Supporting Information for

Link between crustal thickness and Moho transition zone at 9°N East Pacific Rise

Zhikai Wang[1,2], Satish C. Singh[1], and J. Pablo Canales[3]

[1] Institut de Physique du Globe de Paris, Université Paris Cité, 1 Rue Jussieu, 75005 Paris, France.

[2] Presently at School of Ocean and Earth Science, University of Southampton, European Way, SO14 3ZH Southampton, UK.

[3] Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole MA 02543, USA.

Contents of this file

Text S1. Finite-difference waveform modelling

The modelled seismic data are calculated by solving 2-D elastic wave equation using a temporal 2nd-order and spatial 4th-order staggered-grid finite-difference scheme [*Levander*, 1988]. To avoid numerical dispersion, we used 57.5 m and 28.75 m grid spacings, respectively, in the modelling of 3-5 Hz and 3-10 Hz data. These settings ensure five grid points are sampled by the shortest wavelength (wavelength in the water column), satisfying the dispersion condition [*Levander*, 1988]. Time steps of 4 ms and 2 ms were used in the modelling of 3-5 Hz and 3-10 Hz data, respectively, to keep the waveform modelling stable. An absorbing boundary condition [*Clayton and Engquist*, 1977] was used to attenuate the reflections from model boundaries.

Text S2. Source wavelet

An accurate source wavelet is critical for the success of FWI because the errors in seismic waveform difference due to an inaccurate source wavelet will be directly mapped into the velocity model. In this study, we estimated the source wavelet by stacking the near-offset water arrivals. The source wavelets are estimated separately for OBSs and OBHs. Here we detailed the workflow of source estimate for OBHs, and that for OBSs is the same except using OBS dataset. We extracted four near-offset traces from each OBH dataset after predictive gapped deconvolution and aligned these traces to the same starting time (0.05 s). The signals after 0.6 s were muted to mitigate the influence of seismic multiples and later reflections. These traces show high similarity in waveform of direct water arrivals (Figure S3A). We stacked the aligned traces (Figure S3B) and filtered the stacked signal (Figure S3C) using the same band-pass filters (3-5 Hz or 3-10 Hz) as applied to seismic data in FWI. The starting time of the source wavelet is determined by performing finite-difference waveform modelling and comparing the modelled and observed near-offset water wave. Figure S3C shows the source wavelets for OBSs (dashed curves) and OBHs (solid curves). Precise amplitude of the source wavelet is not needed because we normalized the seismic traces in the trace normalized FWI and the whole seismic gather in the true amplitude FWI (see text below). The good match between the modelled and observed near-offset direct water arrivals for OBH and OBS data (Figure S4) indicates that the estimated source wavelets are sufficiently accurate for performing FWI.

Text S3. Starting models

The water velocity is set to 1.5 km/s and is not updated in the FWI. The starting crustal P-wave velocity model is obtained from a ray-based travel time tomography of Pg and PmP arrivals

[*Canales et al.*, 2003]. The starting mantle P-wave velocity is expanded from a one-dimensional velocity profile hanging from the seafloor, where the mantle velocity increases linearly from 7.9 to 8.2 km/s within 5 km depth range. We smoothed the velocity around the tomographically constrained Moho to avoid a sharp boundary between the crust and mantle. The starting P-wave velocity model is shown in Figure S5. The starting S-wave velocity and density are calculated from P-wave velocity using the empirical relations given in [*Brocher*, 2005]. The S-wave velocity and density are not updated in the FWI.

Text S4. FWI

We used a 2-D time domain elastic FWI developed originally by *Shipp and Singh* [2002] for marine streamer data with constant shot and receiver spacings and modified the code to accommodate seismic data recorded by ocean bottom instruments with arbitrary geometry in 2-D space. Starting from an initial estimate of velocity of the subsurface, the elastic FWI iteratively updates the velocity model by reducing the misfit between the observed and modelled seismic data

$$m_{n+1} = m_n + \alpha_n g_n \, , \tag{1}$$

where m is the model parameter, α is a step length, g is the gradient of the misfit function [*Shipp and Singh*, 2002] and n is the iteration number. In this study, we only inverted the P-wave velocity of the subsurface. A constant step-length of 30 m/s was used in all iterations.

We simultaneously inverted the pressure data recorded by OBHs and the vertical component data of OBSs. The conventional elastic FWI directly compares the least-squared difference between the modelled and observed waveforms [*Shipp and Singh*, 2002; *Tarantola*, 1986]. Because the magnitude of amplitudes of the OBH and OBS data are very different, direct comparison of waveforms leads to unbalanced contributions to the gradient for OBH and OBS data. To solve this problem, two FWI approaches comparing normalised seismic data were used in this study.

We performed the trace normalized FWI of *Tao et al.* [2017] in the first stage. In this FWI approach, each trace of the observed and modelled data is normalised by its L_2-norm, and the misfit function is defined as the least-squared difference between the modelled and observed seismic data after trace-by-trace normalisation

$$J_1 = \sum_i^{N_s} \sum_j^{N_r} \left\| \frac{d_{i,j}}{\|d_{i,j}\|} - \frac{u_{i,j}}{\|u_{i,j}\|} \right\|^2 , \tag{2}$$

where d and u represent the observed and modelled seismic data, N_s is the number of seismic gathers and N_r is the number of traces within each seismic gather, and $\| \quad \|$ represents the L_2 norm. From Equation (2), we can see that the trace normalised FWI is insensitive to the amplitude of the seismic data [*Tao et al.*, 2017]. Furthermore, the trace-normalised FWI is capable of inverting triplicated waveforms [*Tao et al.*, 2017], which in our case are the PmP arrivals. However, this method ignores the amplitude variation with offset (AVO) effect and mainly compares the phase information in seismic data, leading to reduced resolution than conventional FWI [*Liu et al.*, 2016]. The seismic residual of the trace normalised FWI is

$$R_{i,j} = \frac{1}{\|d_{i,j}\|\|u_{i,j}\|}\left(\frac{\int d_{i,j,t}\cdot u_{i,j,t}dt}{\|d_{i,j}\|\|u_{i,j}\|}u_{i,j} - d_{i,j}\right), \qquad (3)$$

where $\int dt$ represents the integration over time and $\cdot$ is the multiplication operator.

To ensure the inversion convergence to the global minimum, we applied the multi-scale inversion strategy of *Bunks et al.* [1995] in the trace normalised FWI. We first inverted the seismic data between 3 and 5 Hz, and then this inverted velocity model was used as a starting model for the inversion of 3-10 Hz data. We also applied the multi-stage inversion strategy of *Shipp and Singh* [2002] for each frequency-band, where the near-offset (<20 km) data are inverted first and we increased the offset by 20 km every 7 iterations.

Taking the inverted model of the trace normalised FWI as starting model, we further performed 30 iterations of true amplitude FWI for 3-10 Hz data in the second stage. In the true amplitude, each seismic gather is normalised by the L_2-norm of the whole seismic gather, which scales the amplitude of OBH and OBS data to similar magnitude. The misfit function of the shot-normalised FWI is defined as the least-squared difference between the modelled and observed data after normalisation

$$J_2 = \sum_i^{N_s}\left\|\frac{d_i}{\|d_i\|} - \frac{u_i}{\|u_i\|}\right\|^2. \qquad (4)$$

This FWI approach compares both the amplitude and the phase information, which can further refine the velocity of the subsurface. The seismic residual is defined as follows

$$r_{i,j} = \frac{1}{\|d_i\|\|u_i\|}\left(\frac{\int d_{i,j,t}\cdot u_{i,j,t}dtdr}{\|d_i\|\|u_i\|}u_{i,j} - d_{i,j}\right), \qquad (5)$$

where $\int dtdr$ represents the integration over time and trace and j is the index of trace. The multi-offset inversion strategy of *Shipp and Singh* [2002] is not used in the second stage because no cycle skipping between modelled and observed data is observed after the trace-normalised FWI.

The gradient (g) in the FWI is computed by zero-lag cross-correlating the source generated forward-propagated wavefield and the adjoint source generated wavefield by back projecting the seismic residuals [*Shipp and Singh*, 2002]. We muted the gradient in the water column to avoid updating the velocity of water and we tapered the gradient within 115 m distance from OBHs and OBSs. A conjugate-gradient method [*Scales*, 1987] was used to speed up the convergence. The conjugate gradient was multiplied by square root of depth to partially account for spherical divergence [*Krebs et al.*, 2009], except for the last twenty iterations of the second stage where the conjugate gradient was multiplied by square of depth to further enhance the energy around and below the MTZ. To suppress the artifacts introduced by the sparse distribution of ocean bottom instruments, we applied a 2-D wavenumber domain low-passed filter [*Jian et al.*, 2021] to the velocity gradient. The low-passed filter is defined as

$$\left|\frac{k_x}{k_{cx}}\right| + \left|\frac{k_z}{k_{cz}}\right| = 1, \tag{6}$$

where k_x and k_z are the wavenumbers along horizontal distance and depth. k_{cx} and k_{cz} are the cut-off wavenumbers of the 2-D low-passed filter. k_{cx} is set as the inverse of the minimum distance (8 km) between two neighbouring ocean bottom instruments ($k_{cx} = 0.125$ km^{-1}). We set k_{cz} as the maximum resolvable wavenumber ($k_{z\,max}$) of FWI, which is defined as follows [*Brenders and Pratt*, 2007]

$$k_{cz} = k_{z\,max} = \frac{2\pi f}{v}, \tag{7}$$

where f is the frequency and v is the background velocity. In this study, we used velocity $v =$ 6.5 km/s and the frequency f of 5 and 10 Hz to determine k_{cz} in FWI of 3-5 and 3-10 Hz data, respectively.

Text S5. Checkerboard test

We performed checkerboard tests to assess the resolution of the FWI result. The checkerboard input models are designed by adding 2-D sinusoidal anomalies into the crust and mantle of the FWI model. The maximum velocity perturbation is ±5%, which is the same as that used in the travel time tomography study from *Canales et al.* [2003]. We tested velocity anomalies with size of 0.5 × 10 km (horizontal × vertical), 0.5 × 8 km, 0.3 × 8 km and 0.5 × 5 km (Figure S7-10A). Synthetic seismic data are modelled by performing the finite-difference modelling using the estimated source wavelets of 3-10 Hz and the same source-receiver geometry as that of the field data. We inverted these synthetic seismic data using the same inversion parameters and time window as that for the FWI of the field data, starting from the final model of FWI using field data. We only performed the second stage of true amplitude FWI in the checkerboard tests, because no obvious cycle-skipping is observed. The results show that velocity anomalies of 0.5 × 10 km and 0.5 × 8 km size are completely recovered between 10 and 80 km horizontal distances (Figure S7B and S8B) and the velocity anomalies of 0.3 × 8 km size are recovered with locally reduced recovery in the mid-crust (Figure S9B). In contrast, the velocity anomalies of 0.5 × 5 km size are not recovered (Figure S10B). Therefore, the minimum resolution is ~8 km in the horizontal direction and ~0.3 km in the depth direction, and therefore, we only interpret anomalies larger than these values.

Text S6. Synthetic tests for the recovery of MTZ

We performed synthetic tests (Figure S11-S17) to assess the resolvability of the FWI method for a thick or thin MTZ, following the approach proposed in *Jian et al.* [2021]. A MTZ with velocity increasing linearly from 7.0 km/s to 7.85 km/s with depth is inserted in the final inverted model from FWI of OBS data. The thickness of MTZ is 0.5, 1.0, 1.5, 2.0, 2.5, 3.0 and 3.5 km in these tests (Figure S11-S17(A)), respectively. These designed models (hereafter referred to as 'synthetic true model') were used to generate synthetic seismic data by performing finite-difference modelling using the estimated source wavelets of 3-10 Hz and the same source-receiver geometry as that of the field data. We designed a starting model (Figure S11-S17(B)) for synthetic tests by smoothing the velocity below the top of the MTZ within the synthetic true model. The lateral width of the smoothing window is 8.0 km and the vertical width is twice of the thickness of the inserted MTZ. This is to ensure the starting model is smooth but doesn't lead to cycle-skipping. We only performed the inversion of 3-10 Hz data in the second stages of the FWI workflow. The same inversion parameters as those for FWI of field data were used in the synthetic tests. The final inverted models and comparisons of some 1-D velocity profiles are shown in Figure S11-S17(C). The difference between the synthetic true and starting models is shown in Figure S11-S17(D) and that between the synthetic true and inverted models is shown in Figure S11-S17(E). For the 0.5 km thick MTZ, the velocity of the MTZ is partially recovered between 10 and 80 km horizontal distance (Figure S11). In contrast, the velocity of the inserted MTZ between 10 and 80 km horizontal distance is almost completely recovered when the MTZ is 1.0-3.5 km thick (Figure S12-S17). The recovery of the MTZ at 10-25 km horizontal distance is slightly worse than that to the further north, likely due to sparser instruments deployed in this region.

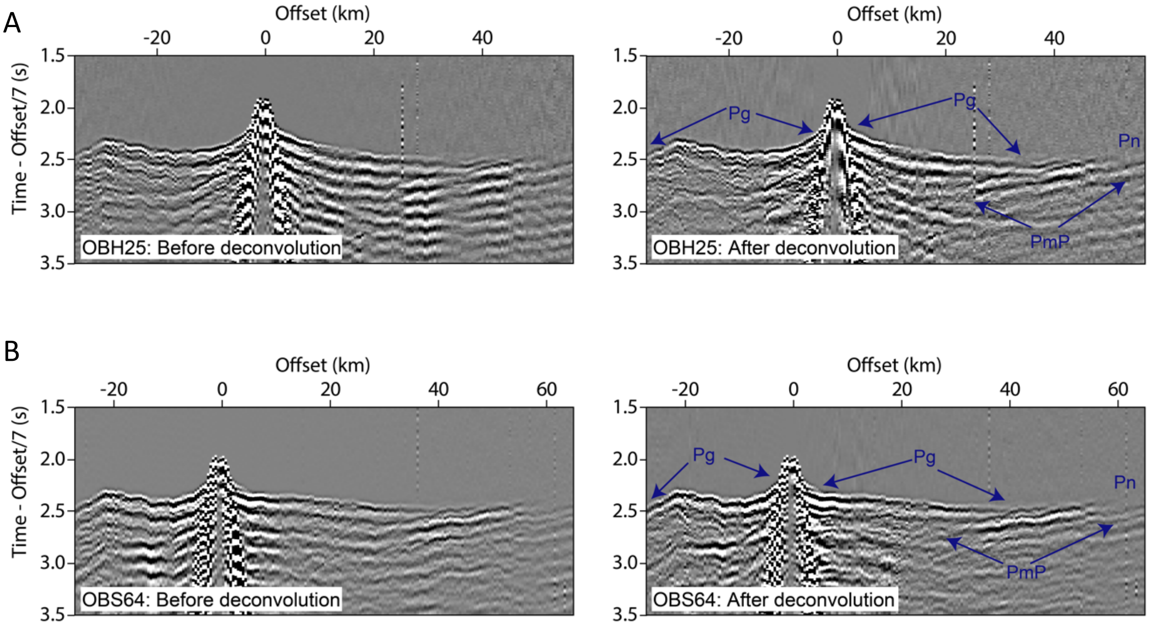


Figure S1. Comparisons of seismic data before (left column) and after (right column) predictive gapped deconvolution. (A) for OBH25 and (B) for OBS64. The travel time of seismic data is reduced using a reduction velocity of 7.0 km/s. The seismic bubble pulses are supressed after the predictive gapped deconvolution, and the crustal refractions (Pg), the Moho reflections (PmP) and the mantle refractions (Pn) are clearer.

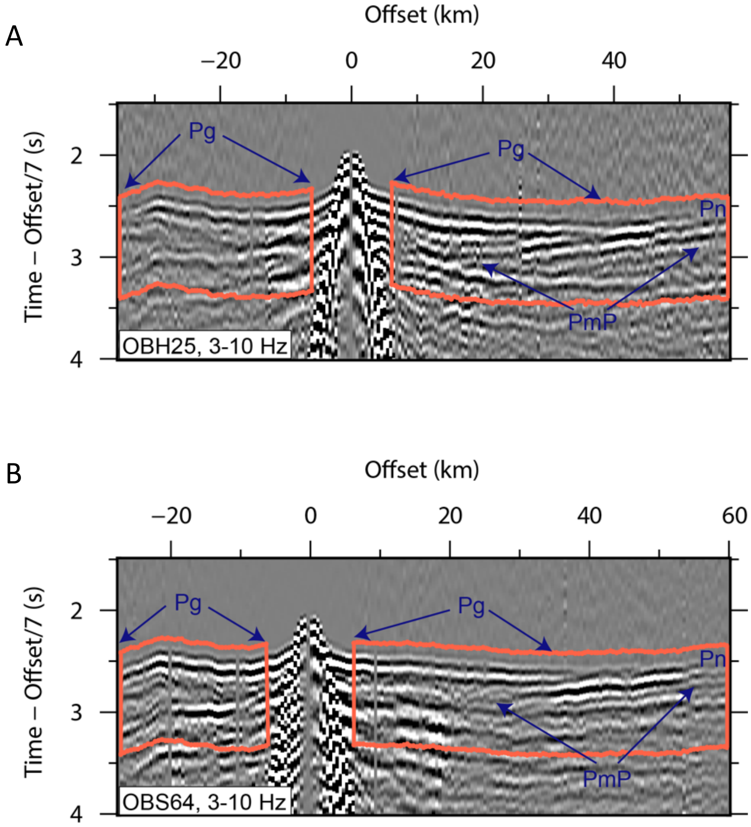


Figure S2. Time window overlapping on seismic plot. (A) for OBH25 and (B) for OBS64. The oranges boxes show the 1.0 s-wide time window used in the FWI of 3-10 Hz data. The waveforms of the Pg, PmP and Pn arrivals are included in the time window. The travel time of seismic data is reduced using a reduction velocity of 7.0 km/s.

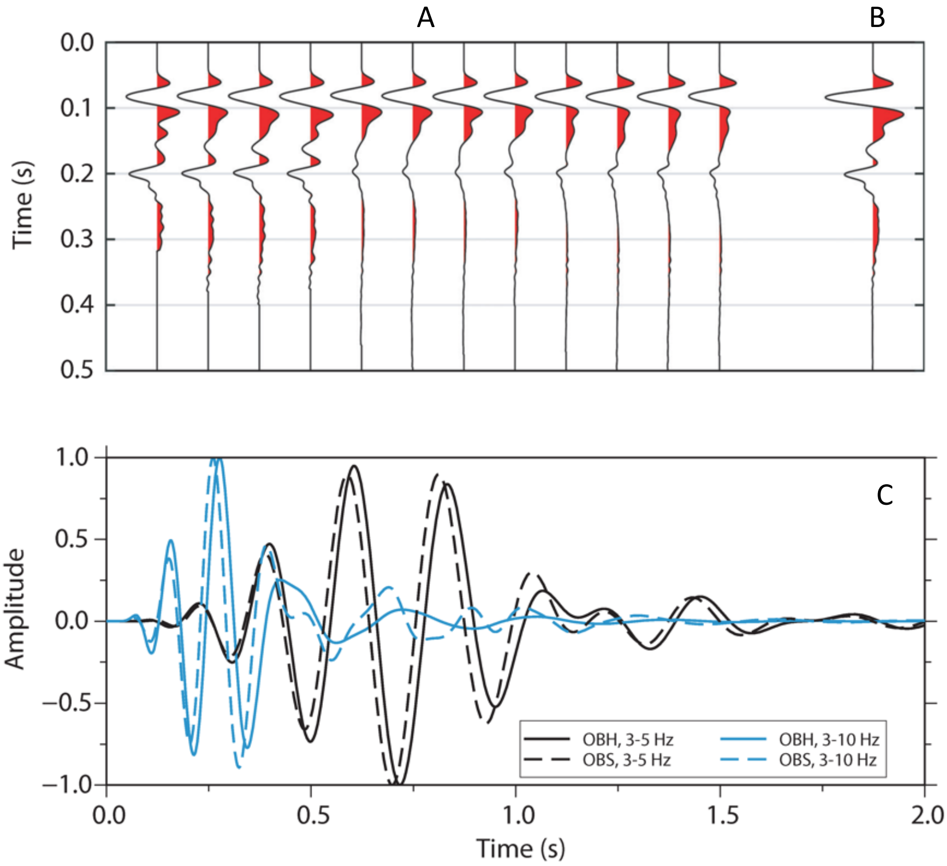


Figure S3. (A) Aligned seismic traces showing the near-offset direct water wave extracted from OBH gathers after filtering between 3-30 Hz. (B) Stack of traces in A. (C) The black and blue solid curves show the source wavelets for modelling of OBH data obtained by filtering the stacked signal in b to 3-5 Hz and 3-10 Hz, respectively. The black and blue dashed curves are the source wavelets for modelling of 3-5 and 3-10 Hz OBS data, respectively.

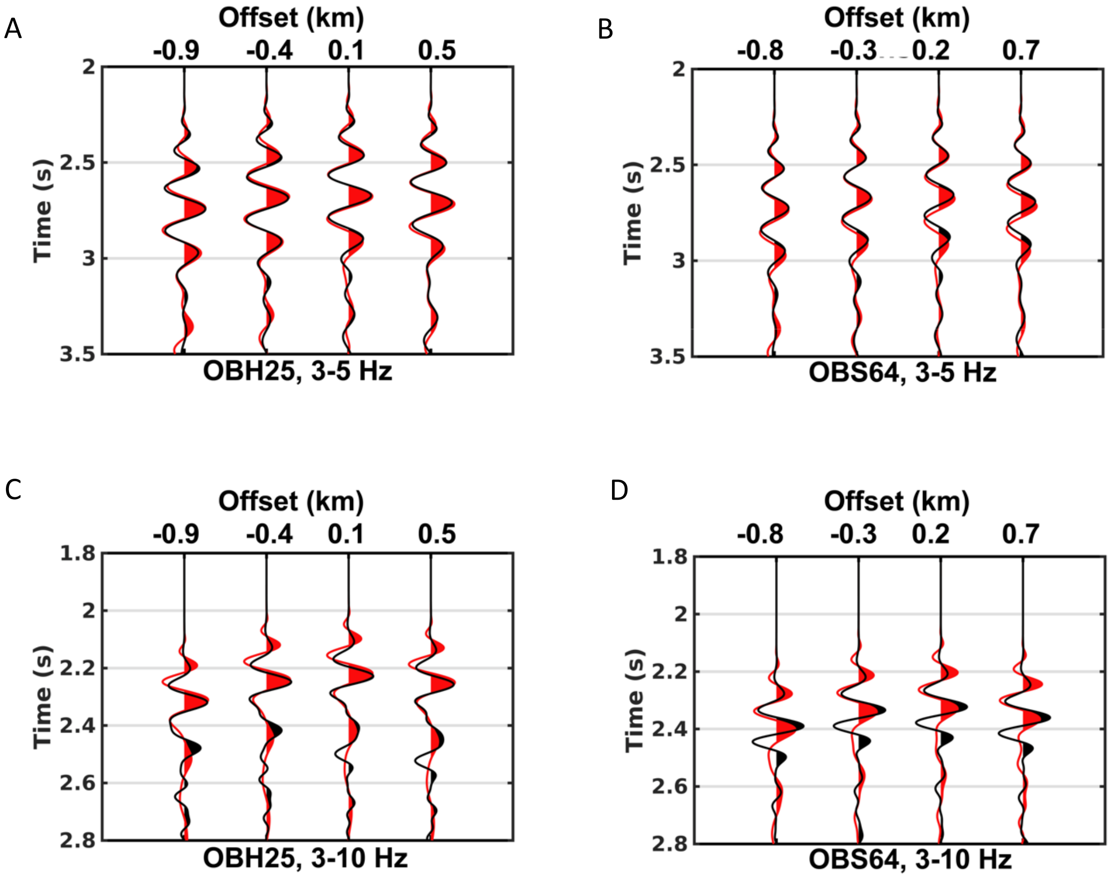


Figure S4. Comparisons of synthetic (in black) and observed (in red) near-offset water wave for OBH25 (A,C) and OBS64 (B,D). The synthetic data shown in A,B and C,D are modelled using the tomographic model using the 3-5 and 3-10 Hz source wavelets, respectively. The source wavelets are shown in Supplementary Fig. 4C. Correspondingly, the observed data are filtered to 3-5 Hz in A,B and to 3-10 Hz in C,D, respectively. The good match between the synthetic and observed data demonstrates the estimated source wavelets in Supplementary Fig. 4C are accurate enough for performing FWI.

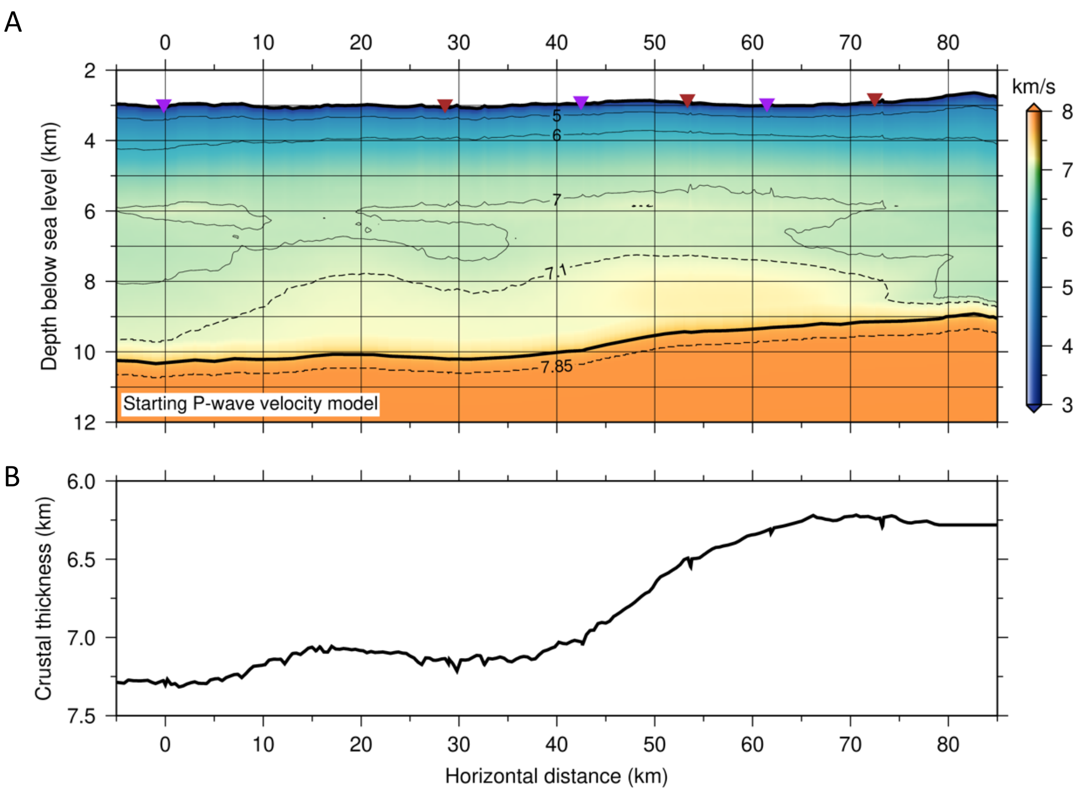
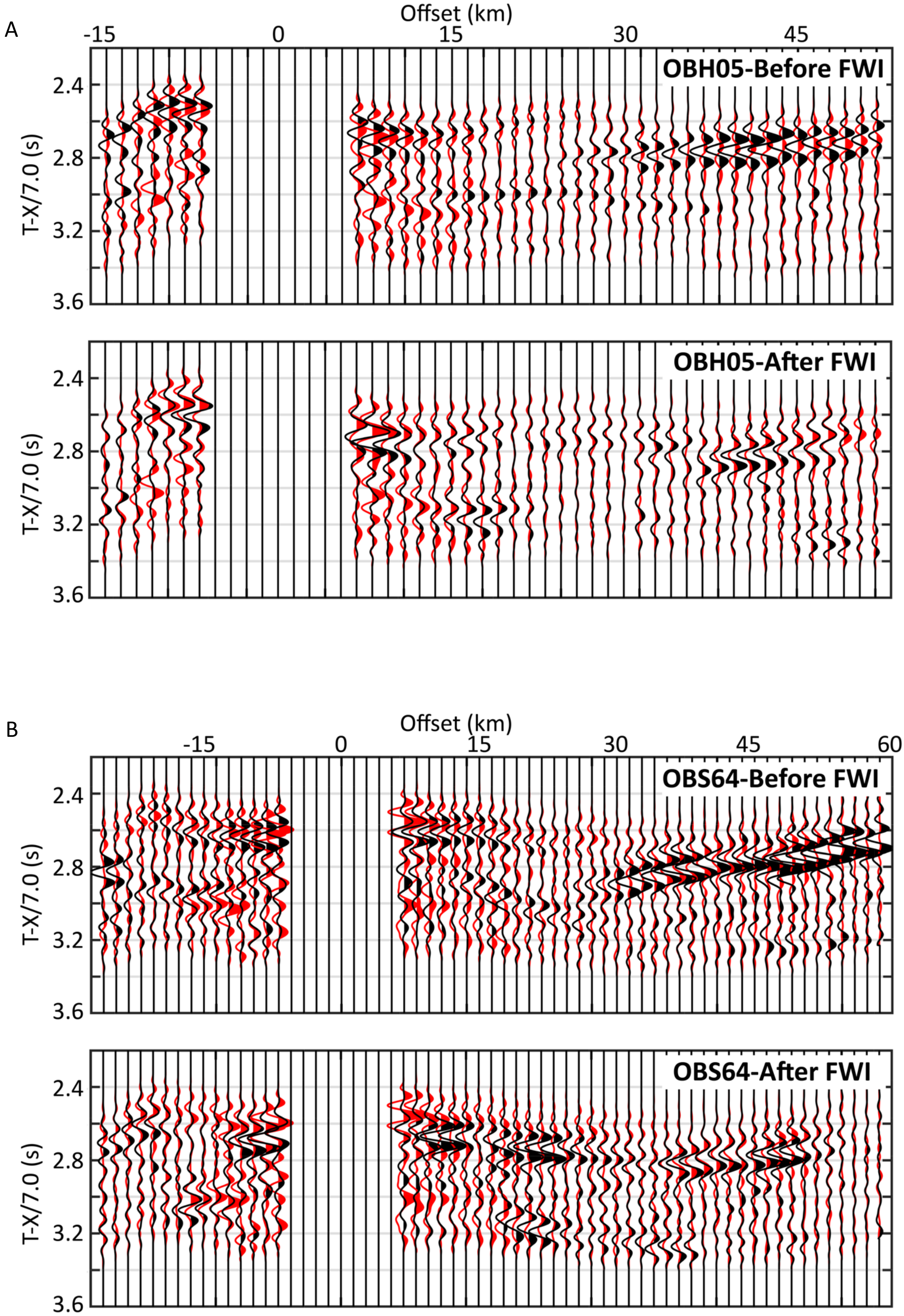
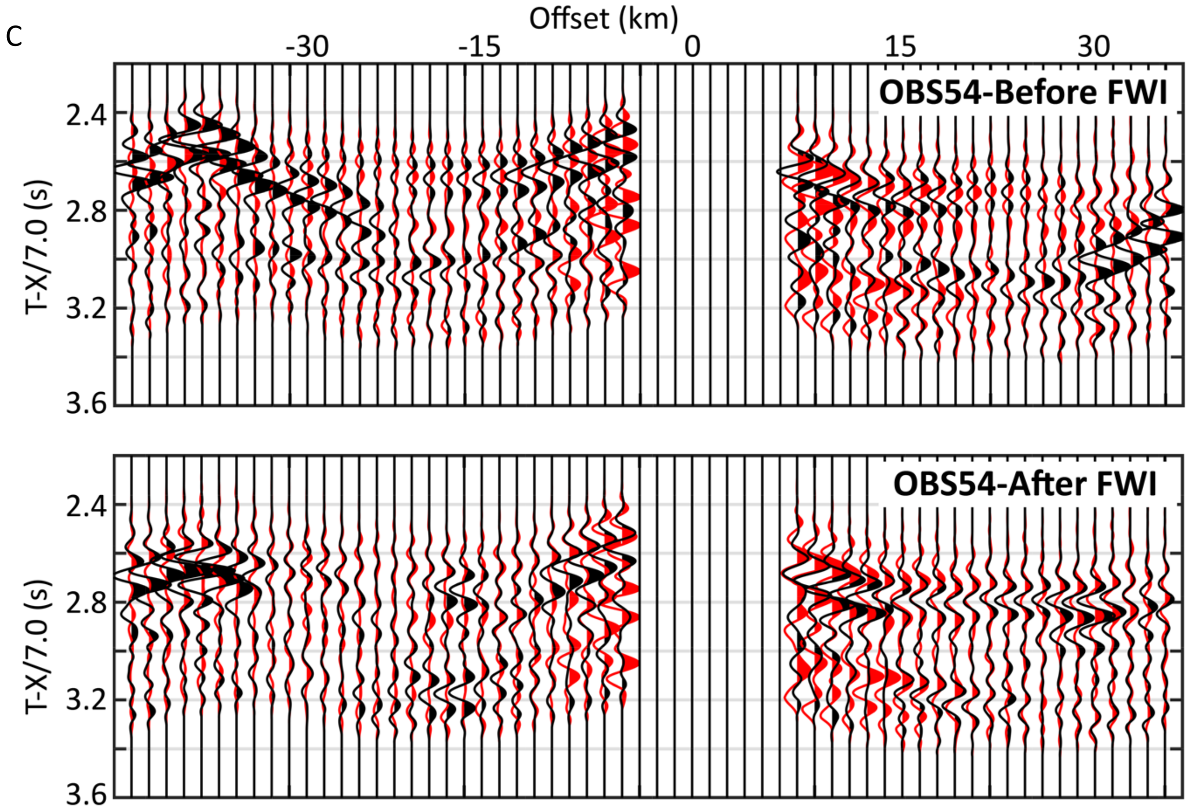
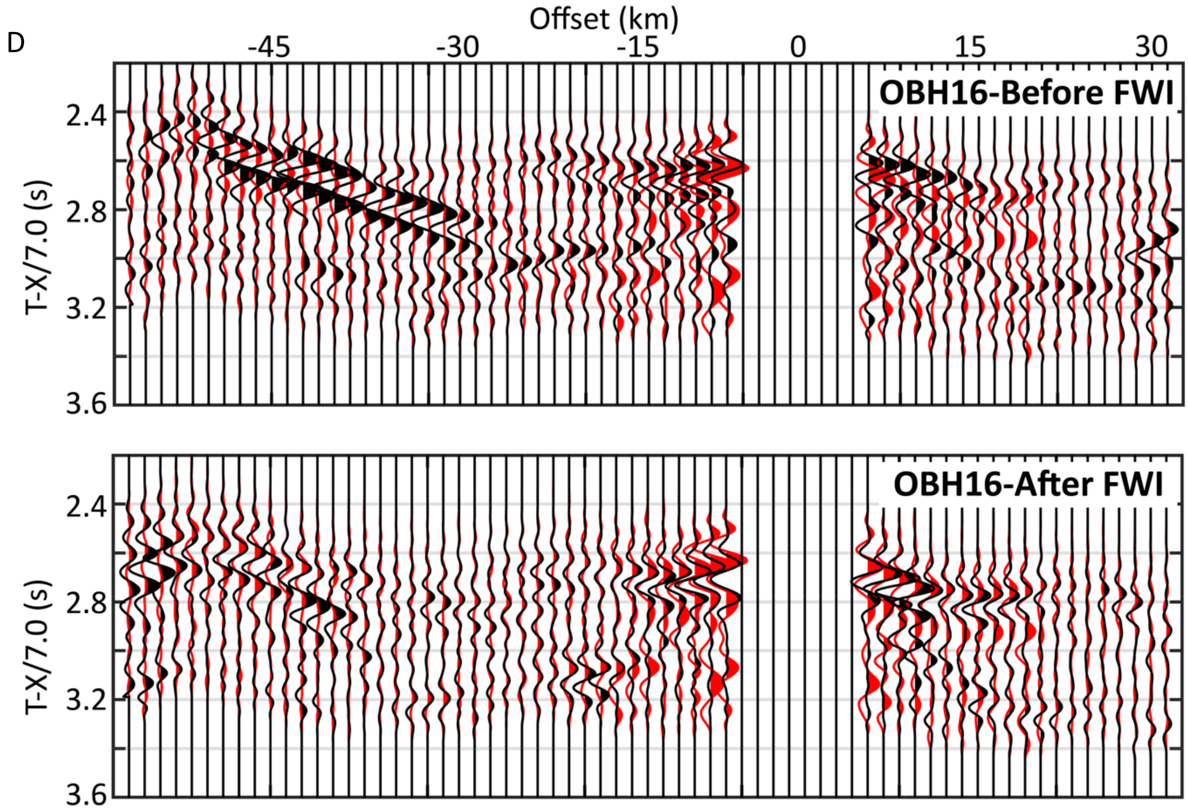


Figure S5. (A) Starting P-wave velocity model for FWI. The thick black curve is the tomographic Moho from *Canales et al.* [2003]. The brown and purple triangles show the locations of OBHs and OBSs, respectively. (B) Variation of crustal thickness obtained from travel time tomography [*Canales et al.*, 2003].


A
Offset (km)
-15
0
15
30
45
OBH05-Before FWI
OBH05-After FWI
T-X/7.0 (s)
2.4
2.8
3.2
3.6
B
Offset (km)
-15
0
15
30
45
60
OBS64-Before FWI
OBS64-After FWI
T-X/7.0 (s)
2.4
2.8
3.2
3.6


C
Offset (km)
-30
-15
0
15
30
OBS54-Before FWI
OBS54-After FWI
T-X/7.0 (s)
2.4
2.8
3.2
3.6


D
Offset (km)
-45
-30
-15
0
15
30
OBH16-Before FWI
OBH16-After FWI
T-X/7.0 (s)
2.4
2.8
3.2
3.6

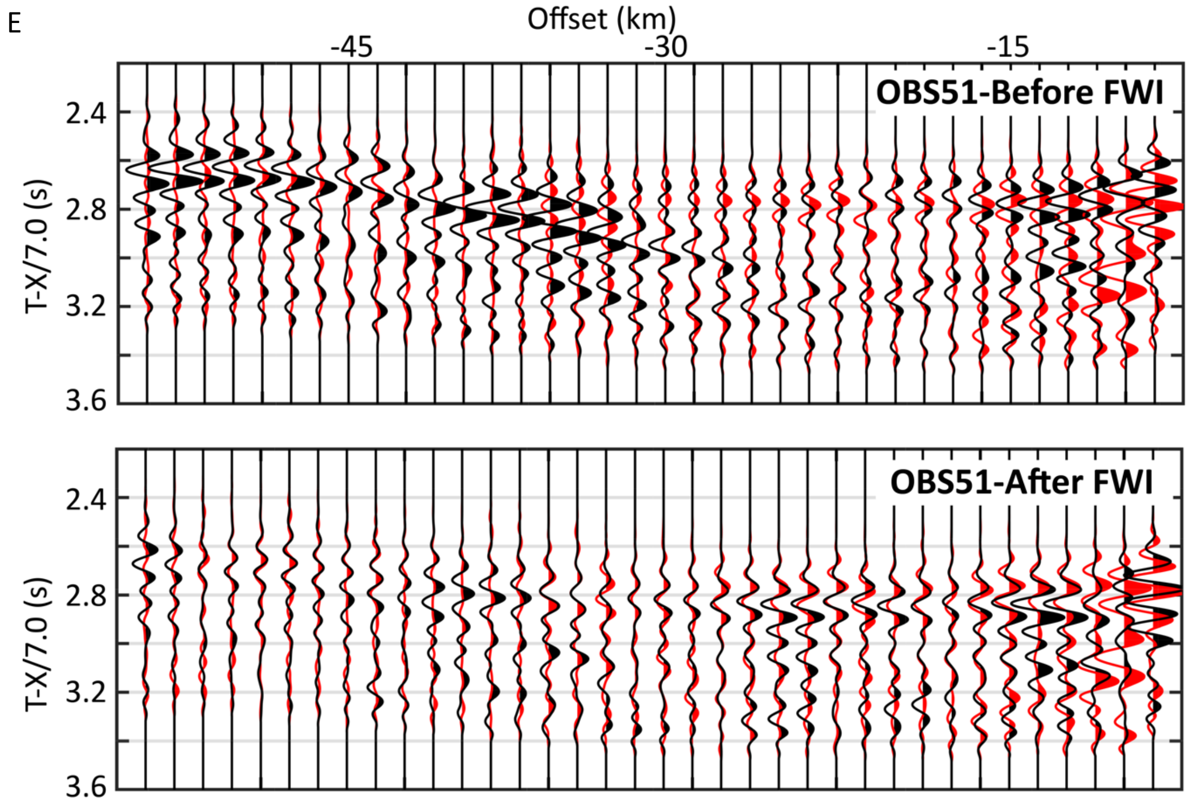


Figure S6. Comparisons of modelled (in black) and observed (in red) seismic data (3-10 Hz) before and after full waveform inversion (FWI). Travel time (T) of seismic data is reduced using a reduction velocity of 7.0 km/s. For better visibility, a scalar weighting factor (1+0.1×X) was multiplied for each trace to enhance the amplitude at large offsets, where X is offset. (A) OBH05; (B) OBS64; (C) OBS54; (D) OBH16; (E) OBS51.

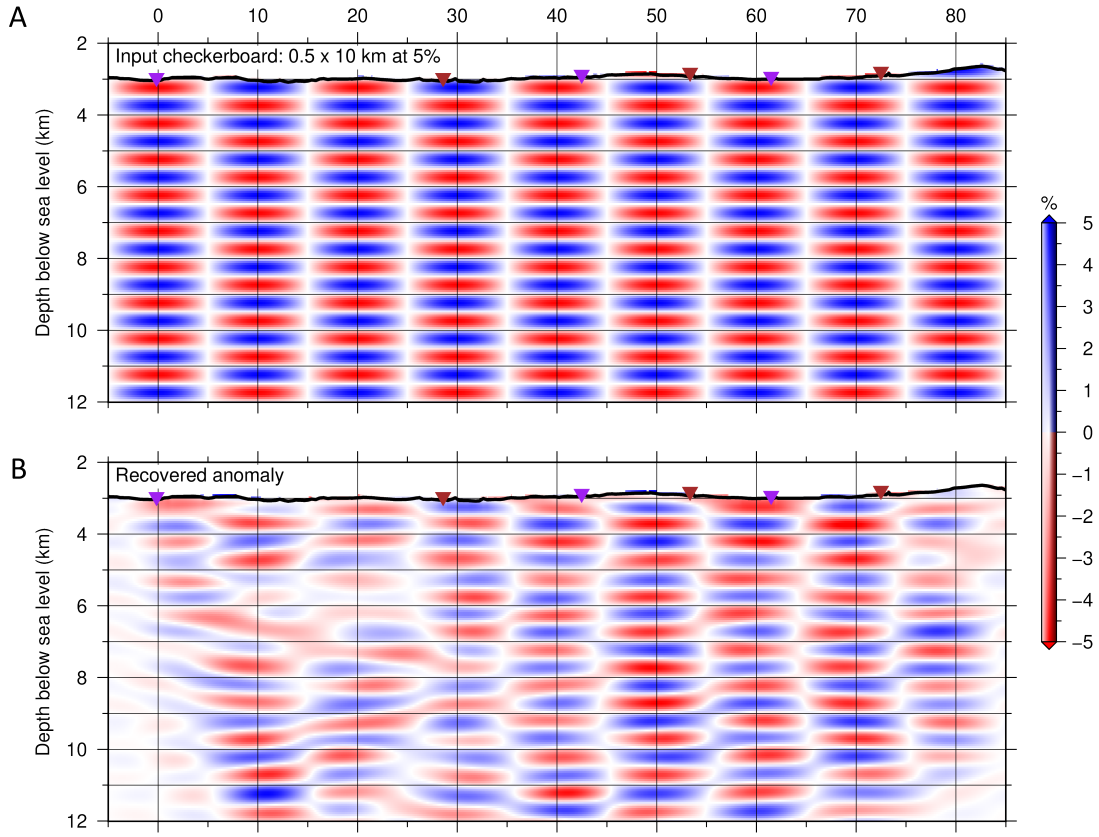


Figure S7. Checkerboard test using 0.5×10 km (vertical × horizontal) checkerboard pattern. Panels (A) and (B) show the input checkerboard pattern and the recovered anomaly, respectively. The maximum velocity perturbation is 5%. The brown and purple triangles show the locations of OBHs and OBSs, respectively.

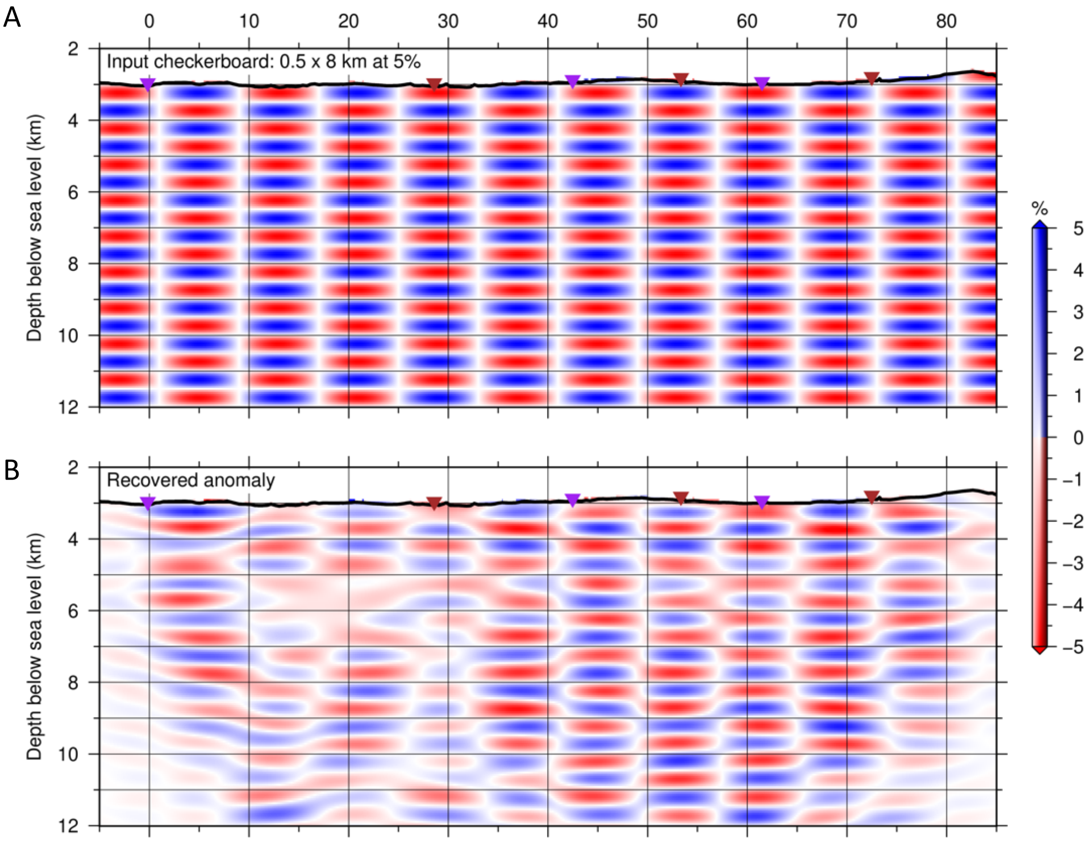


Figure S8. Checkerboard test using 0.5 × 8 km (vertical × horizontal) checkerboard pattern. Panels (A) and (B) show the input checkerboard pattern and the recovered anomaly, respectively. The maximum velocity perturbation is 5%. The brown and purple triangles show the locations of OBHs and OBSs, respectively.

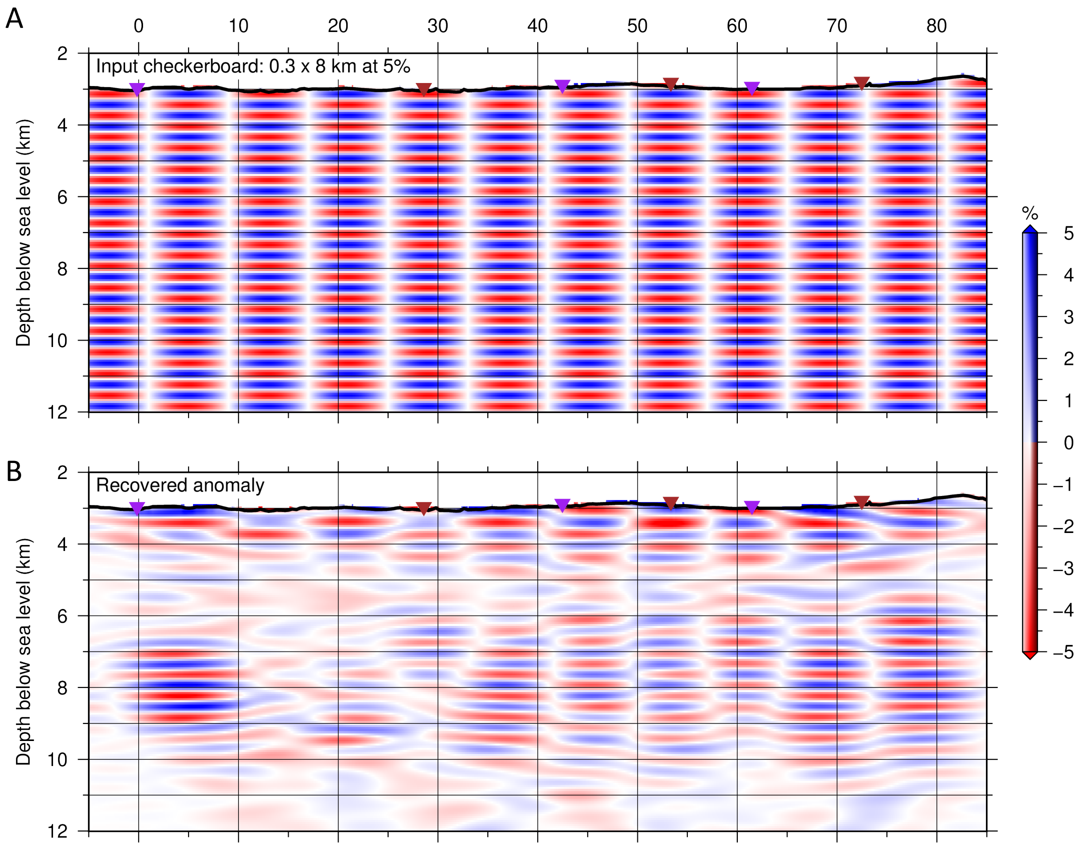


Figure S9. Checkerboard test using 0.3 × 8 km (vertical × horizontal) checkerboard pattern. Panels (A) and (B) show the input checkerboard pattern and the recovered anomaly, respectively. The maximum velocity perturbation is 5%. The brown and purple triangles show the locations of OBHs and OBSs, respectively.

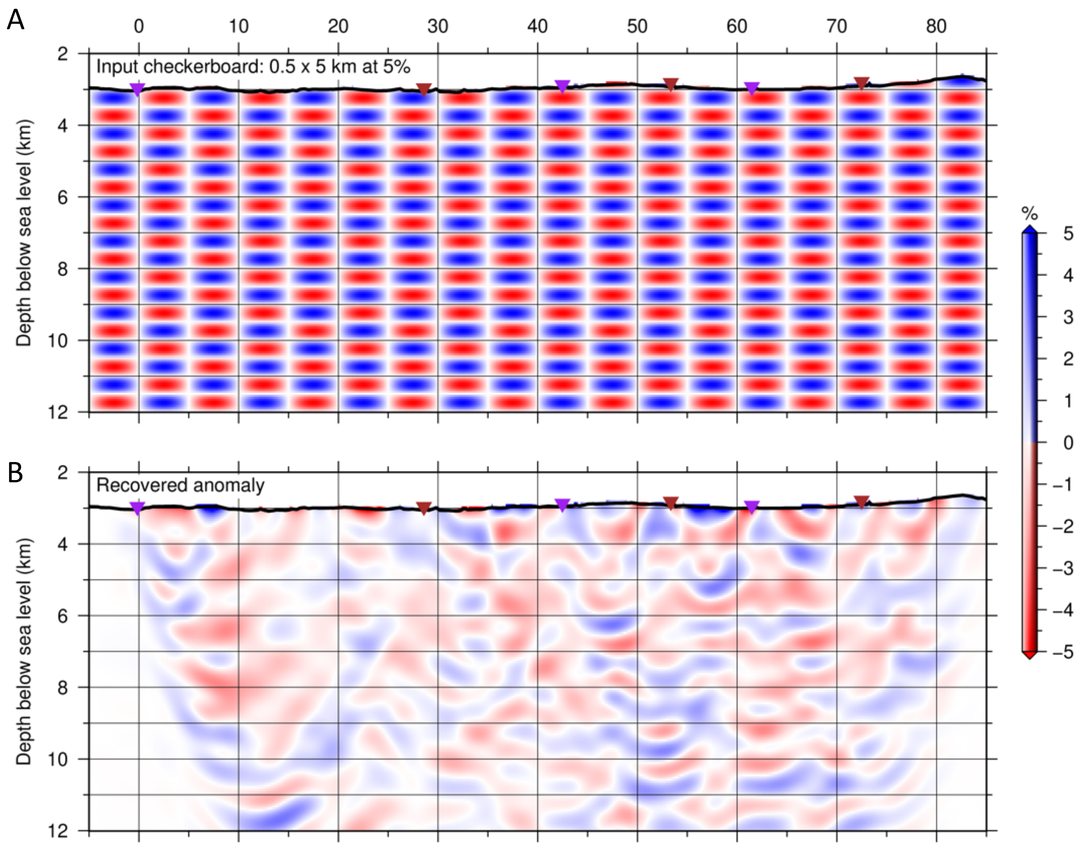


Figure S10. Checkerboard test using 0.5×5 km (vertical × horizontal) checkerboard pattern. Panels (A) and (B) show the input checkerboard pattern and the recovered anomaly, respectively. The maximum velocity perturbation is 5%. The brown and purple triangles show the locations of OBHs and OBSs, respectively.

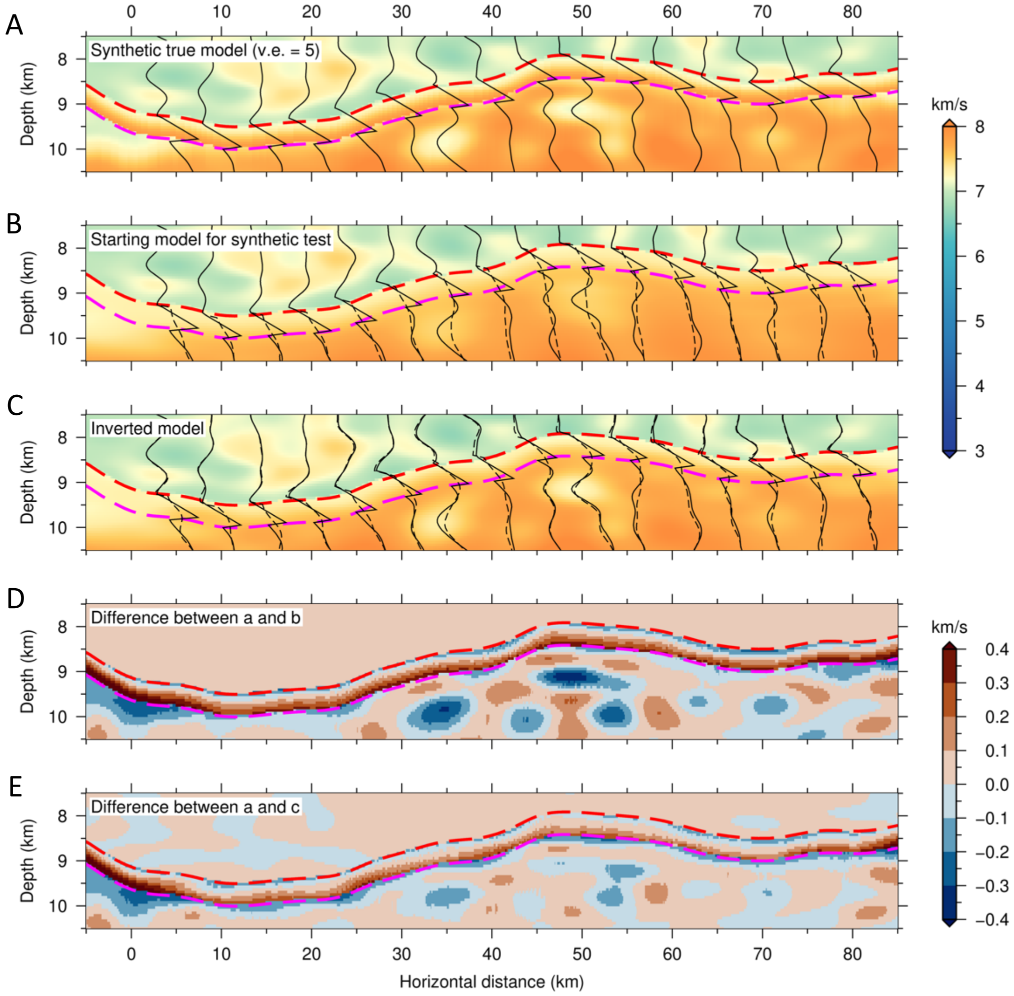


Figure S11. Synthetic test for the recovery of a 0.5 km thick Moho transition zone (MTZ). (A) True model for synthetic modelling which is modified by inserting a 0.5 km thick MTZ into the final model of full waveform inversion (FWI) of field data. Only the portion of the model around the MTZ is shown. The velocity of the inserted MTZ increases linearly with depth from 7.0 to 7.85 km/s. The 1-D profiles show vertical velocity profiles every 5 km between 5 and 80 km horizontal distance. (B) Starting model for synthetic test. The 1-D profiles compare the synthetic true (solid curves) and starting (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (C) Inverted model from FWI. The 1-D profiles compare the synthetic true (solid curves) and inverted (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (D) Difference between the synthetic true model (A) and the starting model (B). (E) Difference between the synthetic true model (A) and the inverted model (C). The red and magenta curves in A-E represent the top and bottom of the inserted 0.5 km thick MTZ, respectively.

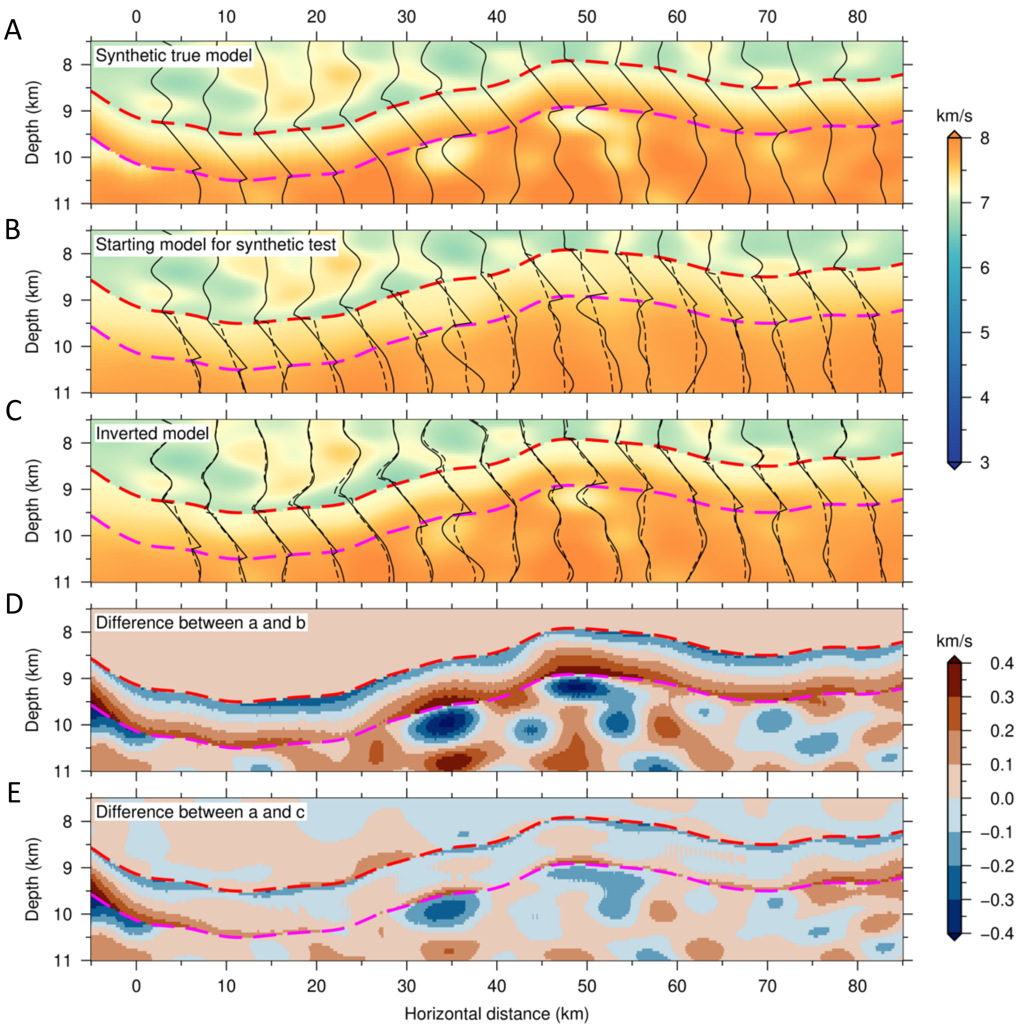


Figure S12. Synthetic test for the recovery of a 1.0 km thick Moho transition zone (MTZ). (A) True model for synthetic modelling which is modified by inserting a 1.0 km thick MTZ into the final model of full waveform inversion (FWI) of field data. Only the portion of the model around the MTZ is shown. The velocity of the inserted MTZ increases linearly with depth from 7.0 to 7.85 km/s. The 1-D profiles show vertical velocity profiles every 5 km between 5 and 80 km horizontal distance. (B) Starting model for synthetic test. The 1-D profiles compare the synthetic true (solid curves) and starting (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (C) Inverted model from FWI. The 1-D profiles compare the synthetic true (solid curves) and inverted (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (D) Difference between the synthetic true model (A) and the starting model (B). (E) Difference between the synthetic true model (A) and the inverted model (C). The red and magenta curves in A-E represent the top and bottom of the inserted 1.0 km thick MTZ, respectively.

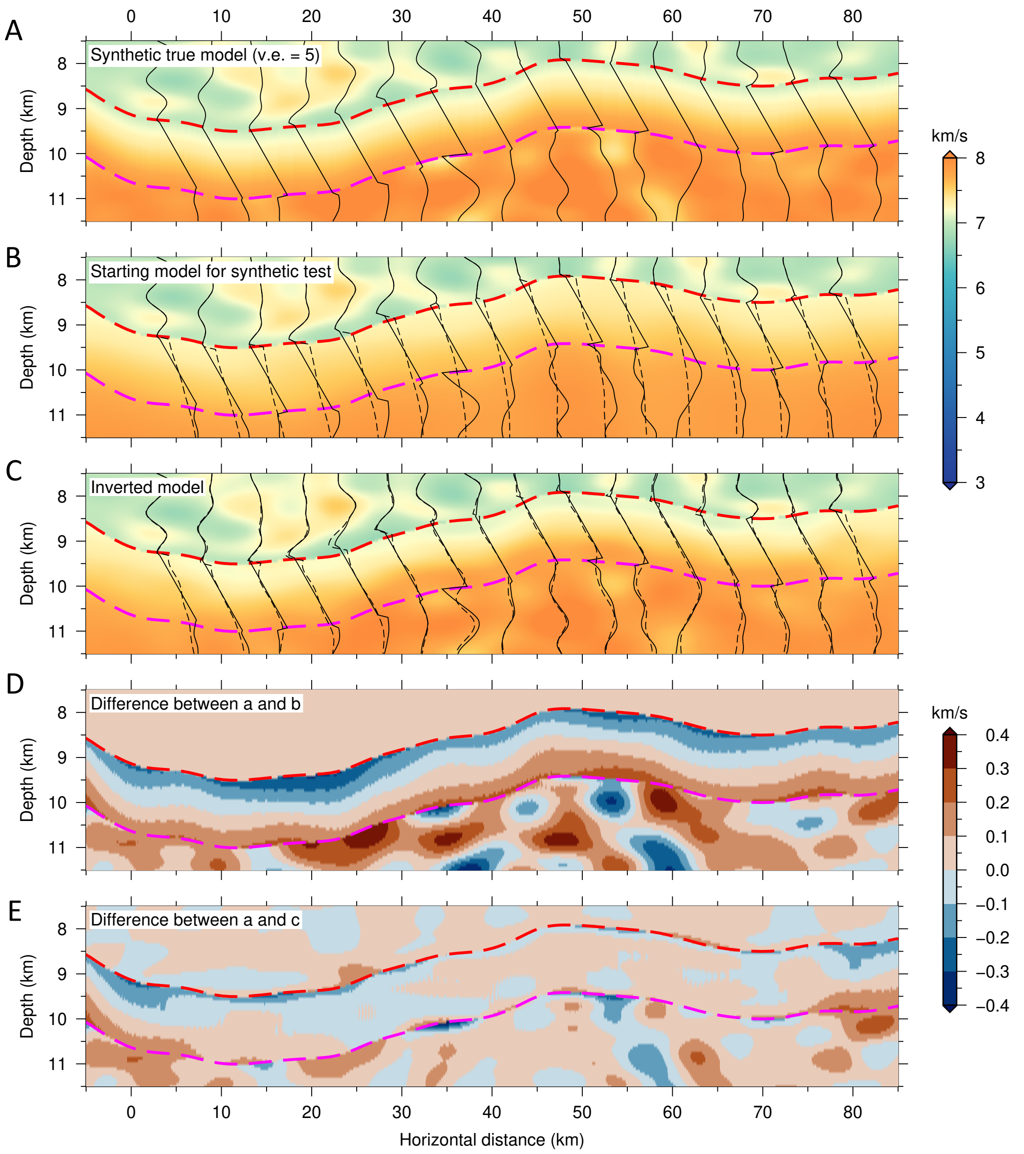


Figure S13. Synthetic test for the recovery of a 1.5 km thick Moho transition zone (MTZ). (A) True model for synthetic modelling which is modified by inserting a 1.5 km thick MTZ into the final model of full waveform inversion (FWI) of field data. Only the portion of the model around the MTZ is shown. The velocity of the inserted MTZ increases linearly with depth from 7.0 to 7.85 km/s. The 1-D profiles show vertical velocity profiles every 5 km between 5 and 80 km horizontal distance. (B) Starting model for synthetic test. The 1-D profiles compare the synthetic true (solid curves) and starting (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (C) Inverted model from FWI. The 1-D profiles compare the synthetic true (solid curves) and inverted (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (D) Difference between the synthetic true model (A) and the starting model (B). e: Difference between the synthetic true model (A) and the inverted model (C). The red and magenta curves in A-E represent the top and bottom of the inserted 1.5 km thick MTZ, respectively.

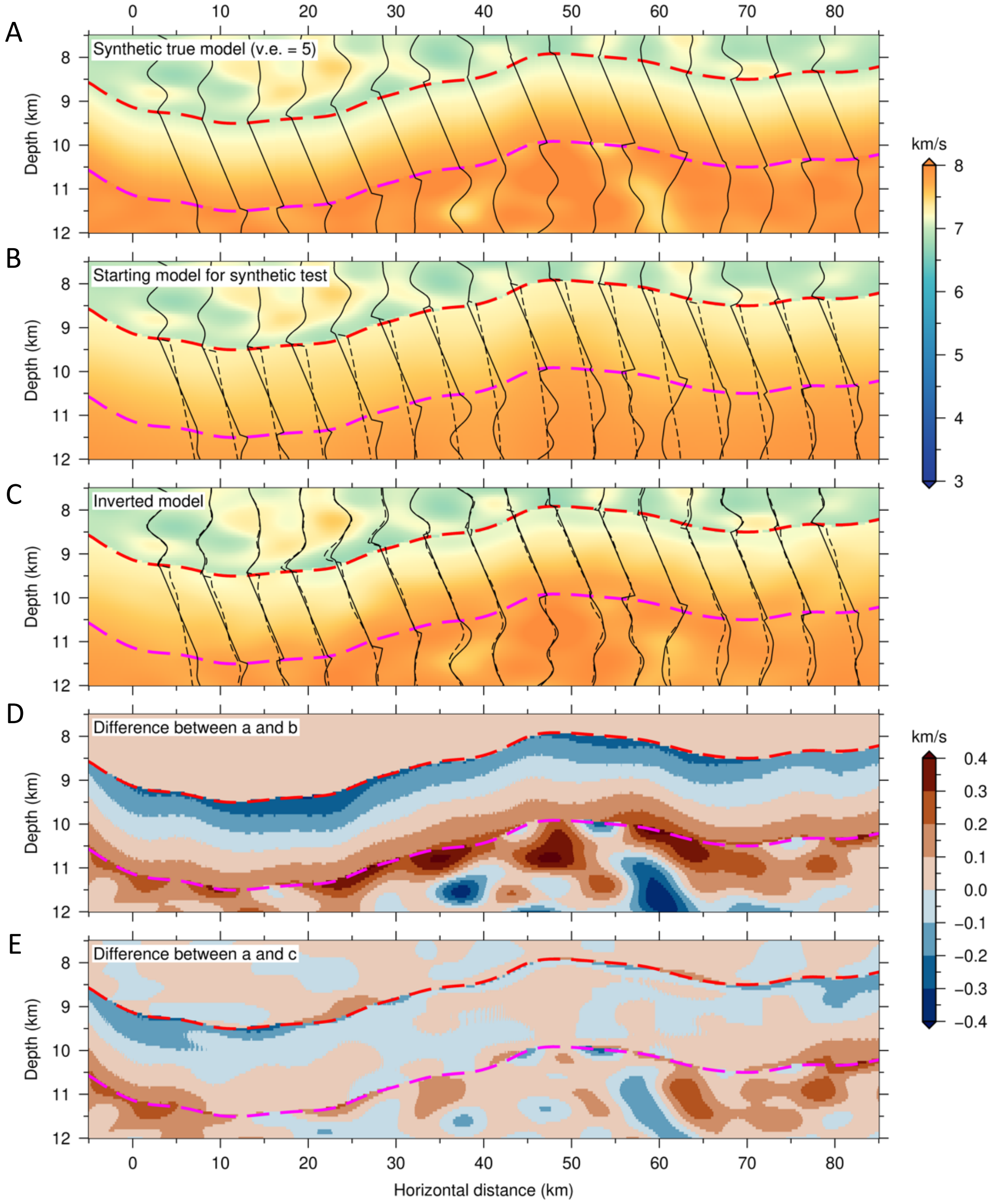


Figure S14. Synthetic test for the recovery of a 2.0 km thick Moho transition zone (MTZ). (A) True model for synthetic modelling which is modified by inserting a 2.0 km thick MTZ into the final model of full waveform inversion (FWI) of field data. Only the portion of the model around the MTZ is shown. The velocity of the inserted MTZ increases linearly with depth from 7.0 to 7.85 km/s. The 1-D profiles show vertical velocity profiles every 5 km between 5 and 80 km horizontal distance. (B) Starting model for synthetic test. The 1-D profiles compare the synthetic true (solid curves) and starting (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (C) Inverted model from FWI. The 1-D profiles compare the synthetic true (solid curves) and inverted (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (D) Difference between the synthetic true model (A) and the starting model (B). (E) Difference between the synthetic true model (A) and the inverted model (C). The red and magenta curves in A-E represent the top and bottom of the inserted 2.0 km thick MTZ, respectively.

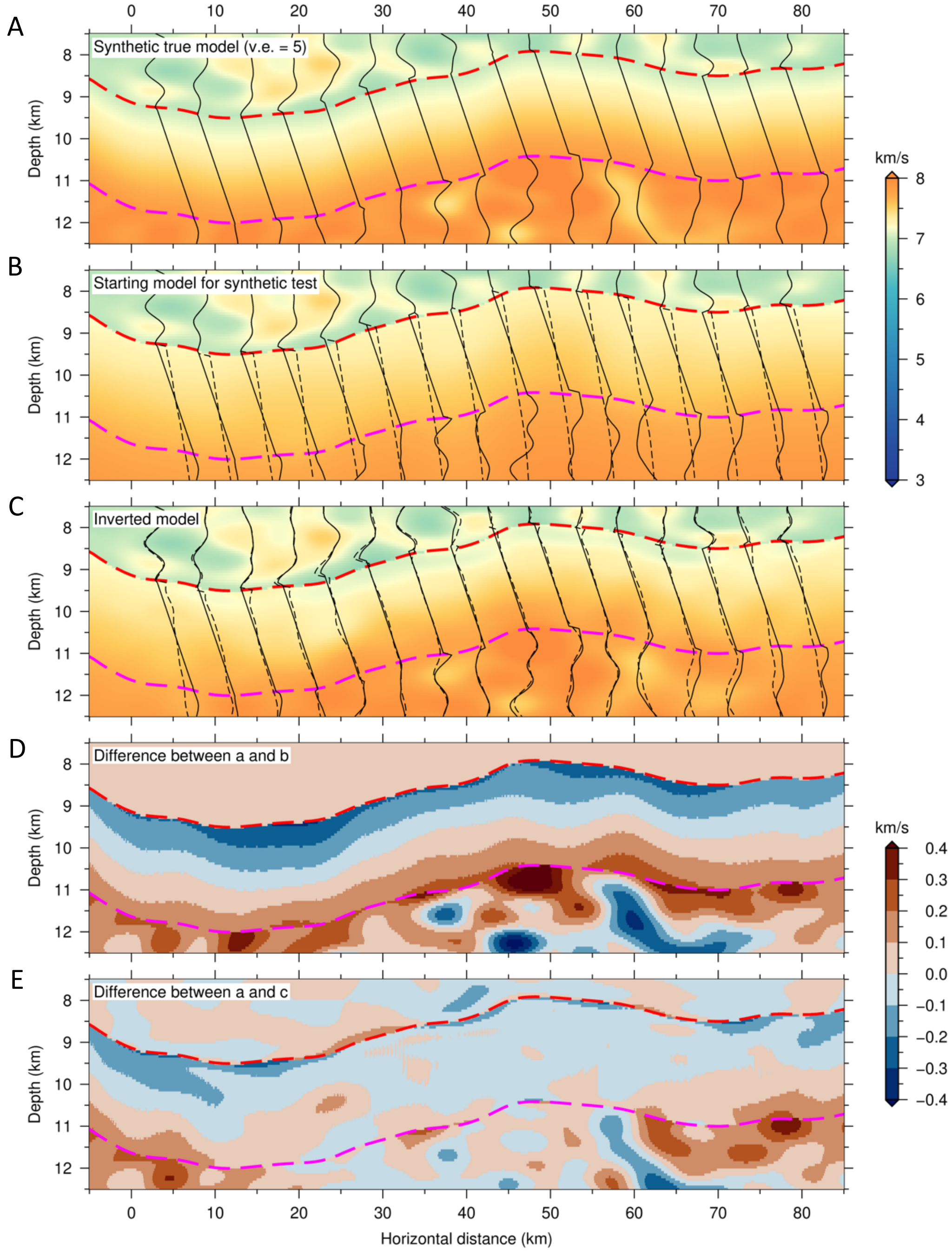


Figure S15. Synthetic test for the recovery of a 2.5 km thick Moho transition zone (MTZ). (A) True model for synthetic modelling which is modified by inserting a 2.5 km thick MTZ into the final model of full waveform inversion (FWI) of field data. Only the portion of the model around the MTZ is shown. The velocity of the inserted MTZ increases linearly with depth from 7.0 to 7.85 km/s. The 1-D profiles show vertical velocity profiles every 5 km between 5 and 80 km horizontal distance. (B) Starting model for synthetic test. The 1-D profiles compare the synthetic true (solid curves) and starting (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (C) Inverted model from FWI. The 1-D profiles compare the synthetic true (solid curves) and inverted (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (D) Difference between the synthetic true model (A) and the starting model (B). (E) Difference between the synthetic true model (A) and the inverted model (C). The red and magenta curves in A-E represent the top and bottom of the inserted 2.5 km thick MTZ, respectively.

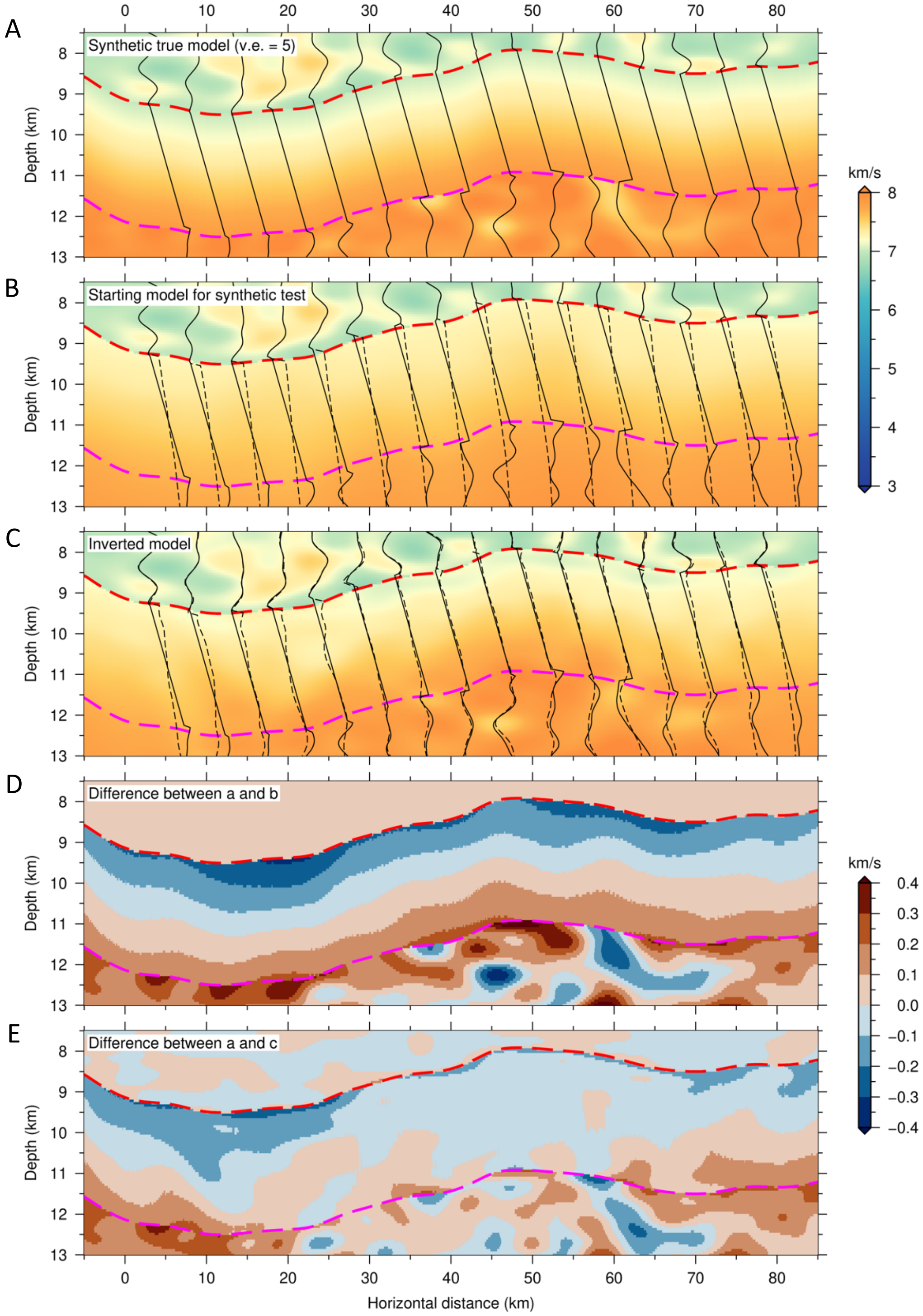


Figure S16: Synthetic test for the recovery of a 3.0 km thick Moho transition zone (MTZ). (A) True model for synthetic modelling which is modified by inserting a 3.0 km thick MTZ into the final model of full waveform inversion (FWI) of field data. Only the portion of the model around the MTZ is shown. The velocity of the inserted MTZ increases linearly with depth from 7.0 to 7.85 km/s. The 1-D profiles show vertical velocity profiles every 5 km between 5 and 80 km horizontal distance. (B) Starting model for synthetic test. The 1-D profiles compare the synthetic true (solid curves) and starting (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (C) Inverted model from FWI. The 1-D profiles compare the synthetic true (solid curves) and inverted (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (D) Difference between the synthetic true model (A) and the starting model (B). (E) Difference between the synthetic true model (A) and the inverted model (C). The red and magenta curves in A-E represent the top and bottom of the inserted 3.0 km thick MTZ, respectively.

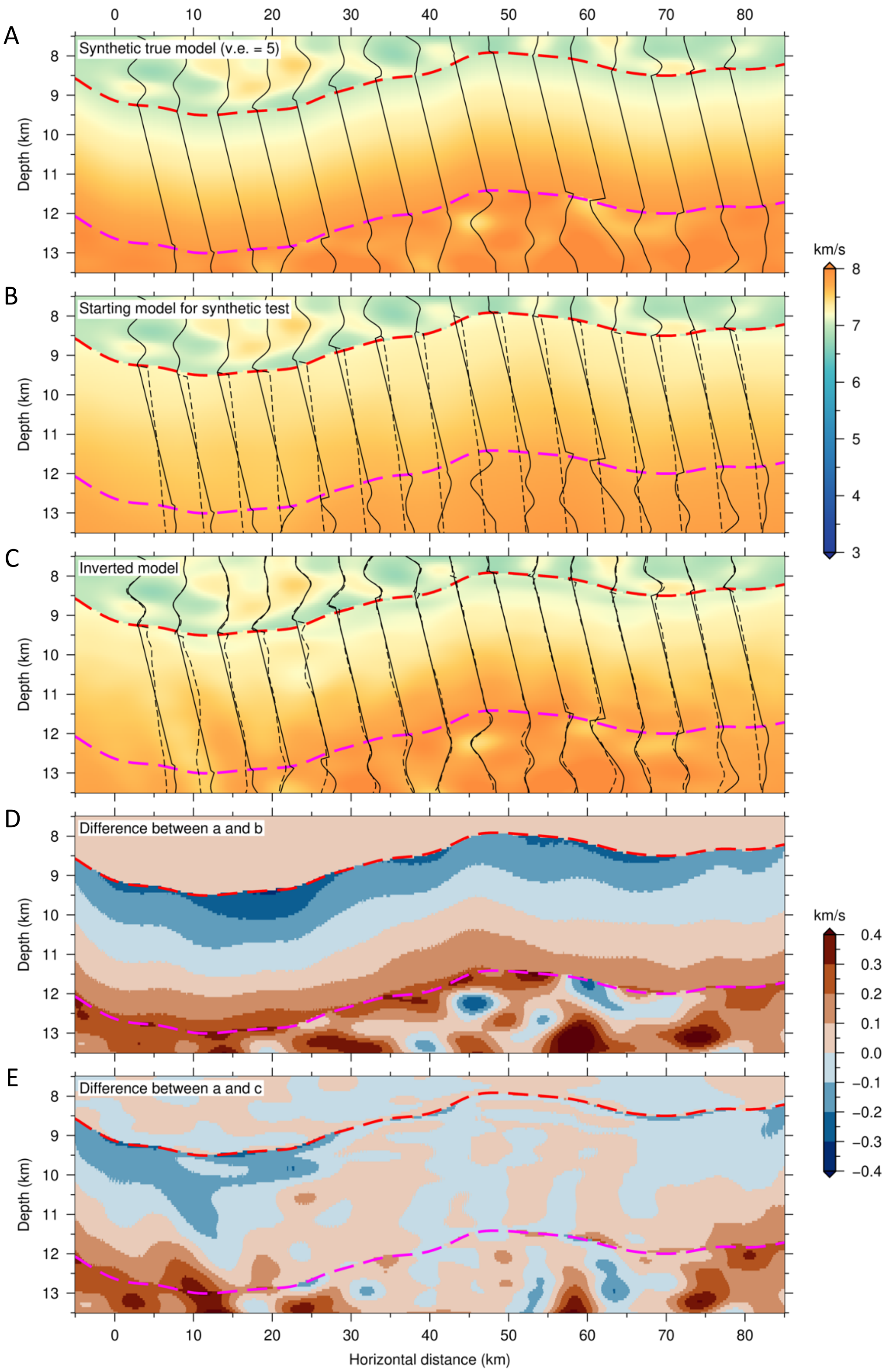


Figure S17. Synthetic test for the recovery of a 3.5 km thick Moho transition zone (MTZ). (A) True model for synthetic modelling which is modified by inserting a 3.5 km thick MTZ into the final model of full waveform inversion (FWI) of field data. Only the portion of the model around the MTZ is shown. The velocity of the inserted MTZ increases linearly with depth from 7.0 to 7.85 km/s. The 1-D profiles show vertical velocity profiles every 5 km between 5 and 80 km horizontal distance. (B) Starting model for synthetic test. The 1-D profiles compare the synthetic true (solid curves) and starting (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (C) Inverted model from FWI. The 1-D profiles compare the synthetic true (solid curves) and inverted (dashed curves) velocities every 5 km between 5 and 80 km horizontal distance. (D) Difference between the synthetic true model (A) and the starting model (B). (E) Difference between the synthetic true model (A) and the inverted model (C). The red and magenta curves in a-e represent the top and bottom of the inserted 3.5 km thick MTZ, respectively.

References

Brenders, A. J., and R. G. Pratt (2007), Efficient waveform tomography for lithospheric imaging: implications for realistic, two-dimensional acquisition geometries and low-frequency data, *Geophys. J. Int.*, *168*(1), 152-170, doi:10.1111/j.1365-246X.2006.03096.x.

Brocher, T. M. (2005), Empirical Relations between Elastic Wavespeeds and Density in the Earth's Crust, *Bull. Seismol. Soc. Am.*, *95*(6), 2081-2092, doi:10.1785/0120050077.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent (1995), Multiscale seismic waveform inversion, *Geophysics*, *60*(5), 1457-1473, doi:doi:10.1190/1.1443880.

Canales, P. J., R. S. Detrick, D. R. Toomey, and W. S. D. Wilcock (2003), Segment-scale variations in the crustal structure of 150–300 kyr old fast spreading oceanic crust (East Pacific Rise, 8°15′N–10°5′N) from wide-angle seismic refraction profiles, *Geophys. J. Int.*, *152*(3), 766-794, doi:10.1046/j.1365-246X.2003.01885.x.

Clayton, R., and B. Engquist (1977), Absorbing boundary conditions for acoustic and elastic wave equations, *Bull. Seismol. Soc. Am.*, *67*(6), 1529-1540.

Jian, H., M. R. Nedimović, J. P. Canales, and K. W. H. Lau (2021), New Insights Into the Rift to Drift Transition Across the Northeastern Nova Scotian Margin From Wide-Angle Seismic Waveform Inversion and Reflection Imaging, *J. Geophys. Res.*, *126*(12), e2021JB022201, doi:https://doi.org/10.1029/2021JB022201.

Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse (2009), Fast full-wavefield seismic inversion using encoded sources, *Geophysics*, *74*(6), WCC177-WCC188, doi:doi:10.1190/1.3230502.

Levander, A. R. (1988), Fourth-order finite-difference P-SV seismograms, *Geophysics*, *53*(11), 1425-1436, doi:10.1190/1.1442422.

Liu, Y., J. Teng, T. Xu, Y. Wang, Q. Liu, and J. Badal (2016), Robust time-domain full waveform inversion with normalized zero-lag cross-correlation objective function, *Geophys. J. Int.*, *209*(1), 106–122, doi:10.1093/gji/ggw485.

Scales, J. A. (1987), Tomographic inversion via the conjugate gradient method, *Geophysics*, *52*(2), 179-185, doi:10.1190/1.1442293.

Shipp, R. M., and S. C. Singh (2002), Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data, *Geophys. J. Int.*, *151*(2), 325-344, doi:10.1046/j.1365-246X.2002.01645.x.

Tao, K., S. P. Grand, and F. Niu (2017), Full-waveform inversion of triplicated data using a normalized-correlation-coefficient-based misfit function, *Geophys. J. Int.*, *210*(3), 1517-1524, doi:10.1093/gji/ggx249.

Tarantola, A. (1986), A strategy for nonlinear elastic inversion of seismic reflection data, *Geophysics*, *51*(10), 1893-1903, doi:10.1190/1.1442046.