Toward robust multicomponent FWI on land data: handling topography and data conditioning

Raul Cova, Bernie K. Law and Kris Innanen

ABSTRACT

Successful full-waveform inversion (FWI) studies using multicomponent marine data have been often reported in the literature. However, FWI applications to multicomponent land data remain limited. Among the challenges for a successful FWI in this setting we can find source repeatability, receiver coupling, rough topography, near-surface heterogeneities and strong elastic and attenuation effects. Due to the difficulty of including all these effects during the inversion process, it is a common practise to minimize their imprint on the data by conditioning the data before the inversion. Here we present a framework to study the effect of this conditioning on the FWI output. We first address the convenience of using a finite difference algorithm to include the topography during the forward modelling and its implications on the modelling of surface-waves. We compared this with the output of a spectral element method (SEM). The latter one showed a more accurate modelling of the surface-wave dispersion profile expected from the data. Also, less backscattered energy resulting from the discretization of the model was obtained. A benchmark dataset was also created to understand the effect of the data conditioning processes on the FWI output. Well log data available along the Hussar 2D-3C seismic line was used to build V_p, V_s and density models. A near-surface V_p model derived from first arrivals tomography was used to include velocity changes in the near-surface. Then, surface-consistent short-wavelength static corrections, amplitude corrections, and deconvolution operators, for the vertical and horizontal components, were derived from the real data and their inverse applied to the synthetic data. The goal was to obtain a synthetic dataset that would include some of the effects that are observed in the real multicomponent land data. This dataset will be used for understanding to what extent the conditioning of the data affects the inversion output and what strategies can be used to minimize their imprint in the inversion process.

INTRODUCTION

In the past years, full waveform inversion (FWI) has become an important tool for building high-resolution velocity models. These models can be used for improving the output of imaging processes and for reservoir characterization. FWI seeks to recover the subsurface model parameters that minimize the misfit between recorded and modelled data. This is achieved by performing iterative updates to the model parameters.

To reduce the computational burden, the seismic data is usually modelled as the result of seismic waves propagating in an acoustic medium. However, complex elastic effects, such as converted-waves, surface-wave propagation and amplitude variations with offset, are invariably recorded on the real world. Elastic FWI provides the means to account for some of these effects during the inversion process, yielding a more complete characterization of the subsurface.

Early applications of FWI using elastic approximations on single component data were

reported by Crase et al. (1990) and Djikpesse and Tarantola (1999). More recently, Sears et al. (2008, 2010), Prieux et al. (2013a,b) and Vigh et al. (2014) used multicomponent data to demonstrated the application of elastic FWI on real data. All of these, used data acquired from marine surveys.

The application of FWI on land data presents a difficult challenge for practitioners. One of the features that difficult FWI on land data is the presence of complex irregular topography. In conventional processing, the effect of topography is removed by applying static corrections. However, this is usually done under the assumption of surface-consistency which might be invalid when near-surface velocities change gradually or where there exist strong variations in elevation. Therefore, proper modelling of seismic data for EFWI should include the topography to properly account for its effect on the traveltime of the events.

Including the topography leads to a second challenge in the application of FWI to land data, which is the introduction of surface-waves. This wave mode accounts for a very large part of the energy recorded on real datasets. Therefore, FWI practitioners are faced with the decision of either removing them from the data and ignore them during the FWI process or use them and exploit the information they carry.

Other characteristics of actual land data that difficult FWI are the poor signal-noise-ratio, unknown and variable source signature and poor receiver coupling to the ground. In this study, we present our initial approximations to address some of these challenges via data conditioning or by including them during the inversion, whenever possible. We first compare the use of a finite difference algorithm versus a spectral element method for the incorporation of the topography in the modelling of multicomponent seismic data. We pay special attention to the modelling of surface waves in each case. We then computed a set of synthetic data using well log data and a tomographic near-surface velocity model. Surface-consistent short-wavelength static corrections, deconvolution operators and amplitude balance corrections were extracted from real seismic data recorded nearby the wells. The inverse of these operators was then applied on the synthetic data to approximate some of the characteristics present on the real data. In future projects, this dataset will be used for further understanding of the effects of data conditioning on the FWI output.

FINITE DIFFERENCE MODELLING

To study the convenience of using a finite difference algorithm to simulate elastic wave propagation from the topography we built the velocity and density models displayed in Figure 1. The actual topography from the Hussar 2D-3C seismic line was used to separate the air layer with parameters $V_p = 310$ m/s, $V_s = 0$ m/s and $\rho = 1.25$ Kg/m^3, from the near-surface layers. The elevation profile along the seismic line varies between 899.5 m and 983.4 m above sea level for an elevation change of 83.9 m. Since the original common midpoint distance of the seismic line was of 5 m we built models with cell sizes submultiple of 5.

The code used for the forward modelling was developed by Bohlen (2002). Despite allowing for 3D modelling and viscoelastic effects in this initial stage of the study we only

used the 2D version and did not consider viscoelastic effects. These will be part of the next stage of this project.

Two very important stability conditions must be met to properly model seismic waves propagation using a finite difference code. The first of them helps to avoid a phenomenon usually referred to as grid dispersion. This requires the grid size (dh) to be between 4 to 12 times smaller than the minimum wavelength we wish to model (Bohlen, 2002). This can be expressed as

$$dh \leq \frac{V_{min}}{n f_{max}}, \tag{1}$$

where V_{min} and f_{max} are the minimum velocity and maximum frequency we wish to model. For the 12-th order space finite difference operator used in this study n must be equal to 4 (Bohlen, 2002). For a minimum subsurface velocity of 500 m/s and maximum frequency of 40 Hz we need a grid size smaller than 3.12 m.

The second stability condition requires that the time step (dt) used for the modelling must be less than the time it would take a wave to travel a distance equal to the grid size. This is expressed as

$$dt \leq \frac{dh}{h\sqrt{2}V_{max}}, \tag{2}$$

where V_{max} is the maximum velocity in the model and h is a factor that depends on the order of the finite difference approximation. For the 12-th order space approximation we used, h is equal to 1.34 for a Taylor finite difference operator (Bohlen, 2002). Considering a maximum velocity of 3000 m/s, and a grid size of 2.5 m the largest time step we could use in this part of the study was 4.4×10^{-4} s. The criterion expressed in equation 2 is called Courant-Friedrichs-Lewy criterion (Courant et al., 1928, 1967).

Figure 2a shows a zoom between the x-coordinates 1200 m and 1300 m. There we can see how the discretization of the model into cells of 2.5×2.5 m difficulted the positioning of the receivers at their actual locations. Moreover, at every elevation change a very large lateral velocity contrast is introduced. This will result in the generation of significant backscattered energy from any wave propagating or arriving at the surface. A common solution for this problem is to smooth the model parameters. Figure 2b shows the result of applying a radial Gaussian smoother with a 5 m half-width. Even though the lateral velocity contrast has been attenuated the positioning of the receivers is still constrained by the cell size. A very expensive solution to this problem is to decrease even further the cell size. In Figure 2c we have decreased the cell size to 1×1 m. Despite allowing a better positioning of the receivers this solution results in a significant increase in the computational cost of the operation. First, decreasing the cell size naturally increases the total number of cells in the model, increasing the number of operations that must be performed at each time step. Secondly, to satisfy the Courant stability criterion the time step also needs to be decreased. Therefore, to cover the desired total propagation time a larger number of time steps will be performed. For the model with the 2.5 m cell size, we obtained a computation time of 110 s, for the simulation of 1 s of wave propagation at a time step of 2×10^{-4} s for a total of 5000 time steps. The computational time for the same record length with the 1 m cell size models resulted in an total time of 1140 s for the simulation of 10000 time
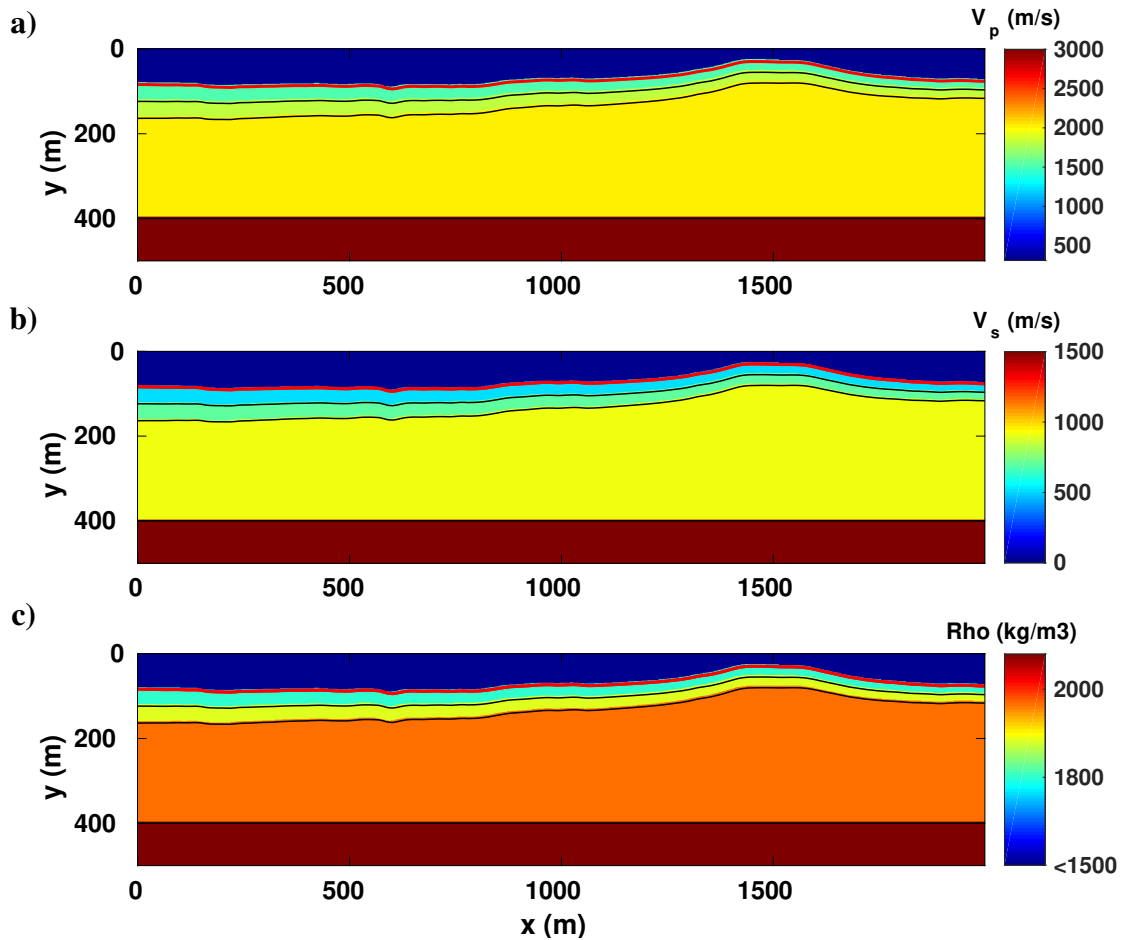
FIG. 1. a) P-wave velocity, b) S-wave velocity and c) density models used for the finite difference modelling of elastic data from the topography. An air layer with V_p=310 m/s, V_s=0 m/s and ρ=1.25 Kg/m^3 was included above the topography to model free-surface conditions.
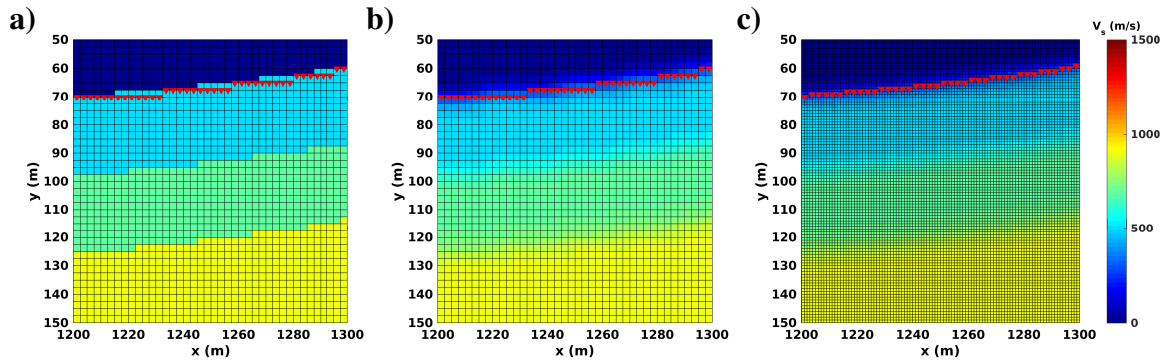
FIG. 2. Zoom around the V_s model used for the finite difference modelling. a) Model with a grid size of 2.5×2.5 m, without smoothing. b) Same grid as a) after applying a 5 m half-width radial Gaussian smoother. c) Refined model a with grid size of 1×1 m and a 5 m half-width Gaussian smoother. The red triangles indicate the receiver locations.

steps at 1×10^{-4} s. As we can see the computational cost increased in a factor of 10 after decreasing the cell size and increasing the number of time steps.

The modelled vertical and horizontal components recorded in each scenario are displayed in Figure 3. As expected, Figure 3a displays very distorted source records. Particularly on the horizontal component, the very large lateral velocity contrasts have introduced very large amplitude variations between receivers. Moreover, significant backscattered energy can be seen in the area of the record covered by the noise cone produced by the surface-waves. This effects are minimized after smoothing the model parameters. The vertical component source gather displays a very clean record. Despite the deformation on the moveout of the events produced by the shape of the topography, these event are very clear and easy to identify. However, the horizontal component still shows some distortions. Due to the very low velocity of S-waves in the near-surface, the 2.5 m constraint in the elevation results in significant artificial static effects. Additionally, the smoothing operation has decreased the S-wave velocities around the receiver locations even further, introducing larger static effects. Finally, Figure 3c displays the source gather recorded with the 1 m cell size models. Under this configuration both components of the seismic data displays a very clean character. Despite some backscattered energy is still present in the noise cone, the PS event recorded on the horizontal component is better defined.

Although smoothing the models reduced the amount of backscattered energy, the near-surface profile was altered. This has an important effect on the dispersion spectrum of the surface-waves. Figure 4 shows the horizontal and vertical component data recorded over a horizontally-layered model with different smoothing degrees. It is clear how smoothing the model introduces different responses from the surface-waves. This is even more clear on the dispersion spectra obtained from the vertical component data from each model (Figure 5). The dashed line represents the analytical solution (Lai, 1998) for the original two-layers near-surface model. In Figure 5a the analytical solution falls very close to the maximum energy of the spectrum. When the model is smoothed (Figure 5b) the dispersion spectrum is significantly changed. The shorter wavelengths (higher frequencies) now travel at lower velocities than before. With a 10 m half-width Gaussian smoother (Figure 5c) the dispersion profile is distorted even further. The effect is similar to the propagation of surface
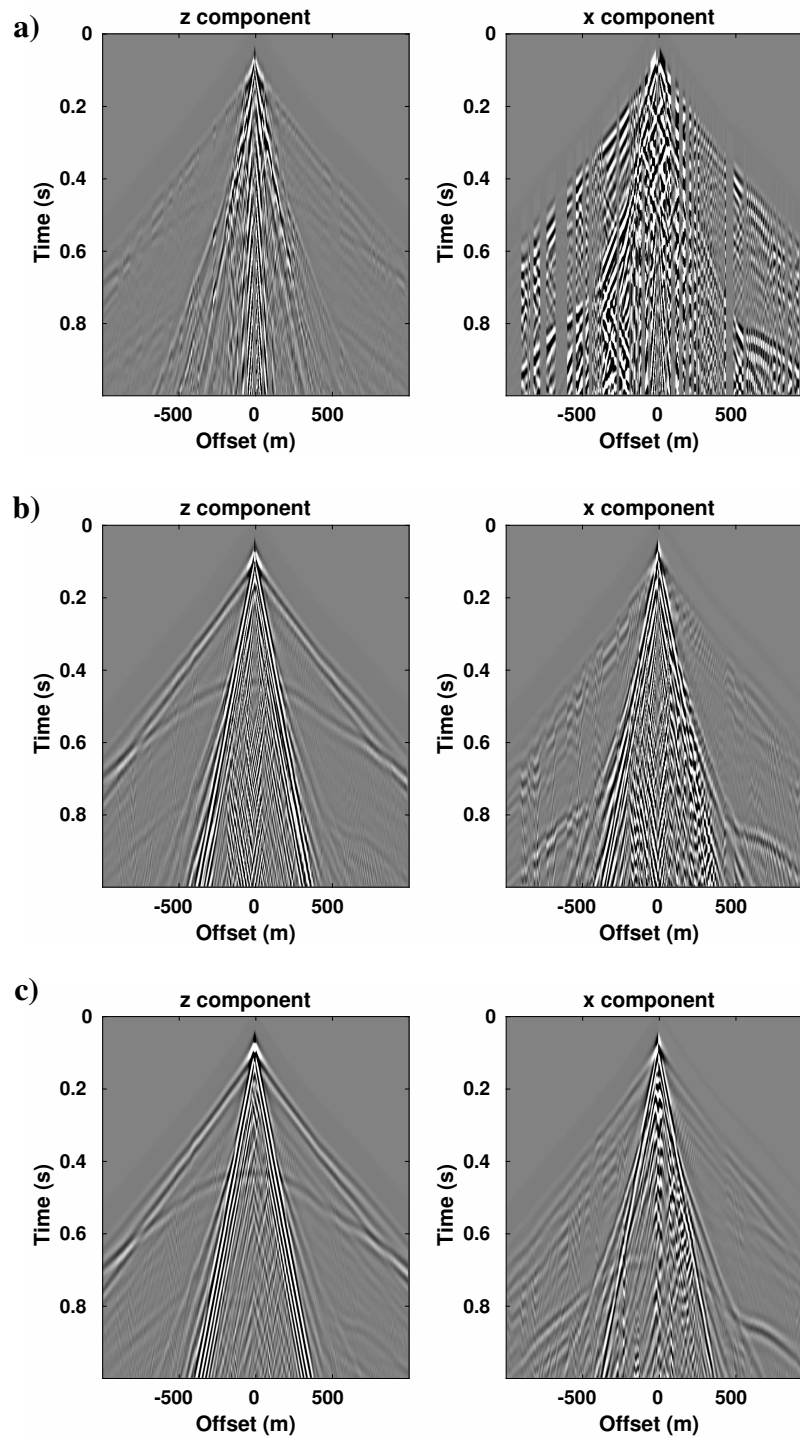
FIG. 3. Vertical and horizontal component data modelled with different grid sizes and smoothing degrees, obtained from a source located in the middle of the model in Figure 1. a) Data obtained with the model with grid size 5×5 m, without smoothing. b) Data obtained after applying a 5 m half-width radial Gaussian smoother. c) Data obtained with the refined model with a 1×1 m grid size and a 5 m half-width Gaussian smoother. Refining and smoothing the parameter models minimizes the amount of backscattered energy resulting from using a rectangular grid to model the topography.

waves with higher modes.

SPECTRAL ELEMENTS MODELLING

Spectral element methods (SEM) offer the possibility to introduce deformed grids for the simulation of waves propagation. The ability to mesh the model conforming defined structural features allows for a better representation of the topography. The SEM algorithm used for this part of the study was developed by Ampuero (2012).

Figure 6 displays the velocity and density models used for the SEM modelling. The model parameter values and the topography are the same used in the previous section. Notice how the mesh adapts to the shape of the topography. As a result, a more accurate placement of the receivers over the topographic profile can be achieved.

Despite the algorithm is computationally more expensive than the FD algorithm, the flexibility provided by the mesh allowed us to obtain results with computing times very close to the previous ones. The mesh we defined had a maximum distance between nodes of 10×10 m with 4 elements between nodes. After reshaping the mesh by including the topography and the geometry of the near-surface layers, the maximum and minimum distances between nodes obtained were of 3.42 m and 0.51 m, respectively. The time step was set at 2.4×10^{-4} s to satisfy the Courant stability condition for a total of 4167 time steps to complete a propagation of 1 s. The average computational time we obtained was of 119 s per source record, which translates in an increase of just 8% respect to the FD algorithm. Notice that despite the total number of time steps was about 16% smaller we still obtained a slighter higher computational cost. However, such increase could be affordable if the improvement in the quality of the data compensates for it.

Figure 7 displays the source gather obtained from the SEM modelling. Notice that without the need of smoothing the model, very clean source records are obtained for both components. The propagation and dispersion of the surface-waves is better defined. Also the PP and PS events are less distorted.

To study the dispersion spectrum of the surface waves obtained using the SEM algorithm we modelled source records over the horizontally-layered medium in Figure 8. The dispersion spectrum obtained from the surface-waves recorded on the vertical component data is displayed in Figure 9. There we can see how the analytical solution falls in the region defined the maximum energy of the spectrum. Notice that the match is closer than the one obtained using the finite difference algorithm. Since no smoothing is needed to avoid amplitude distortions, the dispersion spectrum and the velocity profile are not altered, allowing for a better representation of the original conditions of the model.

The results obtained with the 10 m mesh are comparable with those from the finite difference modelling with the 1 m grid size. More importantly, the cost of this simulation is similar to that obtained with the 2.5 m finite difference grid. Therefore, a more accurate elastic wavefield propagation can be performed with SEM algorithms at reasonable computational cost, without the need of introducing artificial velocity changes in the near-surface.
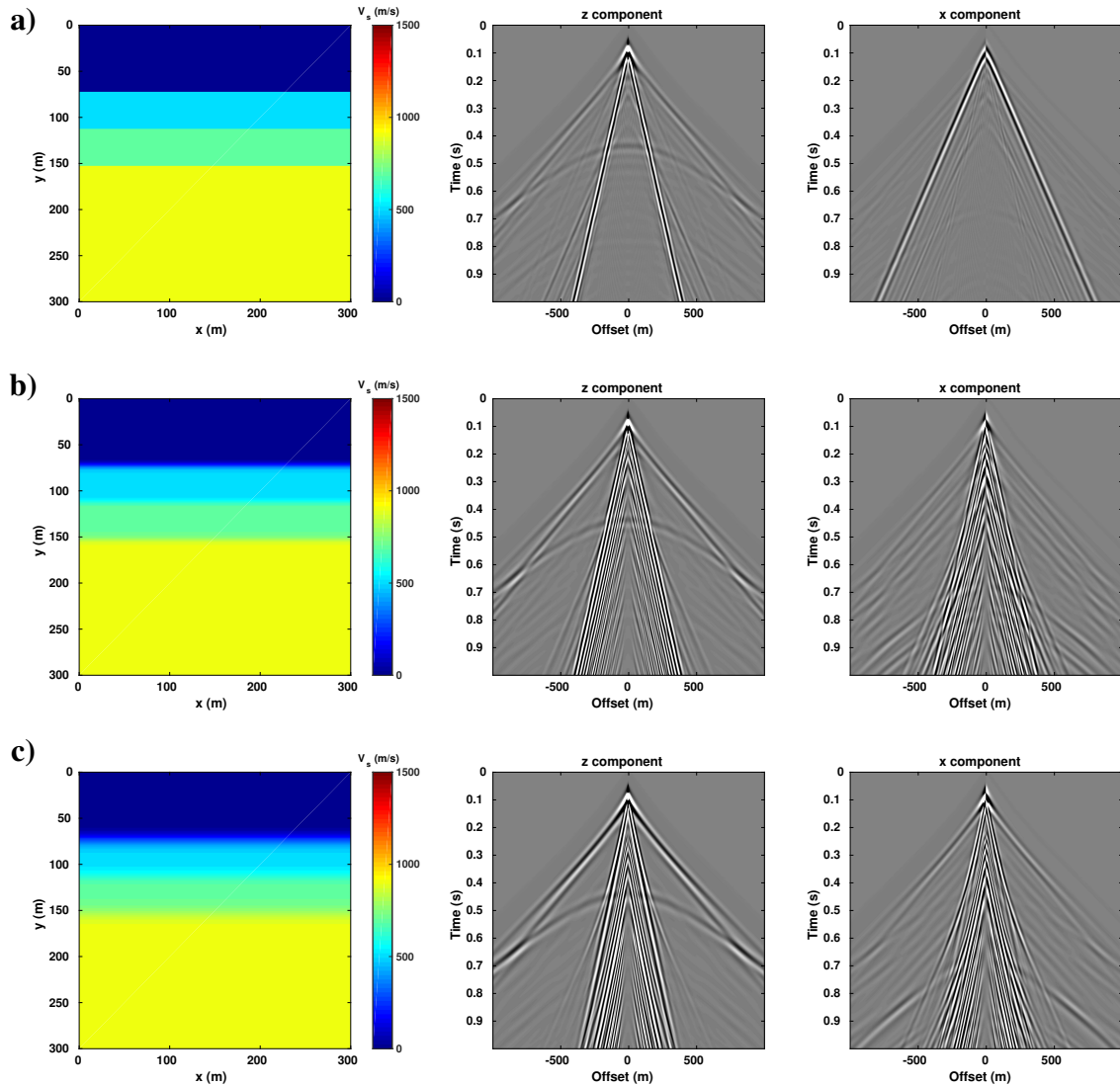
FIG. 4. Resulting data from horizontally layered S-wave velocity models a) without smoothing, b) after applying a 5 m half-width radial Gaussian smoother and c) with a 10 m half-width radial Gaussian smoother. Notice how the dispersive character of the surface-wave data changes as the smoothing increases.
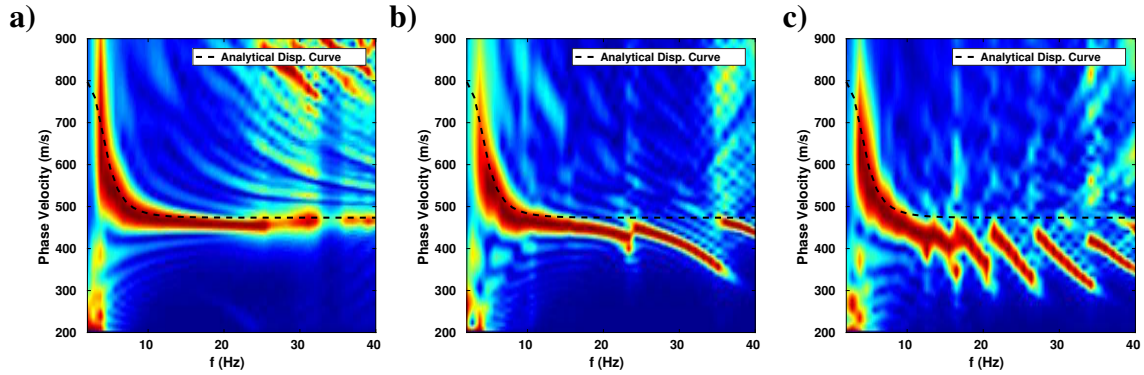
FIG. 5. Dispersion spectra obtained from the vertical component data computed with the horizontally layered models a) without smoothing and after applying a b) 5 m and c) 10 m half-with radial Gaussian smoothers. The dashed line represents the analytical solution for a 1D velocity profile without smoothing. The smoothing operation alters the dispersion spectrum of the surface-waves.
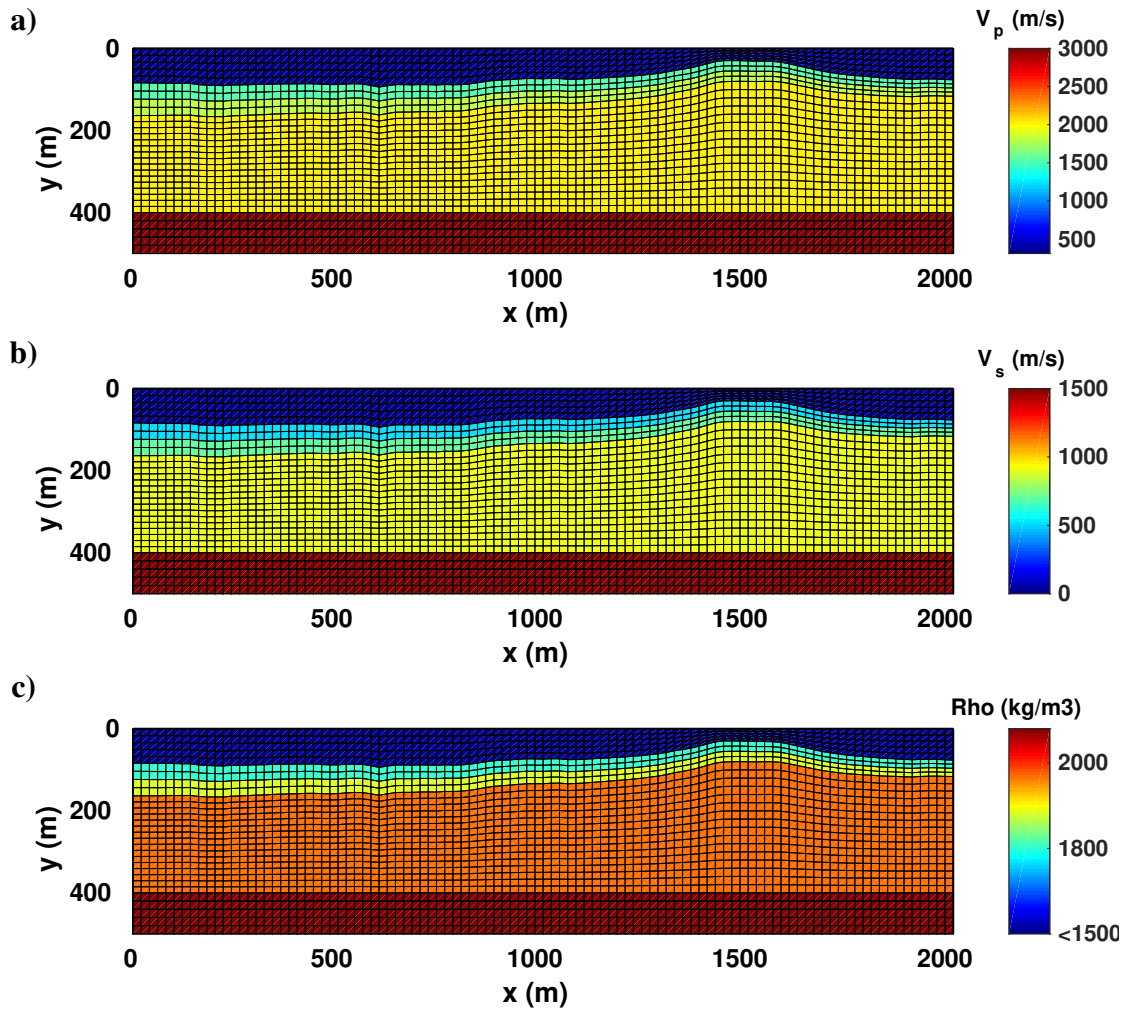
FIG. 6. a) P-wave velocity, b) S-wave velocity and c) density models used for the spectral elements modelling of elastic data from the topography. An air layer with V_p=310 m/s, V_s=0 m/s and ρ=1.25 Kg/m^3 was included above the topography to model free-surface conditions. The mesh has a maximum size of 10 m by 10 m with 4 elements per node.
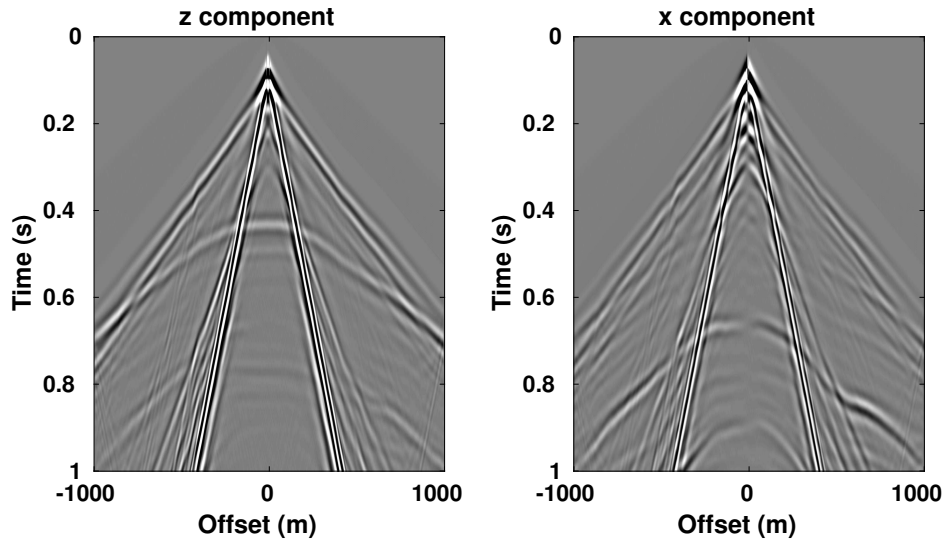
FIG. 7. Vertical and horizontal component data modelled with the spectral elements method. Notice that no backscattered energy is produced by the elevation changes.
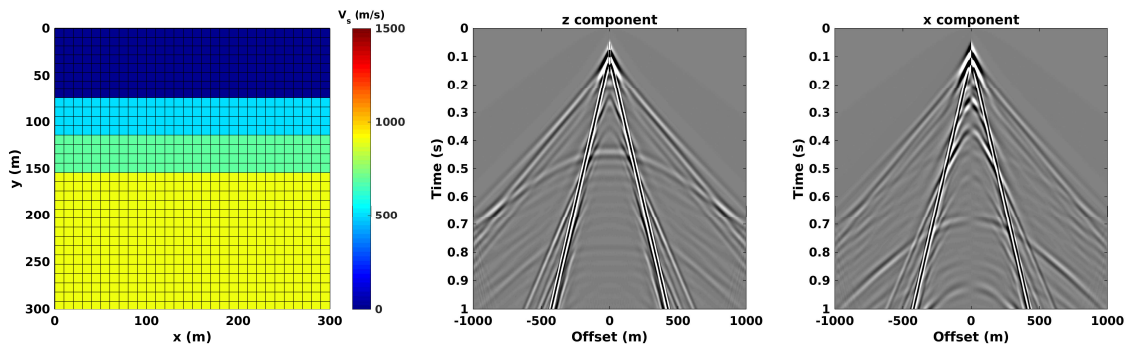

FIG. 8. Horizontally layered S-wave velocity model and the resulting vertical and horizontal component data using the spectral elements method.
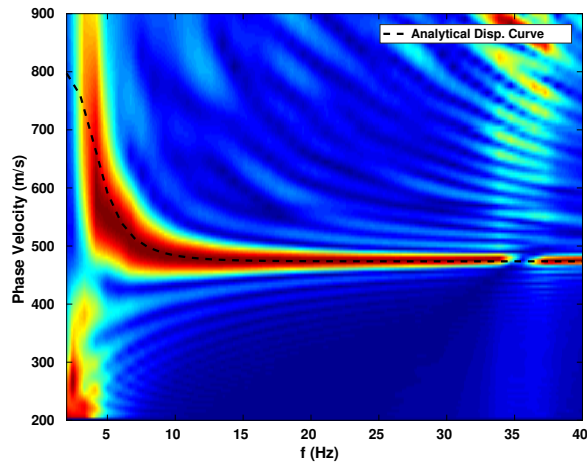

FIG. 9. Dispersion spectrum obtained from the vertical component data using the SEM algorithm and the horizontally-layered models. Notice how the analytical solution (dashed line) falls exactly on top of the area of the spectrum with maximum energy.
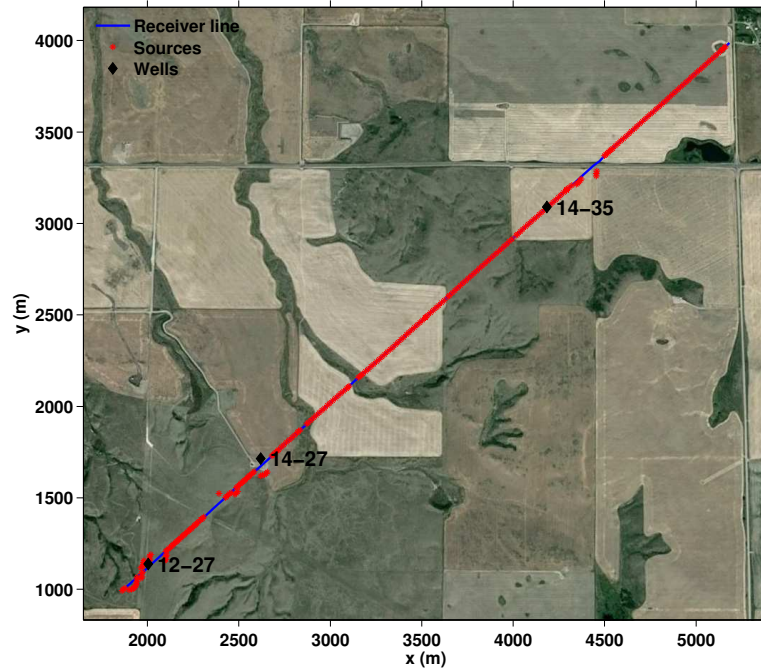
FIG. 10. Hussar 2D-3C seismic line base-map and the location of the wells used in this study.

BUILDING THE BENCHMARK DATASET

To study the effect of the conditioning of the data before performing a FWI we designed a controlled experiment using modified synthetic data. We computed synthetic data using known velocity and density models and then imposed on them some of the effects observed in actual seismic land data.

Figure 10 shows a satellite view of the area around the Hussar 2D-3C seismic line. A total of 257 shots stations, spaced approximately every 20 m and, 448 receiver stations spaced every 10 m were used for the acquisition of the real data. Three wells with velocity and density information are located along the line covering the majority of the sedimentary column. Figure 11 shows the well logs used to compute the subsurface models. A linear transformation between P-wave and S-wave velocities, calibrated with the data from well 12-27, was used to compute the S-wave velocity logs for wells 14-27 and 14-35. The equation obtained for this transformation was $V_s = 0.6074V_p - 420.43$ for V_p and V_s in m/s. Additional well log data from a well located at 7.8 Km from the seismic line was used to extend the model parameters to a maximum depth of 2000 m.

Figure 12 shows the V_p, V_s and density models obtained from the well logs. The tops interpreted in the wells were used to define horizons that guided the interpolation and extrapolation of the well data. A layer above the topography with the elastic parameters of the air ($V_p = 310$ m/s, $V_s = 0$ m/s and $\rho = 1.25$ Kg/m^3) was included to simulate the free-surface conditions. The P-wave velocities in the near surface (< 150 m depth) were obtained from a tomographic velocity model computed from the first arrivals picked on the vertical component data (Law and Trad, 2017). A fixed V_p/V_s ratio of 4 was used to compute the S-wave velocities in the near-surface. The density values in this region was
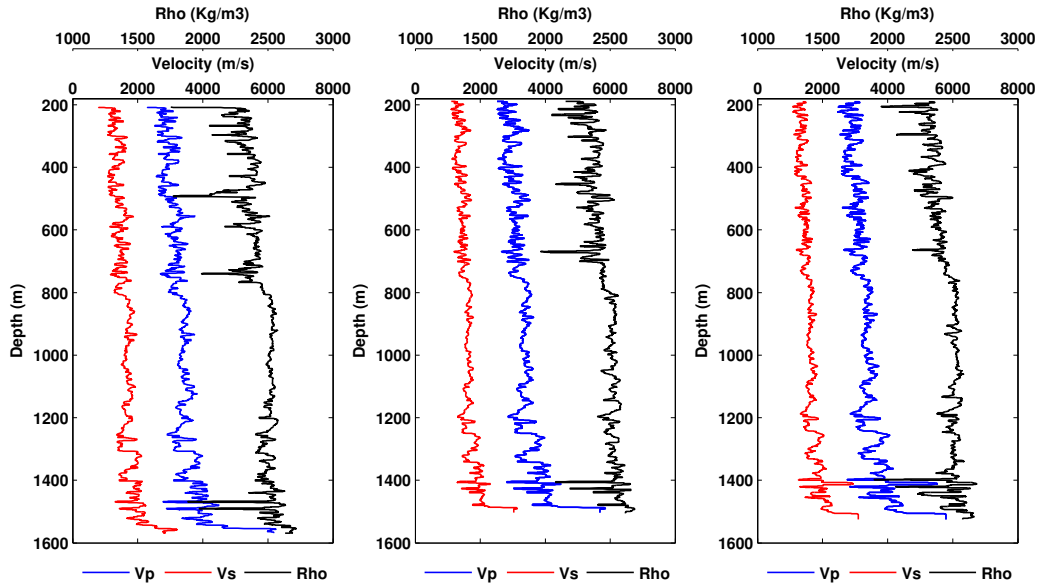
FIG. 11. Well log data used to build the subsurface models.

fixed at 2000 Kg/m^3.

A total of 200 shots spaced every 20 m at a depth of 15 m from the topography were modelled using an elastic finite differences algorithm (Bohlen, 2002). The cell size used to discretize the model was of 1×1 m. A radial Gaussian smoother with a half-width of 5 m was used to avoid the introduction of numerical artifacts during the simulation.

Figure 13 shows the horizontal and vertical component data recorded from a source located at the x-location 1900 m. There, we can see a series of PP reflections on the vertical component data and PS events on the horizontal component data. Notice that the moveout of the events has been clearly deformed by the changes in the receivers elevation. Since S-wave velocities in the near-surface are 4 times slower, their effect is magnified on the PS events. This type of very clean data is the standard input for testing FWI algorithms.

To produce synthetic data with conditions closer to the actual land data in the Hussar 2D-3C seismic line, we extracted all the surface-consistent corrections derived from the actual data and applied their inverse on the synthetic data. Figure 14 shows the extracted short-wavelength static corrections from the actual data. These corrections represent the components of the static solution that could not be explained by the tomographic near-surface velocity model we used to compute the synthetic data. The static effects for the horizontal component were computed assuming a factor of 2 between P-wave and S-wave residual static corrections. These surface-consistent corrections were then used to shift the traces depending on their source-receiver combination.

Similarly, Figure 15 displays the weights obtained for the surface-consistent amplitude balancing of the real data. These weights are meant to capture effects like variations in the energy of the source, different receivers coupling to the ground, local changes in the soil conditions around sources and receiver locations among other possible causes. The inverse of this weights were then applied to scale the synthetic data in a surface-consistent manner.
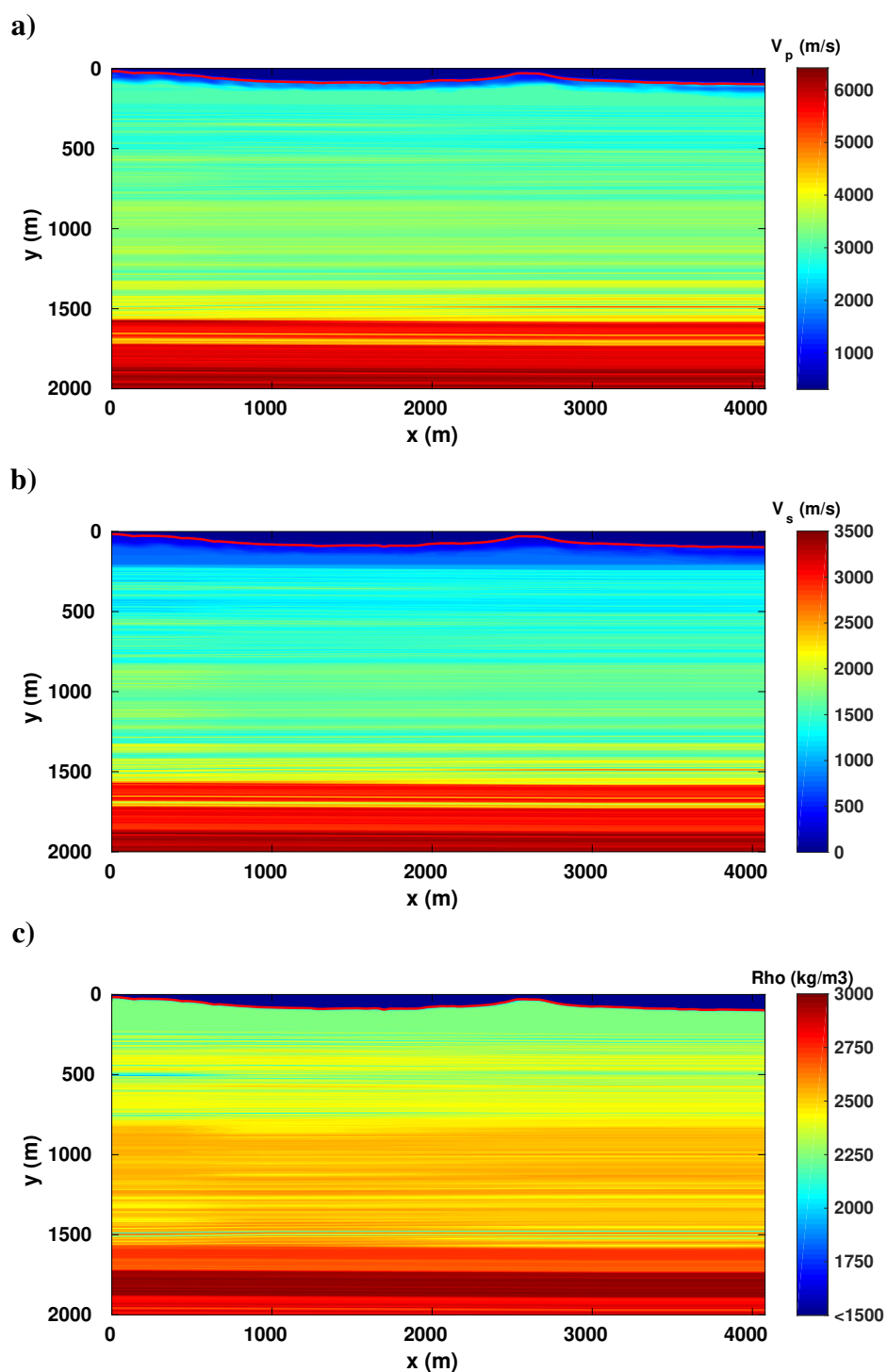
FIG. 12. (a) P-wave velocity, b) S-wave velocity and c) density models used for the finite difference modelling of elastic data along the Hussar 2D-3C seismic line. An air layer with V_p=310 m/s, V_s=0 m/s and ρ=1.25 Kg/m^3 was included above the topography to model free-surface conditions. A grid size of 1×1 m and a radial Gaussian smoother with a half-width of 5 m were used to build the models.
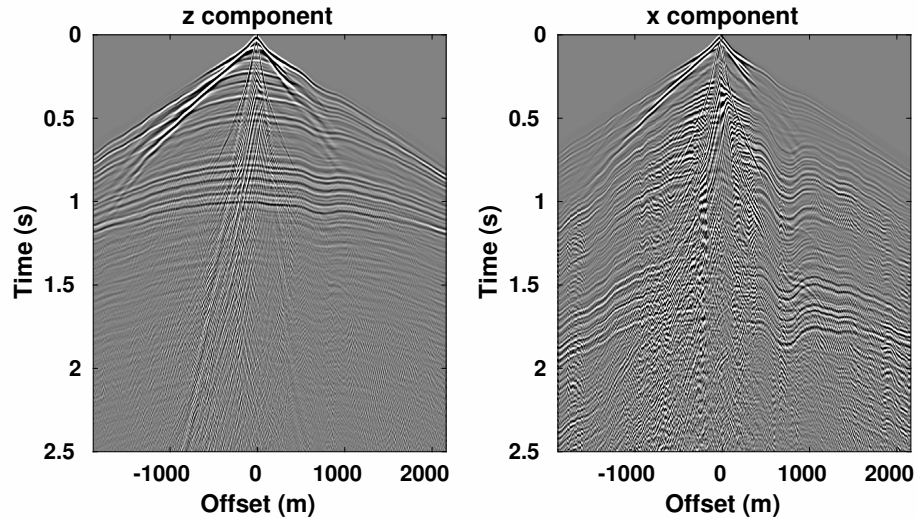
FIG. 13. Vertical and horizontal component data obtained from a source located at $x =$ 1900 m in the models displayed in Figure 12.
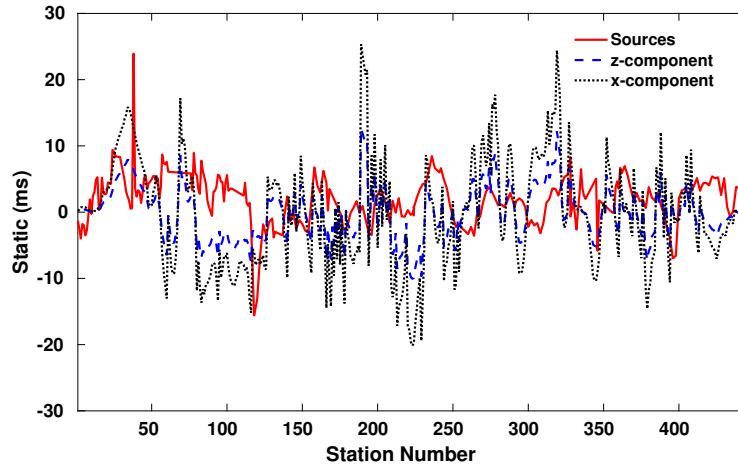

FIG. 14. Surface-consistent short wavelength static corrections extracted from the actual data from the Hussar survey. Here we use a factor of 2 to model the horizontal component static effects.
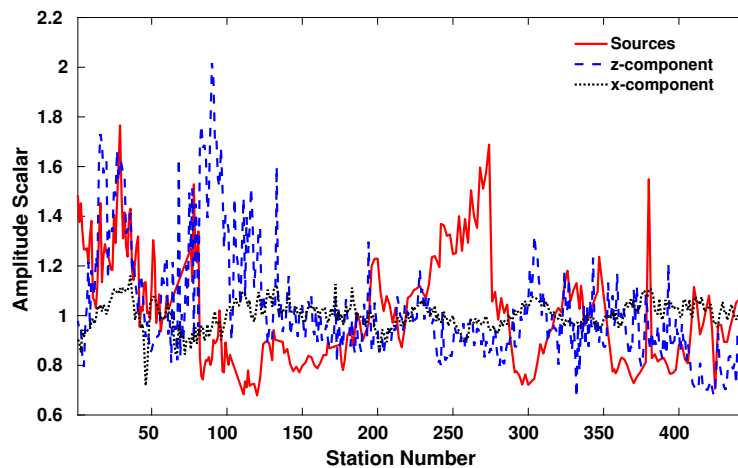

FIG. 15. Surface-consistent amplitude corrections obtained from the Hussar 2D-3C dataset.
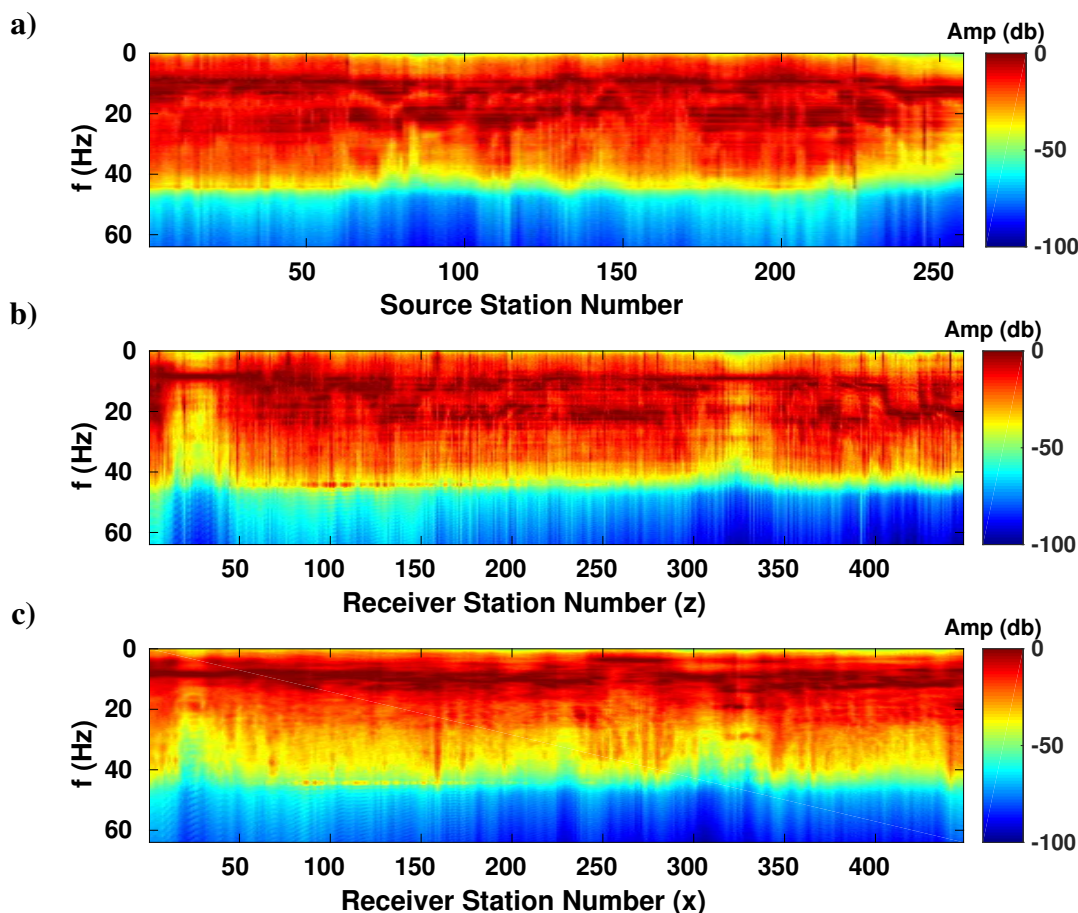
a)

b)

c)

FIG. 16. Spectra of the surface-consistent deconvolution operators computed from the Hussar field data. a) Source, b) vertical component and c) horizontal component deconvolution operators spectra.

Finally, the last set of modifications applied to the synthetic data are presented in Figure 16. There we can see the amplitude spectra of the deconvolution operators obtained from the real data. This operators are meant to capture any spectral response (filter) that is consistent at each source and receiver component. In conventional processing this information would be used to derive inverse filters that will remove these effects. In our study we used these spectra to design minimum phase filters that were directly applied to the data to imprint their effect on the synthetic traces.

The horizontal and vertical component source gathers obtained after applying this series of modifications are displayed on Figure 17a. There we can see how the frequency content of the original data has been degrade and the amplitudes of the traces has been altered. The effect of the residual statics is also manifested in the presence of jitter in the moveout of the events.

For comparison purposes Figure 17b displays the horizontal and vertical component data from the actual seismic line. An amplitude scaling of $\sqrt{t}$ was applied to the field data to approximate the amplitudes modelled in our 2D simulation. Notice that the surface-wave
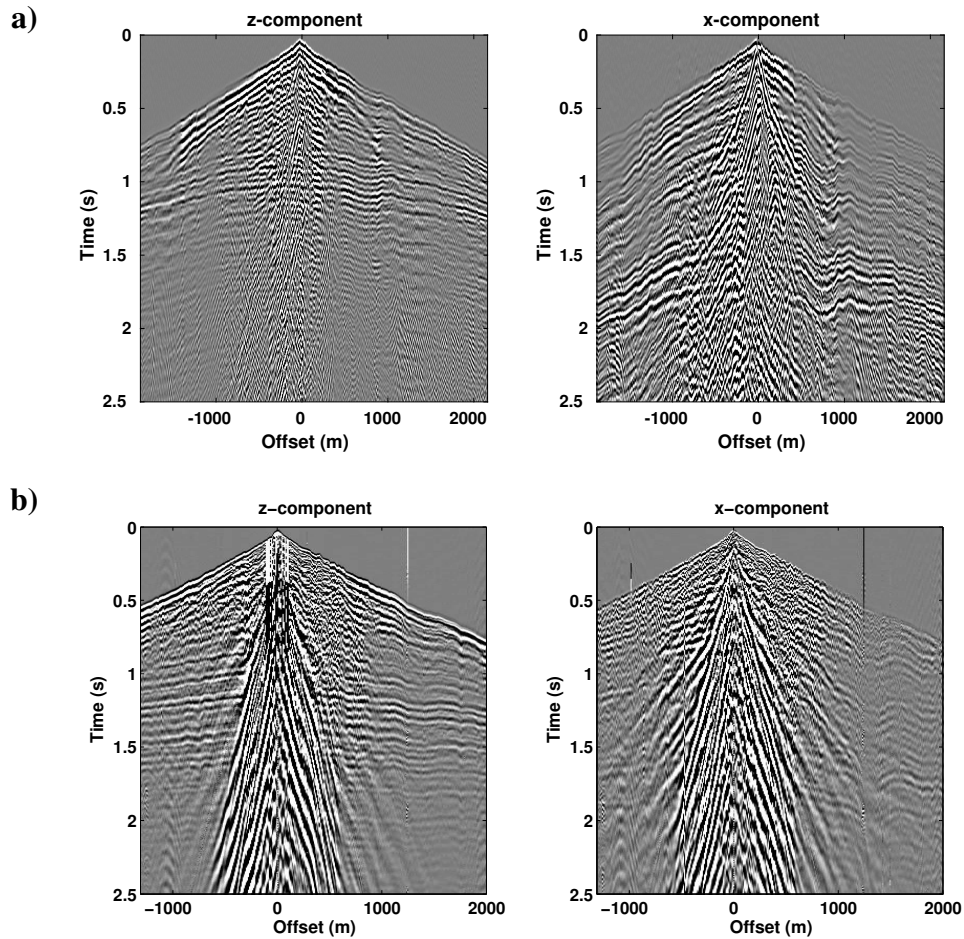
FIG. 17. a) Synthetic vertical and horizontal component data after applying the inverse of the corrections derived from the field data. b) Actual vertical and horizontal component data from the Hussar 2D-3C dataset. Notice how the synthetic data now resemble the character of the field data.

data on the actual records show larger amplitudes than in the modelled data. Even though some of this energy might have been attenuated during the application of the deconvolution operators, when compared to the original synthetic records (Figure 13), the relative amplitude of the surface-waves to the rest of the data does not match the one in the actual data. To properly model the surface-wave data we will need to compute a more accurate S-wave velocity model for the near-surface.

In terms of the moveout of the reflections and frequency content of the data, the simulated land data resemble the field data very closely. There are certainly many physical effects that were not included during the modelling. Particularly, anelastic attenuation may have an important effect on the decay of the recorded amplitudes, specially on the horizontal component data. Notice that in the field data the amplitude of the converted-wave events decays faster than in the modelled data. Overall, the simulated land data now contains some of the characteristics that we can find in the actual data and they can be used to challenge our FWI algorithms.

CONCLUSIONS

Including the topography in the FWI process requires of algorithms able to accommodate elevation changes while honouring the propagation of waves along the surface. Using rectangular grids for the finite difference modelling of elastic waves demands of very fine cell sizes that may render the FWI process prohibitive in terms of computational cost. Smoothing the model parameters partly avoids the need for a finer grid but it introduces artificial velocity changes into the modelling. This introduces significant changes in the modelling of the surface-wave data. On the other hand, the SEM algorithm used in this study provided accurate elastic modelling while honouring the shape of the topography, without introducing significant numerical noise in the data. Moreover, the ability of the spectral element algorithm to adapt to the geometry of the interfaces included in the model allowed us to optimize the computational burden of the modelling. In this study the SEM modelling using a mesh of 10×10 m provided results comparable to those obtained with a 1×1 m finite difference grid, at a computational cost 10 times lower. Therefore, spectral element method seems to be an attractive alternative for the modelling and inversion of multicomponent land data.

We also computed a set of synthetic data that resembles more closely the character of actual land data. We achieved this by extracting short-wavelength static corrections, amplitude balancing scalars and deconvolution operators from the field data and applying their inverse on the synthetic data. The result was a dataset that contains some of the features that are usually removed during the pre-processing of the data. We will use this dataset to design optimal strategies for multicomponent data conditioning in future elastic FWI projects.

ACKNOWLEDGEMENTS

REFERENCES

Ampuero, J. P., 2012, SEM2DPACK - A Spectral Element Method for 2D wave propagation and fracture dynamics, with emphasis on computational seismology and earthquake source dynamics. Version 2.3.8. URL https://sourceforge.net/projects/sem2d

Bohlen, T., 2002, Parallel 3-D viscoelastic finite-difference seismic modelling: Comput. and Geosci., **28**, No. 8, 887–899.

Courant, R., Friedrichs, K., and Lewy, H., 1928, Über die partiellen Differenzengleichungen der mathematischen Physik: Mathematische Annalen, **100**, 32–74.

Courant, R., Friedrichs, K., and Lewy, H., 1967, On the partial difference equations of mathematical physics: IBM Journal, 215–234.

Crase, E., Pica, A., Noble, M., McDonald, J., and Tarantola, A., 1990, Robust elastic nonlinear waveform inversion: Application to real data: Geophysics, **55**, No. 5, 527–538.

Djikpesse, H. A., and Tarantola, A., 1999, Multiparameter l1 norm waveform fitting: Interpretation of gulf of mexico reflection seismograms: Geophysics, **64**, No. 4, 1023–1035.

Lai, C., 1998, Simultaneous inversion of rayleigh phase velocity and attenuation for near-surface site characterization.

Law, B., and Trad, D., 2017, Robust refraction statics solution using feedback from reflection data, 2666–2670.

Prieux, V., Brossier, R., Operto, S., and Virieux, J., 2013a, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the valhall field. part 1: imaging compressional wave speed, density and attenuation: Geophysical Journal International, **194**, No. 3, 1640–1664.

Prieux, V., Brossier, R., Operto, S., and Virieux, J., 2013b, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the valhall field. part 2: imaging compressive-wave and shear-wave velocities: Geophysical Journal International, **194**, No. 3, 1665–1681.

Sears, T., Singh, S., and Barton, P., 2008, Elastic full waveform inversion of multi-component obc seismic data: Geophysical Prospecting, **56**, No. 6, 843–862.

Sears, T. J., Barton, P. J., and Singh, S. C., 2010, Elastic full waveform inversion of multicomponent ocean-bottom cable seismic data: Application to alba field, u. k. north sea: Geophysics, **75**, No. 6, R109–R119.

Vigh, D., Jiao, K., Watts, D., and Sun, D., 2014, Elastic full-waveform inversion application using multicomponent measurements of seismic data collection: Geophysics, **79**, No. 2, R63–R77.