Geophysical Journal International

Multiscale and Layer-Stripping Wave-Equation Dispersion Inversion of Rayleigh Waves

Journal:	*Geophysical Journal International*
Manuscript ID	GJI-S-18-0820
Manuscript Type:	Research Paper
Date Submitted by the Author:	04-Sep-2018
Complete List of Authors:	Liu, Zhaolun Huang, Lianjie; Los Alamos National Laboratory,
Keywords:	Surface waves and free oscillations < SEISMOLOGY, Inverse theory < GEOPHYSICAL METHODS, Waveform inversion < GEOPHYSICAL METHODS

submitted to *Geophys. J. Int.*

Multiscale and Layer-Stripping Wave-Equation Dispersion Inversion of Rayleigh Waves

Zhaolun Liu[1⋆] and Lianjie Huang[1†]

[1] *Los Alamos National Laboratory, Geophysics Group, MS D452, Los Aamos, NM 87545, USA*

SUMMARY

The iterative wave-equation dispersion inversion can suffer from the local minimum problem when inverting seismic data from complex Earth models. We develop a multiscale, layer-stripping method to alleviate the local minimum problem of wave-equation dispersion inversion of Rayleigh waves and improve the inversion robustness. We first invert the high-frequency and near-offset data for the shallow S-velocity model, and gradually incorporate the lower-frequency components of data with longer offsets to reconstruct the deeper regions of the model. We use a synthetic model to illustrate the local minima problem of wave-equation dispersion inversion and how our multiscale and layer-stripping wave-equation dispersion inversion method can mitigate the problem. We demonstrate the efficacy of our new method using field Rayleigh-wave data.

Key words: Layer-stripping, multiscale, surface wave, wave-equation dispersion inversion.

⋆ Corresponding Author: zhaolun.liu@kaust.edu.sa (Z. Liu)
† Corresponding Author: ljh@lanl.gov (L. Huang)

Introduction

Wave-equation dispersion inversion (WD) of Rayleigh waves uses solutions to the 2D or 3D elastic-wave equation to invert the dispersion curves of surface waves for the S-velocity model (Li & Schuster 2016; Li et al. 2016, 2017a,b,c; Liu et al. 2017, 2018a). The advantage of WD over the conventional dispersion inversion method (Haskell 1953; Xia et al. 1999, 2002; Park et al. 1999) is that WD does not assume a layered velocity model and is valid when there are strong lateral gradients in the S-velocity model. The WD method also enjoys robust convergence because the skeletonized data, namely the dispersion curves, are simpler than a trace with many dispersive arrivals. Such traces are used in full waveform inversion (FWI) (Groos et al. 2014; Pérez Solano et al. 2014; Dou & Ajo-Franklin 2014; Groos et al. 2017).

The iterative WD method can suffer from the local minimum problem when inverting seismic data from complex Earth models. One method to tackle this problem is the multiscale method (Masoni et al. 2016). For Body waves, the low-to-high frequency content of data is first used to update the large-scale velocity structure and then the more detailed features of the velocity model are reconstructed (Sirgue & Pratt 2004; Bunks et al. 1995). However, a high-to-low frequency strategy for surface waves is needed because the frequency content of surface waves is directly related to their penetration depth: higher-frequency and shorter-wavelength surface waves sample the top layers of a medium, while lower-frequency and longer-wavelength surface waves sample deeper subsurface regions (Masoni et al. 2016).

The WD method needs to determine the offset range (denoted as R) starting from the near offset for retrieving the dispersion curves of the data using F-K or Radon transforms. A narrow range of offsets corresponding to a small R is not adequate for accurate retrieval of the low-frequency component of dispersion curves (indicated in Figs. 2.3-8 of Yilmaz (2015)), but can provide high lateral resolution in the tomographic image. Conversely, a wide range of offsets is adequate for accurately retrieving the low-frequency dispersion curves but the penalty is that it only provides a low-wavenumber estimate of the velocity model. As a rule of thumb, we choose R to be about three or four times greater than the depth of interest to make sure that WD has enough penetration depth and lateral resolution (Liu et al. 2018a). However, a fixed value of R would result in a loss of either the low-frequency information in the dispersion curves or the lateral resolution of the inverted S-velocity model. Thus, an iterative small-to-large offset range strategy is needed to obtain both high lateral resolution and low-frequency information.

In this paper, we first use the high-frequency surface-wave data with a small-offset range to update the shallow velocity model, and then use the low-frequency surface-wave data with a large-offset range to update the deeper regions of the velocity model. We employ a layer-stripping method (Shi et al.

2015; Masoni et al. 2016) to reconstruct the velocity model from the shallow to deep regions. The layer-stripping method assumes that all layers above a given layer have been inverted using the near-offset and high-frequency surface-wave data. We use the far-offset and low-frequency data to invert for the velocity model of the deep layers. This procedure is repeated until the entire volume of interest is reconstructed.

After the introduction, we describe the theory of WD and the workflow for the layer-stripping approach. Numerical tests on synthetic and field surface-wave data are presented in the third section to demonstrate the improvement of the method, followed by the discussion and conclusions.

Theory

The wave-equation dispersion inversion method inverts for the S-wave velocity model to minimize the dispersion objective function

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

where $\kappa(\omega,\theta)_{pre}$ 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. In the 2D case, the azimuth angles have only two values: $0°$ and $180°$, corresponding to the left and right directions, respectively.

The gradient $\gamma(\boldsymbol{x})$ of ε with respect to the S-wave velocity $v_s(\boldsymbol{x})$ is given by Liu et al. (2018a,b):

$$\gamma(\boldsymbol{x}) = \frac{\partial\varepsilon}{\partial v_s(\boldsymbol{x})} = -\sum_{\omega} 4v_{s0}(\boldsymbol{x})\rho_0(\boldsymbol{x})\Re\Bigg\{$$

$$\overbrace{\int\sum_{\theta}\frac{1}{A(\theta,\omega)}\Delta\kappa(\theta,\omega)\hat{D}(\boldsymbol{g},\theta,\omega)^*_{obs}G_{3k,k}(\boldsymbol{g}|\boldsymbol{x})d\boldsymbol{g}}^{backprojected\ data=B_{k,k}(\boldsymbol{x},\omega)^*}\ \overbrace{D_{j,j}(\boldsymbol{x},\omega)}^{source=f_{j,j}(\boldsymbol{x},\omega)}$$

$$\overbrace{-\frac{1}{2}\int\sum_{\theta}\frac{1}{A(\theta,\omega)}\Delta\kappa(\theta,\omega)\hat{D}(\boldsymbol{g},\theta,\omega)^*_{obs}G_{3n,k}(\boldsymbol{g}|\boldsymbol{x})d\boldsymbol{g}}^{backprojected\ data=B_{n,k}(\boldsymbol{x},\omega)^*}\overbrace{[D_{k,n}(\boldsymbol{x},\omega)+D_{n,k}(\boldsymbol{x},\omega)]}^{source=f_{n,k}(\boldsymbol{x},\omega)}\Bigg\}, \tag{2}$$

where $v_{s0}(\boldsymbol{x})$ and $\rho_0(\boldsymbol{x})$ are the reference S-velocity and density at location $\boldsymbol{x}$, respectively. $A(\theta,\omega)$ is given in Liu et al. (2018b). $D_i(\boldsymbol{x},\omega)$ denotes the i^{th} component of the particle velocity recorded at $\boldsymbol{x}$ resulting from a vertical-component force. The Einstein notation is assumed in equation (2), 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 resulting from a vertical-component source at $\boldsymbol{x}$ in the reference medium. Term $f_{i,j}(\boldsymbol{x},\omega)$ for i and $j \in \{1,2,3\}$ is the downgoing source field at $\boldsymbol{x}$, and $B_{i,j}(\boldsymbol{x},\mathrm{s},\omega)$ for i and $j \in \{1,2,3\}$ is the backprojected scattered field at $\boldsymbol{x}$. $\hat{D}(\boldsymbol{g},\theta,\omega)^*_{obs}$ is the weighted conjugated data function defined as

$$\hat{D}(\boldsymbol{g},\theta,\omega)^*_{obs} = 2\pi i\boldsymbol{g}\cdot\boldsymbol{n}e^{i\boldsymbol{g}\cdot\boldsymbol{n}\Delta\kappa}\int_C D(\boldsymbol{g}',\omega)^*_{obs}d\boldsymbol{g}', \tag{3}$$

where $\boldsymbol{n} = (\cos\theta, \sin\theta)$ and C is the line $(\boldsymbol{g}' - \boldsymbol{g})\cdot\boldsymbol{n} = 0$. The above equation indicates that the gradient is computed using a weighted zero-lag correlation between the source and backward-extrapolated receiver wavefields.

The optimal S-wave velocity model $v_s(\boldsymbol{x})$ is obtained using the steepest-descent formula (Nocedal & Wright 2006)

$$v_s(\boldsymbol{x})^{(k+1)} = v_s(\boldsymbol{x})^{(k)} - \alpha\gamma(\boldsymbol{x}), \tag{4}$$

where α is the step length and the superscript (k) denotes the k^{th} iteration. We use a preconditioned conjugate gradient method to update the S-wave velocity model.

In our multiscale, layer-stripping WD (MSLSWD), we use the high-frequency data with a small offset R to first update the shallow velocity model. Then we assume this shallow velocity model is known and use the low-frequency data with a large offset R to update the deeper regions of the velocity model. For a given frequency band, according to the dispersion curves, we estimate an average wavelength

$$\lambda = 1/\kappa,$$

where κ is the average wavenumber. The penetration depth z is estimated as half of the wavelength λ, and the maximum offset R is estimated as three or four wavelengths.

Workflow

The workflow for implementing our multiscale, layer-stripping WD method is summarized in the following six steps.

1. Determine the frequency range of observed data. Divide the frequency range into several frequency bands for each MSLSWD step.
2. Retrieve the dispersion curves from the whole common-shot gather (CSG) and estimate the range of the average wavenumber k for each frequency band.
3. For a given frequency band, determine the maximum offset R according to the maximum wavelength λ calculated from the range of k values and estimate the observed dispersion curves from seismic traces within the maximum offset R for each CSG.
4. Calculate the gradient according to equation (2). Only the region within a depth window is used to update the S-velocity model. The depth window can be estimated from half of the wavelength range.

5. Use the updated S-velocity model as the initial model to perform WD for the next frequency band.
6. Repeat the last three steps for all frequency bands.

Numerical Results

We first use a synthetic model displayed in Fig. 1a to verify the effectiveness of MSLSWD. We modify the model from Pérez Solano et al. (2014) (Fig. 6d) that was also used by Masoni et al. (2016) (Fig. 4). We generate the observed and predicted data using an O(2,8) time-space-domain solution to the first-order 2D elastic-wave equation with a free-surface boundary condition (Graves 1996). MSLSWD inverts only the S-wave velocity model. We use the actual P-wave velocity model for modeling predicted surface waves. The source wavelet is a Ricker wavelet with a center frequency of 40 Hz, which is assumed to be known during inversion. The P-wave velocity v_p is calculated from the S-wave velocity v_s using the relation $v_p = 2v_s$, with a homogeneous density model of 1000 kg/m^3. We use the fundamental dispersion curves from each common-shot gather (CSG) for inversion along the azimuth angles of $0°$ (toward the right-hand side (RHS)) and $180°$ (toward the left-hand side (LHS)).

Synthetic Model

We generate a total of 40 CSGs for vertical sources located at $z = 0.2$ m below the free surface with a spatial interval of 1.5 m. Each CSG has 150 vertical-component receivers at $z = 0.2$ m below the surface with a spatial interval of 0.2 m. The initial S-velocity model used is a model with a linear gradient in depth (Fig. 1b).

We first study the influence of the receiver spread length R on the penetration depth and lateral resolution in the WD results. We perform the first numerical test by setting R to be 8 m. The resulting inverted S-velocity model is shown in Fig. 1c. We can see that the two high-velocity anomalies are separated clearly. The observed dispersion curves for all the CSGs along the azimuth angles of $0°$ and $180°$ are shown in Figs. 2a and 2b respectively, where the black dashed lines, the magenta dash-dot lines and the red lines represent the contours of the observed, initial and inverted dispersion curves, respectively. The contours of the inverted dispersion curves correlate well with those of the observed data.

We carry out the second numerical test by increasing the offset R to 20 m. The resulting inverted S-velocity model is displayed in Fig. 1d. The contours of the observed, initial and inverted dispersion curves are shown in Figs. 2c and 2d. The contours of the inverted dispersion curves correlate well with the observed ones. However, the two high-velocity anomalies cannot be separated, indicating lower lateral resolution. Figs. 3a and 3b show the vertical-velocity profiles at $X = 20$ m and $X = 38$ m respectively for the true model (blue line), the initial model (black dash-dot line) and the inverted S-velocity models by setting R=8 m (magenta line) and $R = 20$ m (red line). It can be seen that even though the frequency content of the data used in WD inversion is the same, the inverted velocity tomogram with the maximum offset $R = 20$ m is more accurately reconstructed in the deeper region than that when $R = 8$ m. This suggests that more accurate dispersion curves for the low-frequency part can be retrieved by longer offsets.

With the same acquisition parameters as above, we conduct two additional numerical tests on the modified S-velocity model shown in Fig. 4a. We add a low-velocity zone in the shallow region and move the location of the high-velocity anomalies to a deeper depth. The initial model used in inversion is a linear velocity gradient displayed in Fig. 4b. We first apply single-scale WD to the data. Fig. 5 shows the contours of the observed, initial and predicted dispersion curves along the azimuth angles of $\theta = 0°$ and $180°$ for these two numerical tests. There is a poor mismatch between the contours of the inverted and observed dispersion curves, which indicates that the WD is stuck to a local minimum for the modified model. The inverted S-velocity tomograms with a maximum offset of $R = 8$ m and R=20 m are shown in Figs. 4c and 4d, respectively. The high-velocity anomalies are not detected in the inverted tomograms of these two numerical tests. Figs. 6a and 6b shows the vertical-velocity profiles at $X = 20$ m and $X = 38$ m, respectively, for the true (blue line), initial (black dash-dot line) and inverted S-velocity models when using R=8 m (magenta line) and R=20 m (red line). The vertical-velocity profiles show that WD incorrectly updates the low-velocity zones in the shallow region ($z < 1$ m). This suggests that WD converges to a local minimum because of the shallow low-velocity zone.

We apply our multiscale, layer-stripping WD method to the same data as above to alleviate the local minimum problem. The frequency spectrum of the data is shown in Fig. 7, where eleven frequency bands are chosen and each of them is plotted as the horizontal bar with a number tag close to it. We select larger frequency windows for high-frequency bands because of the small amplitude of the frequency spectrum of surface wave data. The ranges of the eleven frequency bands and the wavelength range corresponding to each frequency band used in MSLSWD are listed in Table 1. The maximum offset R is calculated using

$$R \approx 3.5 * \lambda_{max} \tag{5}$$

where λ_{max} is the maximum wavelength. The misfit-change column shows the normalized misfit change after 10 iterations. The updated depth window for each frequency band is determined using half of the wavelength. A taper is used at the top and bottom boundaries of a depth window as shown in Fig. 8. The updated velocity models for all eleven steps are shown in Figs. 9 and 10, where the black dashed lines indicate the location of the high-velocity anomalies. We can see that the deeper region of the model is gradually updated step by step. The contours of the observed, initial and inverted dispersion curves for each step are shown in Figs. 11 and 12. The contours of the inverted dispersion curves correlate well with the observed ones for each step. The vertical-velocity profiles at $X = 20$ m and $X = 38$ m extracted from the inverted tomogram with MSLSWD (red lines) are shown in Fig. 13. They show better agreement with the true ones (blue lines) than those extracted from the tomogram without layer stripping (magenta lines). The results demonstrate that MSLSWD can mitigate the local minimum problem of WD for this model caused by the low-velocity layer in the shallow region.

Surface Seismic Data from the Blue Mountain Geothermal Field

Seven 2D lines of surface seismic data were acquired at the Blue Mountain geothermal field in Nevada, USA, using dynamite sources. We use one 2D line of data for our study. The line consists of 121 receivers with an interval of 33.5 m, and 57 dynamite sources with an interval of 67 m. One of the CSGs is shown in Fig. 14a, which clearly shows three modes of surface waves. These three modes are also shown in the dispersion images (Fig. 14b) calculated using the frequency-sweeping method (Park et al. 1998) with the maximum offset R =500 m. We pick only the dispersion curves of the fundamental-mode surface waves (Fig. 14c). The observed dispersion curves for all CSGs are shown in Fig. 15, where the black dashed lines indicate the contours of the observed dispersion curves. The initial S-velocity model is shown in Fig. 16a, and the S-velocity tomogram obtained using the conventional WD is displayed in Fig. 16b. The contours of the initial and predicted dispersion curves for all CSGs are plotted in Fig. 15 using the cyan and red dash-dot solid lines, respectively. It shows that only the dispersion contours from the high-frequency components and the CSGs NO. 1-30 match well, which indicates that WD is stuck to a local minimum. To alleviate the local minima problem, we apply our multiscale and layer-stripping WD method to the data.

We use three frequency bands for MSLSWD: (a) 7-10 Hz, (b) 5-8 Hz and (c) 2-6 Hz. The corresponding depth windows are 0-45 m, 45-100 m and 100-250 m. The initial S-velocity model, gradient update and the inverted model for each step are shown in Figs. 17, 18 and 19. The comparison of the S-velocity tomograms without and with using the layer stripping approach are shown in Figs. 20a and 20b. It can been seen that the deeper regions are significantly updated using layer-stripping WD.

The predicted dispersion contours obtained with and without the layer stripping approach are displayed in Fig. 21, where the black dashed lines, the cyan lines and the red dash-dot lines represent the contours of the observed dispersion curves, the predicted dispersion curves without and with layer stripping, respectively. The dispersion contours calculated using layer stripping WD for the low-frequency components and the CSGs No.31-56 correlate better with the observed ones compared to those calculated using the conventional WD. The S-velocity tomogram is also consistent with the P-wave tomogram shown in Fig. 20c.

To further test the accuracy of the layer-stripping WD method. Fig.22 shows the comparison between the observed (red) and synthetic (blue) traces from the S-velocity tomogram without (LHS panels) and with (RHS panels) layer-stripping methods for CSG No. 20 in (a) and (b), and CSG No. 30 in (c) and (d). We use a matched filter to reshape the synthetic waveform. The predicted fundamental-mode surface waves closely match the observed ones. Figs.23 and 24 show the common offset gathers (COGs) with offsets of 335 m and 670 m from the S-velocity tomogram (a) without and (b) with layer-stripping methods, respectively. The blue and red wiggles represent the observed and predicted COGs, respectively. The synthetic COGs computed from the S-velocity tomogram inverted using the MSLSWD more closely agree with the observed ones compared with to those inverted with WD without layer stripping.

Discussion

The lateral resolution of the WD tomogram is related to the length of the receiver spread. Different receiver-spread lengths 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, but can provide a deep penetration depth (Foti et al. 2014). Our results of synthetic surface seismic data demonstrate that layer-stripping wave-equation dispersion inversion can provide better depth penetration and higher lateral resolution than the conventional wave-equation dispersion inversion without layer stripping.

In our field data example, the interval of the geophones is so large that only few geophones are involved when inverting the high-frequency-band data if we set the offset according to the maximum wavelength (equation 5). It is hard to pick the dispersion curves because of low signal-to-noise ratio when using only a few geophones. Thus, we use the same receiver-spread length (500 m) for all frequency bands, which is approximately three times the length of the wavelength of the lowest frequency data (3 Hz). It will decrease the lateral resolution of the shallow part of the tomogram.

We assume that the effects of attenuation and topography on dispersion curves are insignificant in our field data test. However, if the attenuation and topography is important, the effects can be accounted for by solving the visco-elastic wave equation with an irregular free surface to compute the theoretical dispersion curves and perform the inversion (Li et al. 2017a,b,c).

We use only the fundamental-mode Rayleigh waves for MSLSWD inversion. Nevertheless, the higher-mode Rayleigh wave data with the same wavelength can have deeper penetration depth, and higher-mode data can increase the resolution of the S-velocity tomogram (Xia et al. 2003). Rather than inverting only the fundamental-mode surface waves, our multiscale and layer-stripping WD method can be extended to invert both fundamental- and higher-mode surface waves.

The main limits for layer stripping WD is to determine the accurate relationship between the frequency bands and the depth windows. When there are strong lateral gradients in the S-velocity model, the penetration depth of different shot gathers can have a dramatic lateral variation for the same frequency bands. In this case, it is inappropriate to use the same depth windows for all the shot gathers. The depth windows can be designed according to the sensitivity kernels, because they can provide a good estimation of the penetration depth (Masoni et al. 2016).

Conclusions

We have developed a new multiscale and layer-stripping wave-equation dispersion inversion method for Rayleigh waves. In this method, the high-frequency and near-offset data are first used to invert for the shallow S-velocity model, and the lower-frequency data with longer offsets are gradually incorporated to invert for the deeper regions of the model. Numerical results of both synthetic and field seismic data demonstrate that the wave-equation dispersion inversion can suffer from the local minima problem when inverting data from a complex earth model, and our multiscale and layer-stripping wave-equation dispersion inversion method can mitigate the local minima problem and enhance convergence to the global minimum.

Acknowledgments

This work was supported by U.S. Department of Energy through contract DE-AC52-06NA25396 to Los Alamos National Laboratory (LANL). We thank AltaRock Energy, Inc. and Dr. Trenton Cladouhos for providing surface seismic data from the Blue Mountain geothermal field. Zhaolun Liu thank King Abdullah University of Science and Technology (KAUST) for funding his graduate studies. The computation was performed using super-computers of LANL's Institutional Computing Program. Additional computational resources were made available through the KAUST Supercomputing Laboratory (KSL).

REFERENCES

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

Bunks, C., Saleck, F. M., Zaleski, S., & Chavent, G., 1995. Multiscale seismic waveform inversion, *Geophysics*, **60**(5), 1457–1473.

Dou, S. & Ajo-Franklin, J. B., 2014. Full-wavefield inversion of surface waves for mapping embedded low-velocity zones in permafrost, *Geophysics*, **79**(6), EN107–EN124.

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

Graves, R. W., 1996. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences, *Bulletin of the Seismological Society of America*, **86**(4), 1091–1106.

Groos, L., Schäfer, M., Forbriger, T., & Bohlen, T., 2014. The role of attenuation in 2D full-waveform inversion of shallow-seismic body and Rayleigh waves, *Geophysics*, **79**(6), R247–R261.

Groos, L., Schäfer, M., Forbriger, T., & Bohlen, T., 2017. Application of a complete workflow for 2D elastic full-waveform inversion to recorded shallow-seismic Rayleigh waves, *Geophysics*, **82(2)**, R109–R117.

Haskell, N. A., 1953. The dispersion of surface waves on multilayered media, *Bulletin of the seismological Society of America*, **43**(1), 17–34.

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

Li, J., Feng, Z., & Schuster, G., 2016. Wave-equation dispersion inversion, *Geophysical Journal International*, p. ggw465.

Li, J., Dutta, G., & Schuster, G., 2017a. Wave-equation qs inversion of skeletonized surface waves, *Geophysical Journal International*, **209**(2), 979–991.

Li, J., Dutta, G., & Schuster, G. T., 2017b. Skeletonized wave-equation qs tomography using surface waves, in *SEG Technical Program Expanded Abstracts 2017*, pp. 2726–2731, Society of Exploration Geophysicists.

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

Liu, Z., Li, J., & Schuster, G., 2017. 3D wave-equation dispersion inversion of surface waves, in *SEG 2017 Workshop: Full-waveform Inversion and Beyond, Beijing, China, 20-22 November 2017*, pp. 22–26.

Liu, Z., Li, J., Hanafy, S. M., & Schuster, G. T., 2018a. 3D wave-equation dispersion inversion of Rayleigh waves, *Geophysics*, p. submit for publication.

Liu, Z., Li, J., Hanafy, S. M., & Schuster, G. T., 2018b. 3D wave-equation dispersion inversion of surface waves, in *SEG Technical Program Expanded Abstracts 2018*.

Masoni, I., Boelle, J.-L., Brossier, R., & Virieux, J., 2016. Layer stripping FWI for surface waves, in *SEG Technical Program Expanded Abstracts 2016*, pp. 1369–1373.

Mi, B., Xia, J., Shen, C., Wang, L., Hu, Y., & Cheng, F., 2017. Horizontal resolution of multichannel analysis of surface waves, *Geophysics*, **82**(3), EN51–EN66.

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

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

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

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

Shi, T., Zhang, J., Huang, Z., & Jin, C., 2015. A layer-stripping method for 3D near-surface velocity model building using seismic first-arrival times, *Journal of Earth Science*, **26**(4), 502–507.

Sirgue, L. & Pratt, R. G., 2004. Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies, *Geophysics*, **69**(1), 231–248.

Xia, J., Miller, R. D., & Park, C. B., 1999. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves, *Geophysics*, **64**(3), 691–700.

Xia, J., Miller, R. D., Park, C. B., & Tian, G., 2003. Inversion of high frequency surface waves with fundamental and higher modes, *Journal of Applied Geophysics*, **52**(1), 45–57.

Xia, J., Miller, R., Park, C., & Tian, G., 2002. Determining Q of near-surface materials from Rayleigh waves, *Journal of Applied Geophysics*, **51**(2-4), 121–129.

Yilmaz, ., 2015. *Engineering Seismology with Applications to Geotechnical Engineering*, Society of Exploration Geophysicists.

LIST OF FIGURES

23 Comparison between the observed (blue) and synthetic (red) common-offset gathers (COGs) with the offset of 335 m from the S-velocity tomogram without (a) and with (b) layer-stripping method.

24 Same as Fig. 23 except the offset is 670 m.

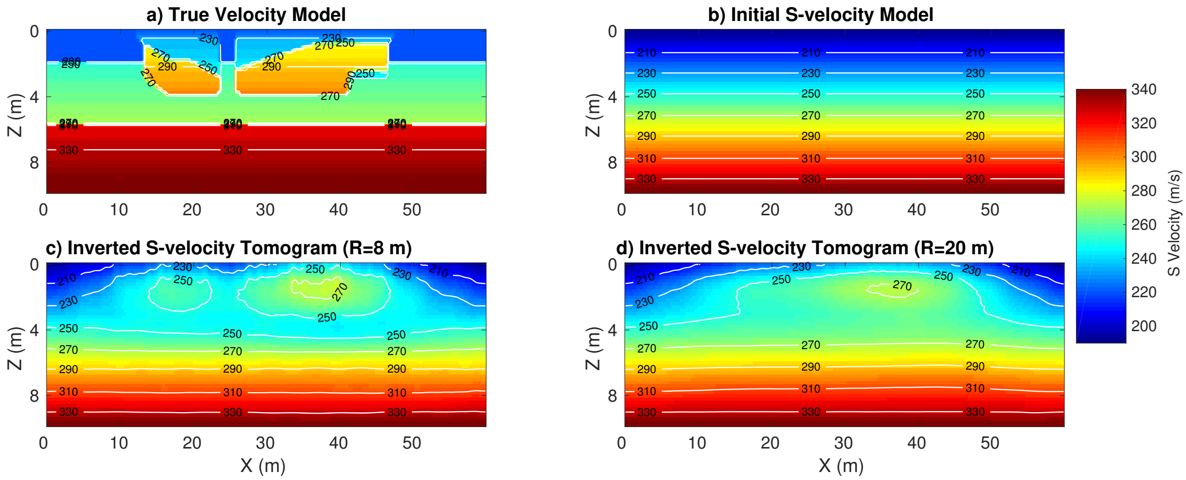


Figure 1. True (a) and initial (b) S-velocity models together with the S-velocity tomograms obtained using WD with maximum offsets of (c) $R = 8$ m and (d) $R = 20$ m.

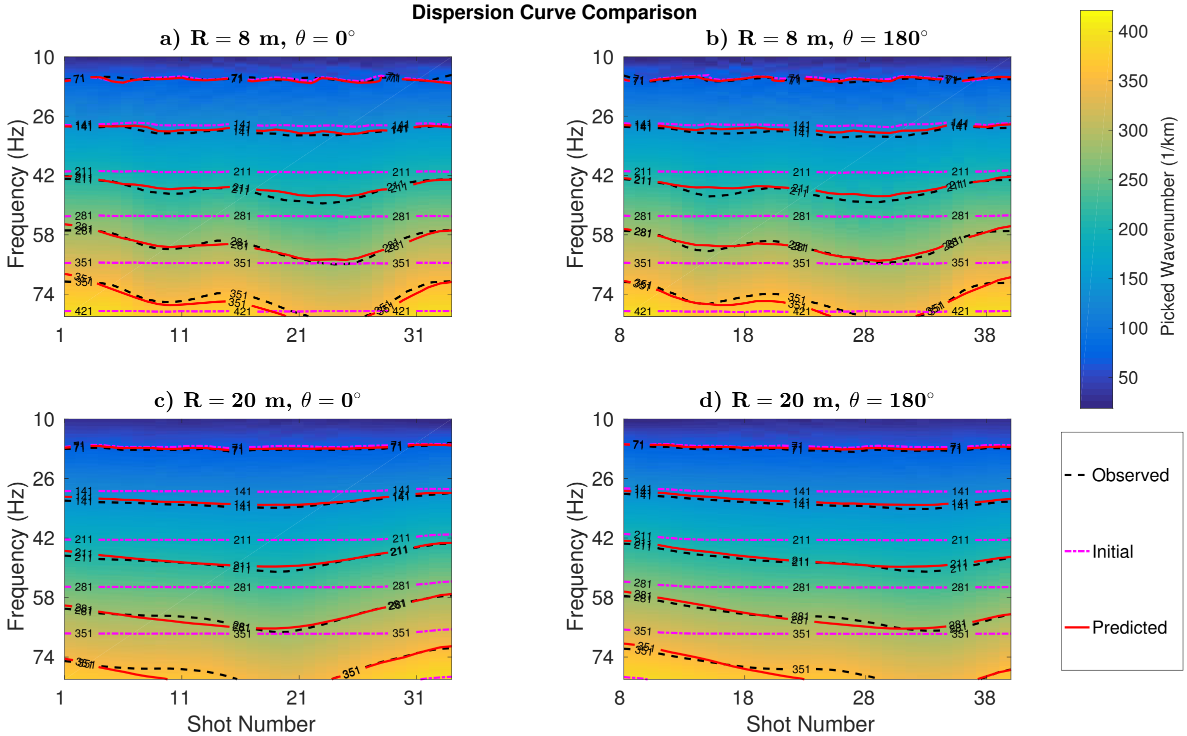


Figure 2. Observed dispersion curves for (a) azimuth angle $\theta = 0^\circ$ with the maximum offset $R = 8$ m, (b) $\theta = 180^\circ$ with $R = 8$ m, (c) $\theta = 0^\circ$ with $R = 20$ m, and (d) $\theta = 180^\circ$ with $R = 20$ m, where the black dashed lines, the magenta dash-dot lines and the red lines represent the contours of the observed, initial and inverted dispersion curves, respectively.

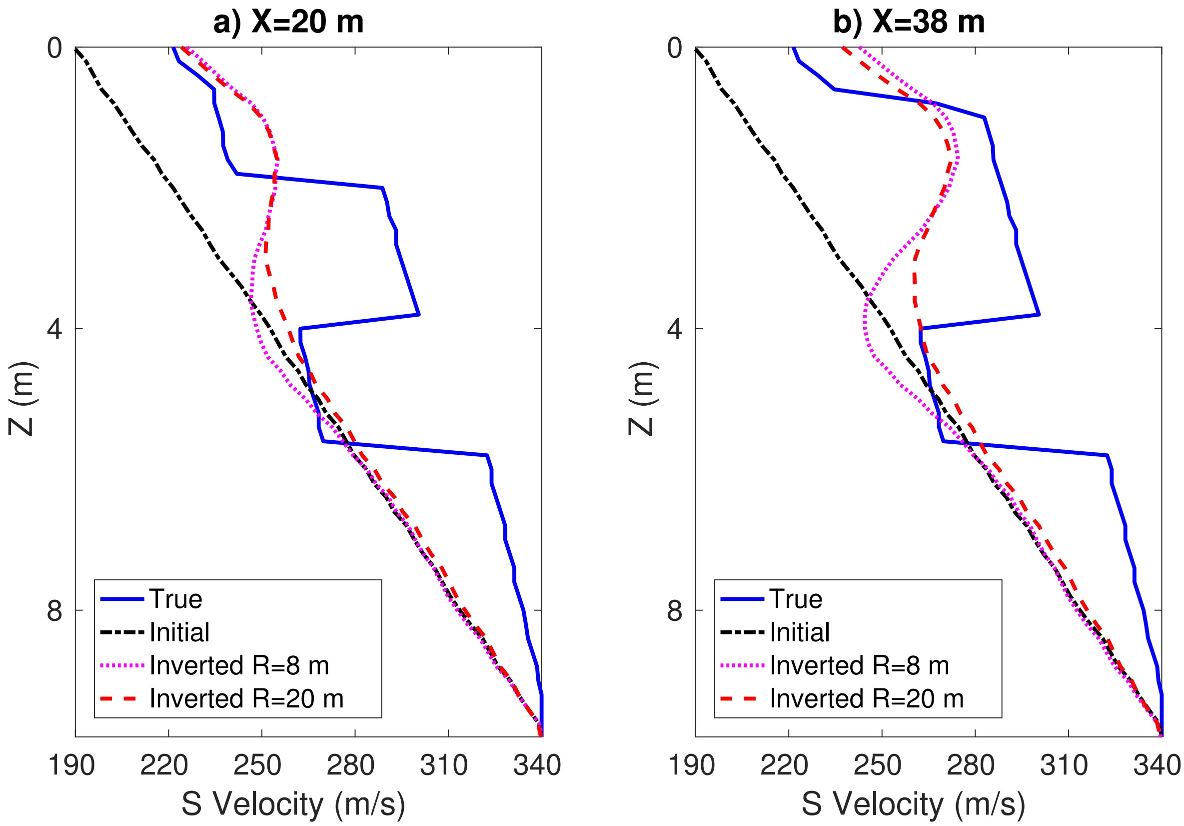


Figure 3. Vertical-velocity profiles at (a) $X = 20$ m and (b) $X = 38$ m for the true model (blue line), the initial model (black dash-dot line) and the inverted S-velocity tomograms when R=8 m (magenta line) and R=20 m (red line) shown in Fig. 1.

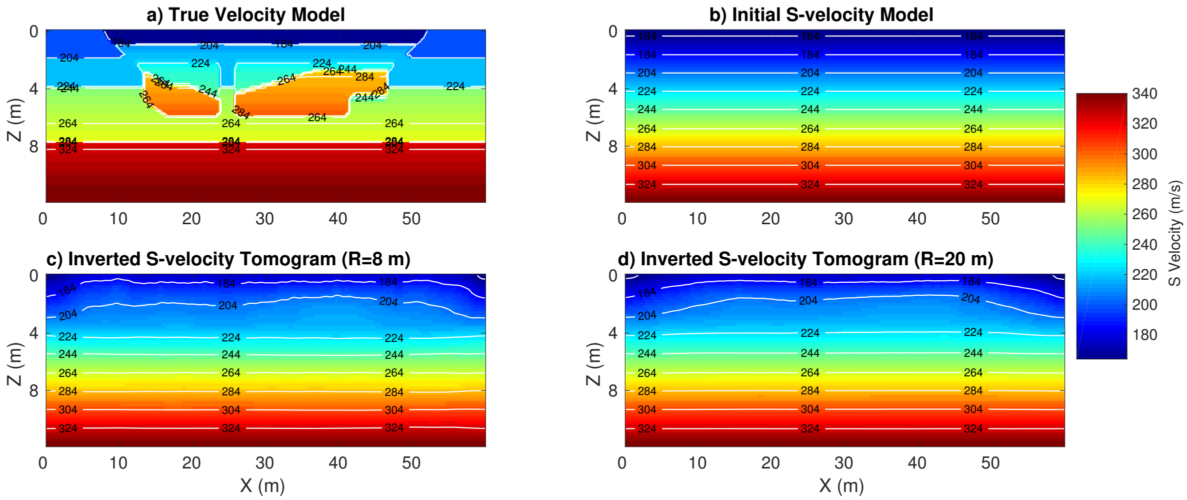


Figure 4. True (a) and initial (b) S-velocity models together with the S-velocity tomograms obtained using WD with maximum offsets of (c) $R = 8$ m and (d) $R = 20$ m. The high-velocity anomalies in (a) are 2 m deeper than the one shown in Fig. 1a.

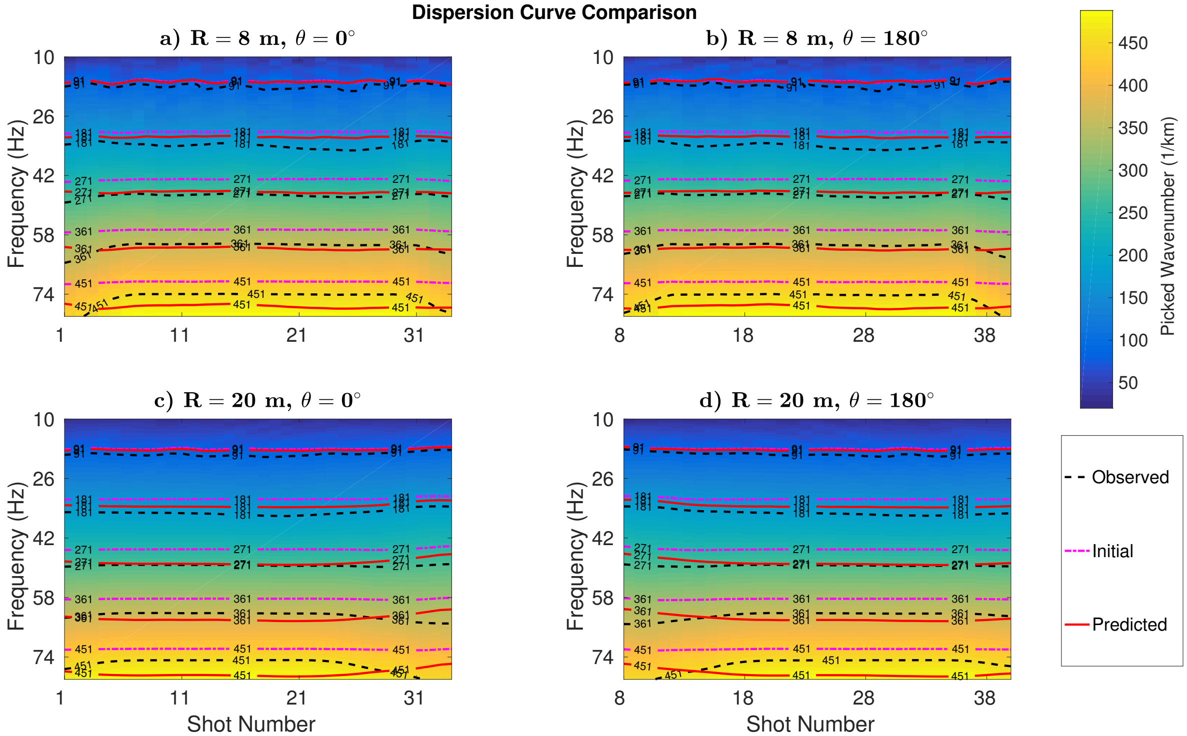


Figure 5. Observed dispersion curves for (a) azimuth angle $\theta = 0°$ with the maximum offset $R = 8$ m, (b) $\theta = 180°$ with $R = 8$ m, (c) $\theta = 0°$ with $R = 20$ m, and (d) $\theta = 180°$ with $R = 20$ m. The black dashed, magenta dash-dot and red lines represent the contours of the observed, initial and inverted dispersion curves, respectively.

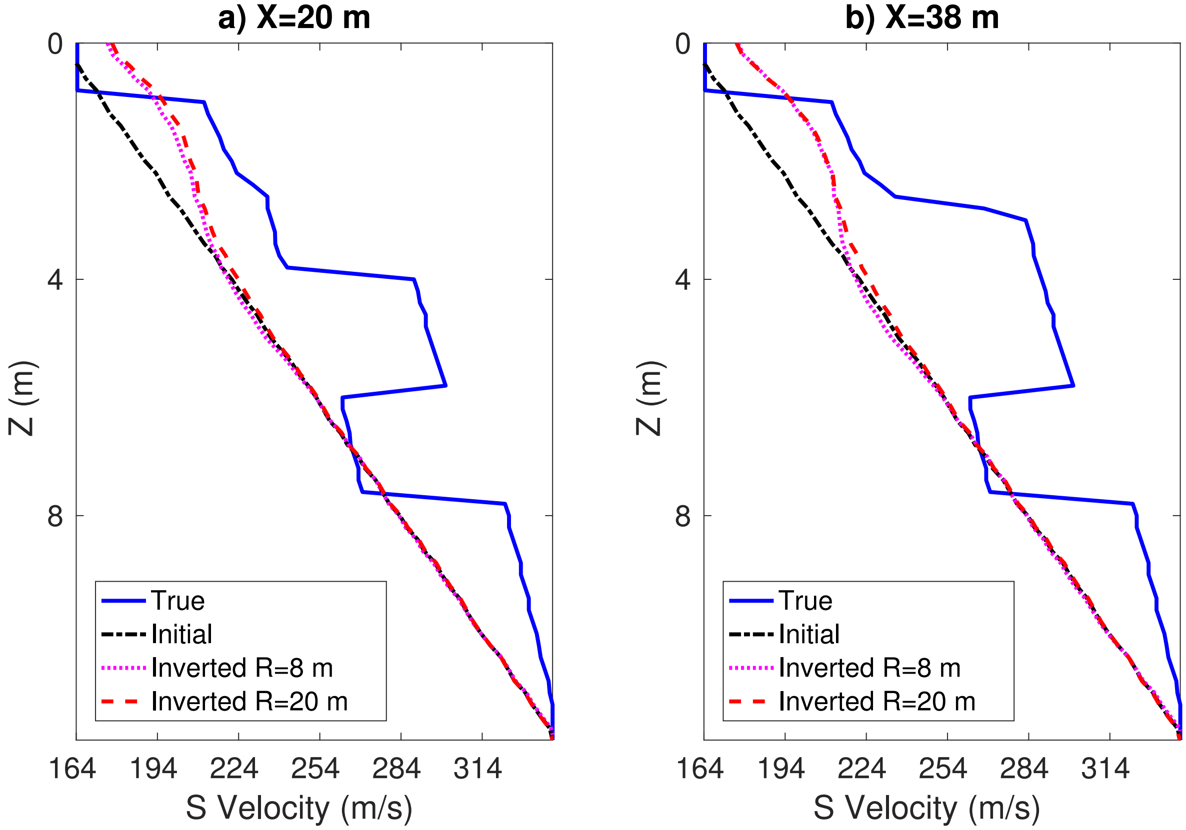


Figure 6. Vertical-velocity profiles at (a) $X = 20$ m and (b) $X = 38$ m for the true model (blue line), the initial model (black dash-dot line) and the S-velocity tomograms by setting R=8 m (magenta line) and R=20 m (red line) shown in Fig. 4.

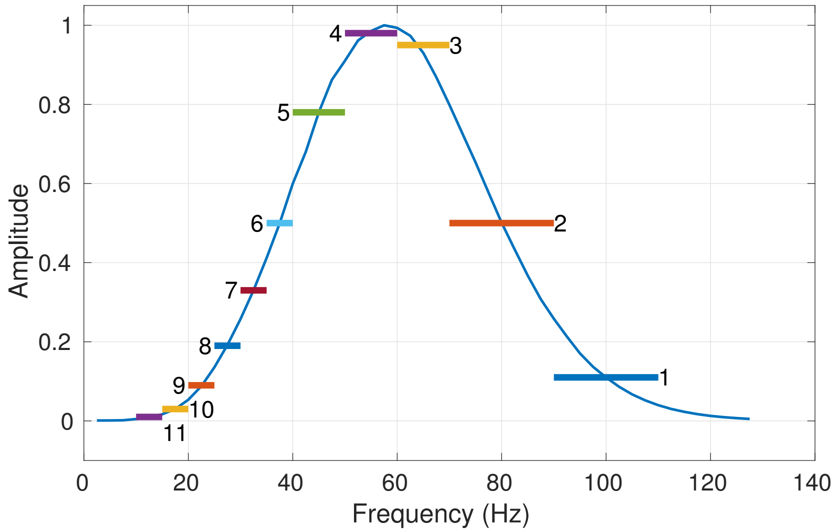


Figure 7. Frequency spectrum of the observed data, which are divided into eleven frequency bands. The frequency bands are plotted as horizontal bars with their corresponding number tags.

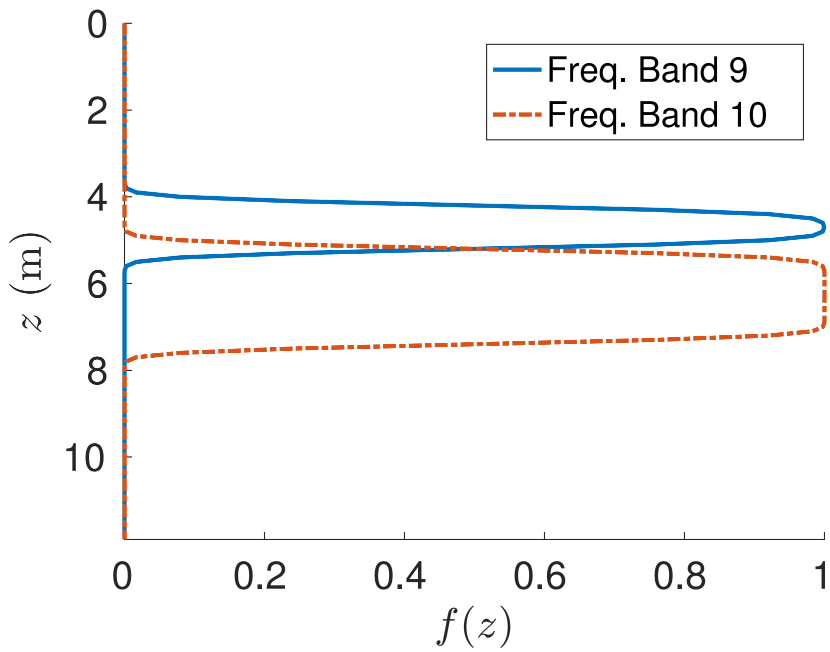


Figure 8. Depth windows for frequency bands 9 (blue solid line) and 10 (red dashed line).

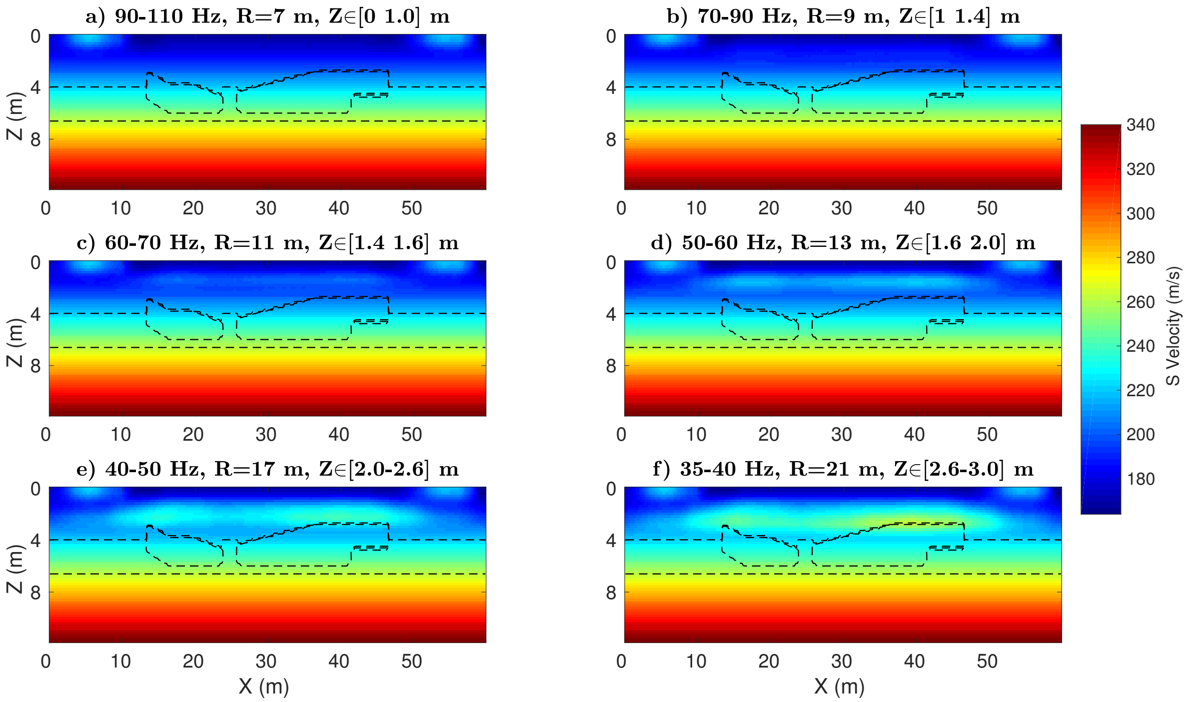


Figure 9. S-velocity tomograms for Steps 1 to 6 (Table 1).

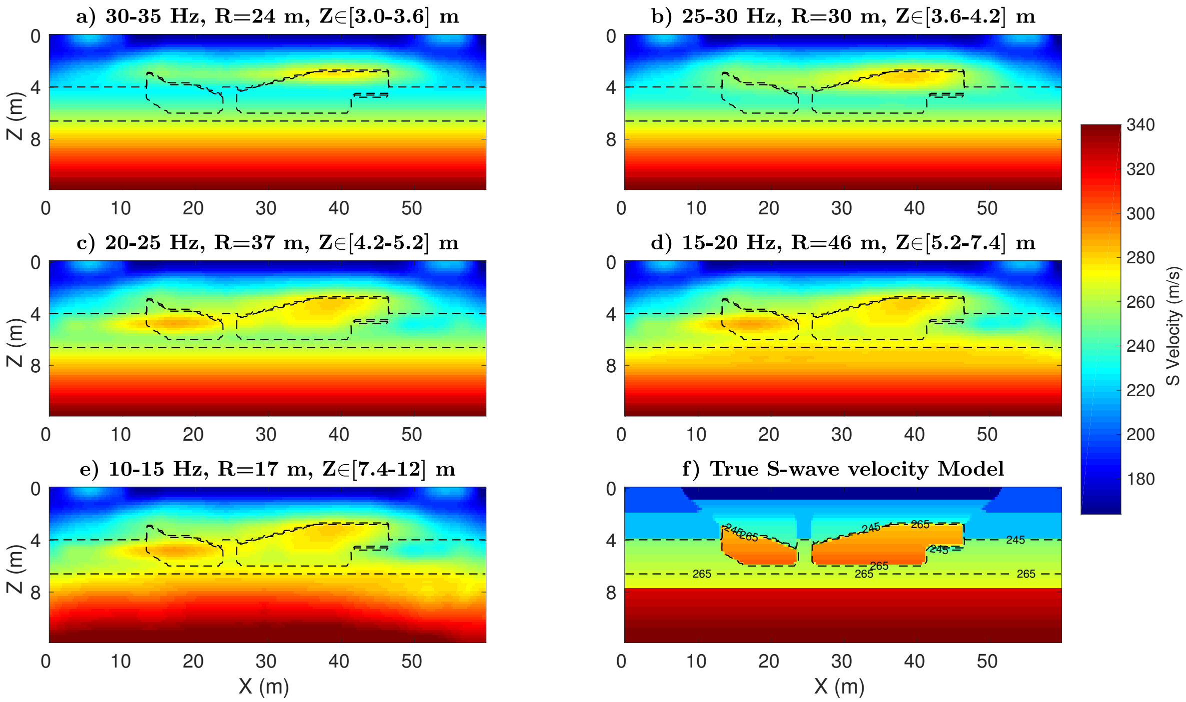


Figure 10. (a)-(e): The S-velocity tomograms for Steps 7 to 11 (Table 1). (f): The true S-velocity model.

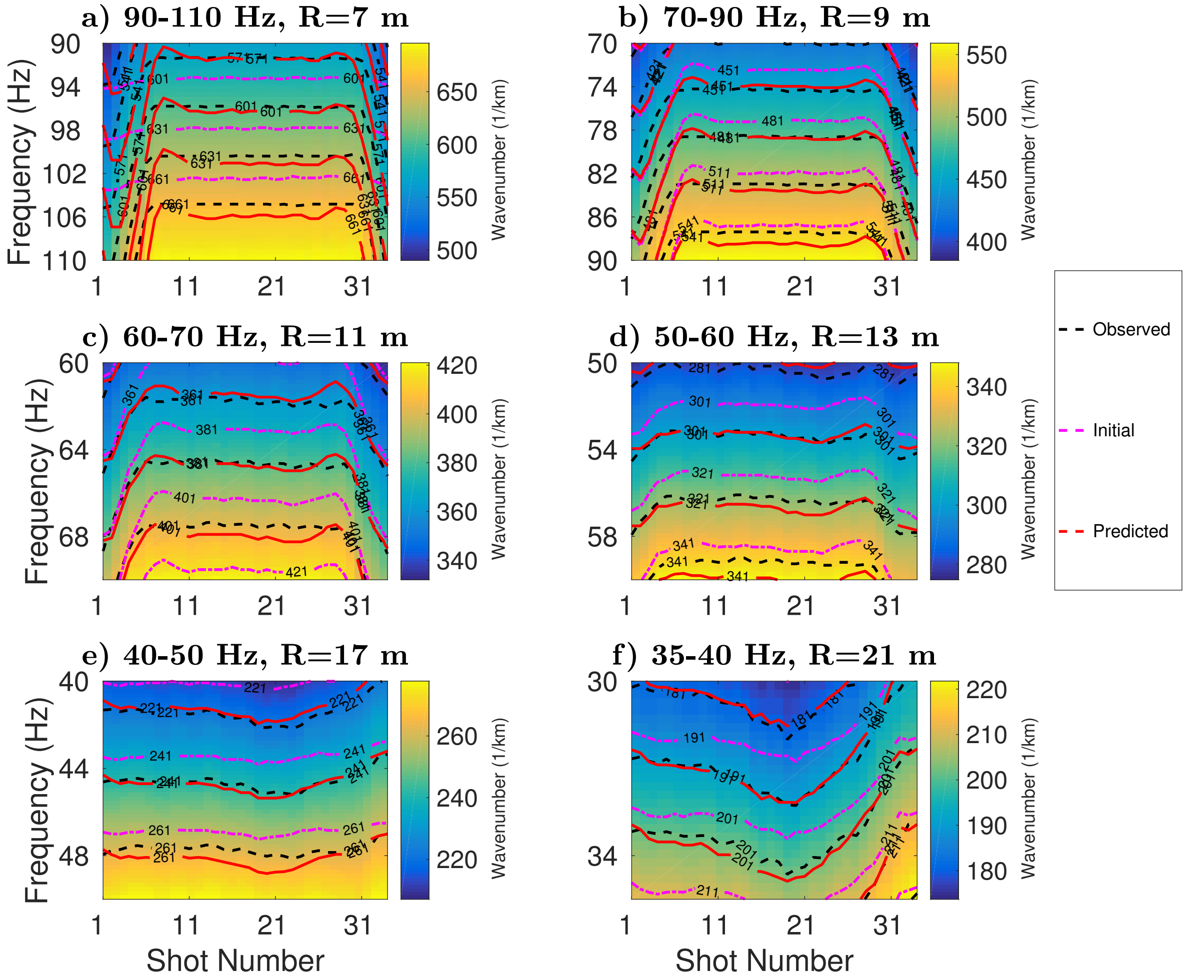


Figure 11. Observed dispersion curves (azimuth angle $\theta = 0°$) for Steps 1 to 6 listed in Table 1, where the black dashed, magenta dash-dot and red lines represent the contours of the observed, initial and inverted dispersion curves, respectively.

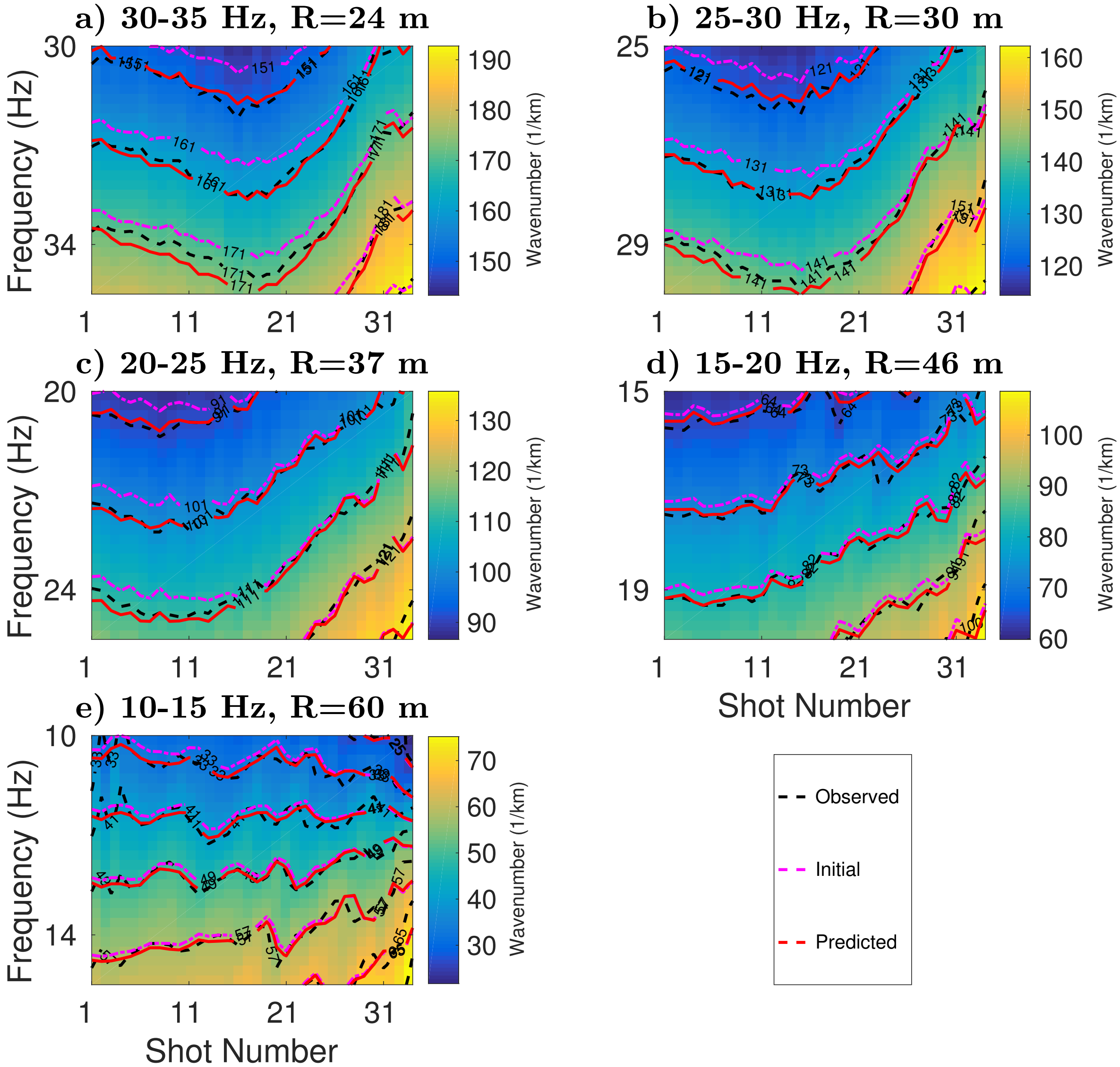


Figure 12. Same as Fig. 11 except we use steps 7-11.

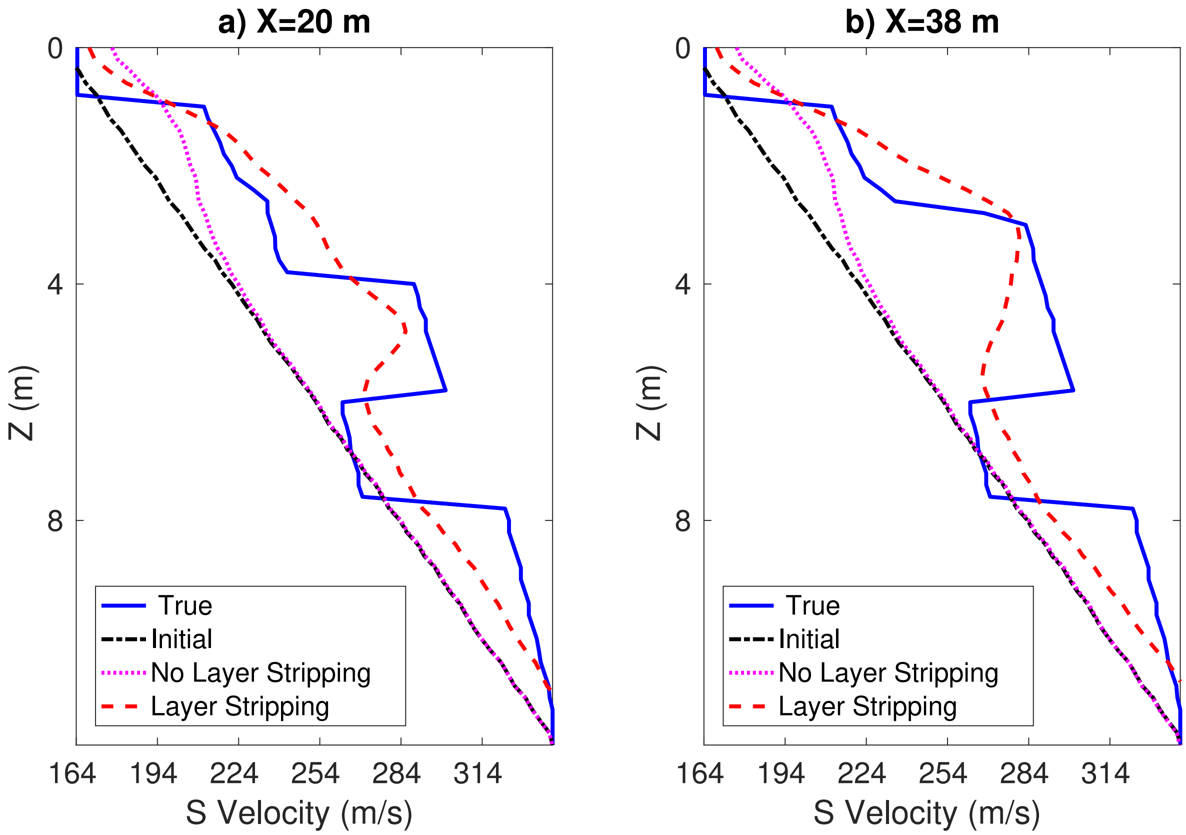
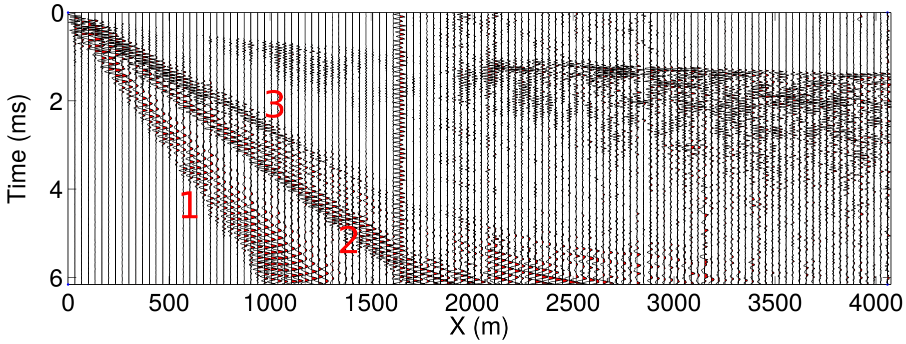
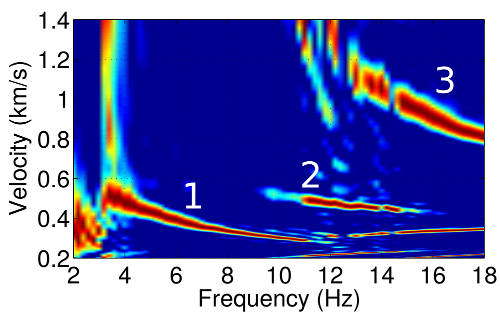


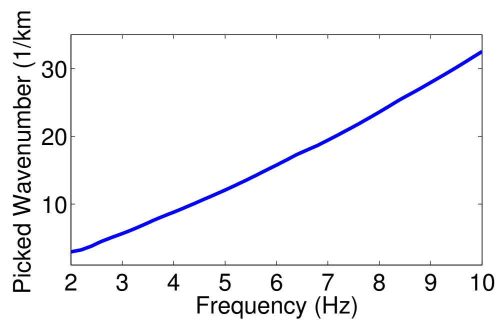
Figure 13. Vertical-velocity profiles at (a) $X = 20$ m and (b) $X = 38$ m for the true model (blue lines), the initial model (black lines), the inverted tomograms with (red lines) and without (magenta lines) layer stripping.



(a) **First CSG**

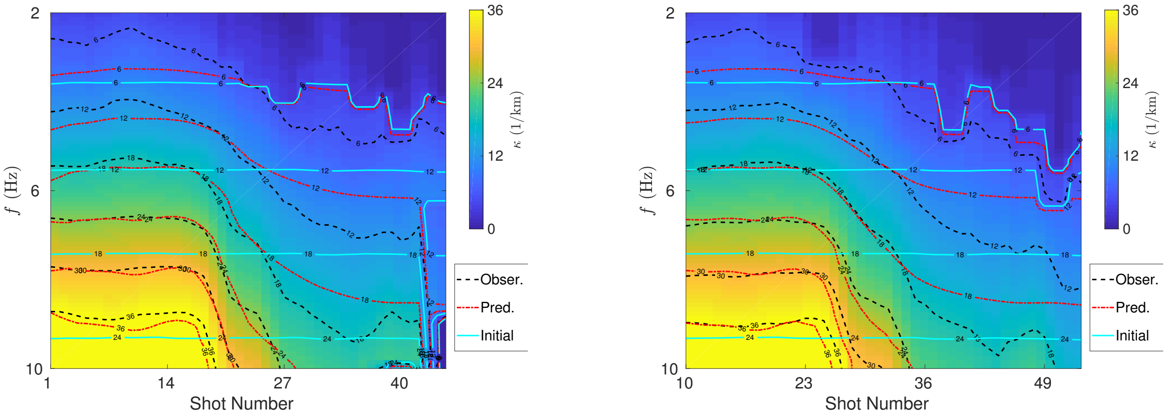


(b) **Dispersion Images**



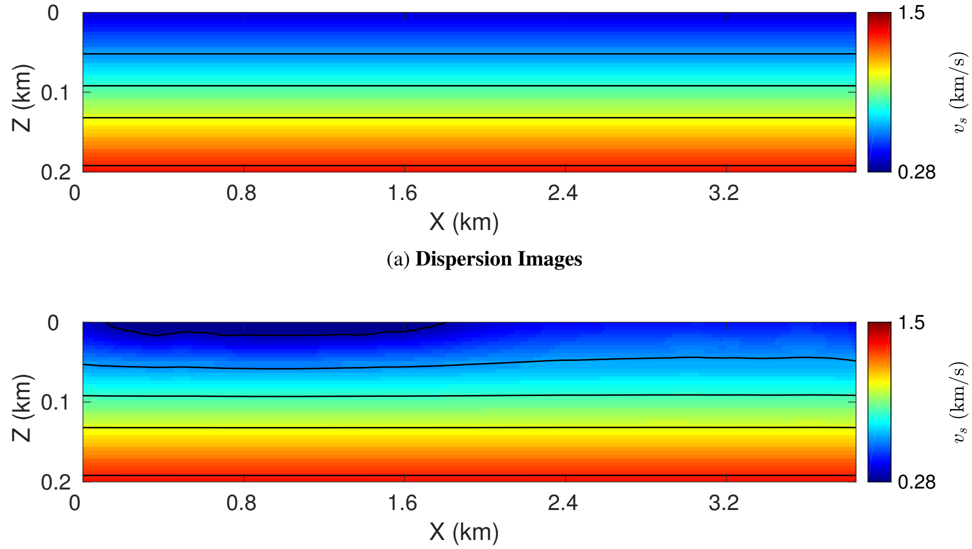
(c) **Picked Dispersion Curve**

Figure 14. (a) First CSG, (b) its dispersion image with the maximum offset R=500 m, and (c) the picked dispersion curve for the fundamental-mode surface waves.



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

Figure 15. Observed dispersion curves along the azimuth angles of (a) $\theta = 0°$ and (b) $\theta = 0°$, where the black dashed, cyan and red dash-dot lines represent the contours of the observed, initial and inverted dispersion curves, respectively.



(a) **Dispersion Images**

(b) **Picked Dispersion Curve**

Figure 16. (a) The initial S-velocity Model; (b) the inverted S-velocity tomogram by WD.

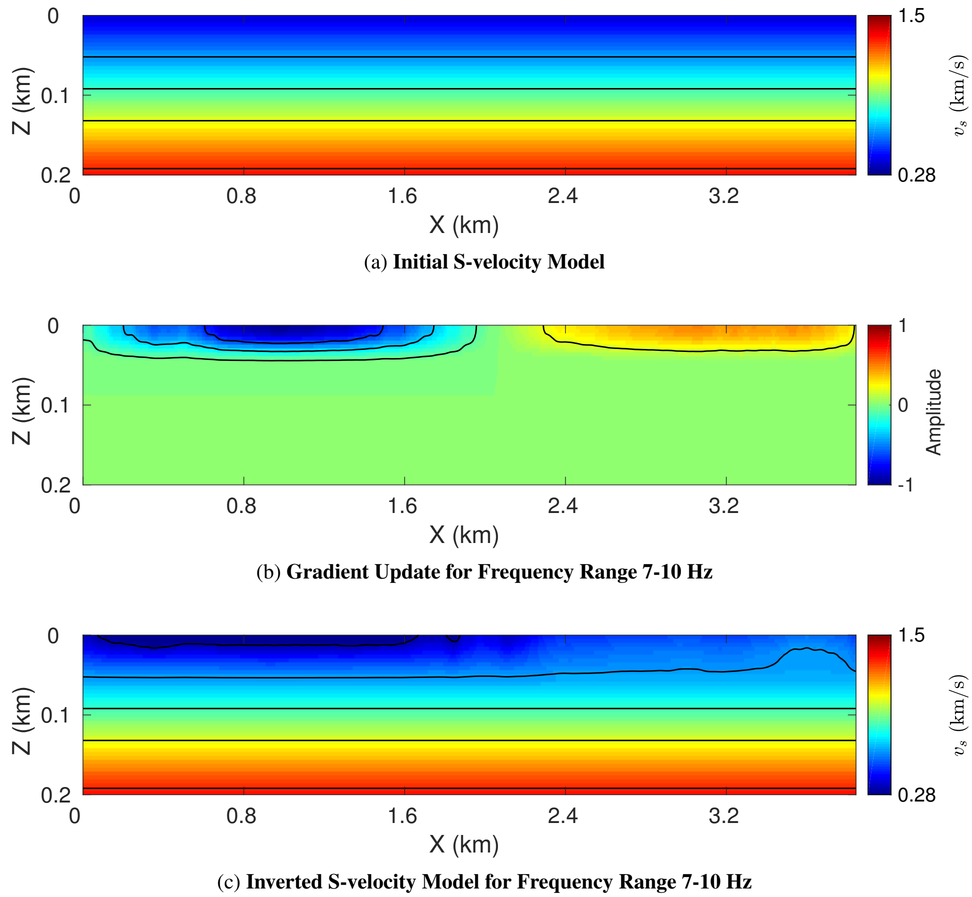
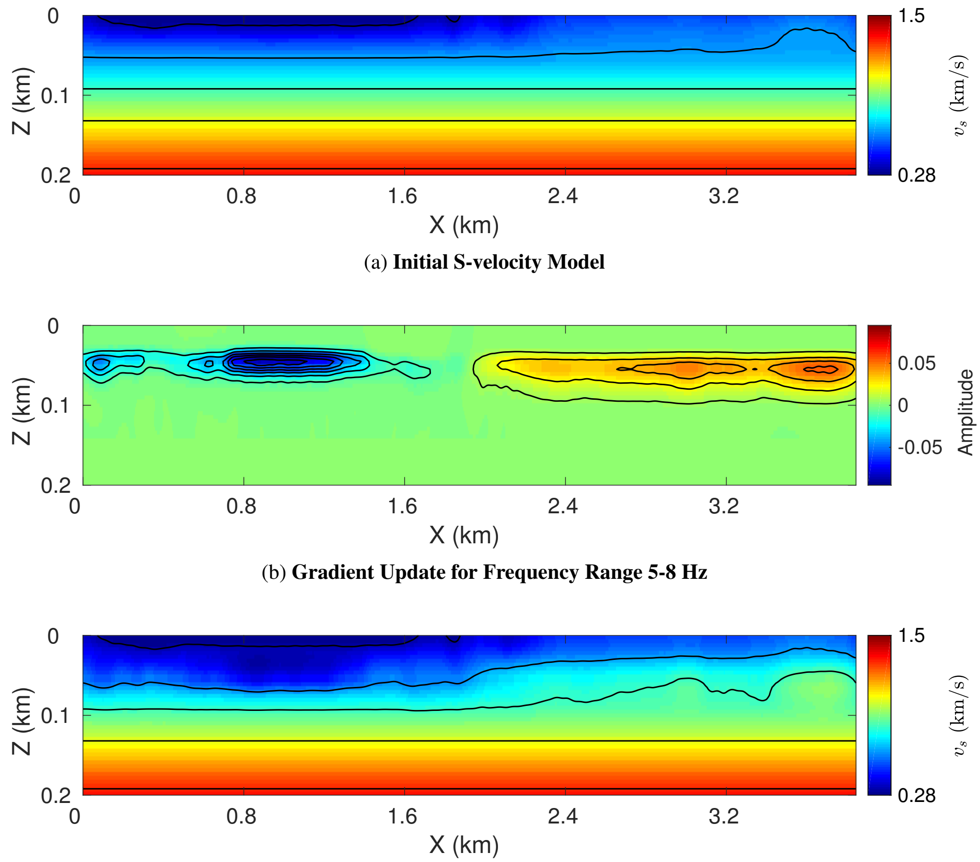


Figure 17. (a) The initial S-velocity model; (b) the gradient update for the first frequency band; (c) the inverted S-velocity model for the first frequency band.

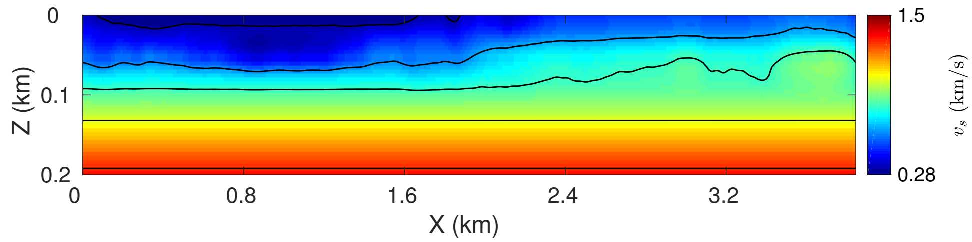


(a) **Initial S-velocity Model**

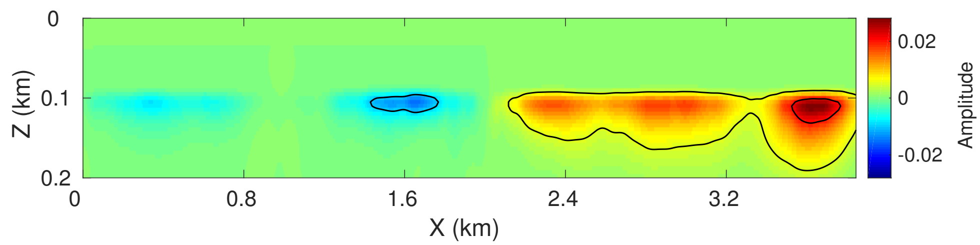
(b) **Gradient Update for Frequency Range 5-8 Hz**

(c) **S-velocity Tomogram for Frequency Range 5-8 Hz**

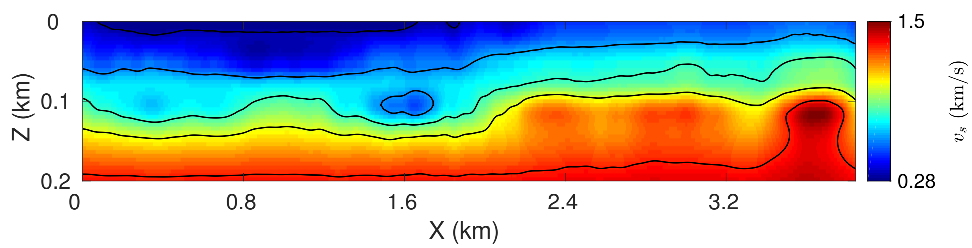
Figure 18. (a)Initial S-velocity model inverted using the first frequency band (Fig. 17); (b) the gradient update for the second frequency band; (c) S-velocity tomogram obtained using the second frequency band.



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

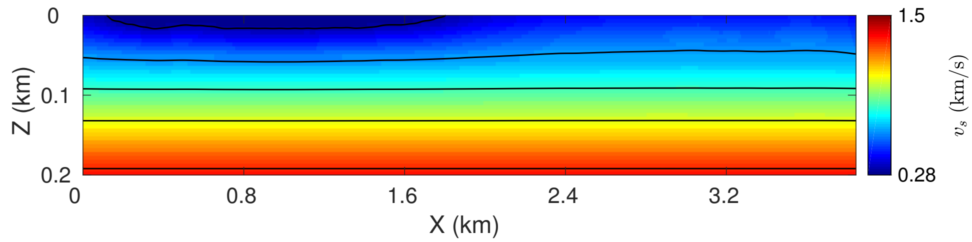


(b) **Gradient Update for Frequency Range 2-6 Hz**

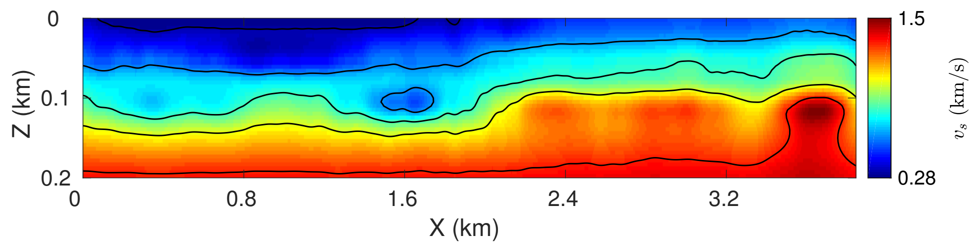


(c) **S-velocity Tomogram for Frequency Range 2-6 Hz**

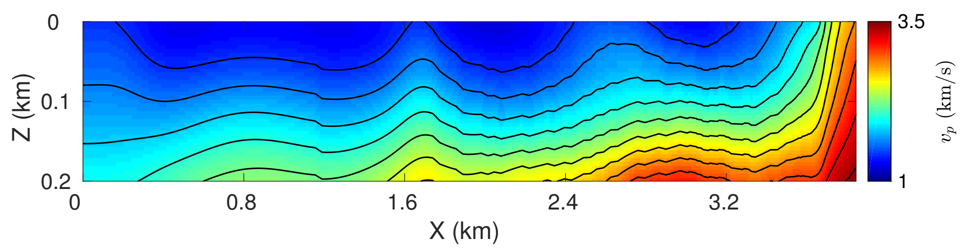
Figure 19. (a) Initial S-velocity tomogram inverted using the second frequency band (Fig. 18); (b) the gradient update for the third frequency band; (c) the inverted S-velocity model obtained using the third frequency band.



(a) **S-velocity Tomogram without Layer-stripping**

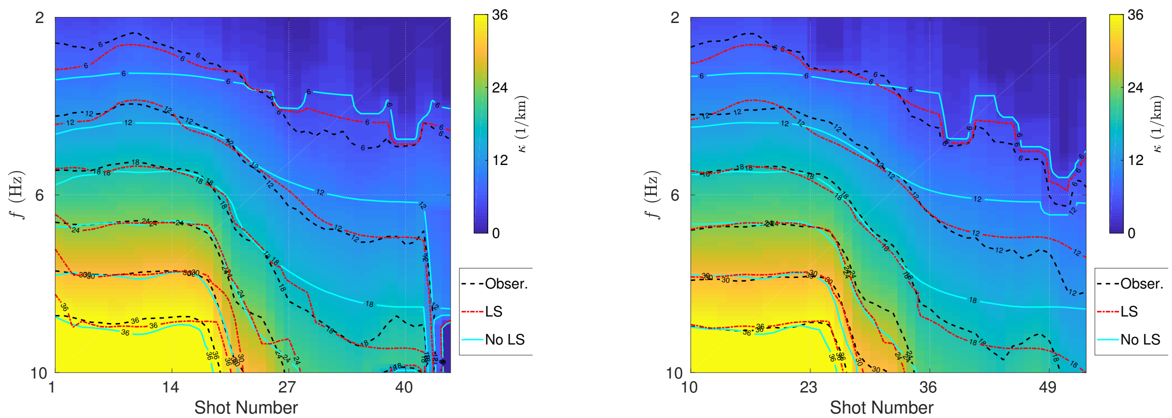


(b) **S-velocity Tomogram with Layer-stripping**



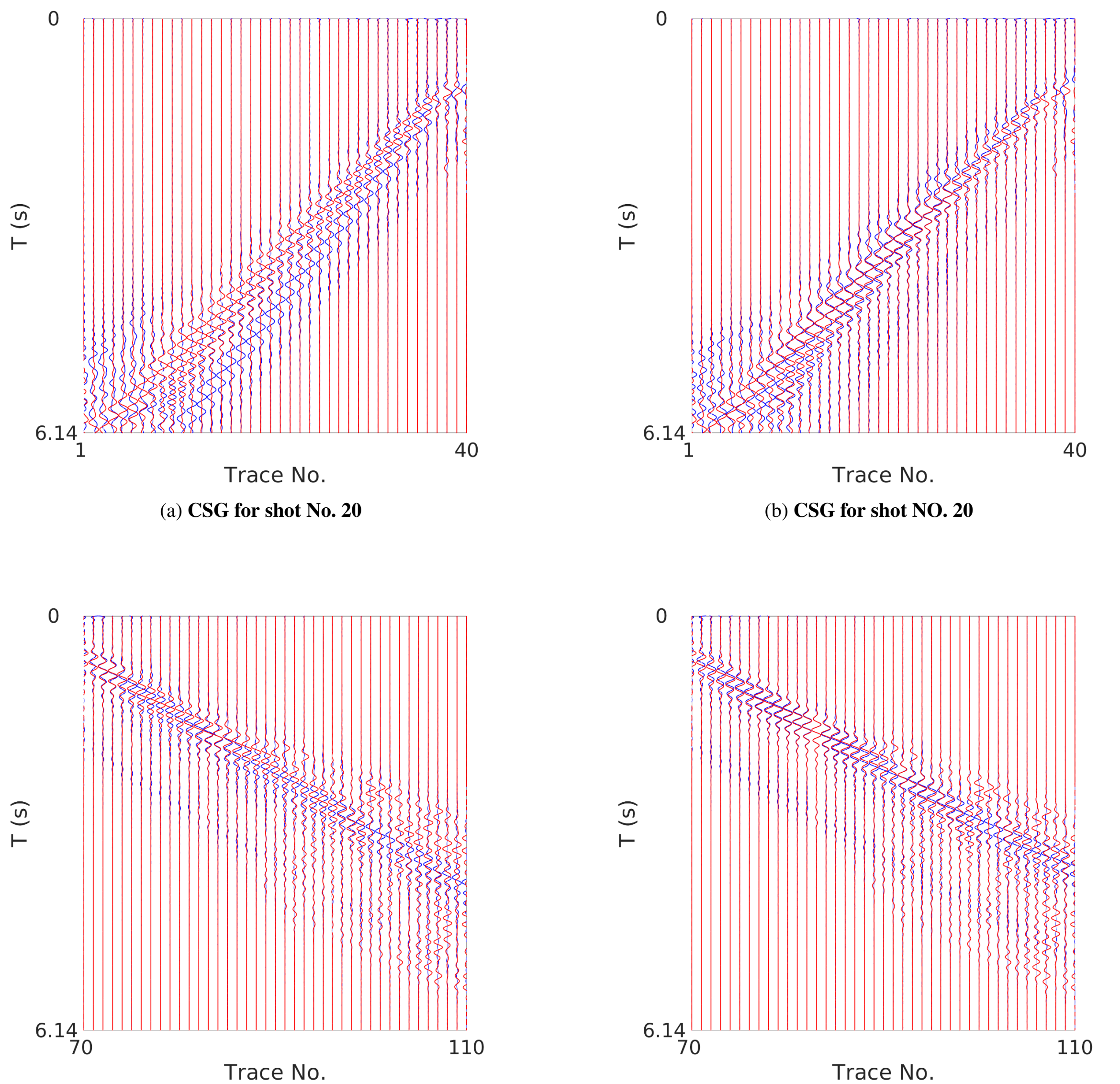
(c) **P-velocity Tomogram**

Figure 20. S-velocity tomograms inverted using the WD methods (a) without and (b) with layer stripping; (c) the P-velocity tomogram.



(a) **Dispersion Curves for Azimuth 0°** (b) **Dispersion Curves for Azimuth 180°**

Figure 21. Observed dispersion curves for (a) $\theta = 0°$ and (b) $\theta = 180°$, where the black dashed lines, the cyan lines and the red dash-dot lines represent the contours of the observed dispersion curves, the predicted dispersion curves obtained without and with layer stripping, respectively.



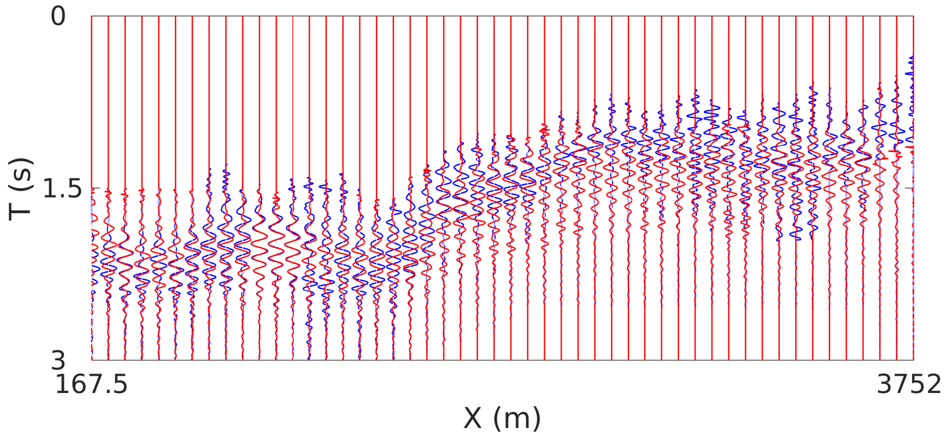
(a) **CSG for shot No. 20**

(b) **CSG for shot NO. 20**

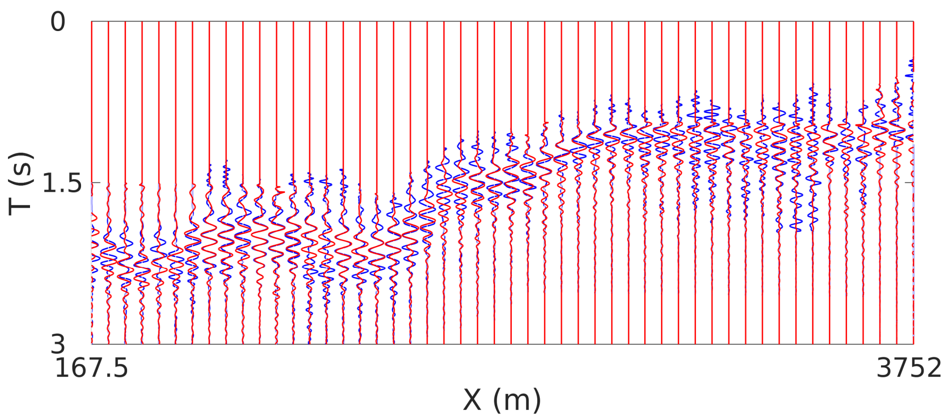
(c) **CSG for shot NO. 30**

(d) **CSG for shot NO. 30**

Figure 22. Comparison between the observed (red) and synthetic (blue) traces from the S-velocity tomogram without (LHS panels) and with (RHS panels) layer-stripping methods for CSG for shot No. 20 in (a) and (b), and CSG for shot No. 30 in (c) and (d).

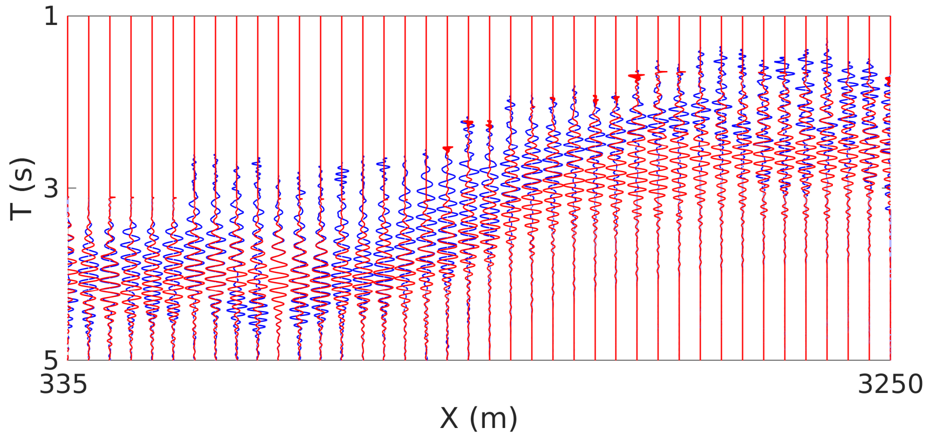


(a) **COG with Offset of 335 m**

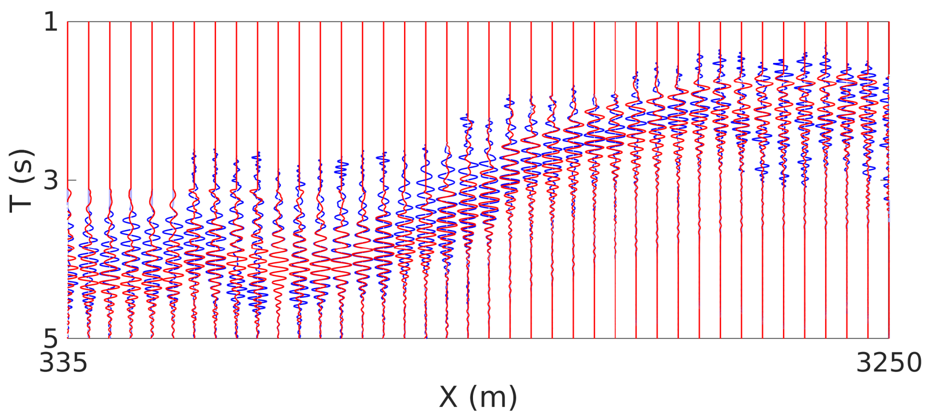


(b) **COG with Offset of 335 m**

Figure 23. Comparison between the observed (blue) and synthetic (red) common-offset gathers (COGs) with the offset of 335 m from the S-velocity tomogram without (a) and with (b) layer-stripping method.



(a) **COG with Offset of 670 m**



(b) **COG with Offset 670 m**

Figure 24. Same as Fig. 23 except the offset is 670 m.

LIST OF TABLES

Table 1. Eleven frequency bands used for MSLSWD, where the wavelength λ is estimated from the dispersion curves; the maximum offset are determined by $R = 3.5\lambda_{max}$; the depth range is calculated using half of the wavelength range with a taper of 0.2 m at both ends.

No.	Freq. Band (Hz)	λ Range (m)	Max. Offset (m)	Depth Range (m)	Misfit Change
1	90-110	1.4-1.9	7	0-1.0	1→0.148
2	70-90	1.9-2.4	9	1-1.4	1→0.4
3	60-70	2.4-2.9	11	1.4-1.6	1→0.1
4	50-60	2.9-3.6	13	1.6-2.0	1→0.13
5	40-50	3.6-4.7	17	2.0-2.6	1→0.125
6	35-40	4.7-5.45	21	2.6-3.0	1→0.1
7	30-35	5.45-6.6	24	3.0-3.6	1→0.17
8	25-30	6.6-8.26	30	3.6-4.2	1→0.25
9	20-25	8.26-10.87	37	4.2-5.2	1→0.35
10	15-20	10.87-15.74	46	5.2-7.4	1→0.7
11	10-15	15.74-33	60	7.4-12	1→0.72