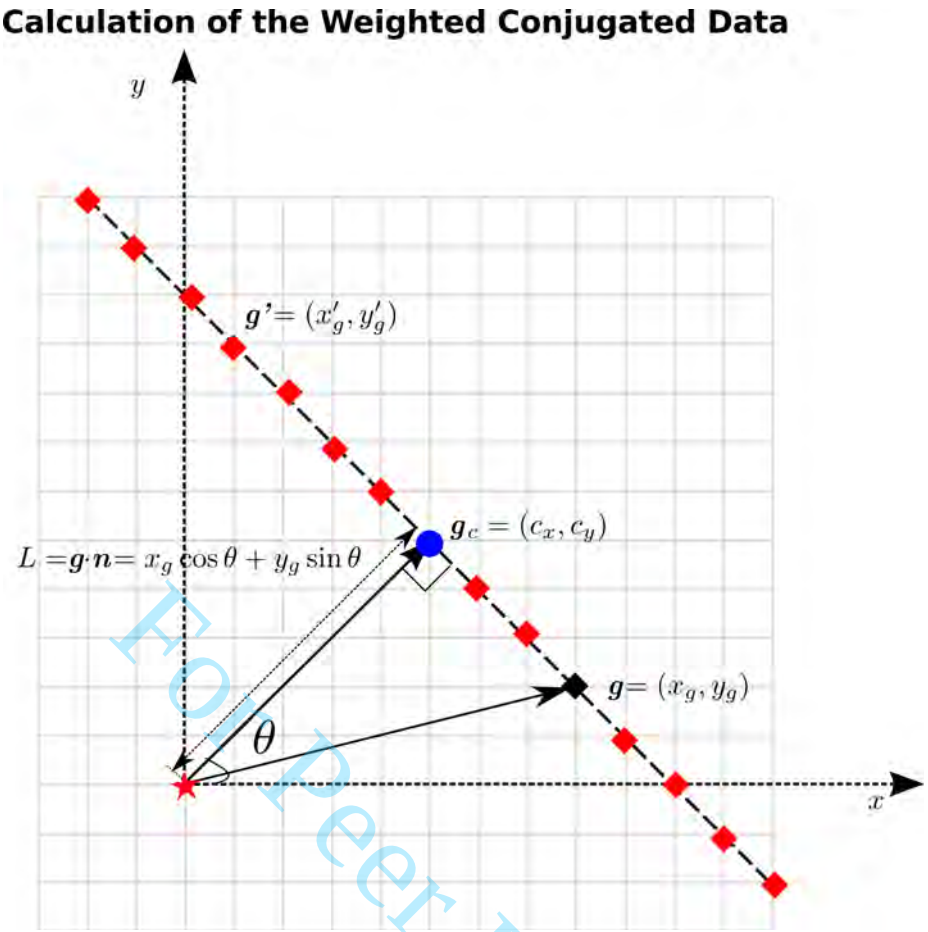


Figure 1: Schematic diagram showing how to calculate the weighted conjugated data $\hat{D}(\boldsymbol{g}, \theta, \omega)^*_{obs}$, where the red star represents the source, the black solid square shows the geophone location at $\boldsymbol{g}$ and the red solid squares represent the geophones along the line C which satisfies $(\boldsymbol{g}' - \boldsymbol{g}) \cdot \boldsymbol{n} = 0$. For the azimuth θ and position $\boldsymbol{g}$, $\hat{D}(\boldsymbol{g}', \theta, \omega)^*_{obs}$ is integrated along the dashed line with the weighting term $2\pi i L e^{i\Delta\kappa L}$, where $L = \boldsymbol{g} \cdot \boldsymbol{n}$. The blue dot at $\boldsymbol{g}_c$ is the stationary point for a homogeneous half-space, and the line integral in equation 13 can be approximated by $D(\boldsymbol{g}_c, \omega)_{obs}$ (see Appendix A for the detailed derivation).

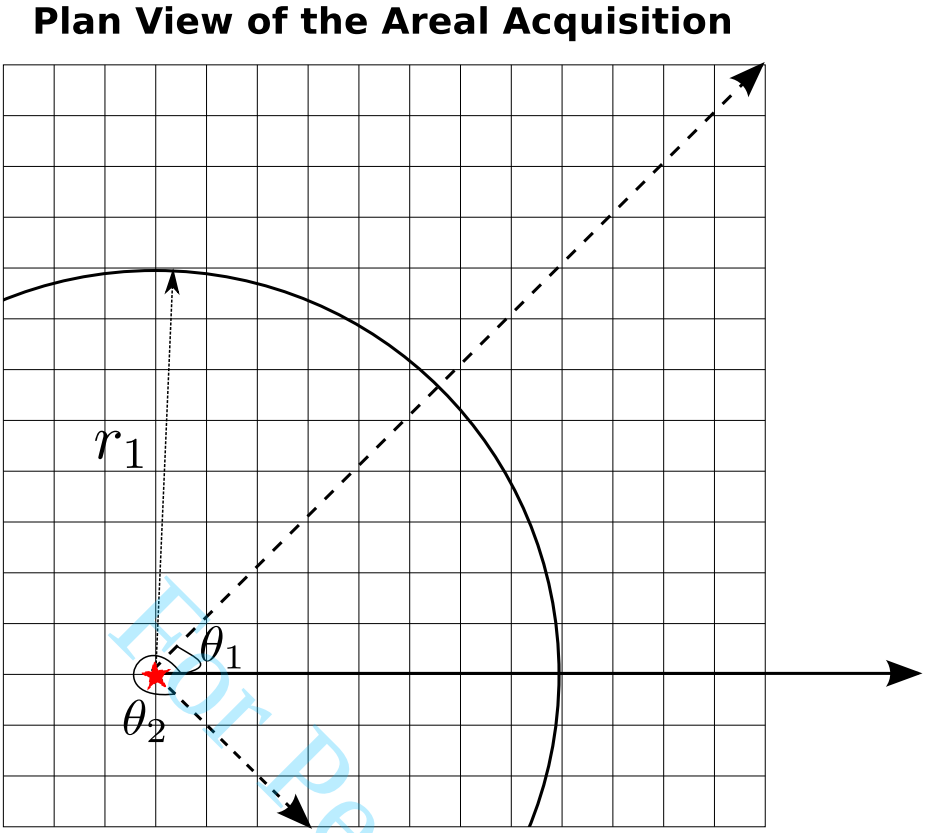
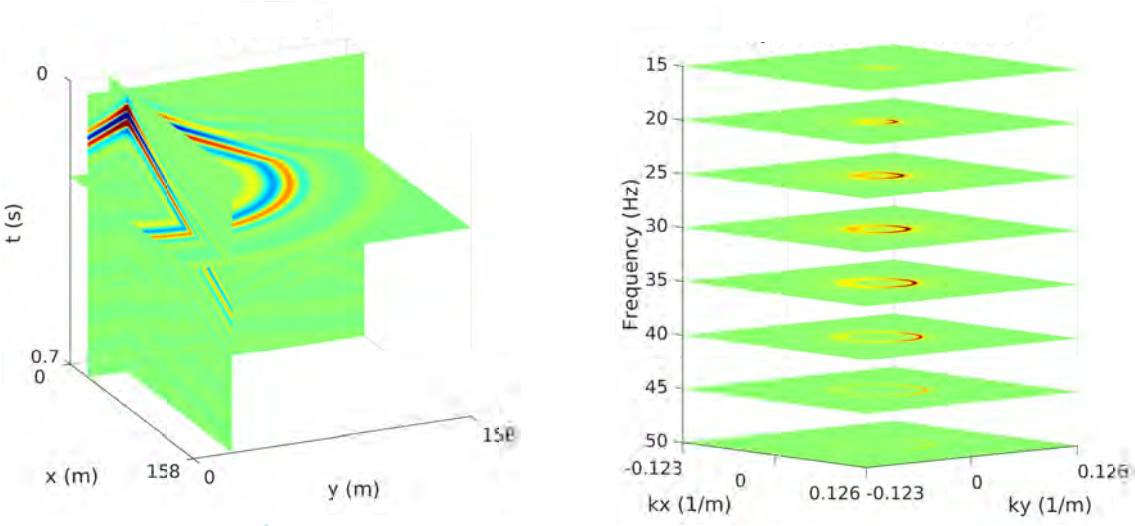


Figure 2: Plan view of the areal acquisition, where the red star represents the source, and the grid points at the line crossings represent the locations of geophones.

(a) **3D CSG**

(b) **Spectrum of 3D CSG**



(c) **Frequency Slice at 50 Hz**

(d) **Picked 3D** $\kappa(k_x, k_y)$

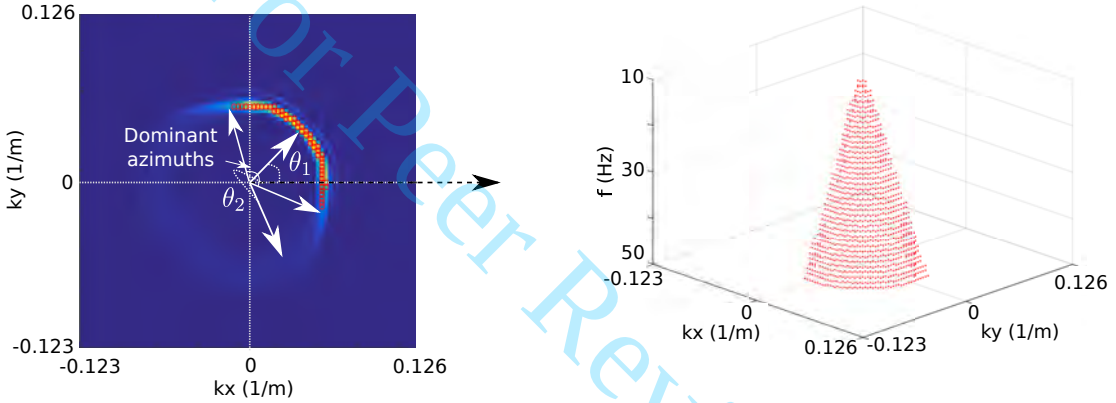


Figure 3: (a) 3D common shot gather, (b) its spectrum, and (c) frequency slice of the magnitude spectrum at 50 Hz from (b). (d) Picked dispersion surface according to the dominant amplitudes of the spectrum.

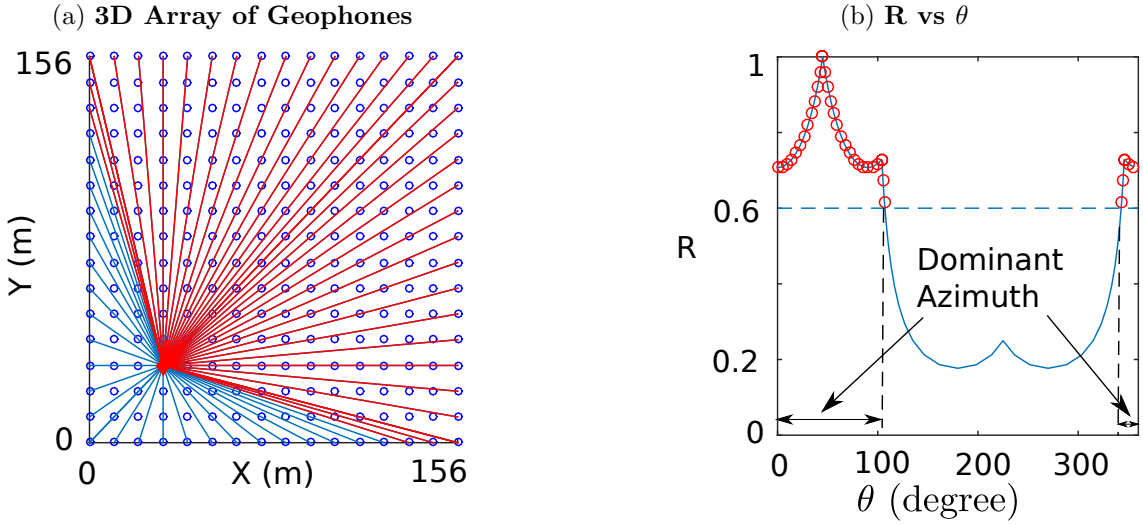


Figure 4: (a) Lines from the source point located at (30 m, 30 m) to the geophones along the boundary, and (b) $R(\theta)$ plotted against the azimuth angles.

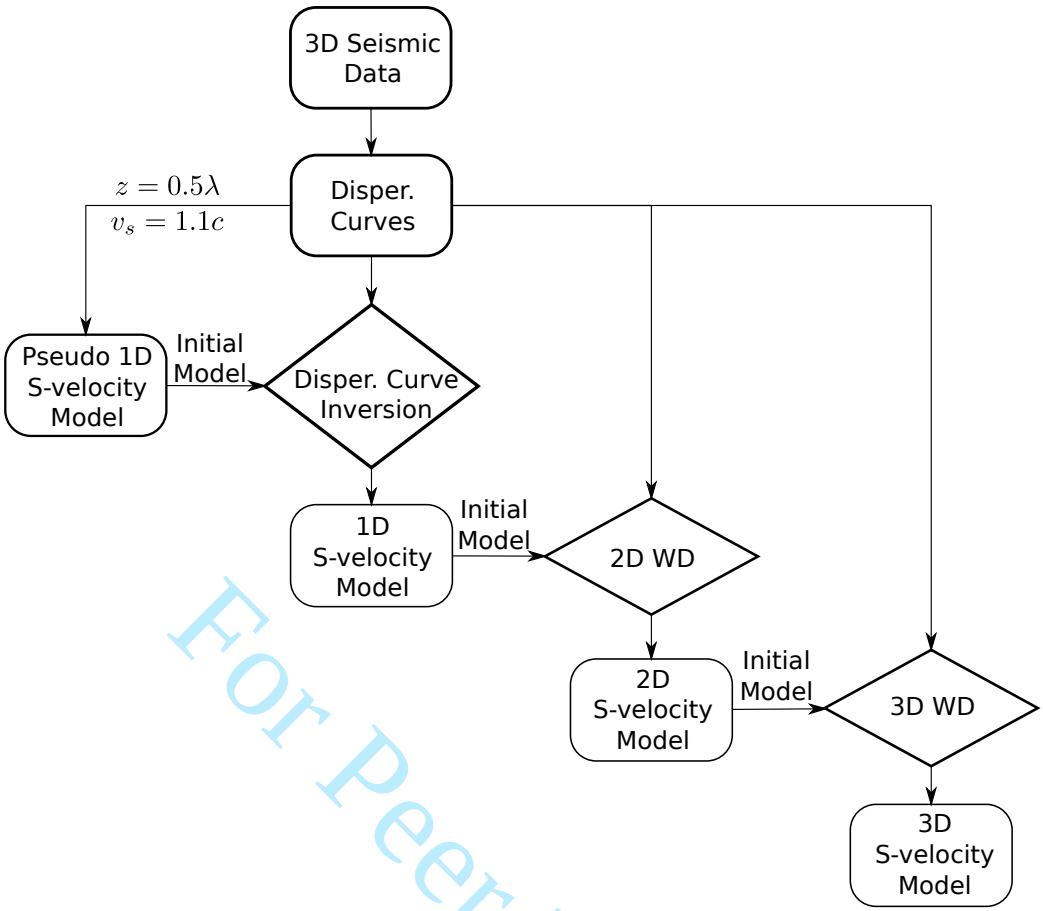
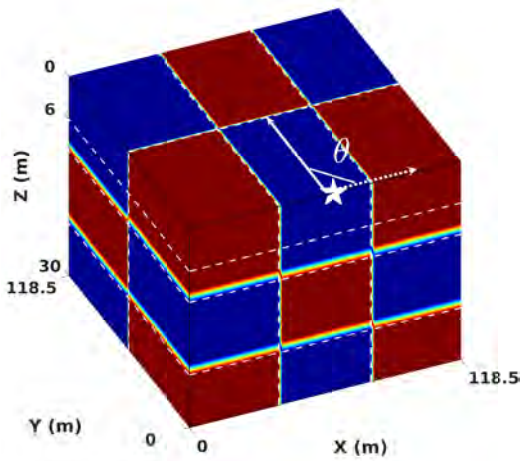
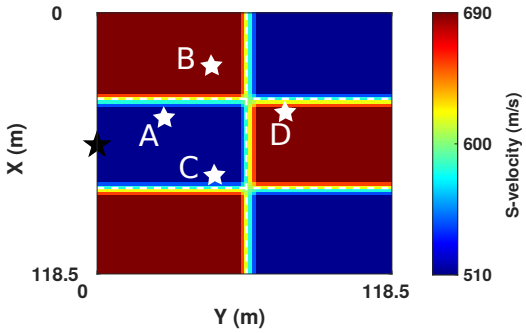


Figure 5: Workflow for calculating the initial S-velocity model for 3D WD.

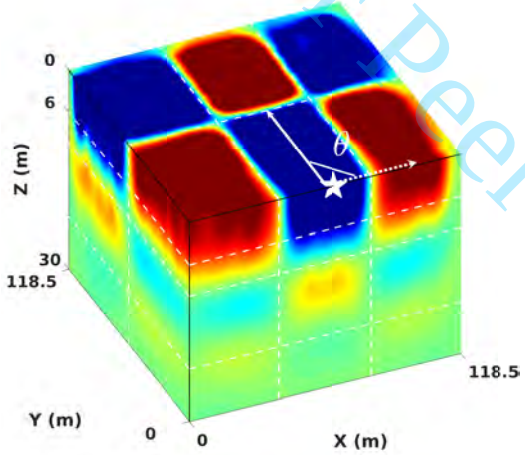
(a) **True S-velocity Model**



(b) **Depth Slice at $z = 6$ m of Model (a)**



(c) **Inverted S-velocity Tomogram**



(d) **Depth Slice at $z = 6$ m of the (c) Tomogram**

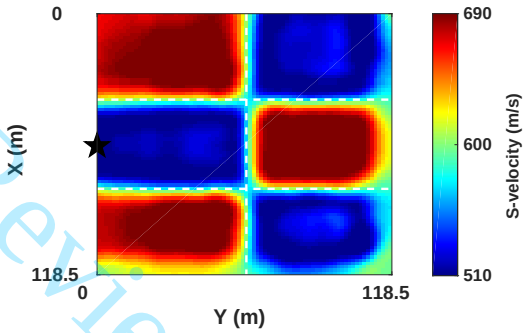
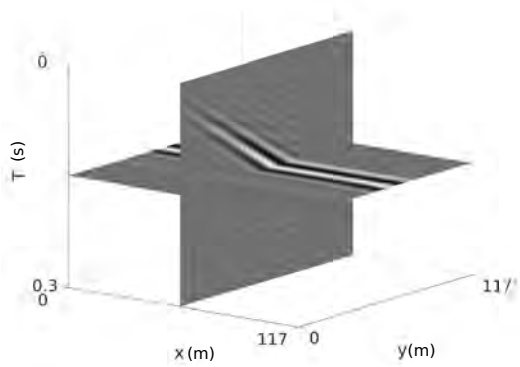
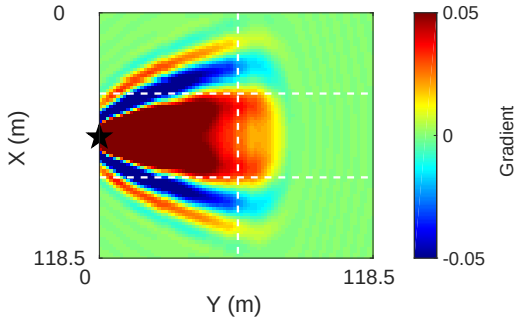


Figure 6: (a) True S-velocity model and its (b) depth slice at $z = 6$ m, (c) inverted S-velocity tomogram and (d) depth slice at $z = 6$ m.

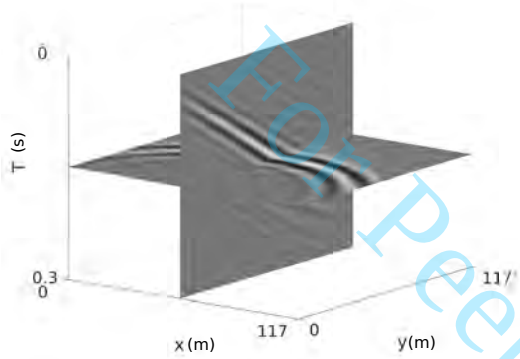
(a) **Adjoint Source for** $\theta = 90°$



(b) **Gradient Slice for** $\theta = 90°$



(c) **Adjoint Source for** θ **from** $0°$ **to** $180°$



(d) **Gradient Slice for** θ **from** $0°$ **to** $180°$

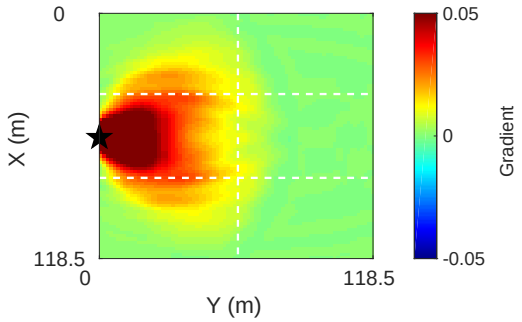


Figure 7: (a) Wavefields of the adjoint source for $\theta = 90°$ and (b) its gradient at the depth slice $z = 6$ m; (c) stacked wavefields of the adjoint sources for θ from $0°$ to $180°$ and (d) the gradient at the depth slice $z = 6$ m, where the maximum source-receiver offset is r_1=80 m and the source is located at $\boldsymbol{s} = (60, 0, 0)$ m.

(a) **Gradient Slice for $\theta = 90°$, $r_1 = 40$ m** (b) **Gradient Slice for $\theta = 90°$, $r_1 = 120$ m**

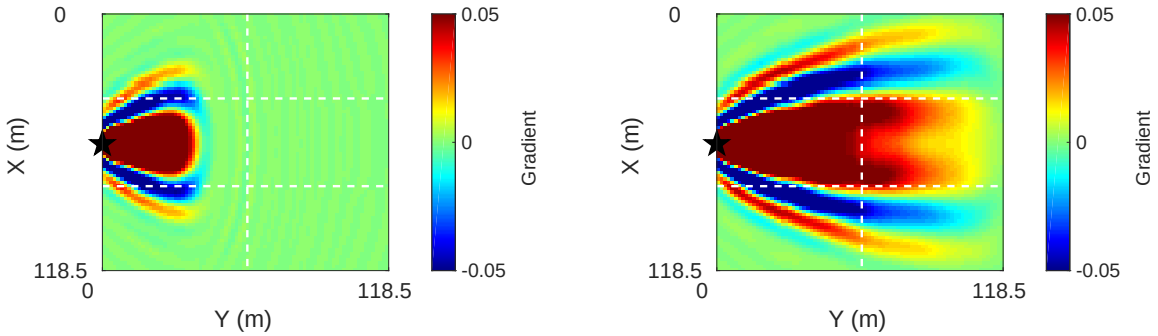
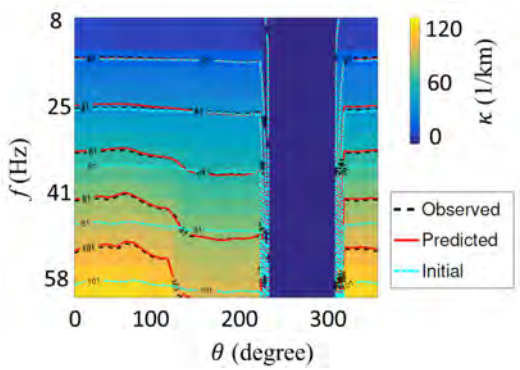
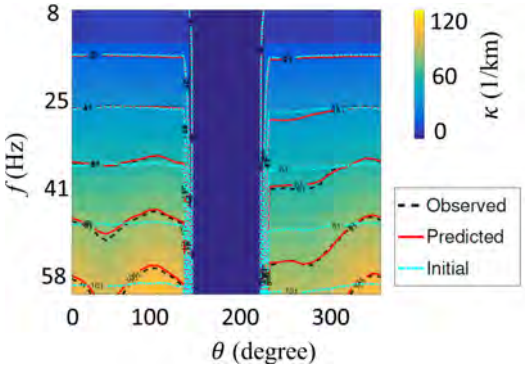


Figure 8: Slices of the gradient at $z = 6$ m for (a) $r_1 = 40$ m and (b) $r_1 = 120$ m.

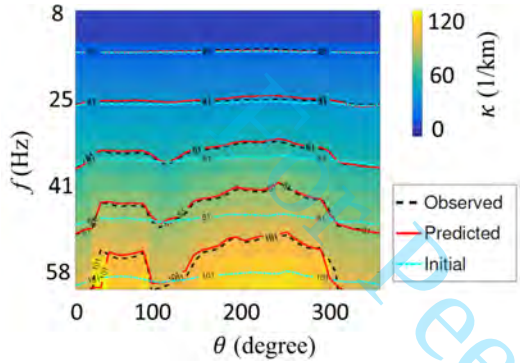
(a) **Dispersion Curves for A**



(b) **Dispersion Curves for B**



(c) **Dispersion Curves for C**



(d) **Dispersion Curves for D**

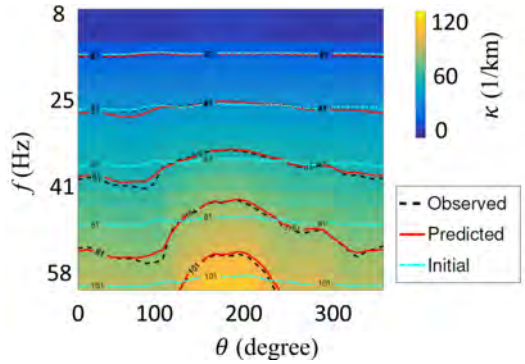
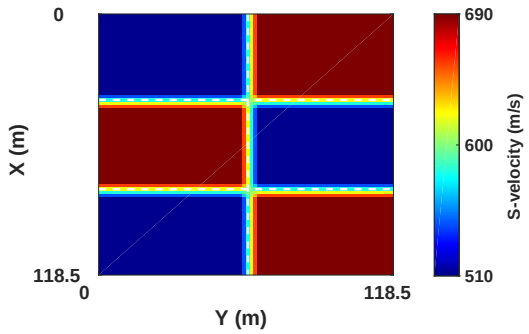
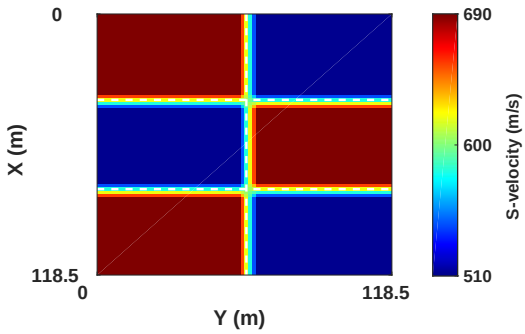


Figure 9: The observed dispersion curves for sources (a) A, (b) B, (c) C and (d) D marked in Figure 6b, where the black dashed lines, the cyan dash-dot lines and the red lines represent the contours of the observed, initial and inverted dispersion curves, respectively.

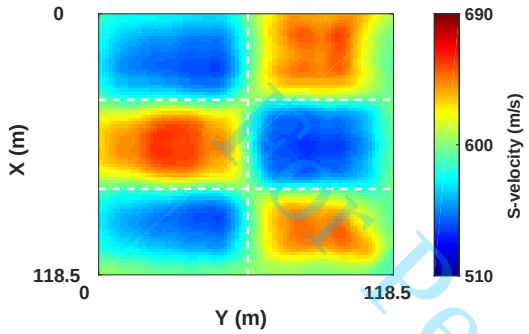
(a) **True Depth Slice at** $z = 15$ **m**



(b) **True Depth Slice at** $z = 24$ **m**



(c) **Inverted Depth Slice at** $z = 15$ **m**



(d) **Inverted Depth Slice at** $z = 24$ **m**

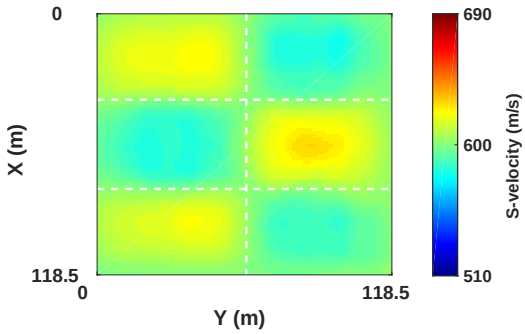
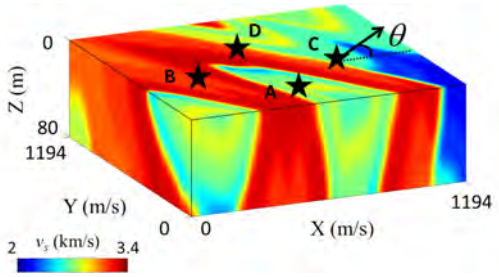
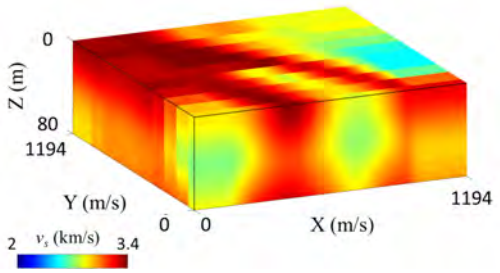


Figure 10: True S-velocity depth slices at (a) $z = 15$ m and (b) $z = 24$ m; inverted S-velocity depth slices at (c) $z = 15$ m and (d) $z = 24$ m.

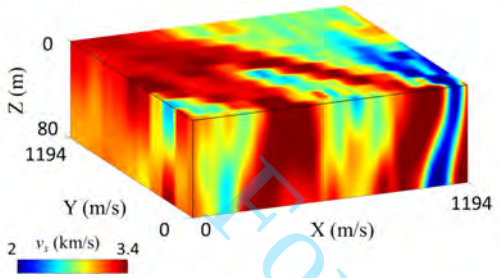
(a) **True Model**



(b) **Inverted Model by 1D Method**



(c) **Inverted Model by 2D WD**



(d) **Inverted Model by 3D WD**

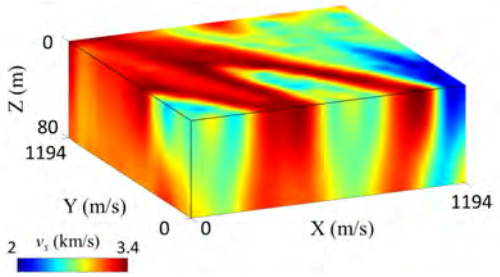
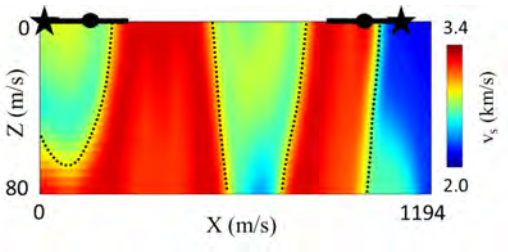
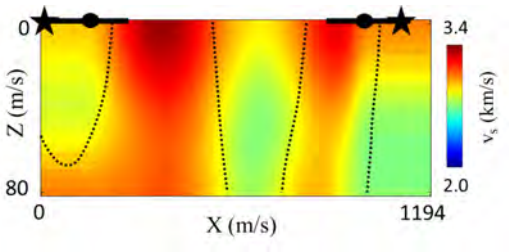


Figure 11: (a) True S-velocity model, and inverted models by the (b) 1D inversion, (c) 2D WD, and (d) 3D WD methods.

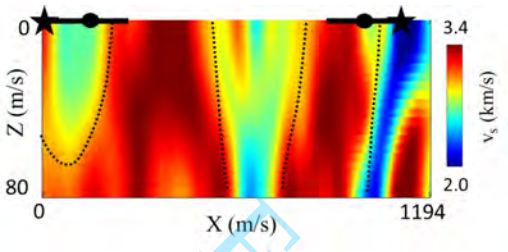
(a) **Slice of True Model at** $y = 120$ **m**



(b) **Inverted 1D Model at** $y = 120$ **m**



(c) **Inverted 2D Model at** $y = 120$ **m**



(d) **Inverted 3D Model at** $y = 120$ **m**

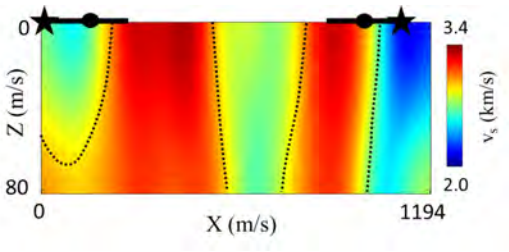


Figure 12: Slices of the (a) true, (b) 1D inversion, (c) 2D WD and (d) 3D WD S-velocity models at $y = 120$ m, where the black dashed lines indicate the large velocity contrast boundaries.

(a) **Phase Velocity Comparison**

(b) **Initial and Inverted S-velocity Profiles**

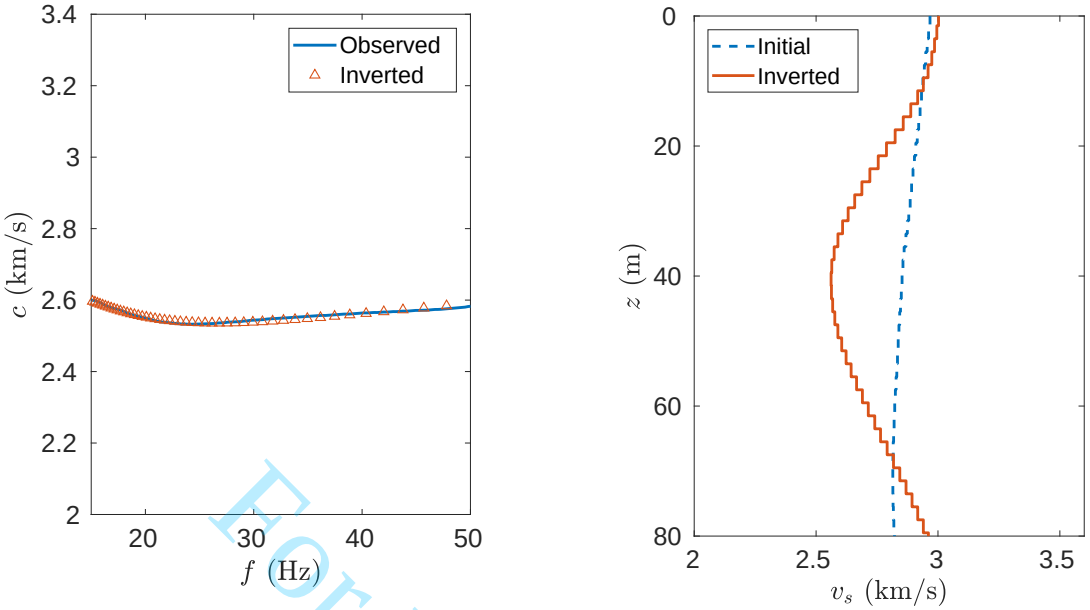


Figure 13: 1D dispersion curve inversion results by SURF96 (Herrmann, 2013): (a) the observed (blue line) and the predicted (red triangles) dispersion curves for CSG No. 30; (b) the initial (blue dashed line) and the inverted (red solid line) S-velocity profiles.

(a) **Dispersion Curves,** $\theta = 0°$ (b) **Dispersion Curves,** $\theta = 180°$

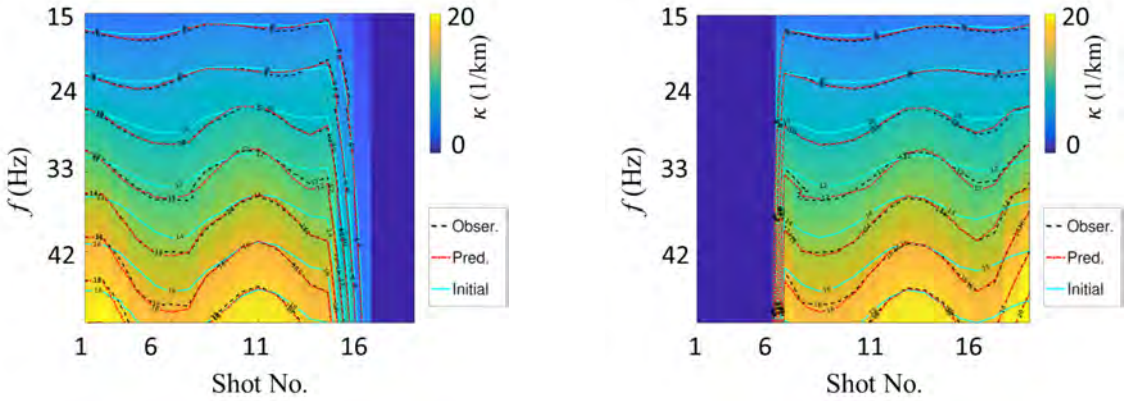
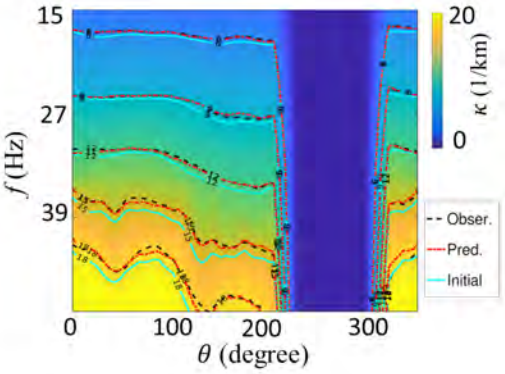
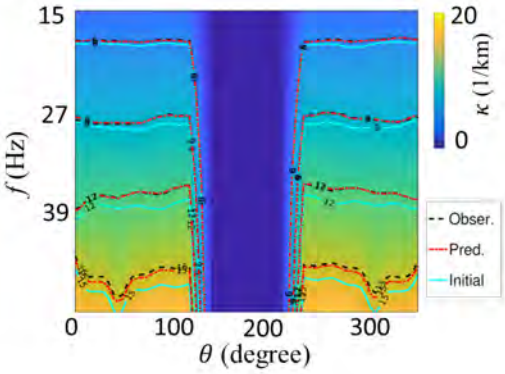


Figure 14: Observed dispersion curves along the azimuth angles of (a) $\theta = 0°$ and (b) $\theta = 180°$ for all the 2D CSGs located at $y = 120$ m, where the black dashed lines, the cyan lines and the red dash-dot lines represent the contours of the observed, initial and inverted dispersion curves, respectively.

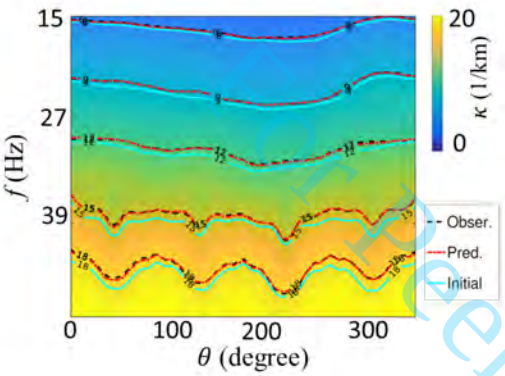
(a) **Dispersion Curves for Source A**



(b) **Dispersion Curves for Source B**



(c) **Dispersion Curves for Source C**



(d) **Dispersion Curves for Source D**

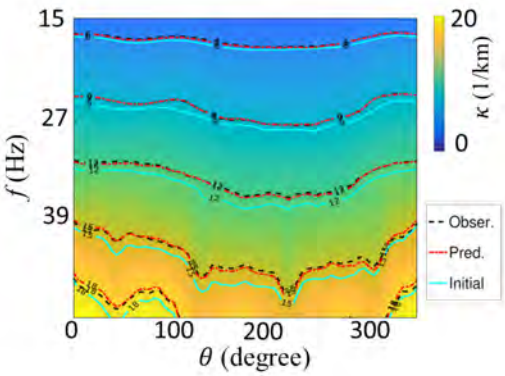
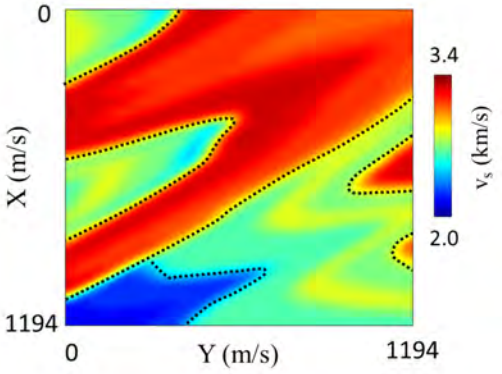
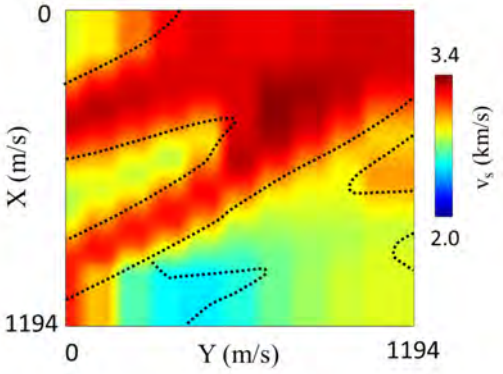


Figure 15: Observed dispersion curves for sources (a) A, (b) B, (c) C and (d) D as indicated in Figure 11a, where the black dashed lines, the cyan lines and the red dash-dot lines represent the contours of the observed, initial and inverted dispersion curves, respectively.

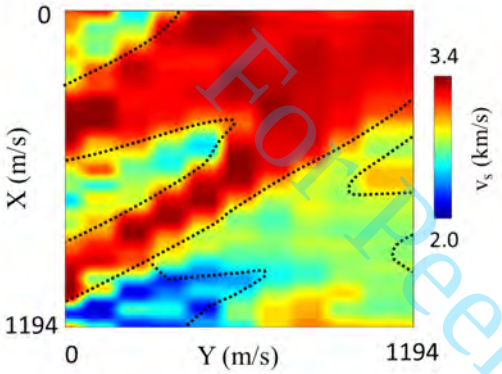
(a) **True S-velocity Model**



(b) **S-velocity 1D Tomogram**



(c) **S-velocity 2D Tomogram**



(d) **S-velocity 3D Tomogram**

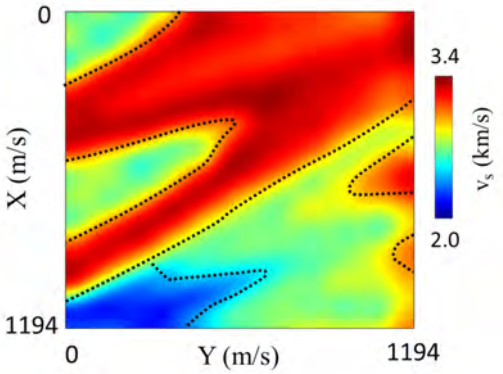


Figure 16: Depth slices at $z = 20$ m of (a) the true S-velocity model and the inverted tomograms computed by the (b) 1D inversion, (c) 2D WD and (d) 3D WD methods, where the black dashed lines indicate the large velocity contrast boundaries.

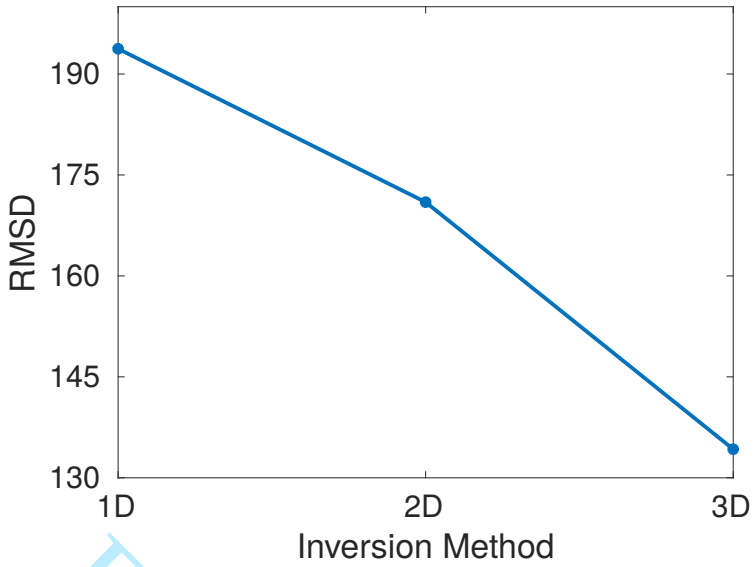
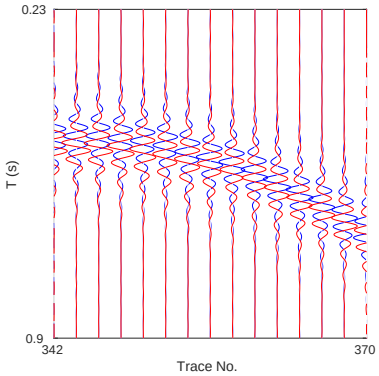
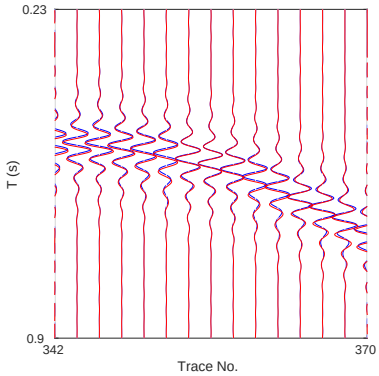


Figure 17: Root-mean-square deviation (RMSD) between the inverted S-velocity models by the 1D inversion, 2D WD and 3D WD methods and the true S-velocity model.

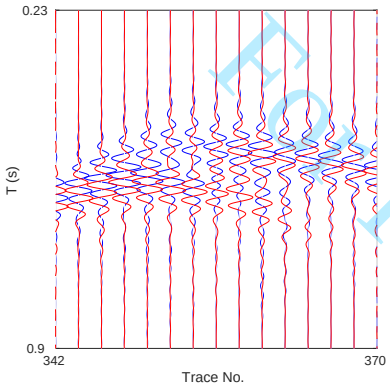
(a) **CSG No. 1**



(b) **CSG No. 1**



(c) **CSG No. 15**



(d) **CSG No. 15**

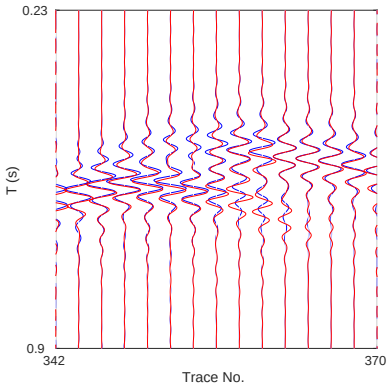


Figure 18: Comparison between the observed (red) and synthetic (blue) traces at far offsets predicted from the initial model (LHS panels) and 3D tomogram (RHS panels) for CSG No.1 in (a) and (b), and CSG No.15 in (c) and (d).

(a) **Initial S-velocity Model** (b) **S-velocity Tomogram**

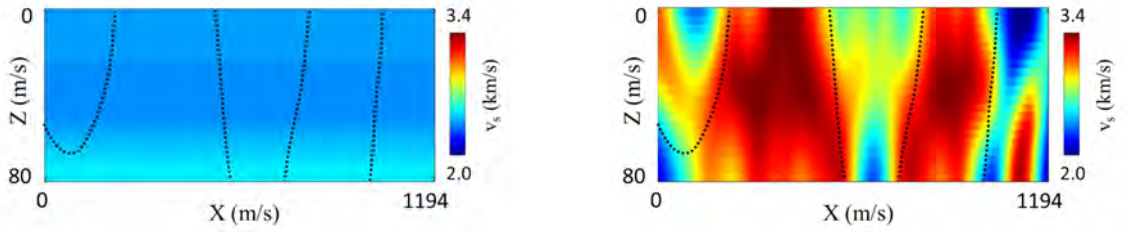


Figure 19: (a) Initial and (b) inverted 2D S-velocity models. The corresponding true model is shown in Figure 12a. Here, the black dashed lines indicate the large velocity contrast boundaries which are the same as those in Figure 12.

(a) **CSG No. 1** (b) **CSG No. 1**

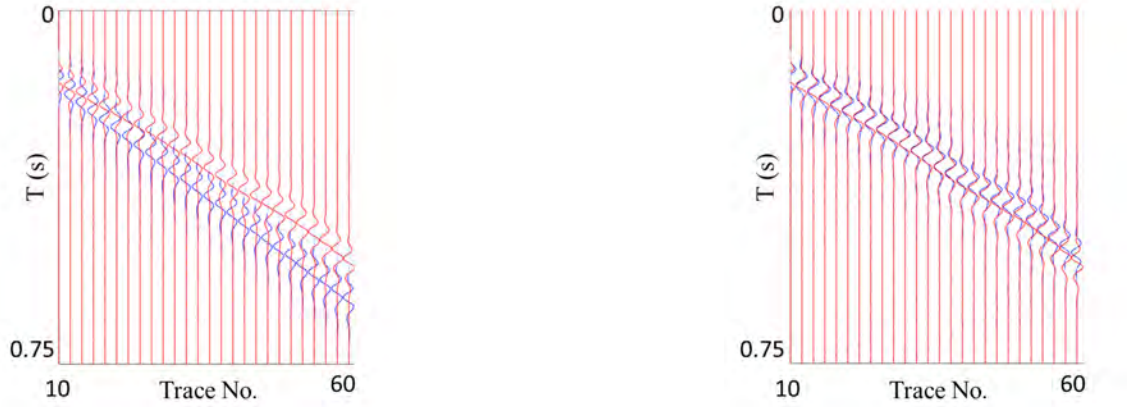


Figure 20: Comparison between the observed (red) and synthetic (blue) traces predicted from the initial model (LHS panels, Figure 19a) and 2D tomogram (RHS panels, Figure 19b) for CSG No.1.

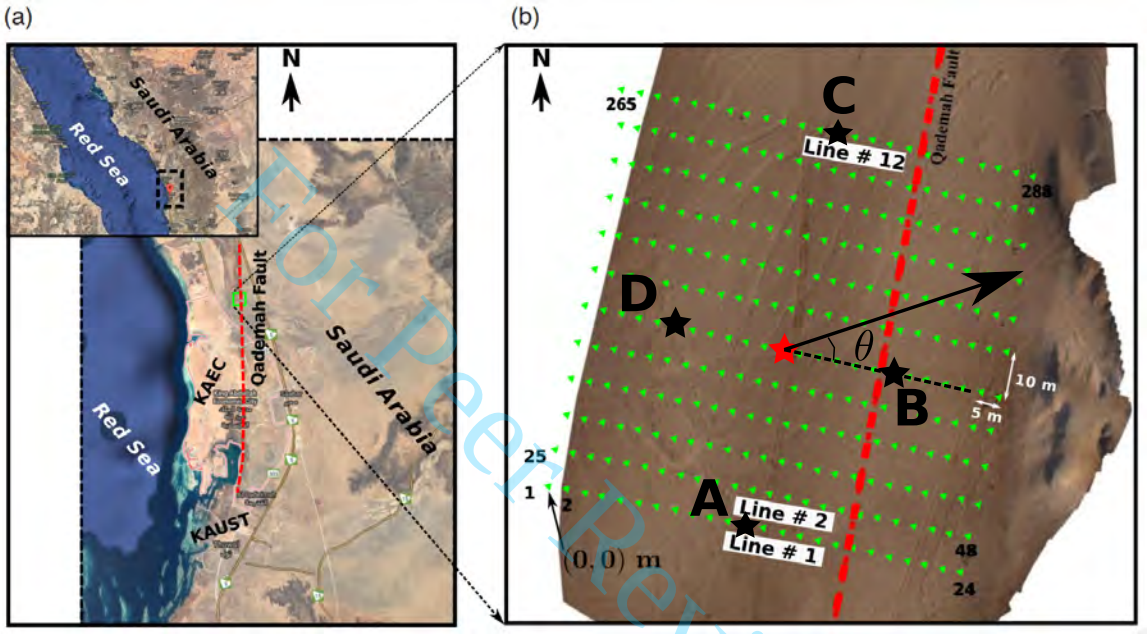
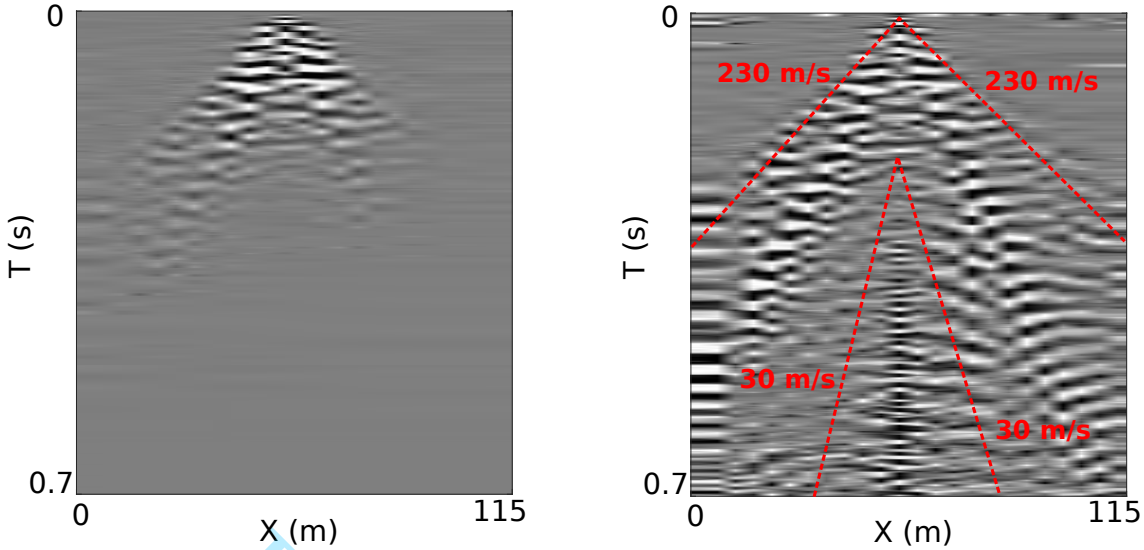


Figure 21: (a) Google map showing the location of the Qademah-fault seismic experiment (Fu et al., 2018b). (b) Receiver geometry for the Qademah-fault data. The Green triangles represent the locations of receivers, where the shots are located at each receiver. The red star represents the location of source No. 132 and the black stars indicate the locations of sources A, B, C and D on the surface. θ is the azimuth angle with respect to the acquisition line of source No. 132.

(a) **Traces of CSG No. 12 at the 1st Line** (b) **After Amplitude Compensation**



(c) **Dispersion Image for $\theta = 0°$** (d) **Dispersion Image for $\theta = 180°$**

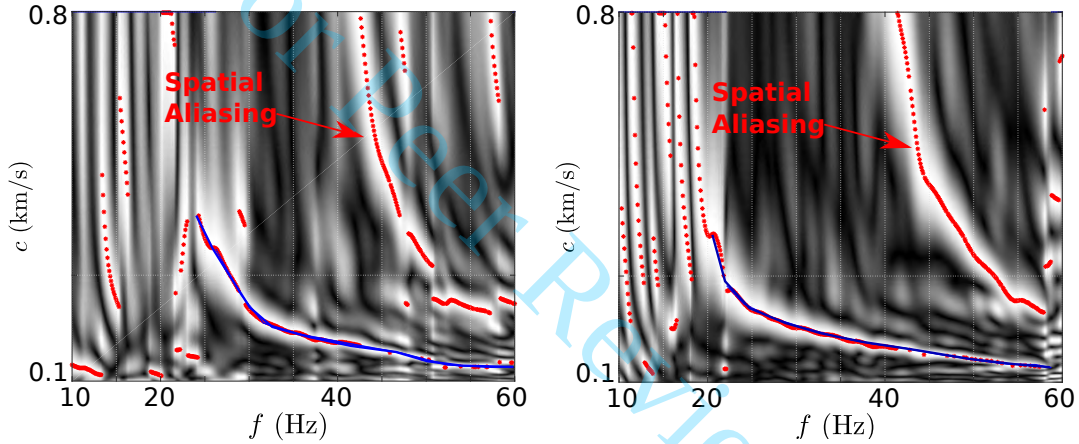


Figure 22: Seismic traces of CSG No. 12 at the first line (a) before and (b) after amplitude compensation; and its dispersion images for (c) $\theta = 0°$ and (d) $\theta = 180°$.

(a) **Picked Dispersion Curve for** $\theta = 0°$ (b) **Picked Dispersion Curve for** $\theta = 180°$

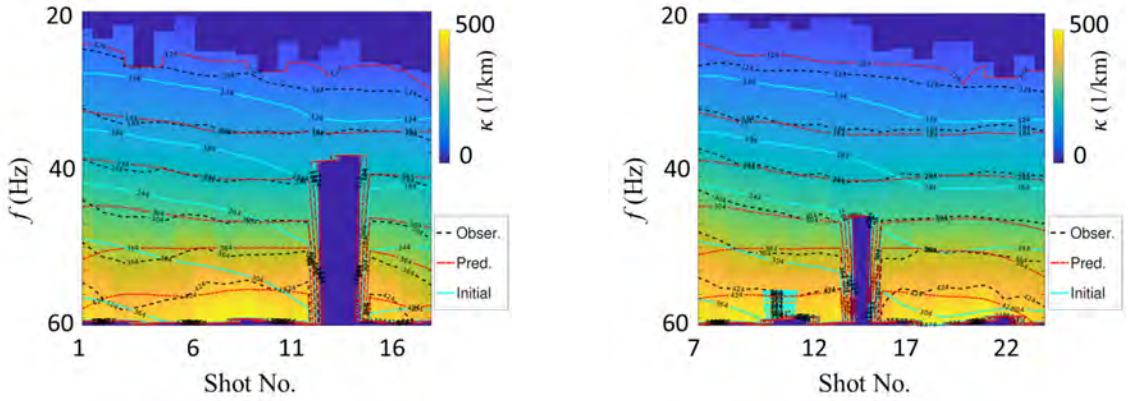


Figure 23: Observed dispersion curves for (a) $\theta = 0°$ and (b) $\theta = 180°$ from all the 2D CSGs in the first line, where the black dashed lines, the cyan lines and the red dash-dot lines represent the contours of the observed, initial and inverted dispersion curves, respectively.

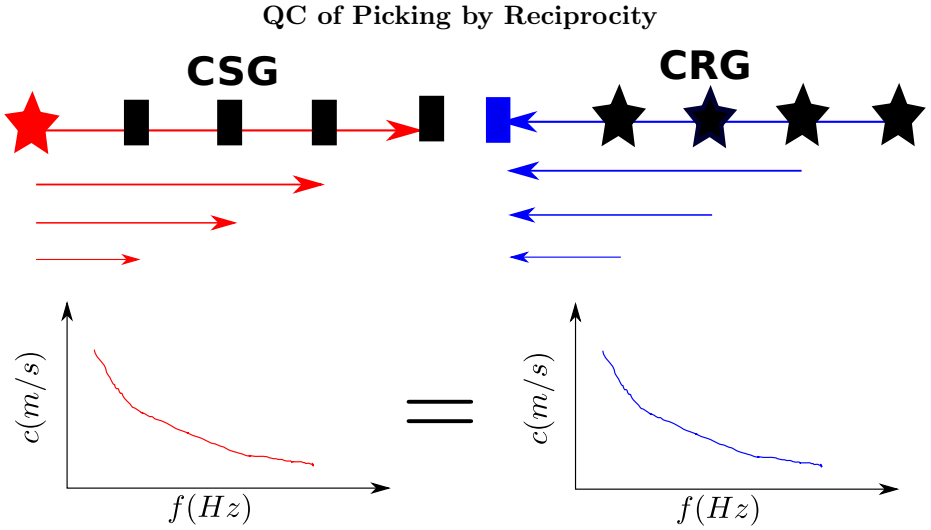


Figure 24: Quality control of the picked dispersion curves by reciprocity, where the stars represent the sources, and the rectangles represent the receivers. If the dispersion curves (red) of the CSG are the same as those (blue) computed from the CRG at the same location, it passes the reciprocity test.

(a) **Phase Velocity Comparison**

(b) **Initial and Inverted S-velocity Profiles**

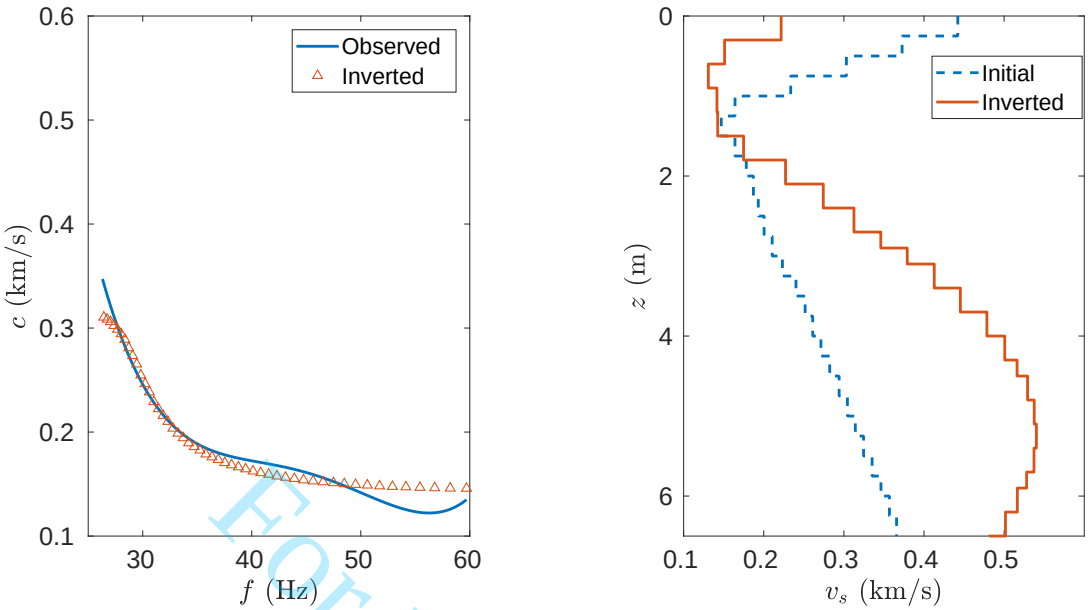


Figure 25: 1D dispersion curve inversion results by SURF96 (Herrmann, 2013): (a) the observed (blue line) and the predicted (red triangles) dispersion curves for CSG No. 12 (see Figure 22c); (b) the initial (blue dashed line) and the inverted (red solid line) S-velocity profiles.

(a) **S-velocity Tomogram by 1D Method** (b) **S-velocity Tomogram by 2D WD**

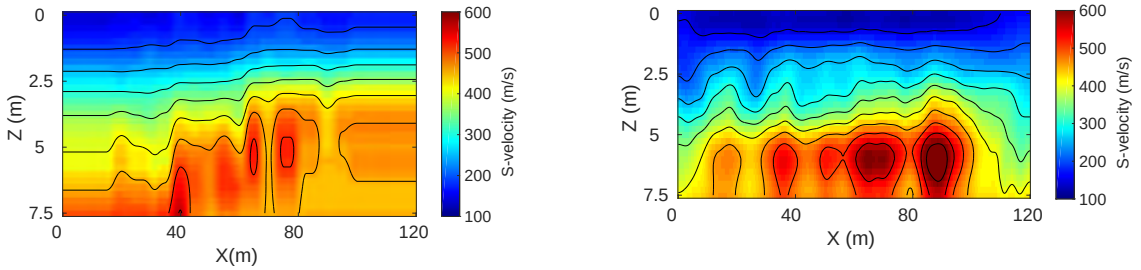
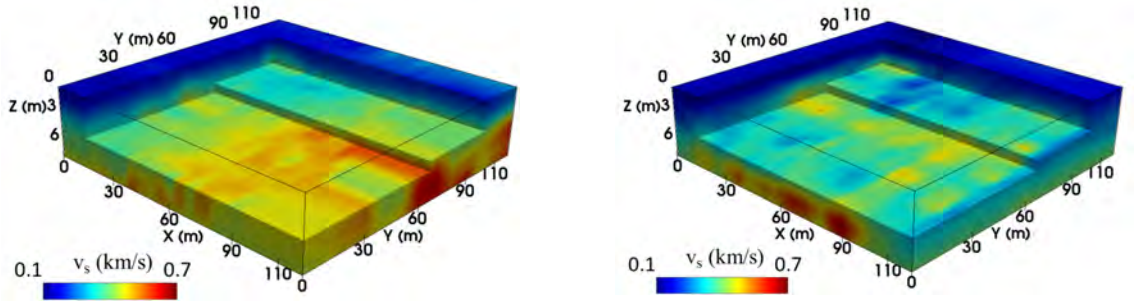


Figure 26: S-velocity tomograms from the 2D CSGs beneath the first line by the (a) 1D inversion and (b) 2D WD methods.

(a) **S-velocity Tomogram by 1D Method**

(b) **S-velocity Tomogram by 2D WD**



(c) **S-velocity Tomogram by 3D WD**

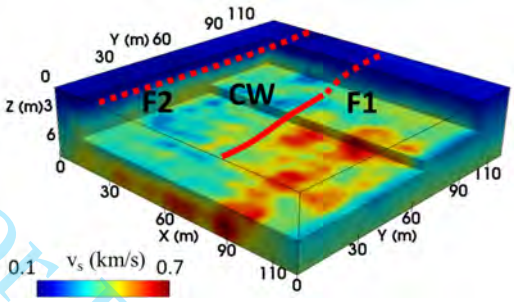


Figure 27: S-velocity tomograms inverted by the (a) 1D inversion, (b) 2D WD, and (c) 3D WD methods. The red solid line labeled by "F1" indicates the location of the conjectured Qademah fault and the dashed red line labeled by "F2" is conjectured to be a small antithetic fault. The low-velocity anomaly between faults "F1" and "F2" is the conjectured colluvial wedge labeled by "CW".

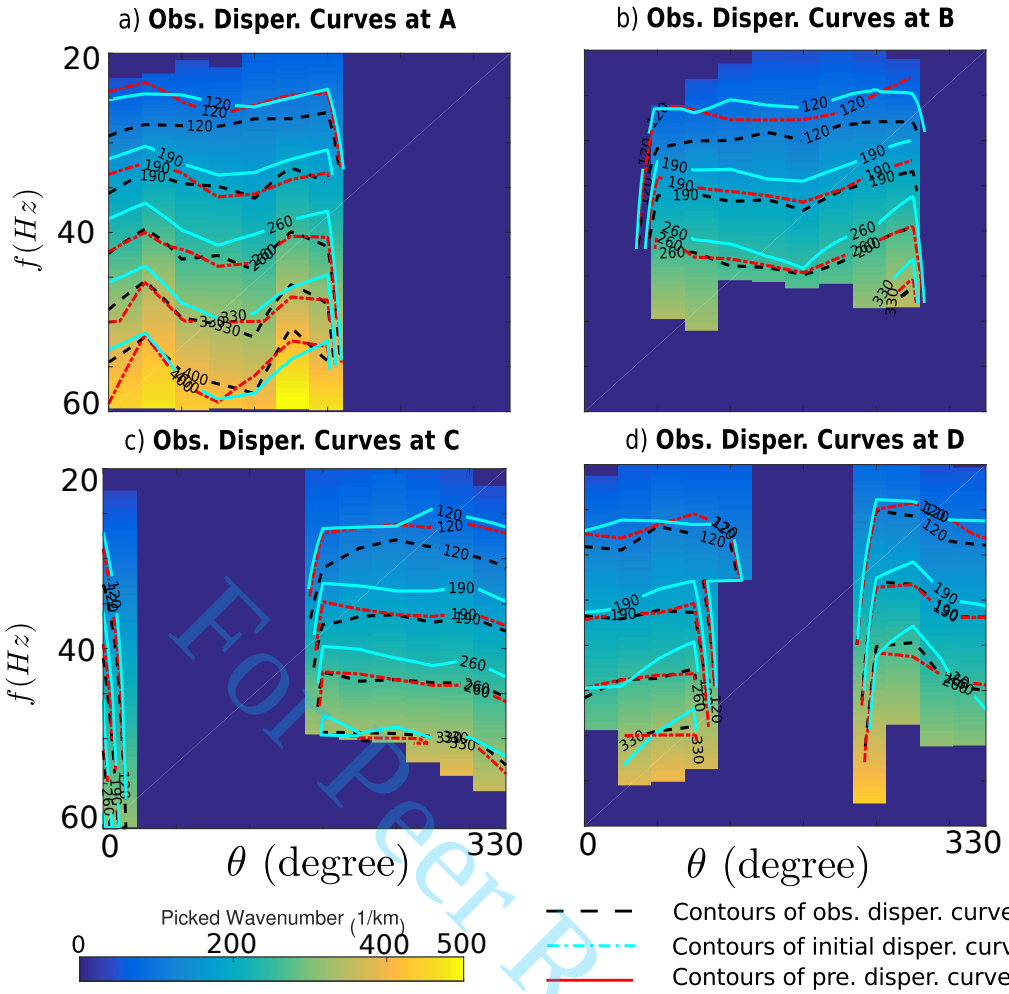
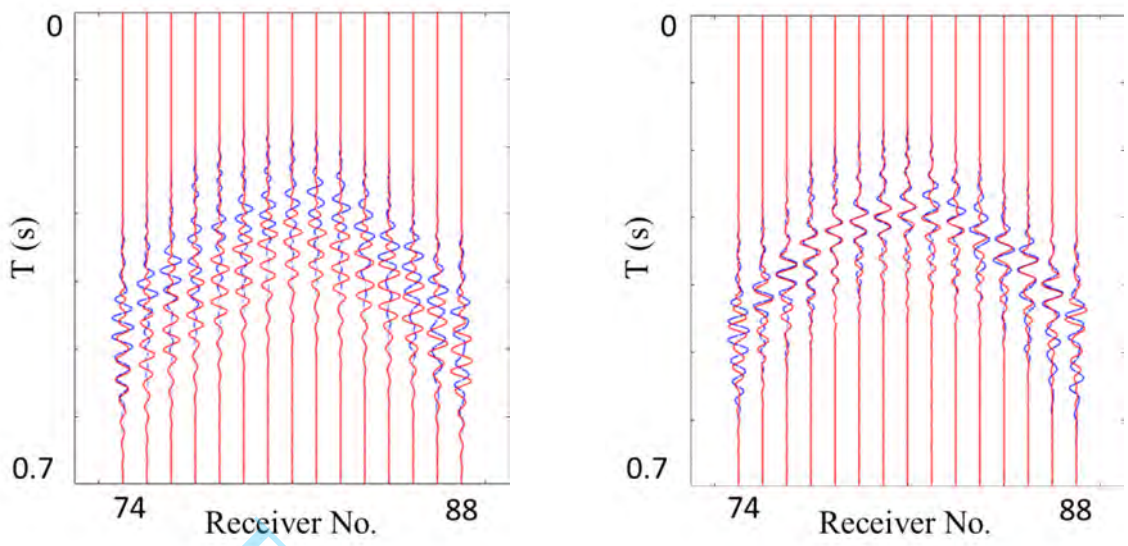


Figure 28: Observed dispersion curves for sources (a) A, (b) B, (c) C and (d) D indicated in Figure 21. The black dashed lines, the cyan lines and the red dash-dot lines represent the contours of the observed, initial and inverted dispersion curves, respectively.

(a) **CSG No. 9** (b) **CSG No. 9**



(c) **CSG No. 105** (d) **CSG No. 105**

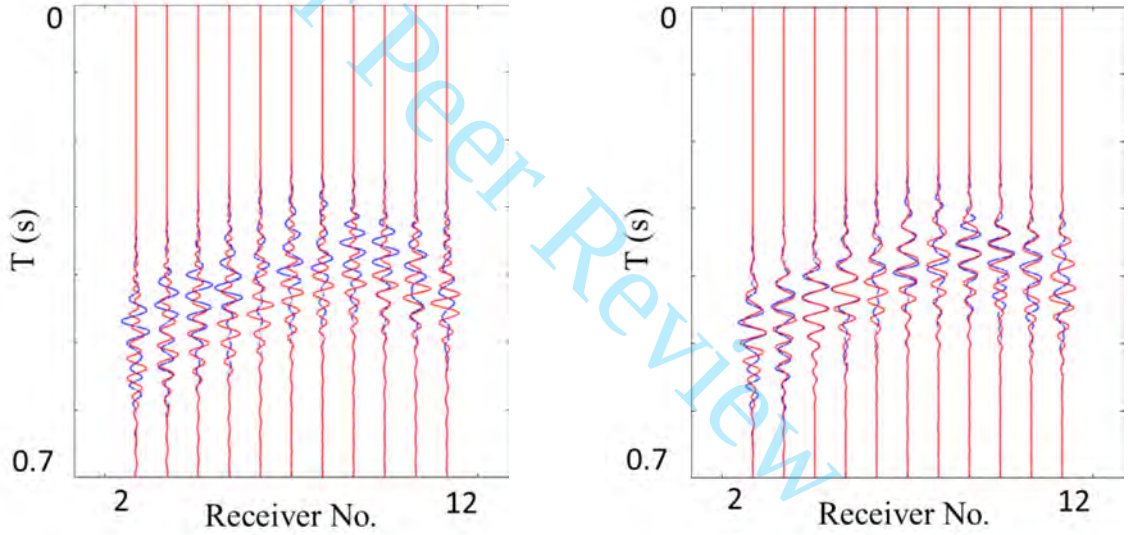
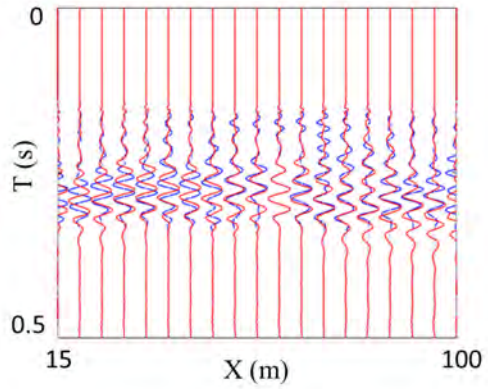
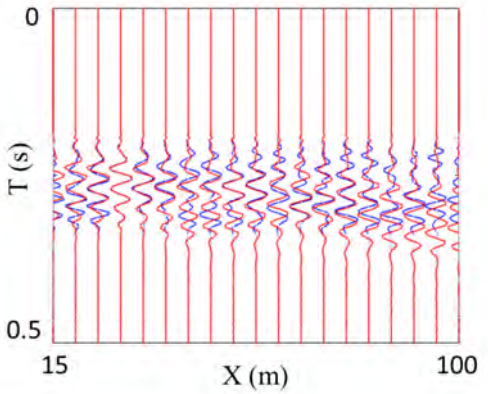


Figure 29: Comparison between the observed (blue) and synthetic (red) traces at far source-receiver offsets predicted from the initial model (LHS panels) and 3D WD tomogram (RHS panels) for CSG No.9 in (a) and (b), and CSG No.105 in (c) and (d). Here, a matched filter is applied to reshape the waveform of the synthetic data.

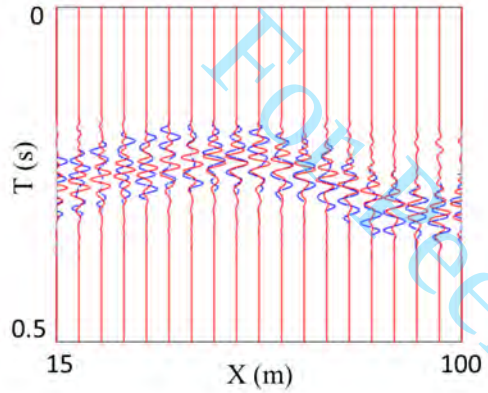
(a) **COG Line 1**



(b) **COG Line 4**



(c) **COG Line 6**



(d) **COG Line 10**

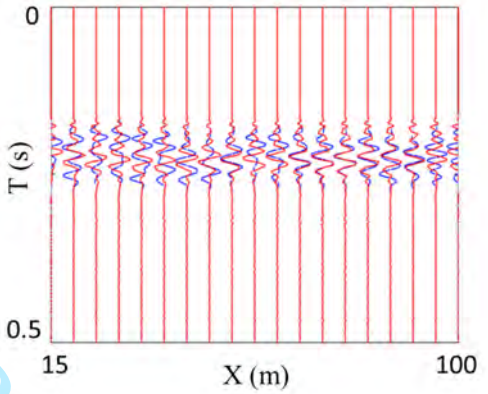


Figure 30: COGs with the offset of 30 m for the selected lines, where the blue and red wiggles represent the observed and predicted COGs, respectively.

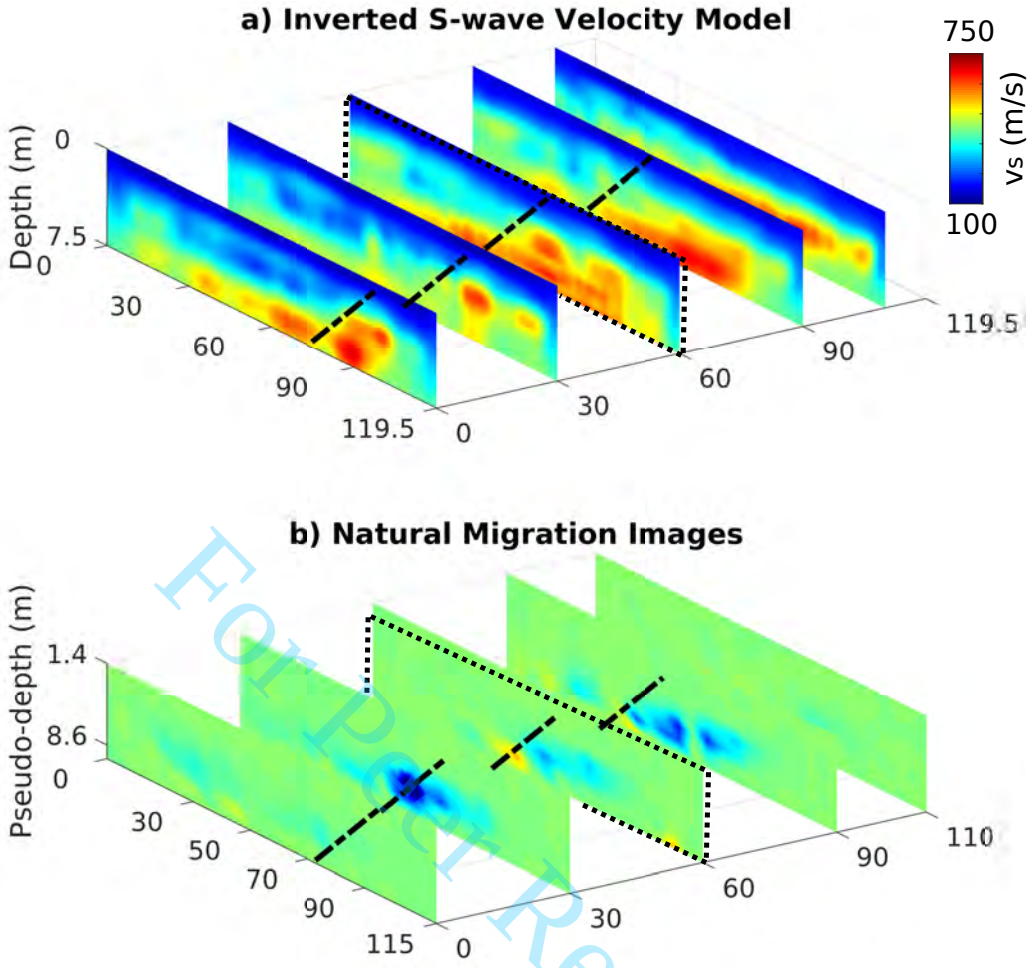


Figure 31: Slices of (a) the inverted S-wave velocity model, and (b) natural migration images (Liu et al., 2017). The dashed lines indicate the location of the interpreted Qademah fault.

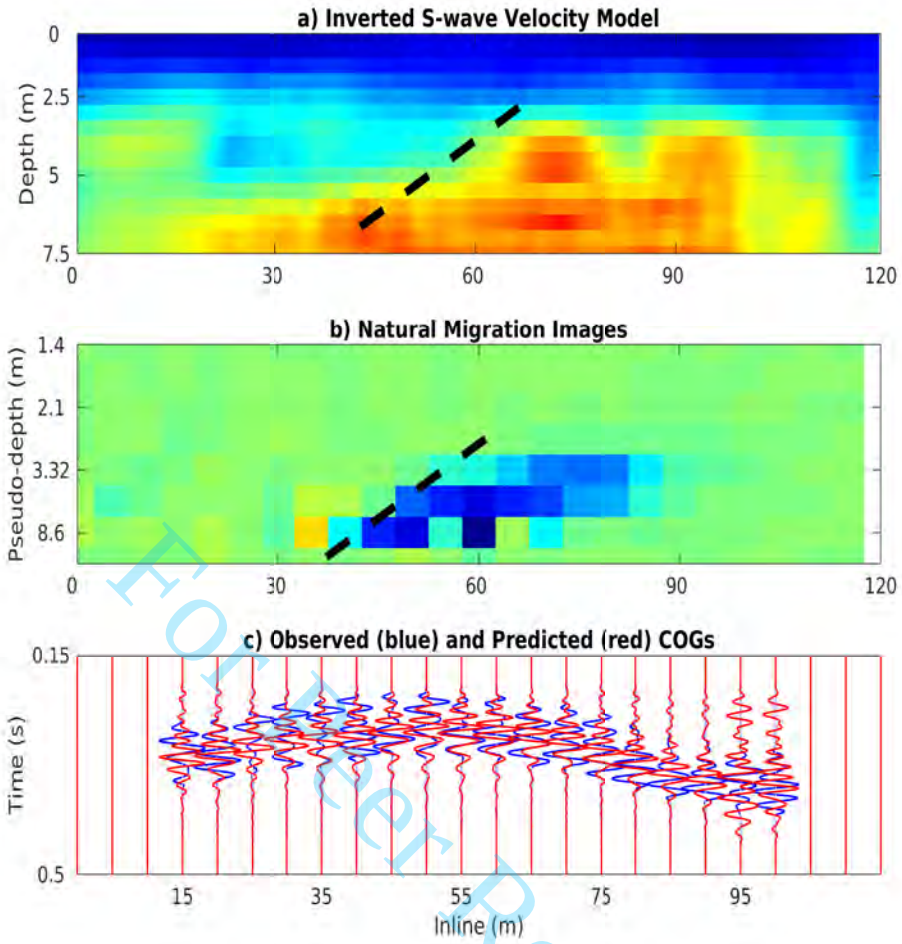


Figure 32: (a) and (b): 2D zoom view of the dashed panels in Figure 31, compared with (c) the COGs.