
Walkaway VSP data conditioning for FWI

Raul Cova, Kris Innanen and Marianne Rauch-Davies

ABSTRACT

Full waveform inversion (FWI) applications on land seismic data remain very limited. The presence of strong anelastic effects, near-surface heterogeneities, unknown source and receiver signature and poor signal-to-noise ratio, among other reasons, challenge the capabilities of most modelling and inversion algorithms. Here, we perform an elastic FWI using land VSP data acquired in a walkaway configuration. We pre-process the data with the intent of improving the signal-to-noise ratio and removing undesired anelastic effects. Elevation differences among source locations were accounted for by applying elevation static corrections. Signal-to-noise ratio was improved by using a predictive filter in the FX domain. Two datasets with different deconvolution conditions were generated. A deterministic deconvolution using the recorded downgoing wavefield was applied to one of the datasets to remove the source signature. Even though this process partially accounts for changes in the wavelet with depth, a single operator is used for all the events recorded on a given trace. For this reason, we also computed a Gabor deconvolution to account for non-stationarity in the source signature. Then, we performed an elastic FWI using a multi-scale approach, with four frequency bands (4-8 Hz, 4-12 Hz, 4-16 Hz and 4-20 Hz) and three different depth windows (250-1000 m, 750-2250 m and 2000-3500 m). The FWI performed on the data deconvolved with the deterministic operators converged toward a solution that was closer to the sonic logs available in the well. Despite providing a wider frequency spectrum, the FWI using the Gabor deconvolved data did not converge toward an optimal solution. A closer examination of the input data revealed that in addition to removing some of the multiples, the deterministic deconvolution resurfaced some downgoing S-wave events that were not evident before. Providing data with less complexity and enhancing prominent events provided us with a more robust initialization of the inversion problem.

INTRODUCTION

Full waveform inversion (FWI) can be considered nowadays a well established method for computing high resolution earth models particularly in marine data processing. Successful results can be found in the literature that illustrate the potential of FWI in this environment (Ratcliffe et al., 2011; Operto et al., 2015; Routh et al., 2017). On land data, this goal is significantly more difficult. The reasons include, but are not limited to, very complex near-surface effects, unknown source and receiver signature, strong anelasticity and poor signal-to-noise ratio particularly for the frequencies in the low end of the spectrum.

Vertical seismic profiles (VSP) provide data where the effects of some of these challenges are reduced. Since in this type of survey receivers are located at depth within the borehole and sources are activated at the surface, near-surface heterogeneities only affect the source side of the wavefield propagation. Also, since receivers are placed in the borehole they are protected against most of the surface-related noise sources providing good signal-to-noise ratios. However, VSP data lack the ample spatial coverage of surface seis-

mic, and other sources of noise such as tube-waves can also contaminate the data. Similarly, a poor coupling of the receivers to the borehole walls can result in very weak signals.

Podgornova et al. (2014) and Pan et al. (2018) report successful results in performing elastic FWI over land VSP datasets. Here, we focus on the pre-processing needed for the FWI to provide accurate results. In particular, the dataset we use presents very large and frequent velocity contrasts that result in very energetic short-wavelength multiples. Contamination by tube-waves and poor signal-to-noise ratio are other issues that are addressed in this report.

THEORY

FWI is an iterative inversion method that aims at recovering the subsurface model parameters (\mathbf{m}) that reproduce a set of multicomponent seismic data ($\mathbf{d}(\mathbf{x}_r, t)$) recorded at N stations \mathbf{x}_r with $r = 1, \dots, N$. Different metrics can be used to determine the goodness of fit of the observed data to the data reproduced with the inverted model parameters ($\mathbf{s}(\mathbf{x}_r, t, \mathbf{m})$). Here, we use the waveform misfit function (Nolet, 1987) defined as the L_2 norm of the data residuals as the objective function $\Phi(\mathbf{m})$ to be minimized,

$$\Phi(\mathbf{m}) = \frac{1}{2} \sum_{r=1}^N \int_0^T \|\mathbf{s}(\mathbf{x}_r, t, \mathbf{m}) - \mathbf{d}(\mathbf{x}_r, t)\|^2 dt. \quad (1)$$

This objective function is minimized by computing successive updates to the model parameters. At each k iteration a new model is computed as,

$$\mathbf{m}_{k+1} = \mathbf{k} + \mu \Delta \mathbf{m}_k \quad (2)$$

where μ is referred to as the step length and $\Delta \mathbf{m}$ is the search direction. The step length can be obtained with line search methods (Nocedal and Wright, 2006) and the search direction can be computed as,

$$\Delta \mathbf{m}_k = -\mathbf{H}_k^{-1} \nabla_{\mathbf{m}} \Phi_k \quad (3)$$

where \mathbf{H} and $\nabla_{\mathbf{m}}$ are the Hessian and gradient of the misfit function for the current model \mathbf{m} . Since the Hessian contains the second derivatives of the objective function with respect to each model parameter its computational cost is extremely expensive. Here we use the L-BFGS method (Nocedal and Wright, 2006) to compute a low-rank approximation of the Hessian.

In this study we use the FWI implementation developed by Yuan et al. (2016) and Pan et al. (2018) which exploits the computation of the misfit kernels and forward modelling capabilities of the *Specfem2d* package using the spectral element method (Komatitsch and Tromp, 1999). The expressions for the gradient and the approximate Hessian are developed and presented in Tromp et al. (2005) and Shin et al. (2001), respectively.

FIELD DATA

The data used in this study consists of 62 source gathers recorded using a VSP walk-away/offset configuration. Only 61 source gathers were acquired in a north-south walk-

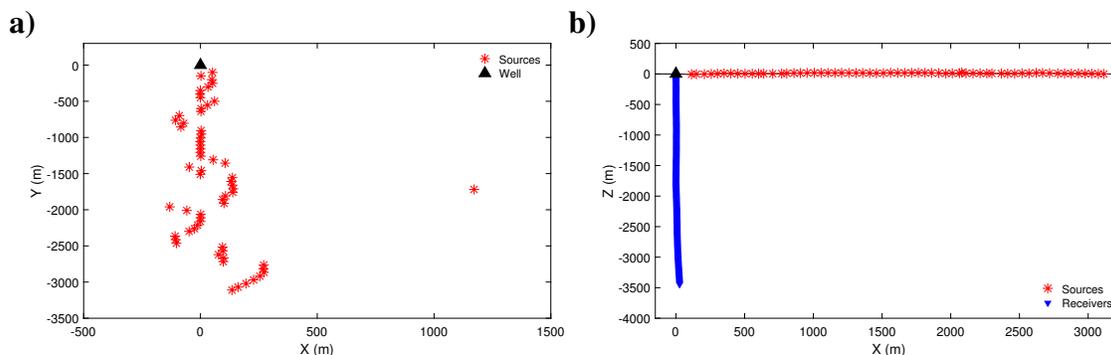


FIG. 1. (a) Map and (b) profile view of the walkaway VSP acquisition geometry. Only the receivers between 2302.7 m and 3445.6 m depth were live for all the shots.

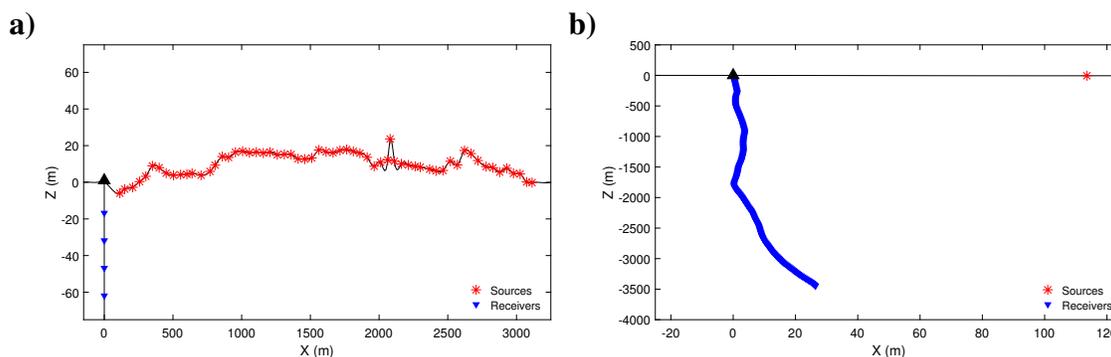


FIG. 2. Zoom around the (a) source and (b) receiver locations. The spike around $X = 2000$ m correspond to the elevation of the off-line source location

away line and one extra source gather was acquired further east at an azimuth of 34° with respect to the walk-away line (Figure 1a). Source points in the walk-away line were spaced every 50 m with a maximum source-well offset of 3113.5 m. Figure 1b displays a profile view of the receivers deployed in the well. Only the source points 1 (113 m offset), 35 (1812.5 m) and 62 (2080.3 m offset, off-line) were recorded using 229 receiver positions spaced every 15 m from 16.7 m to 3445.6 m depth. The rest of the walk-away source points were recorded only in the bottom section of the well between 2302.7 and 3445.6 m for a total of 77 geophone levels.

The source used was a vibroseis truck with a linear sweep between 2 Hz and 140 Hz with a length of 16 s. The listening time was 6 s for a total record length of 22 s sampled at 1 ms. The receivers consisted of an array of 43 three-component digital geophones (DS-150) that were deployed at 6 different levels to cover the total length of the well.

Figure 2 displays a zoom around the source and receiver locations. The maximum elevation difference between the source points and the well elevation is of 17.8 m for the sources in the walk-away line. The spike in the elevation profile around the offset 2000 m corresponds to the off-line source point with an elevation of 23.6 m. The average elevation in the walk-away line is of only 9.7 m.

In Figure 2b the receivers layout is displayed. Notice how starting at 1750 m depth the receiver locations start to deviate from the nearly vertical trajectory reaching a maximum

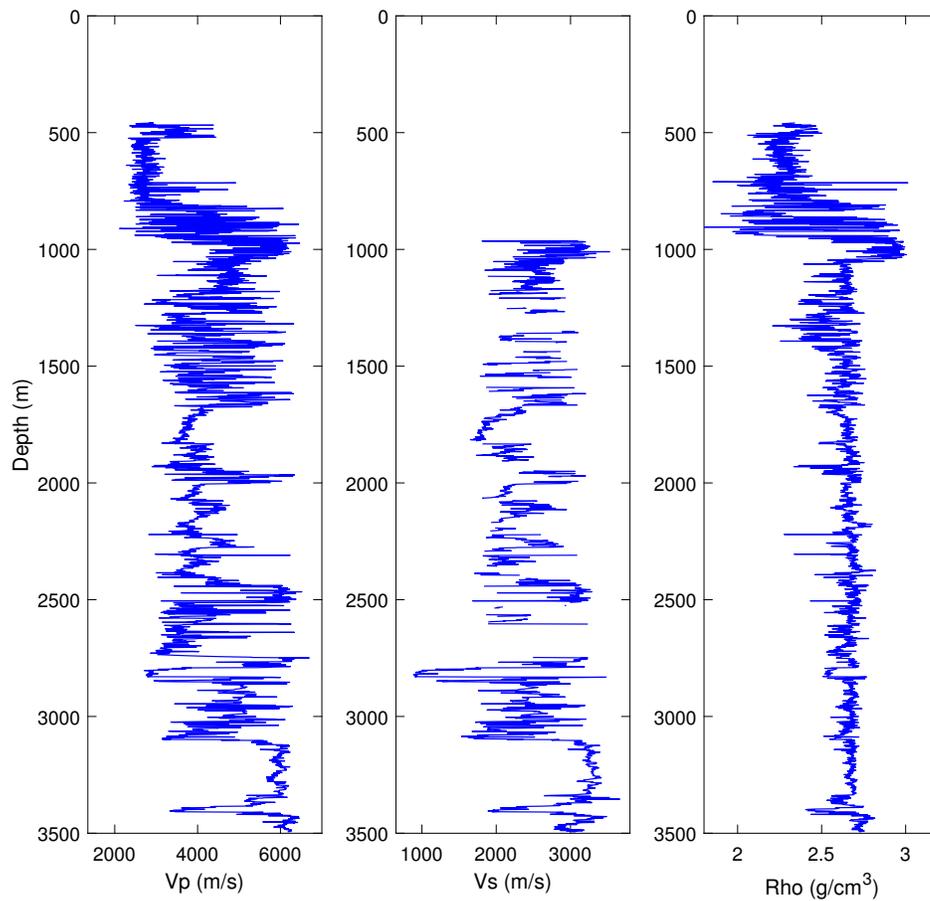


FIG. 3. (a) P-wave velocity, (b) S-wave velocity and (c) density logs available in the well.

deviation of 26.4 m registered at the deepest receiver location.

Figure 3 displays the set of logs available in the well. Most of the logs cover the section between 478 m and 3500 m. However, the shallow section between 478 and 900 m is missing in the Vs log as well as on other smaller intervals along the well. From the P-wave velocity log it is clear that there are rapidly alternating high and low velocity sediments in the stratigraphic column, particularly between 1200 m and 1600 m depth. There is also a significant increase in the velocity and density values around 1000 m which is related to the presence of an anhydrite layer in the sedimentary column.

Receiver components rotations

The first step in the processing workflow consisted of projecting the amplitudes recorded by the horizontal components of the geophones toward the source-receiver plane. This step aims to correct differences in the polarity and magnitudes of the recorded data due to the random orientation of the horizontal components of the geophones in the wellbore. The rotation angle is computed by analyzing the hodograms resulting from the amplitudes

recorded within a window defined around the first arrival picked on the vertical component data. Figure 4 displays a subset of source gathers before the rotation of the H_1 and H_2 components of the data. Notice the random character of the amplitudes in panels b) and c) of Figure 4. The data after rotation is displayed in Figure 5. The rotated horizontal components are now referred to as H_{max} and H_{min} . Very coherent P and S-wave energy can now be seen in the H_{max} component. On the other hand, the H_{min} component mostly contains off-plane S-wave energy. Since the inversion algorithm used in this study is 2D, only the energy contained in the source-receiver plane can be modelled, therefore we only use the vertical and H_{max} components for the inversion.

Lastly, traces were multiplied by \sqrt{t} to compensate for extra 3D spherical divergence that is not modelled by 2D algorithms. Also, a convolution with $1/\sqrt{t}$ was performed to move the data from a point source to a line source condition (Pica et al., 1990).

Near-surface corrections

Elevation corrections were computed at each source location using the elevation profile displayed in Figure 2a. The replacement velocity used was of 1816 m/s, which was the shallowest velocity measured with the VSP first arrivals.

Figure 6 shows the magnitude of the computed elevation corrections. Notice that they range between -10 ms and 3 ms. In the light of these results we applied these corrections to the data and decided to proceed with the inversion assuming a flat acquisition surface. We made this choice because the maximum elevation difference in the data is smaller than the element size used for the inversion. In our opinion, the magnitudes of the elevation correction, neither in time nor distance, were large enough to justify the increase in computational cost of using an irregular topography for the inversion.

After computing elevation corrections we also computed a set of residual static corrections. For this, we organized the data into receiver gathers and applied an NMO correction. Figure 7 shows the results for the shallowest receiver (receiver 153 at 2302 m) that was live for all the source locations. As we can see, even after applying elevation corrections the data still display frequent time shifts between consecutive traces.

To correct for this effect we windowed the data around the NMO-corrected first arrivals (Figure 8a) and created a reference or pilot wavefield by performing a lateral moving average operation. The window length of the moving average was set at only 5 traces to ensure that we only removed the short-wavelength components of the time shifts. Figure 8b displays the resulting pilot traces. Notice that there still exists some long wavelength travel time effects which might be the results of lateral velocity changes or due to the presence of anisotropy.

Finally, the NMO-corrected first arrivals and the pilot traces are crosscorrelated (Figure 8c). The time lag of the peak of the crosscorrelation functions is extracted and interpreted to represent the residual static correction for each source location included in the receiver gather.

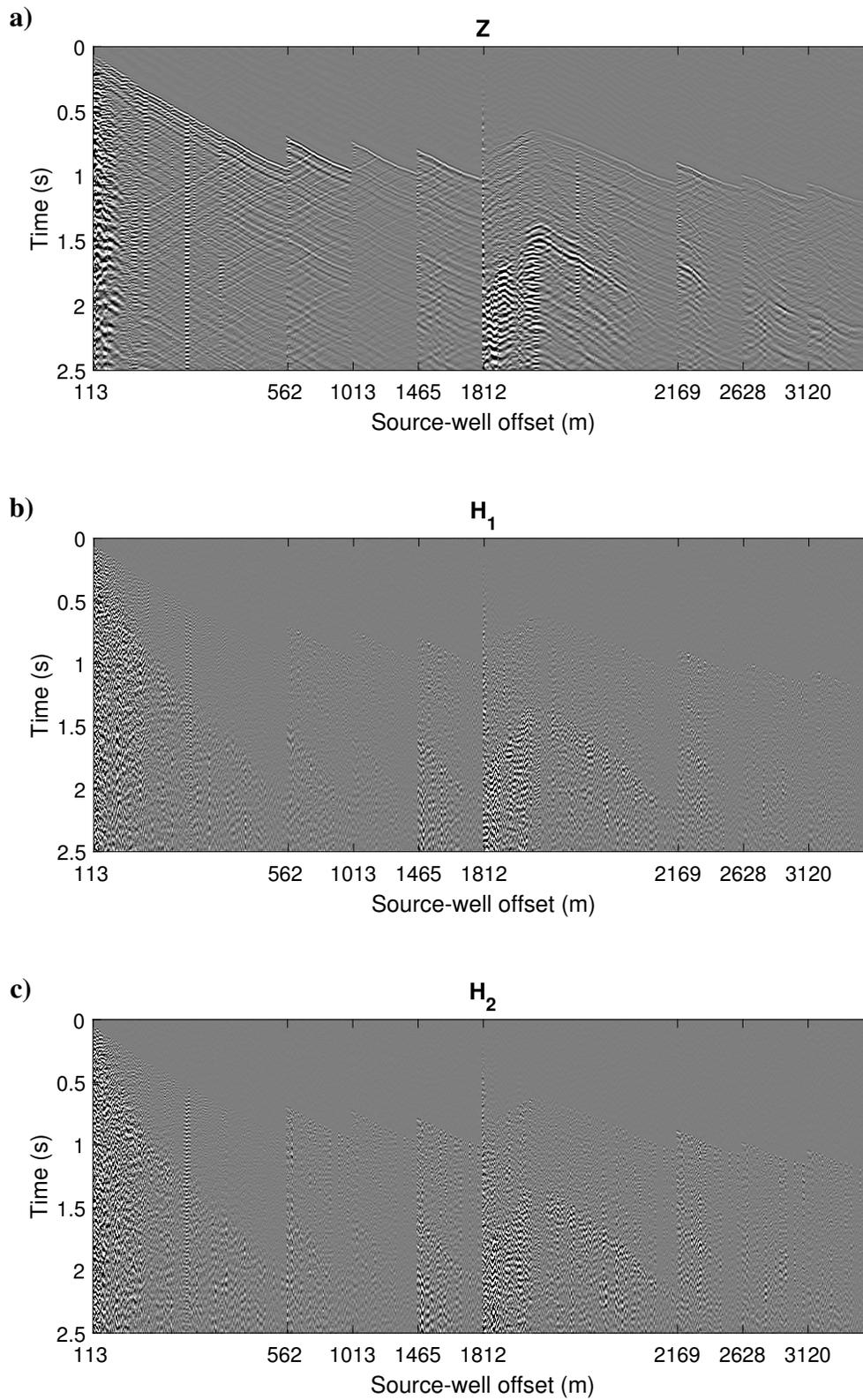


FIG. 4. Data before rotations. (a) Vertical, (b) Horizontal-1 (c) Horizontal-2 component.

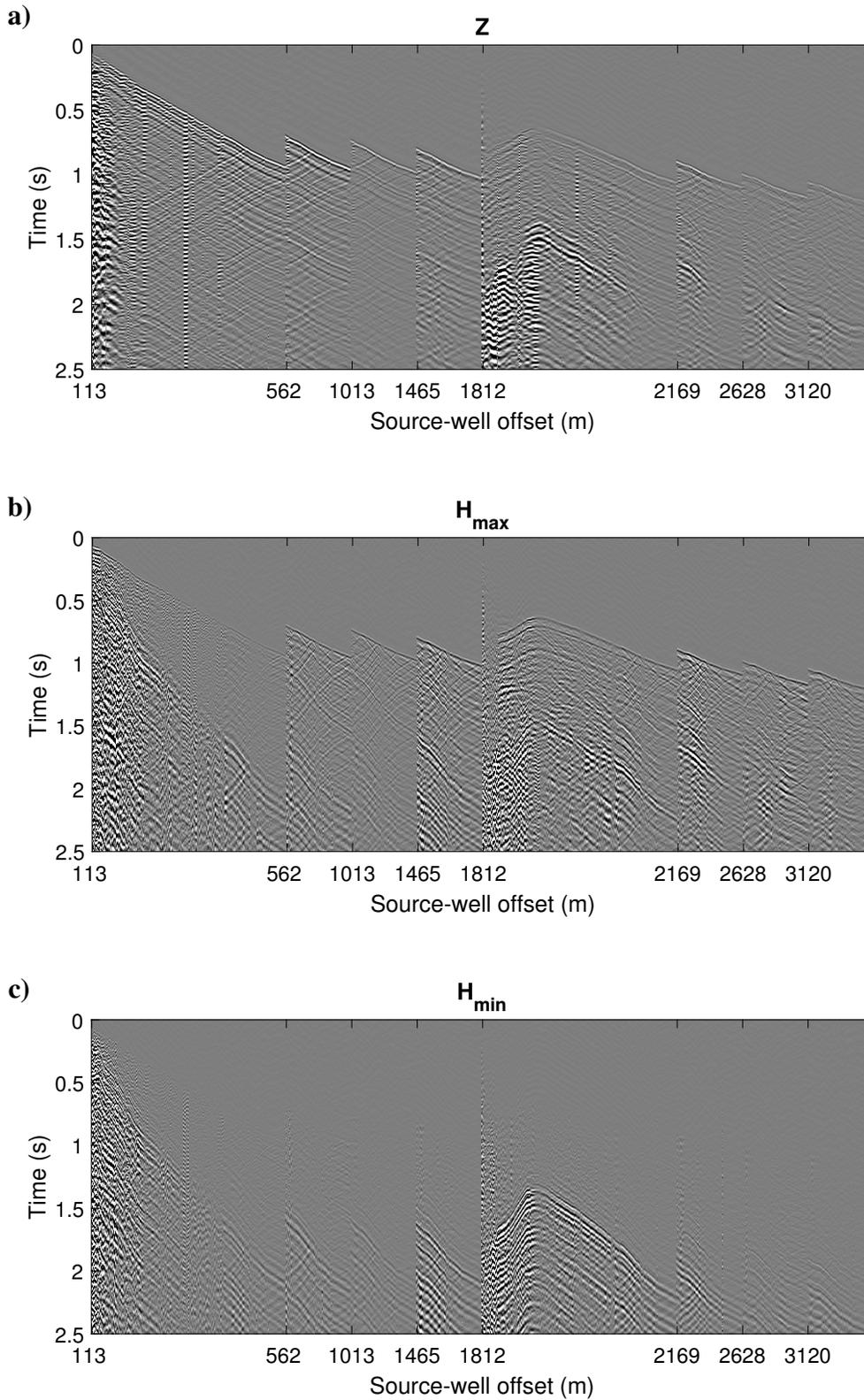


FIG. 5. Data after rotations. (a) Vertical, (b) maximum and (c) minimum energy horizontal components.

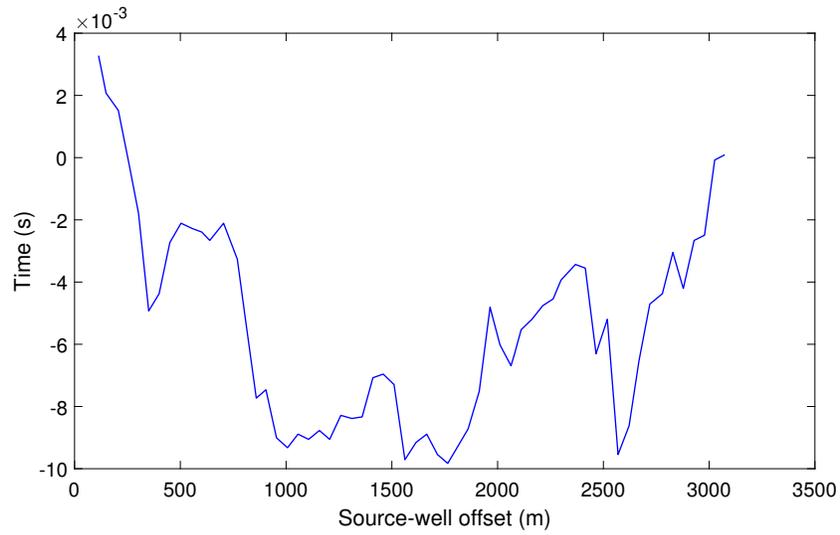


FIG. 6. Elevation statics (replacement velocity = 1816 m/s).

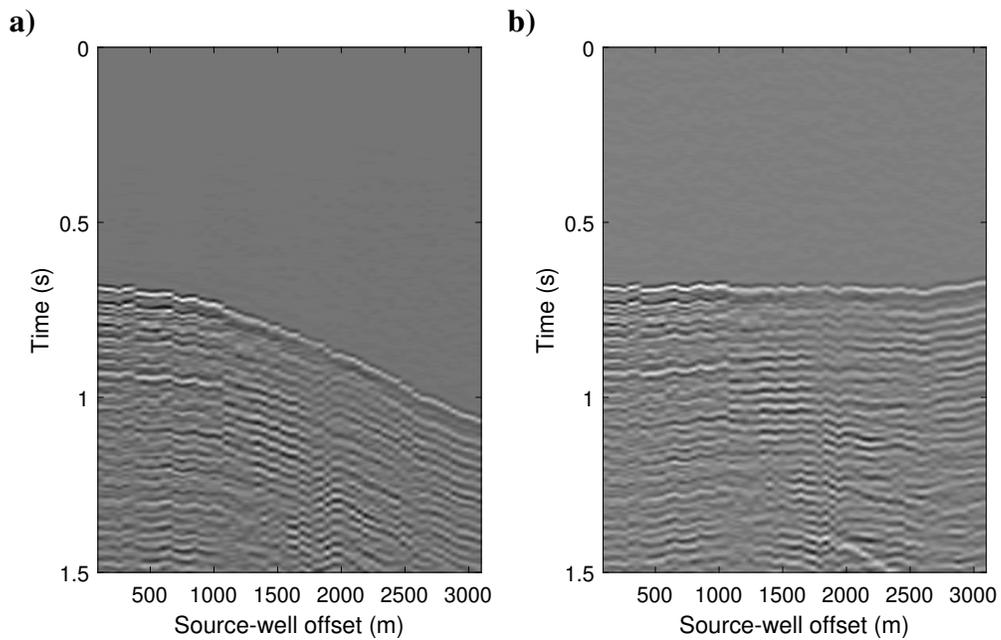


FIG. 7. Rec gather (a) before and (b) after NMO.

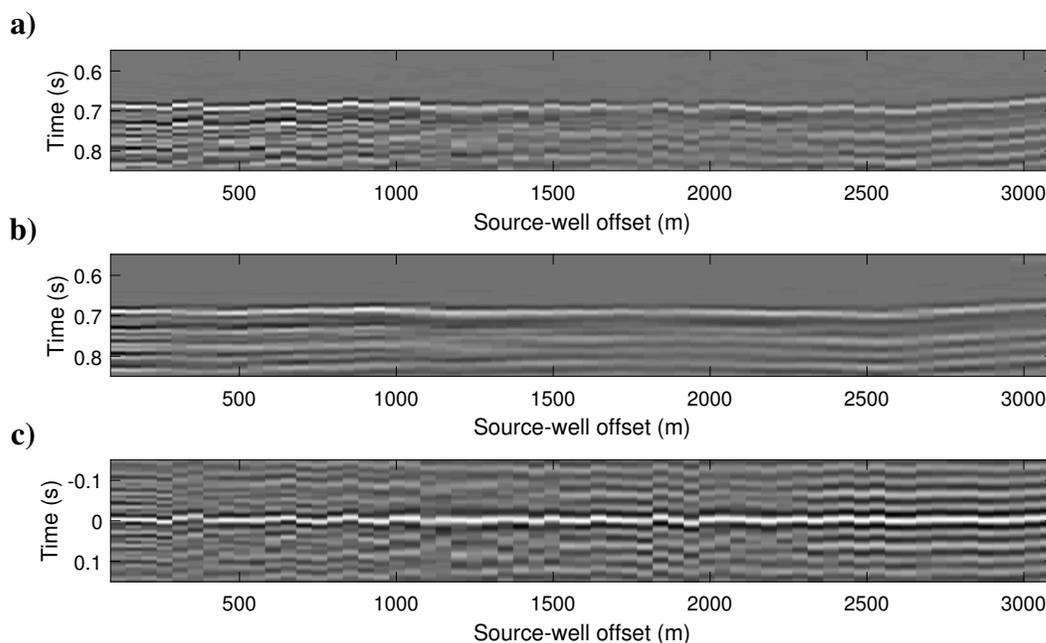


FIG. 8. (a) Data window around NMO-corrected first arrivals. (b) Pilot traces. (c) Crosscorrelation between (a) and (b).

To study whether the magnitude of these residual statics presented some degree of non-stationarity we repeated the experiment for the deepest receiver location (receiver 229 at 3445.6). Figure 9 compares the residual statics measured at the shallowest and deepest receiver locations. There we can see that the static corrections are very similar independently of the receiver depth. The absolute difference between the corrections obtained at both levels amounts to 1 ms in average. Therefore, we computed the average of both solutions and used the resulting statics as the final corrections that we applied to the data.

Noise attenuation

For improving the signal-to-noise ratio for P- and S-wave events in a single pass we applied a prediction filter in the FX domain (Canales, 2005). This type of filter assumes that the signal to be filtered consists of linear events that can be predicted in the spatial direction. This makes this filter a good alternative to the use of median filters for denoising VSP data. Since we are interested in recovering both P- and S-wave events, using a median filter would require picking downgoing P- and S-wave direct arrivals. Even though P-wave arrivals were readily available, there was not a clear and continuous downgoing S-wave event available in the near-offset data. For this reason we chose to proceed with the FX filter.

Figures 10 and 11 show the performance of this filter on the vertical and horizontal component data of three different source gathers. Notice how after applying the FX filter the anomalous amplitudes and frequencies present in the data have been removed. Particularly in shot number 3 for times later than 2 s there are several traces with this problem. On the horizontal component of shot number 3 (Figure 11a), a trace with anomalous ampli-

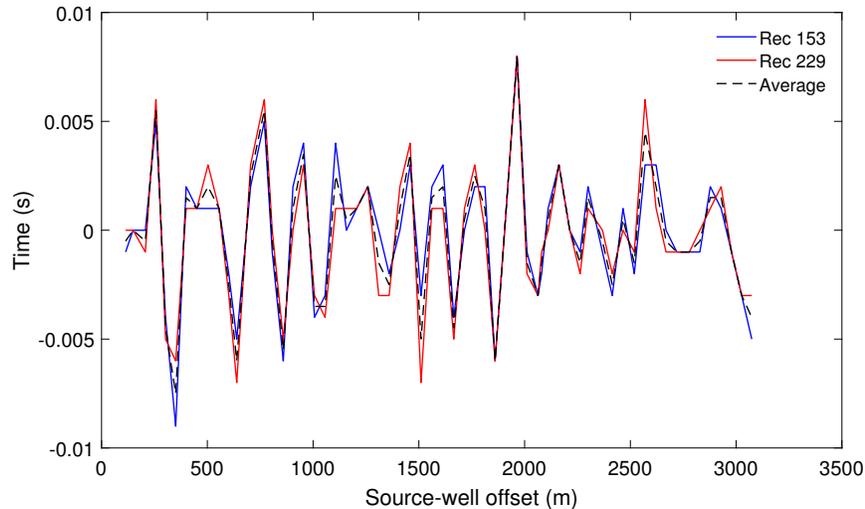


FIG. 9. Residual statics for the shallowest and deepest receivers in the walkaway data.

tudes for almost the full length of the record is present. After the FX filter application most of the anomalous amplitudes have been removed. It is also clear from the frequency spectra of the filtered data (Figure 10d and Figure 11d) that the energy on the low frequency end of the spectra is relatively higher after filtering.

Another important effect of the FX filter was the successful attenuation of the tube-waves present in the data. This wave mode is very clear over the bottom half of shot number 3 as indicated by the arrows on Figure 12a. Tube-waves can be identified in VSP data as events with very low velocity and high frequency. Their expression in the FK domain can be identified in Figure 12b. Very often these signals are filtered in the FK domain based on their clear separation from the more energetic downgoing and upgoing events. However, after applying the FX filter most of the energy of this wave mode was attenuated. Figures 12c and 12d show the FK spectra of the data before and after FX filtering, respectively. There, we can confirm that most of the energy of the tube-waves and their aliased components have been strongly attenuated.

Deconvolution

Deterministic Deconvolution

Two types of deconvolution were tested on the data. First, we tried a conventional deterministic deconvolution using the downgoing wavefield. To obtain this wavefield we flattened the data using the first arrivals and then applied a median filter with a length of 15 traces. This process is meant to reinforce the energy of the downgoing wavefield and remove the upgoing wavefield from the data. Then, a window of 300 ms was defined around the flattened first arrivals and a 1% prewhitening amplitudes were added to the spectrum of the downgoing wavefield to compute the deconvolution operators.

A subset of vertical and horizontal component data, before and after deconvolution, are displayed on Figure 13 and Figure 14. There we can see how the character of the events

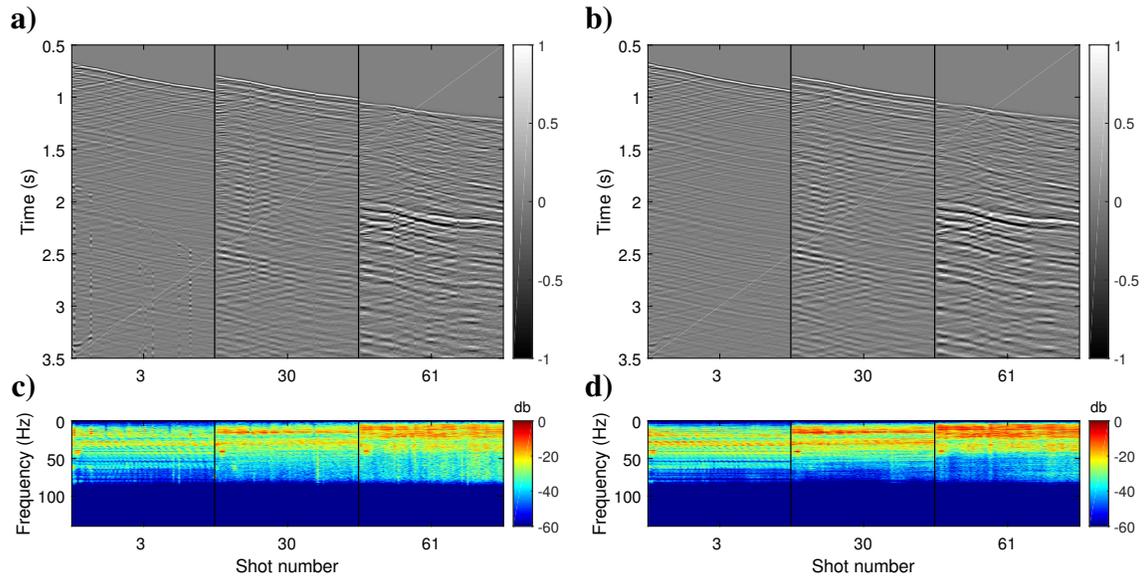


FIG. 10. Vertical component source gathers (a) before and (b) after FX filtering and their corresponding frequency spectra (c) and (d).

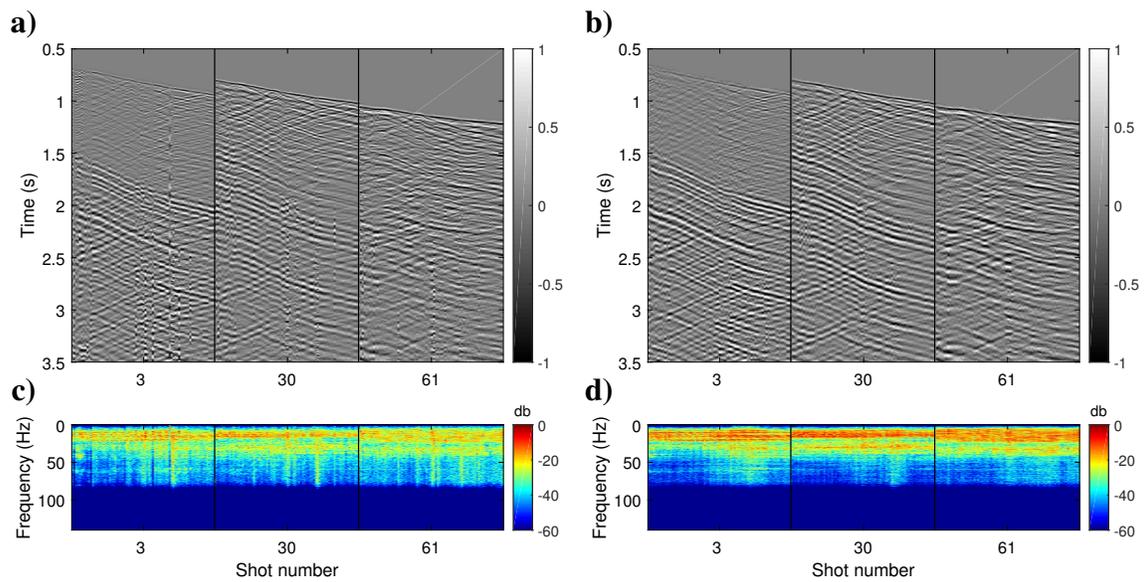


FIG. 11. Horizontal component source gathers (a) before and (b) after FX filtering and their corresponding frequency spectra (c) and (d).

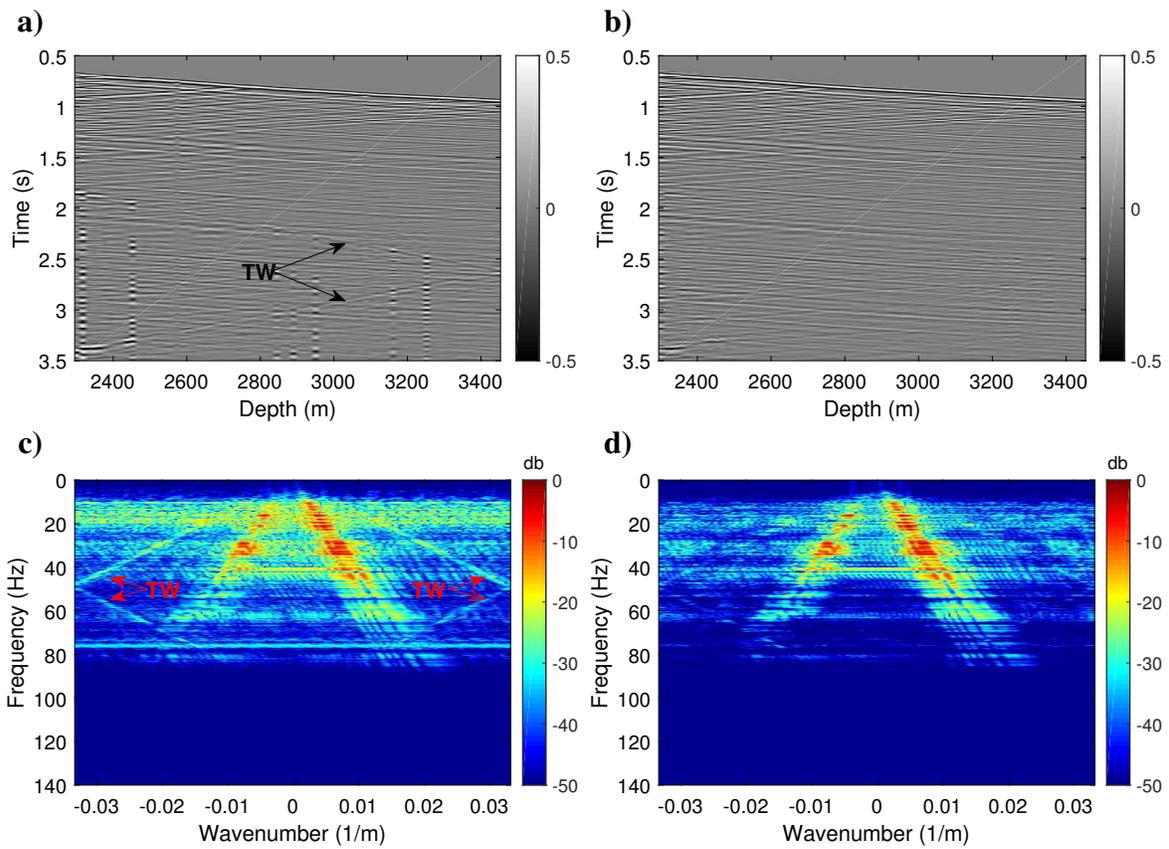


FIG. 12. Source gather (a) before and (b) after tube wave attenuation and their corresponding FK spectra (c) and (d).

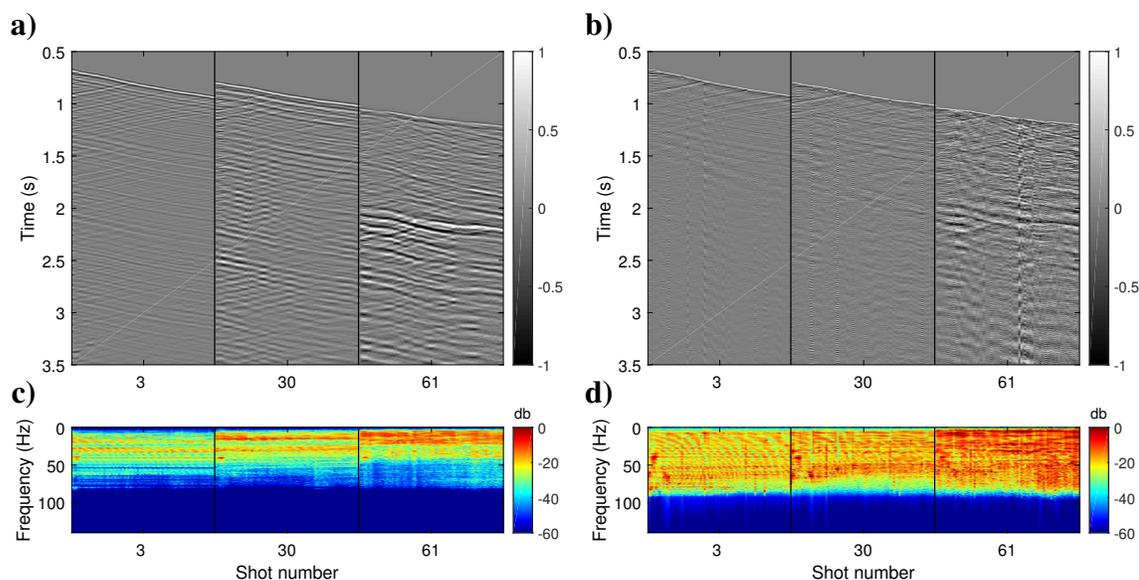


FIG. 13. Vertical component source gathers (a) before and (b) after deterministic deconvolution and their corresponding frequency spectra (c) and (d).

has been sharpened and the frequency spectra have been equalized for the band between 4 Hz and 80 Hz. This can be confirmed in the average frequency spectra presented in Figure 15. Of particular interest for FWI is the improvement of the frequency content on the lower end of the spectrum on both components. In this case, after deconvolution we have gained about 20 db on the frequencies between 4 Hz and 10 Hz. Despite the gain in bandwidth, the deconvolution operation has also introduced or re-energized some of the noise that was removed in the previous step. To attenuate this effect a second pass of the FX filter was performed. Figure 16 and Figure 17 compare the input data before deconvolution and the newly filtered and deconvolved data. There, it is clear how the spectra have been whitened and most of the noise has been attenuated. The average spectra displayed in Figure 18 confirm the gain on the frequency band between 4 Hz and 10 Hz for the vertical component data. However, the spectrum of the horizontal component data displays less impressive gains.

Gabor Deconvolution

In order to compensate for anelastic effects we also considered applying a Gabor deconvolution (Margrave and Lamoureux, 2001) to the data. In contrast to the previous deterministic deconvolution where a fixed source signature is used to deconvolve all the data recorded by a given trace, with the Gabor deconvolution we aim to remove time-dependent source wavelet variations. The deconvolution was performed using windows of 0.2 s in increments of 0.01 s and temporal and frequency smoothers of 0.4s and 4 Hz, respectively.

The horizontal and vertical component data before and after Gabor deconvolution are shown in Figure 19 and Figure 20. Similar to the deterministic deconvolution the Gabor deconvolved data resulted in a wider and more equalized frequency spectrum. Moreover,

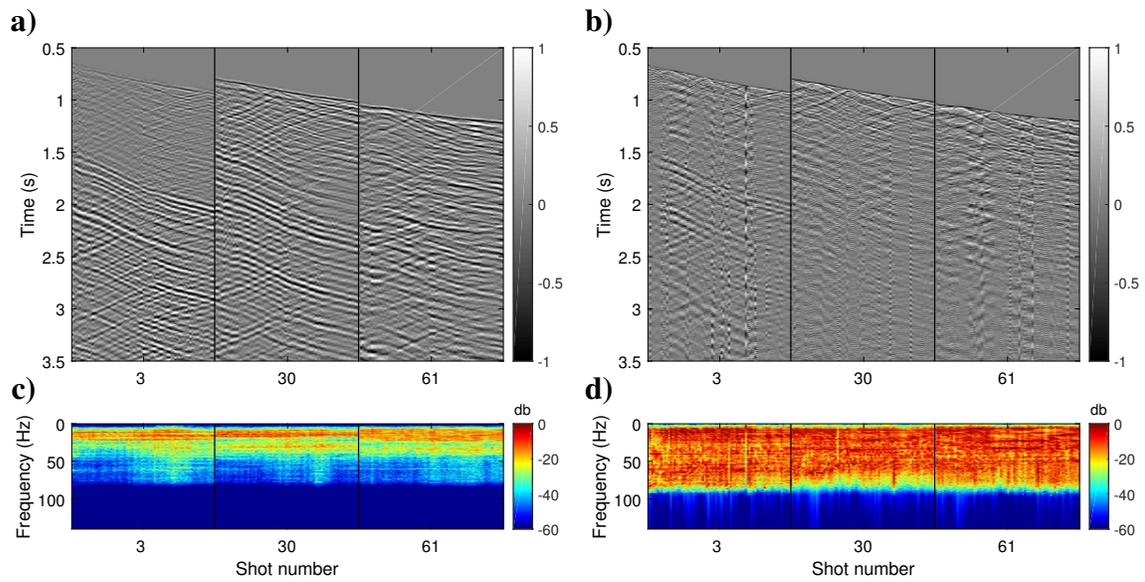


FIG. 14. Horizontal component source gathers (a) before and (b) after deterministic deconvolution and their corresponding frequency spectra (c) and (d).

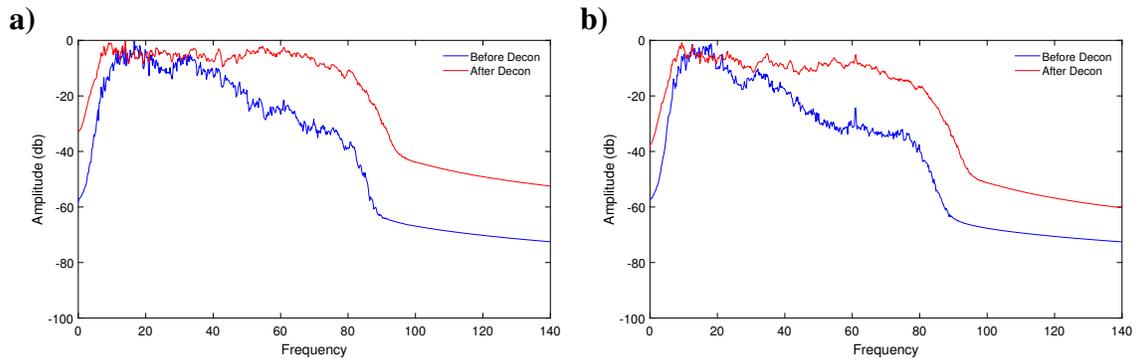


FIG. 15. Frequency spectra before and after deterministic deconvolution for the (a) vertical and (b) horizontal component data.

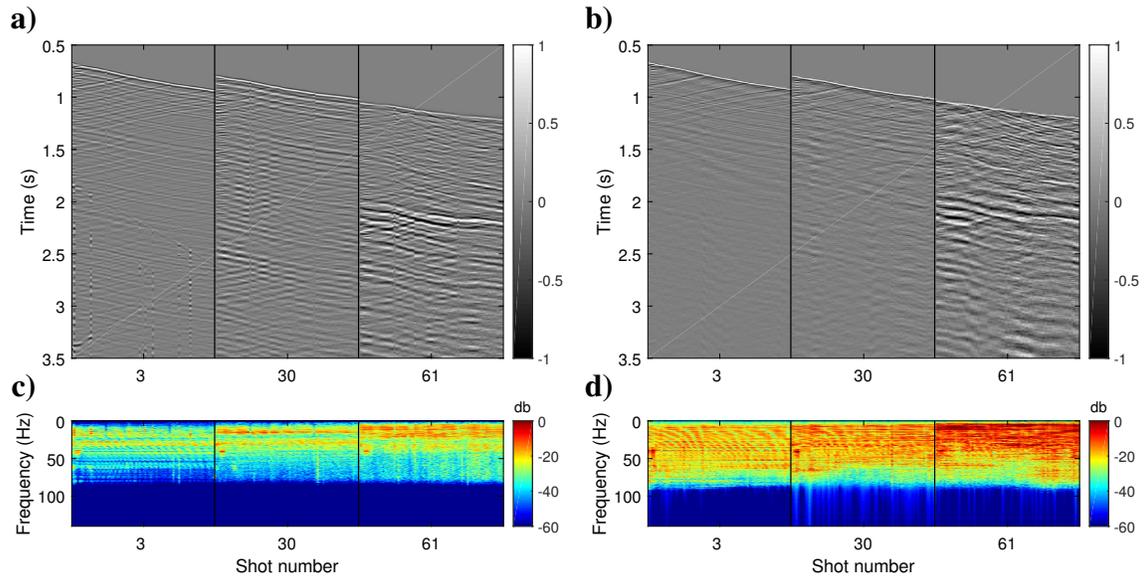


FIG. 16. Vertical component data (a) before and (b) after data preprocessing and their corresponding frequency spectra (c) and (d) using a deterministic deconvolution.

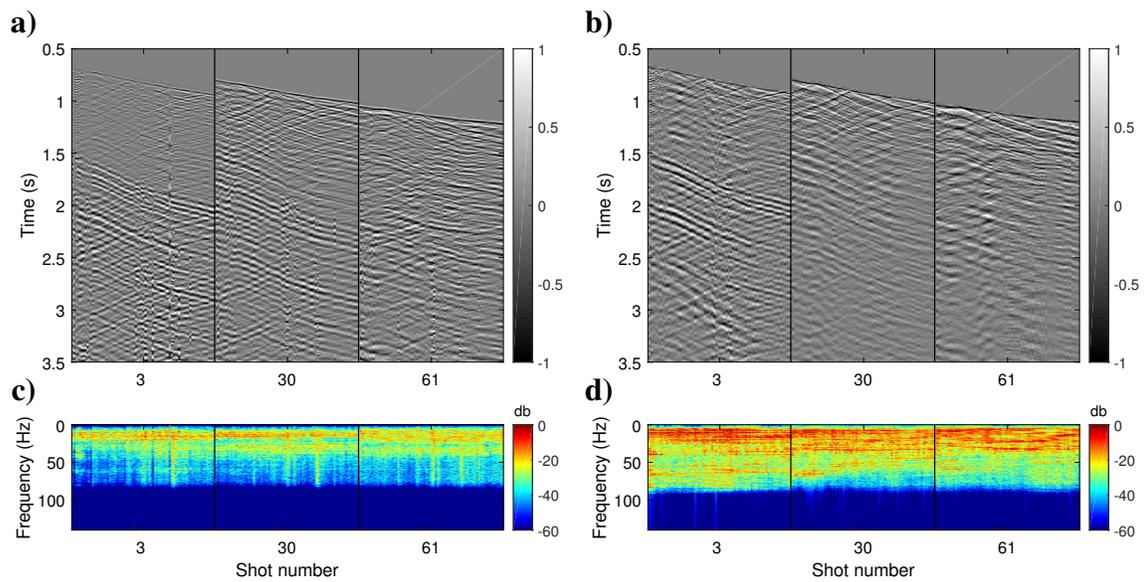


FIG. 17. Horizontal component data (a) before and (b) after data preprocessing and their corresponding frequency spectra (c) and (d) using a deterministic deconvolution.

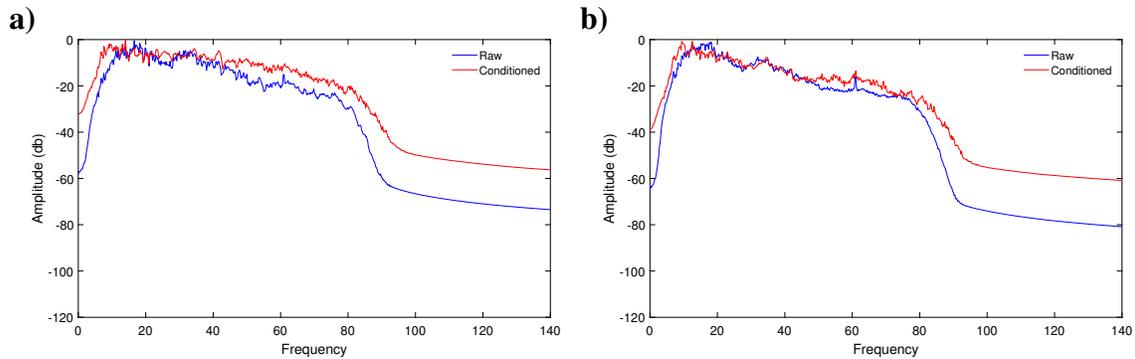


FIG. 18. Frequency spectra before and after preprocessing for the (a) vertical and (b) horizontal component data using a deterministic deconvolution.

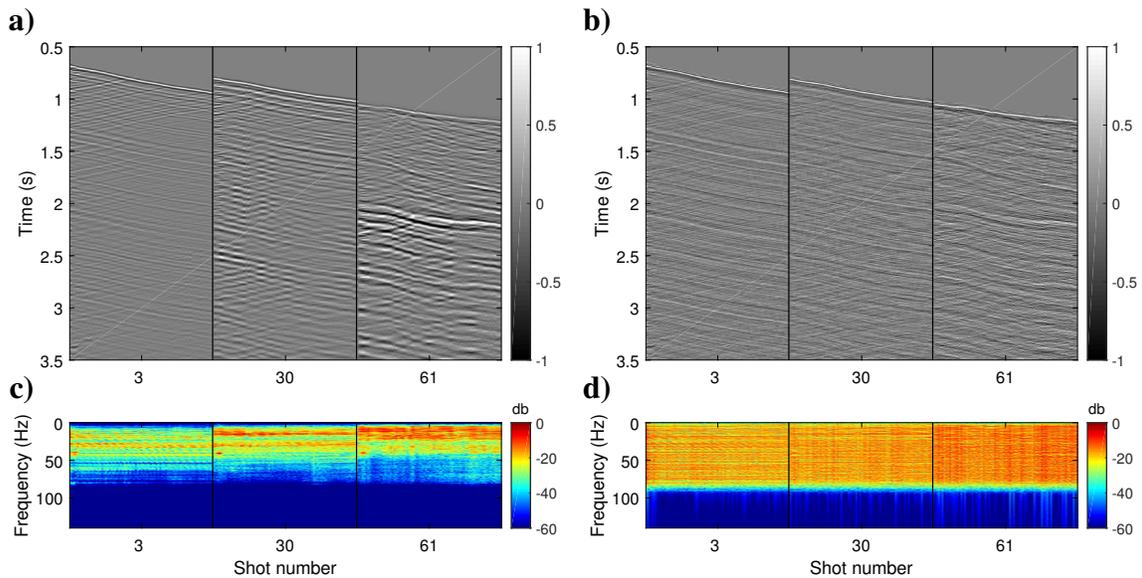


FIG. 19. Vertical component source gathers (a) before and (b) after Gabor deconvolution and their corresponding frequency spectra (c) and (d).

the events after Gabor deconvolution display better coherency and the amount of noise introduced by this deconvolution is much lower.

The flat character of the frequency spectrum between 4 Hz and 80 Hz is very clear in Figure 21. Both the horizontal and vertical component data display significant gains in the low end of the spectrum.

To attenuate any noise introduced by the deconvolution we performed a second pass of the FX prediction filter. A comparison between the input data before deconvolution and the newly processed data is displayed in Figure 22 and Figure 23. There we can see how the processes applied not only sharpened the character of the events but it also improved their coherency.

Average frequency spectra after pre-processing are shown in Figure 24. The gains on

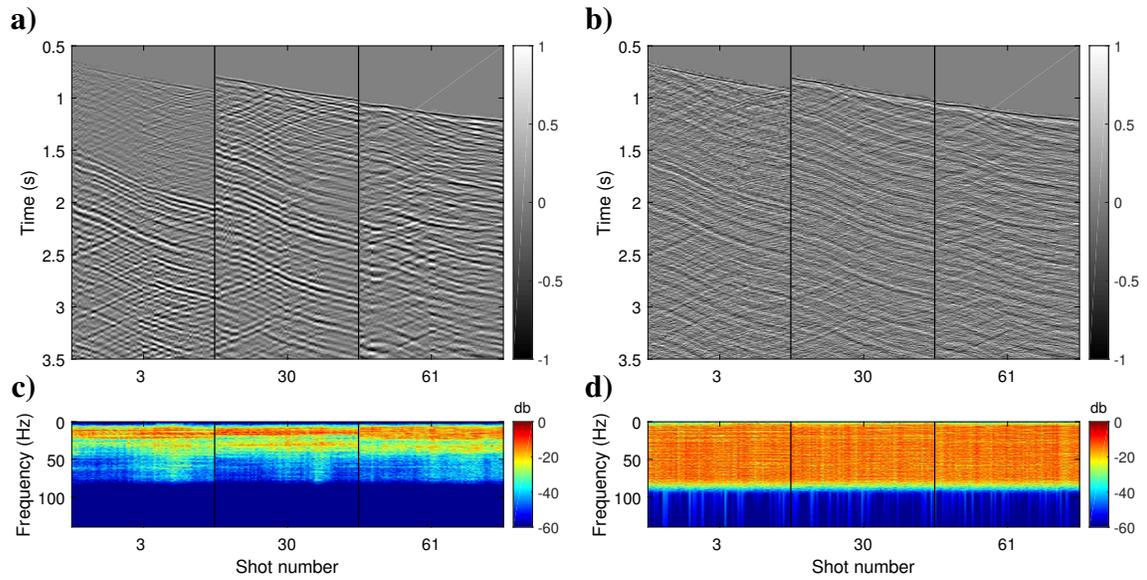


FIG. 20. Horizontal component source gathers (a) before and (b) after Gabor deconvolution and their corresponding frequency spectra (c) and (d).

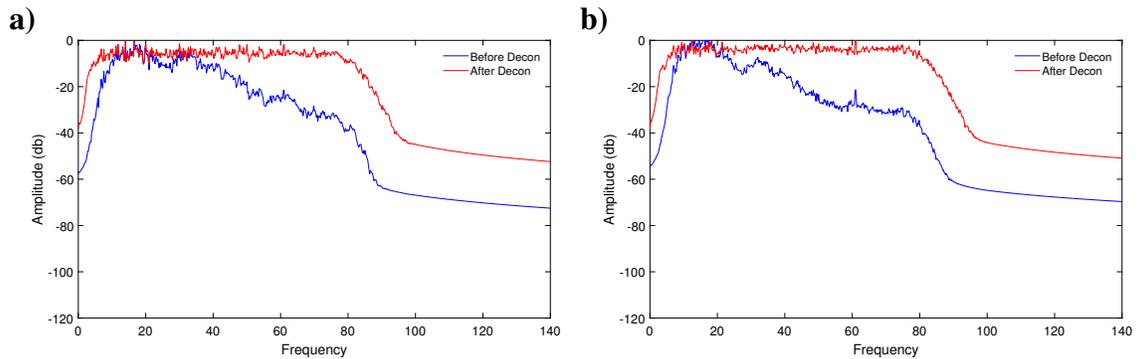


FIG. 21. Frequency spectra before and after Gabor deconvolution for the (a) vertical and (b) horizontal component data.

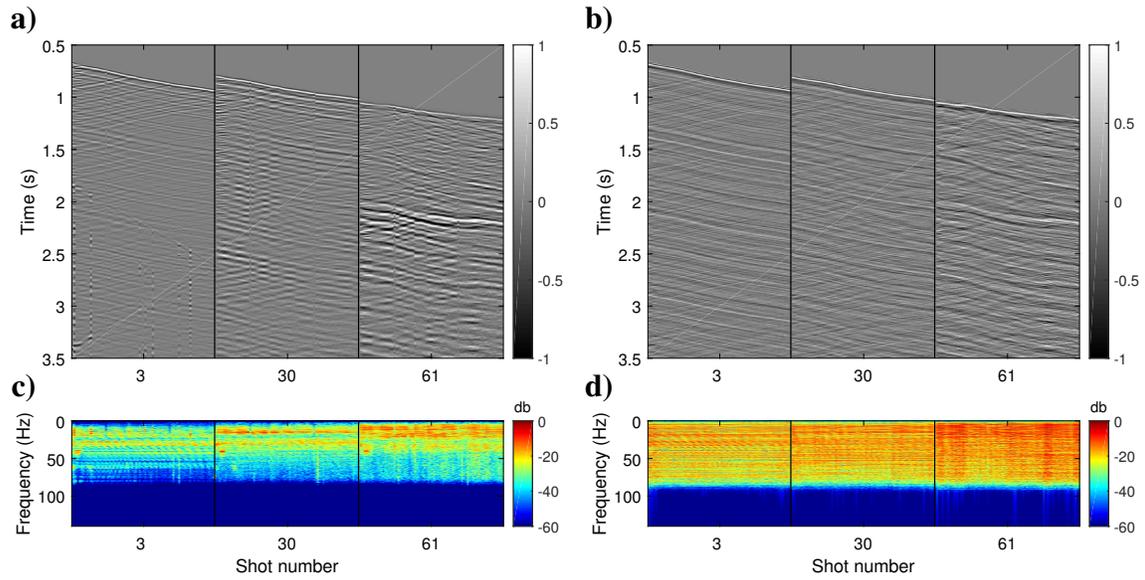


FIG. 22. Vertical component data (a) before and (b) after data preprocessing and their corresponding frequency spectra (c) and (d) using a Gabor deconvolution.

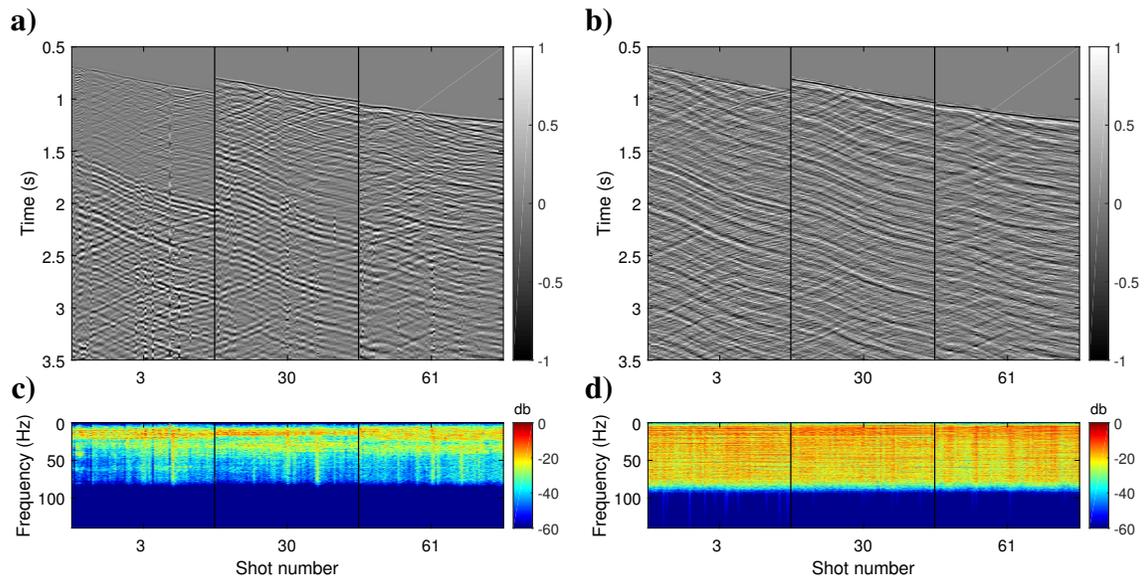


FIG. 23. Horizontal component data (a) before and (b) after data preprocessing and their corresponding frequency spectra (c) and (d) using a Gabor deconvolution.

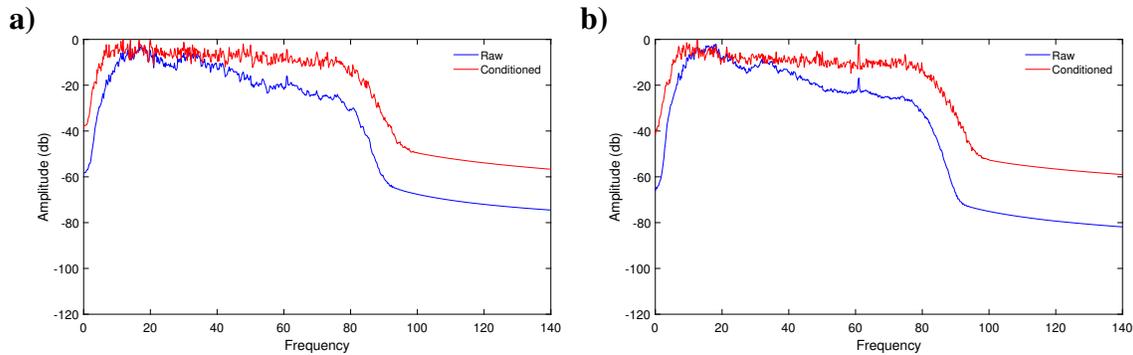


FIG. 24. Frequency spectra before and after preprocessing for the (a) vertical and (b) horizontal component data.

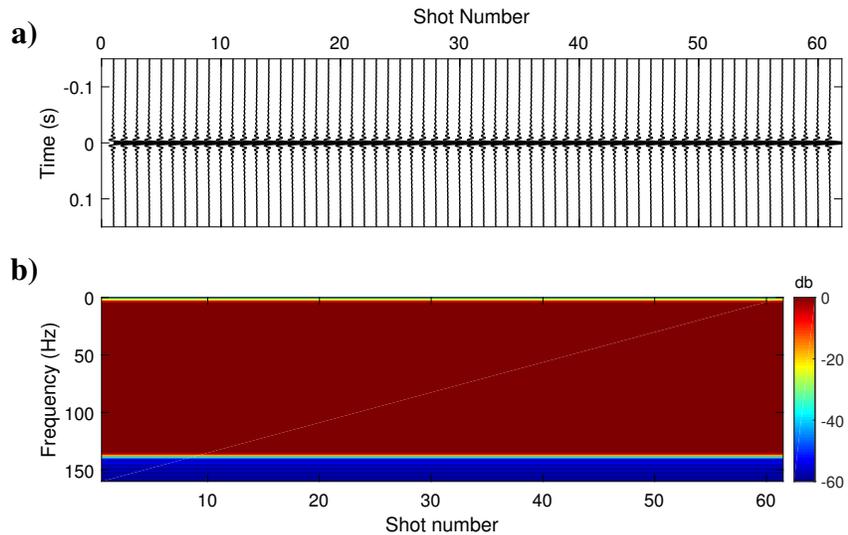


FIG. 25. Source wavelets computed from the autocorrelation of the vibe sweeps and their corresponding frequency spectra (b)

the low frequency end of the spectrum between 4 Hz and 10 Hz amount to approximately 20 db.

Source wavelet computation

To initialize the FWI process a source wavelet must be provided to the algorithm. Since every process applied to the data modifies the underlying wavelet we computed three different sets of source wavelets. First, we computed a set of wavelets assuming the autocorrelation of the sweep as the source wavelet. Since the sweep had the same parameters for all the source locations this wavelet is constant and it contains the full bandwidth of the sweep (Figure 25).

Second, we extracted a source wavelet for each source gather after applying the deterministic deconvolution. For this, we flattened the deconvolved data and applied a median filter with a window length of 15 traces to remove the upgoing wavefield. We then defined a window of 300 ms around the flattened first arrivals and computed the average amplitude spectrum of these data. Finally, we invert for the shape of the zero-phase filter that

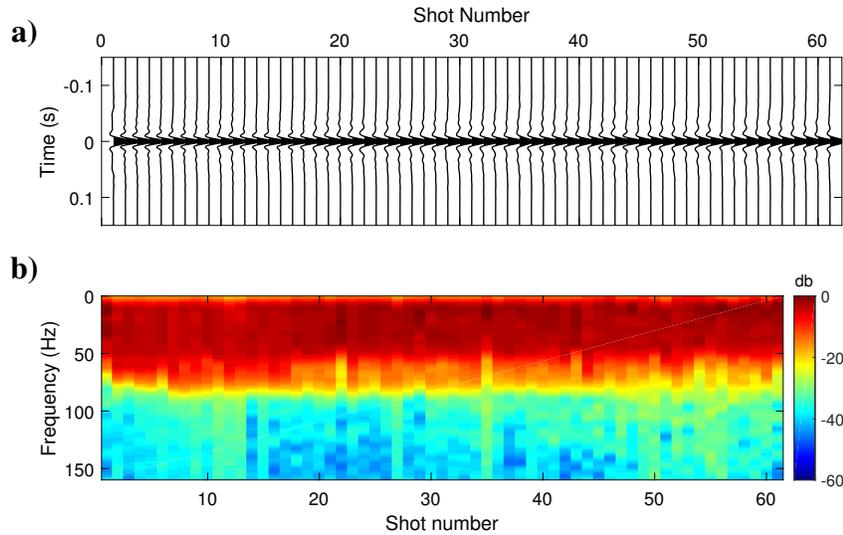


FIG. 26. (a) Source wavelets extracted from the data preprocessed using deterministic deconvolution and their corresponding frequency spectra (b)

reproduces the same amplitude spectrum. A Gaussian smoother with a width of 10 Hz and prewhitening of 1% are applied to the frequency spectrum of the data before the inversion. This process is performed on a trace-by-trace basis for each trace in every source gather. Then, all the source wavelets obtained within a given source gather are averaged and saved as the source wavelet corresponding to that specific source location. Figure 26 shows the computed source wavelets and their corresponding frequency spectra. There we can see that most of energy is concentrated in the band between 2 Hz and 70 Hz and the spectra is very stable across most of the sources.

Following the same process we extracted source wavelets for the data deconvolved using the Gabor deconvolution. Results are displayed in Figure 27. Even though the bandwidth of these wavelets is similar to the one in Figure 26b the spectra are not as stable as before.

Figure 28 compares the extracted wavelets for the near-offset location. There we can observe how the wavelet extracted from both of the deconvolved datasets show smaller and shorter lobes that the wavelet derived from the sweep of the vibe truck.

FULL WAVEFORM INVERSION

We used a time domain elastic FWI algorithm based on spectral elements modelling (Komatitsch and Tromp, 1999). We defined a mesh with an element size of 25 m and 5 Gauss-Lobatto-Legendre (GLL) points per element. This resulted in a minimum distance between points of 4.31 m. The modelling time step was set at 3.25×10^{-4} s to satisfy the CFL stability condition given the velocities, mesh size and frequencies to be used in the modelling.

V_p , V_s and density models were built from the available well log data. Missing V_s and density values were completed from the available V_p measurements using a linear regression for V_s ($V_s = 0.53V_p - 57.96$) and an exponential relationship ($\rho = 0.45V_p^{0.2}$)

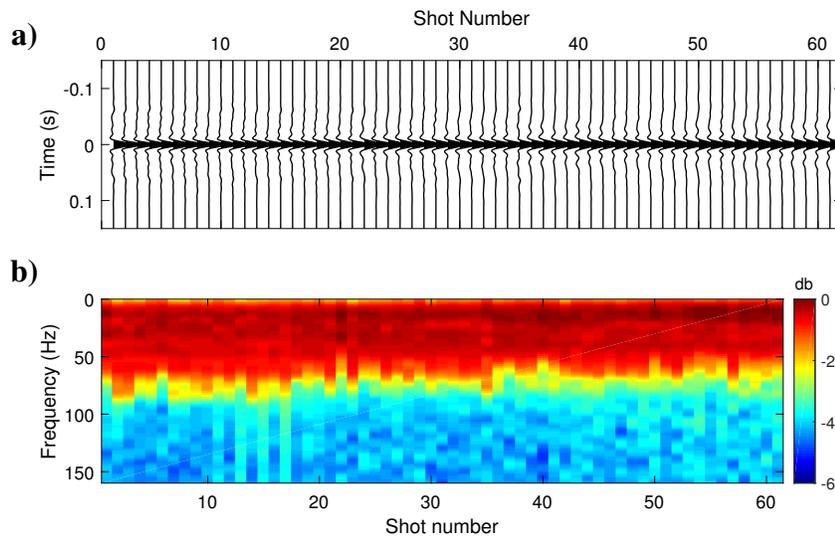


FIG. 27. (a) Source wavelets extracted from the data preprocessed using Gabor deconvolution and their corresponding frequency spectra (b).

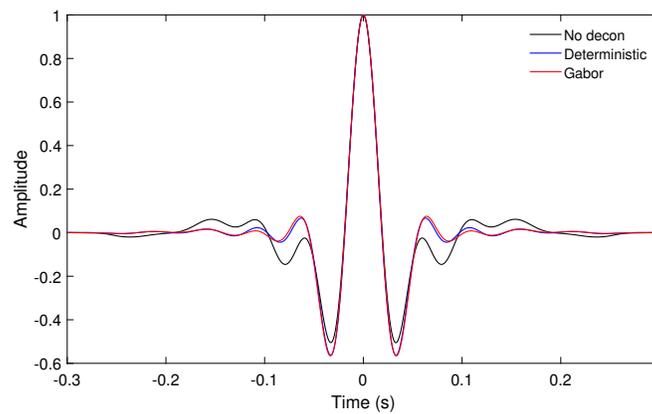


FIG. 28. Source wavelets computed for the near-offset source.

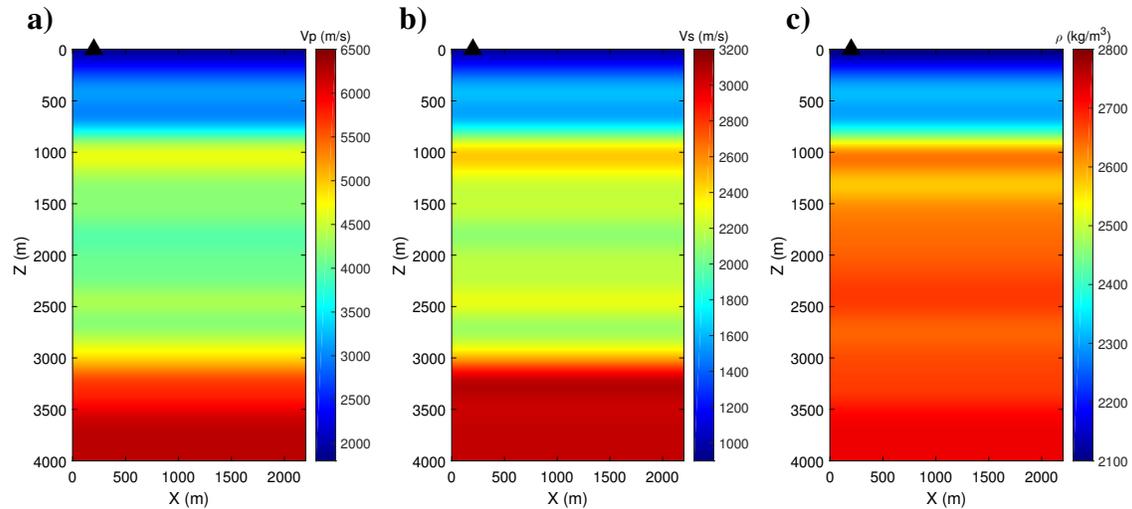


FIG. 29. Initial (a) P-wave velocity, (b) S-wave velocity and (c) density models used in the inversion.

for the density. The coefficients used in both regressions were computed from the intervals where all the logs were available. Finally, well logs were smoothed using a Gaussian filter with a half-width of 100 m to create the initial velocity and density models used in the inversion (Figure 29).

The inversion was carried out in three depth windows going from [250 m, 1000 m] to [750 m, 2250 m] and [2000 m, 3500 m]. At each depth window the inversion is performed using a multi-scale approach with four expanding frequency bands from [4 Hz, 8 Hz] to [4 Hz, 12 Hz], [4 Hz, 16 Hz] and [4 Hz, 20 Hz]. A maximum of ten iterations per frequency band is allowed. Only eight source locations were used in the inversion ranging between 113 m and 1812.5 m offset from the well. These are the data recorded between the near-offset source and the only far-offset source along the walkaway line that was recorded over the full length of the well. The other six source location used for the inversion were recorded only around the target window (2302.7 m to 3445.6 m depth).

Figure 30 shows the near-offset data for the first scale and first depth window. The effects of the pre-processing are very clear in this example. On the data deconvolved with the deterministic operators (Figure 30b) it is possible to identify a downgoing S-wave arrival that was not evident in the data before deconvolution (Figure 30a) and is not observable in the Gabor-deconvolved data (Figure 30c). Even though the latter exhibited wider amplitude spectra it was only the dominant P-wave energy which benefited the most, overwhelming the lower amplitude S-wave events that were present. On the other hand, the deterministic deconvolution, by trying to deconvolve the downgoing wavelet, collapsed most of the downgoing P-wave energy around the direct arrivals revealing the downgoing S-wave energy present in the data.

To test the performance of the inversion on both scenarios we first inverted the downgoing wavefield of the near-offset source with both types of deconvolution. Since the near-offset data with the Gabor deconvolution showed no evident downgoing S-wave energy we only inverted for V_p values and kept fixed the V_s and density models. The downgoing

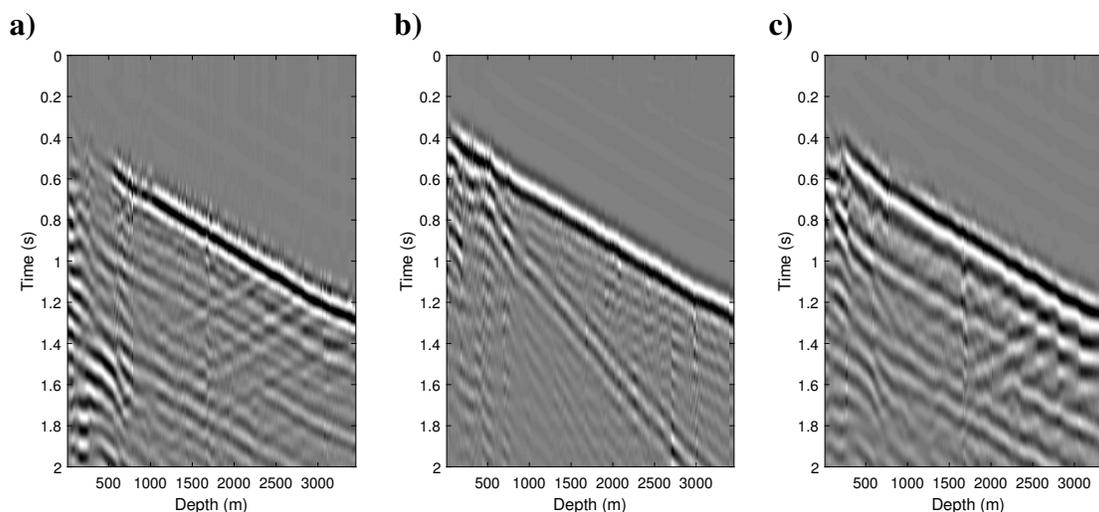


FIG. 30. Data used for the inversion of the first depth window and first frequency band, (a) without deconvolution and after (b) deterministic and (c) Gabor deconvolution.

wavefield was separated from the full wavefield on-the-fly by using an FK filter, suppressing all the negative wavenumbers. Only the data within a 500 ms window centered around the first arrivals were used for the inversion, the rest of the data were muted. The inverted V_p values for this test using the first depth window are shown in Figure 31. There, we can observe that the inversion performed with the Gabor deconvolved data is diverging significantly from the initial V_p model. Using data without any deconvolution also resulted in a divergent solution. Only the data deconvolved with deterministic operators show a stable solution, providing reasonable model updates around the initial V_p model. These observations can be confirmed on the data space (Figure 32) where only the data modelled after the inversion with the deterministic deconvolution closely resembles the observed data. The results obtained with the data without deconvolution and the Gabor deconvolved data are far from the observed datasets.

One of the reasons that may explain the previous results can be related to the effect of multiples in the data. In particular, the presence of fine layering with large velocity contrasts that results in very short-period internal multiples, introduces a coda in the downgoing wavefield that is very difficult to explain with a smooth velocity model. Therefore, the inversion is not able to reproduce this energy, resulting in large residuals that will not be minimized but that will keep producing incorrect model updates.

In addition to revealing some of the downgoing S-wave energy present in the data, the deterministic deconvolution collapsed most of the energy around the downgoing wavefield into a single sharper arrival. As a result, our initial velocity model was able to model very closely the input data providing proper model updates.

Based on the previous results we chose to proceed with the inversion using the data with deterministic deconvolution only. We only inverted for V_p and V_s , keeping the density model fixed, given that there is not a regular offset distribution for all the receiver depths. The results at the location of the well are shown in Figure 33. Overall, the inverted V_p values follow very closely the expected values according to the well log. However, the

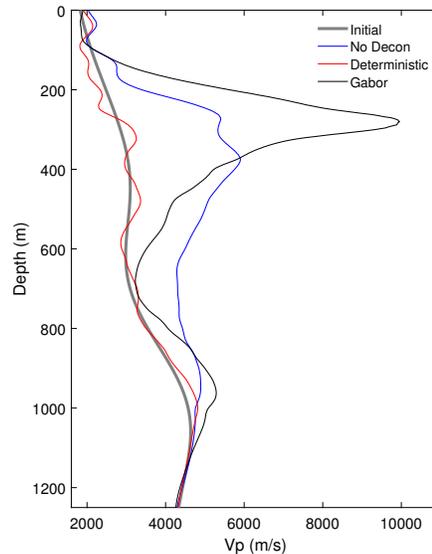


FIG. 31. Inverted Vp profiles at the well location for the first depth window.

inversion underestimated the actual Vp values in the section around 1500 m depth. The interval between 2250 m and 3250 m shows a good agreement with the log data. Notice that this is the only depth interval that was recorded at all the source points, therefore more data was available in this window. The results for Vs also show a good agreement in the shallow part of the section. However, for depths under 2000 m the results are mixed. We would expect these results to improve by including farther source locations with more energetic S-wave arrivals in the inversion.

Figure 34 shows the 2D inverted Vp and Vs models. There we can see that most of the updates were focused in an area of 500 m around the well. One of the drawbacks of attenuating the multiples before the inversion is that we lose the extra illumination provided by these events. FWI algorithms oriented toward exploiting these events remain to be explored.

Figure 35 compares the observed data and the data modelled with the inverted model for a near- and far-offset locations. There we can see a good agreement between the modelled and observed downgoing wavefields, particularly for the vertical component data on both offsets. Even some of the downgoing multiples that were attenuated in the data used for the inversion, have been modelled. On the horizontal components the results are mixed. For the near-offset data there is very little coherent energy above 1.5 s. There is a large amplitude S-wave event around 1.75 s that is partially matched on the synthetic data. However, their frequency content are quite different. On the observed far-offset data S-wave energy displays better coherency. There, we can observe a better agreement among the downgoing events on both datasets, even though the observed horizontal component data still display some data quality problems. Overall, there is still a slight mismatch regarding the source wavelet. The modelled data display a lower frequency content than the observed data for all the components. Updating the source wavelet after each iteration might help to reduce this mismatch.

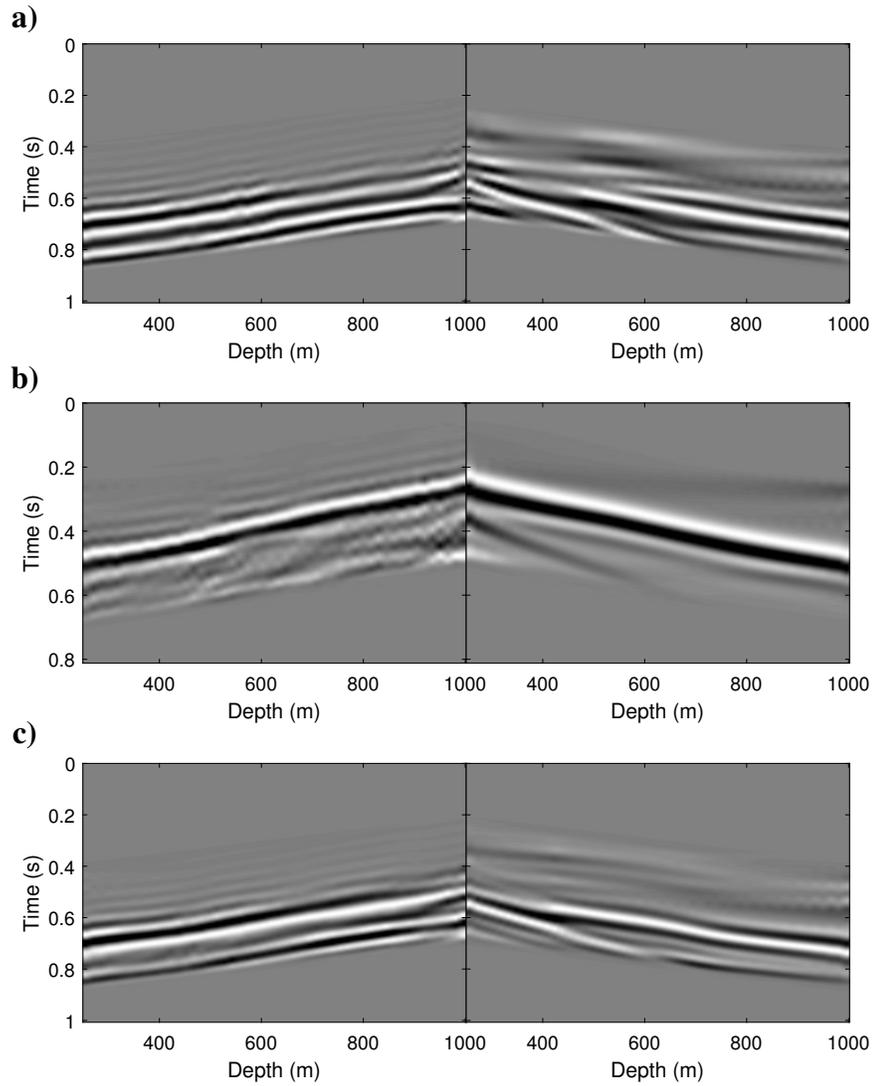


FIG. 32. (left) Observed and (right) modelled downgoing wavefields after FWI using data (a) without deconvolution, (b) with deterministic deconvolution and (c) with Gabor deconvolution.

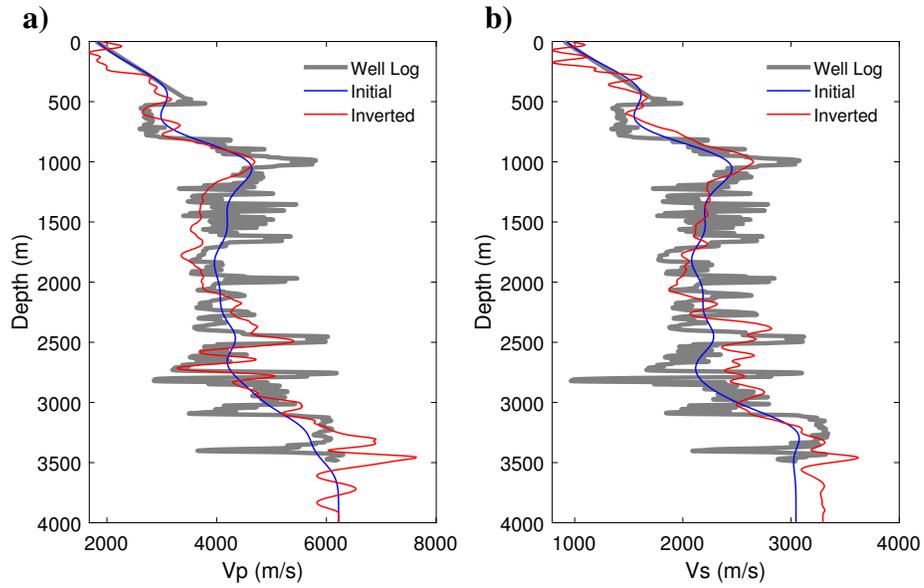


FIG. 33. (a) V_p and (b) V_s velocity values obtained at the well location.

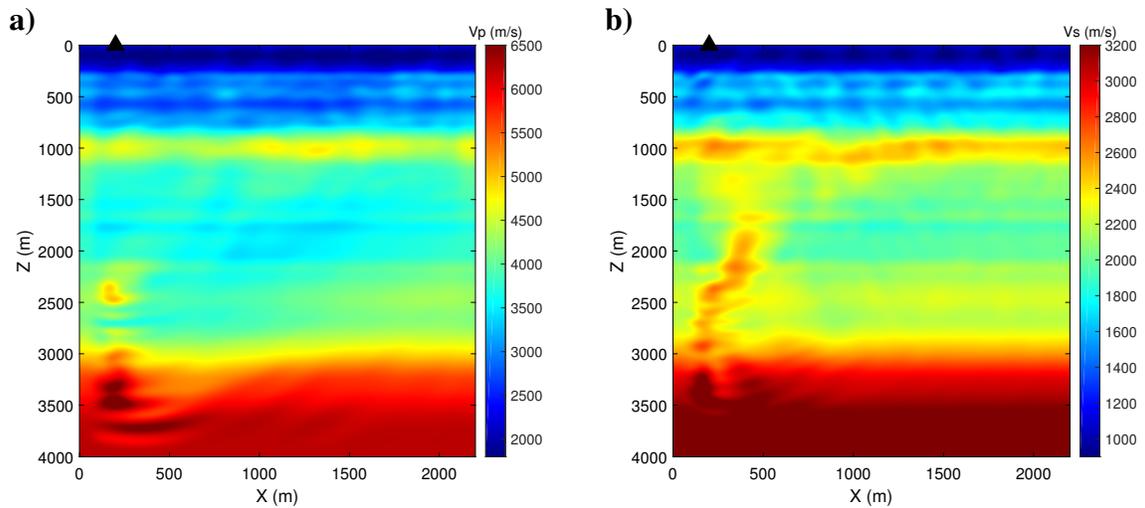


FIG. 34. (a) V_p and (b) V_s velocity models obtained after FWI

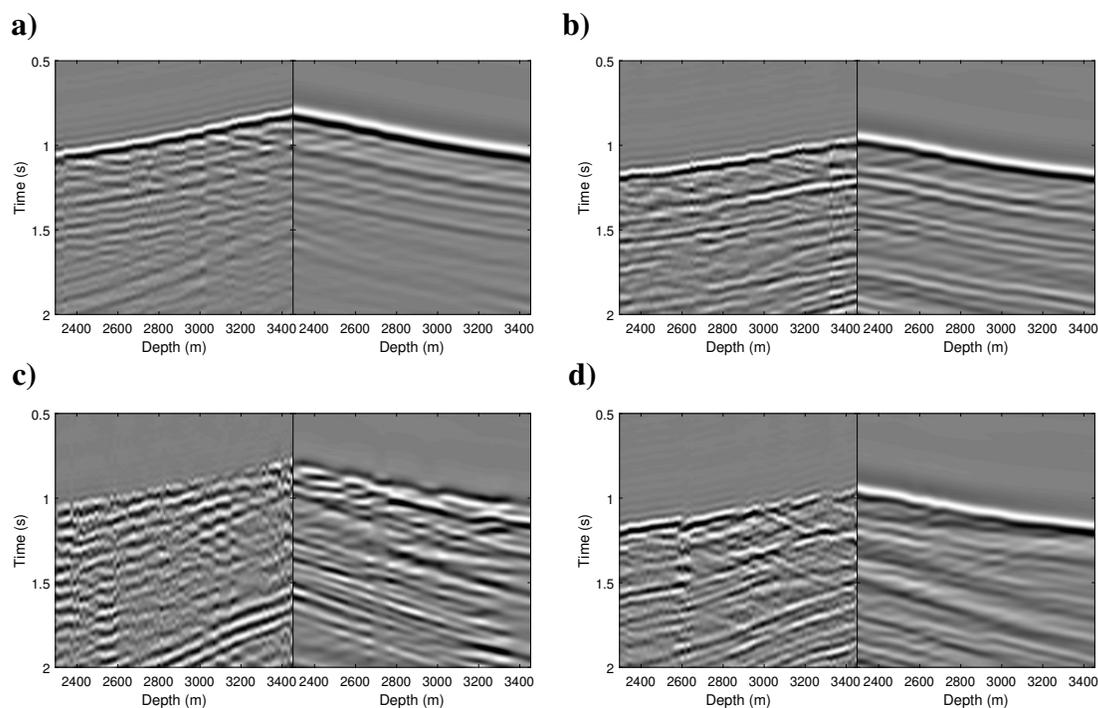


FIG. 35. (a) Near and (b) far offset vertical component data before and after FWI. (c) Near and (d) far offset horizontal component data before and after FWI.

CONCLUSIONS

The non-linearity of the FWI problem can lead to very different solutions given small perturbations, not only in the model space but also in the data domain. We inverted two datasets pre-processed with different deconvolution methods that were trying to account for the missing physics (anelasticity) in our FWI algorithm. The results obtained with a simple deterministic deconvolution were superior than the ones obtained with a Gabor deconvolution. We argue that by collapsing and attenuating some of the multiples present in the data, the deterministic deconvolution provided an easier to model dataset given a smooth initial velocity model. Moreover, by attenuating the multiples the deterministic deconvolution was able to reveal some low energy downgoing S-wave events that were not evident before. The Gabor deconvolution certainly provided a much wider amplitude spectrum but it left untouched the energy resulting from multiple events resulting in a more complex dataset. In this case, a more complex initial subsurface model might be needed for the inversion to converge toward a reasonable model.

Similar to what is proposed by the multi-scale approach where lower frequencies are inverted first to avoid cycle skipping, here we argue that providing data with only first-order events at the early stages of an FWI can produce similar effects. Our results with the data deconvolved using deterministic operators were more stable than with the Gabor deconvolution. However, the data modelled with our final inverted model did show the presence of multiples. Continuing the inversion with the data without deconvolution or with the Gabor deconvolved data might help to refine our results. One of the drawbacks of

not including the multiples earlier in the inversion is that the extra illumination that they provide, particularly for VSP configurations, is not exploited.

Updating or inverting for the source wavelet after each iteration can also help to improve our inversion. The results show that even though the timing of the events on both the observed and modelled dataset are in agreement, the shape of the waveforms display some mismatches.

ACKNOWLEDGEMENTS

The authors thank the sponsors of CREWES for continued support in particular to Devon Energy for facilitating the data used in this study. This work was funded by CREWES industrial sponsors, CFREF (Canada First Research Excellence Fund) and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13. We are also grateful to Compute Canada for providing the computational resources used in this study.

REFERENCES

- Canales, L. L., 2005, Random noise reduction: SEG Technical Program Expanded Abstracts 1984, 525–527.
- Komatitsch, D., and Tromp, J., 1999, Introduction to the spectral-element method for 3-D seismic wave propagation: *Geophysical Journal International*, **139**, No. 3, 806–822.
- Margrave, G., and Lamoureux, M., 2001, Gabor deconvolution: CREWES Research Report, **13**, 241–276.
- Nocedal, J., and Wright, S. J., 2006, Numerical optimization: Springer.
- Nolet, G., 1987, Seismic Tomography : with applications in global seismology and exploration geophysics, Seismology and exploration geophysics: Reidel, Dordrecht, the Netherlands.
- Operto, S., Miniussi, A., Brossier, R., Combe, L., Metivier, L., Monteiller, V., Ribodetti, A., and Virieux, J., 2015, Efficient 3-d frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to valhall in the visco-acoustic vertical transverse isotropic approximation: *Geophysical Journal International*, **202**, No. 2, 1362–1391.
- Pan, W., Innanen, K. A., and Geng, Y., 2018, Elastic full-waveform inversion and parametrization analysis applied to walk-away vertical seismic profile data for unconventional (heavy oil) reservoir characterization: *Geophysical Journal International*, **213**, No. 3, 1934–1968.
- Pica, A., Diet, J. P., and Tarantola, A., 1990, Nonlinear inversion of seismic reflection data in a laterally invariant medium: *Geophysics*, **55**, No. 3, 284–292.
- Podgornova, O., Leaney, S., Charara, M., and Lunen, E. V., 2014, Elastic full waveform inversion for land walkaway VSP data: CSEG Geoconvention 2014 Expanded abstracts.
- Ratcliffe, A., Win, C., Vinje, V., Conroy, G., Warner, M., Umpleby, A., Stekl, I., Nangoo, T., and Bertrand, A., 2011, Full waveform inversion: A north sea obc case study: SEG Technical Program Expanded Abstracts 2011, 2384–2388.
- Routh, P., Neelamani, R., Lu, R., Lazaratos, S., Braaksma, H., Hughes, S., Saltzer, R., Stewart, J., Naidu, K., Averill, H., Gottumukkula, V., Homonko, P., Reilly, J., and Leslie, D., 2017, Impact of high-resolution FWI in the western black sea: Revealing overburden and reservoir complexity: *The Leading Edge*, **36**, No. 1, 60–66.
- Shin, C., Jang, S., and Min, D.-J., 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: *Geophysical Prospecting*, **49**, No. 5, 592–606.

Tromp, J., Tape, C., and Liu, Q., 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: *Geophysical Journal International*, **160**, No. 1, 195–216.

Yuan, Y. O., Simons, F. J., and Tromp, J., 2016, Double-difference adjoint seismic tomography: *Geophysical Journal International*, **206**, No. 3, 1599–1618.