Answers to referee comments on "Full waveform inversion of short-offset, band-limited seismic data in the Alboran basin (SE Iberia)" by C. Gras et al.

We would like to thank the reviewer for her thorough revision and insightful comments. She clearly understood the goals and contributions of the presented work and the comments will surely help to improve the manuscript. Our answers to her questions are nested below (in blue).

Page 2:

- *Referee*: Lines 20-21. The sentence is a bit redundant with respect to what you state on lines 10-11.

- *Answer*: We changed this by a new sentence in Line 17: "In deep water settings, the critical distance for refracted waves is often beyond this distance."

- *Referee*: Line 25: part of the sentence ". . ., so there is typically no signal above noise below this frequency." This construction is a bit strange. Perhaps adding ". . .above noise level.."?

- *Answer*: We changed this by a new sentence in Line 23: "The most commonly used sources in MCS systems are airguns, which are typically not capable of generating energy below 4-5 Hz, so in those cases the signal to noise ratio is very low."

- *Referee*: Line 29: There is one right bracket extra, please remove. In addition, it would be good to acknowledge the work of authors who did significant work on implementing re-datuming technique and applying it to real data: Arnulf et al., 2011(GRL); 2013 (GJI); 2014 (JGR); Henig et al., 2014 (G-cubed) etc. In fact Qin and Singh (2017; 2018) use the code introduced by Arnulf et al. (2011).

- *Answer*: We removed the extra bracket. The sentence has now changed to include new references in Line 27: "A common approach to mitigate the above-mentioned issues and apply TTT+FWI to marine MCS data in deep water settings is by re-datuming the collected data (Arnulf et al., 2011, 2014; Henig et al., 2012) to a virtual recording surface first (Qin and Singh, 2017),..."

Page 3:

- *Referee*: Line 22: It would be better to say: "to reveal refraction signal" instead of "to modify recordings"

- *Answer*: We changed it as: "to recover refracted signals" (Line 20).

- *Referee*: Lines 26-28: These three sentences could be summarized into one.

- *Answer*: The sentence is now reformulated in Line 24: "This procedure is easy to implement and can deal with spatial velocity variations within the water column and irregular datum surfaces, which are often located at the seafloor (Fig 2c)."

Page 4:

- *Referee*: Line 25: It would be good that you list all of the processing steps that you apply to the data before re-datuming. You mention later in the text that you did some filtering, but it would be good make it clear what steps you applied prior to DC.

- *Answer*: Processing steps are now listed in page 4 Line 23: "Prior to re-datuming, we: (1) deleted noisy channels, (2) applied Butterworth 1/3-40/60 Hz filtering and (3) muted direct arrival. We did not apply amplitude corrections. "

- *Referee*: Question: Do you consider 3-D effect in downward continuation? You mention this as an important effect for FWI. Perhaps as you are doing only TTT using DC data this may not be relevant.

- *Answer*: We do not consider 3-D effects in the downward continuation. As you mention, these are not relevant to perform TTT because we only use travel times. These effects would be relevant to perform waveform inversion of the downward continuation results. A comment clarifying this issue has now been added (see answer to previous question).

Page 5:

- *Referee*: It would be good to have a short paragraph that compares your re-datuming approach with respect to Arnulf et al. (2011).

- *Answer*: A short paragraph is now included in Page 5 Line 13: "Unlike re-datuming approaches that use the Kirchhoff implementation [Berryhill 1979, 1984; Shtivelman and Caning, 1988; Arnulf et al., 2011], the computation of extrapolated data at intermediate depth levels is needed. Despite the fact that it increases the computational complexity as compared with the Kirchhoff method [Arnulf et al., 2014], the wave equation imaging allows including laterally and vertically variable water velocity, in contrast with other techniques in which the water velocity is considered constant [Cho et al., 2016] or need a specific ray tracing [Shtivelman and Caning, 1988]. Recursive extrapolation makes that spatial velocity variations can be handled properly...".

- *Referee*: Lines 19-21: This statement is not completely true. Arnulf et al. (2011) chose a flat surface to extrapolate the data; it is a choice of the authors not a limitation of the method (Qin and Singh use the same method to re-datum to follow surface close to seafloor).

- *Answer*: Thanks for the clarification. This part of the statement (..., as is the case of other methods such as the one proposed by Arnulf et al. (2014)) is now removed.

- *Referee*: Question: Do you DC the data exactly to seafloor surface? There are authors that argue that the extrapolations surface should be at least several tens of meters above seafloor surface (e.g., Harding et al., 2016 - G-cubed)? It would be good to comment this.

- *Answer*: Yes, DC data had sources and receivers located at the node of the grid that is closest to the

seafloor. Harding et al. [2016] downward continue the data to a horizon 75 m above the seafloor to use the near-offset reflection as check on the accuracy of the DC parameters and the seafloor bathymetry. In this or similar studies where the extrapolation surface is several meters above the seafloor surface, the DC data is used to perform waveform inversion. So, it might be recommended to avoid the generation of surface waves during the modeling. In contrast, we only use the travel time information of the first arrivals for travel time tomography, so we do not need to do it. A specific comment on this issue is now included in Page 5 Line 22: "Moreover, the new virtual positions do not necessarily have to be in a flat surface or at several meters above seafloor surface as in other studies [e.g. Harding et al.,2016], here the new virtual surface is the seafloor relieve itself extracted from bathymetric data. In contrast to Harding et al. [2016], which use the DC data to perform FWI, we only use first arrival travel times to perform TTT, so surface-related effects do not affect here the inversion results."

- *Referee*: Question: You mention that you use variable water velocity. Have you quantified the effect of using constant vs. variable water velocity in downward continuation?

- *Answer*: The effect of a variable water velocity in downward continuation in our study area is minor for travel times. Velocity profiles obtained from oceanographic measurements ranges from 1.507 to 1.526 km/s giving: (1) for a 1 km of water depth, travel times from 0.664 s to 0.655 s, or (2) for a 2 km of water depth, from 1.327 s to 1.311 s. These are smaller than the travel time picking uncertainty (±0.035 s). It would probably be good to consider the effects more carefully if the objective is performing waveform inversion of the DC results. A specific comment about this topic is now included in Page 5 Line 20: "We have also tested the effect of the water velocity model used to DC the shotgathers versus the travel time picks. In our area of study the velocity changes in the water layer are small (<20 km/s) so considering either a realistic or an homogeneous model have a minor effect in the first arrival travel time picks, smaller than the travel time picking uncertainty (±0.035 s)."

- *Referee*: Lines 28-32: This is common to all DC techniques, it may not be really necessary to mention it after you provided detailed explanation of the technique.

- *Answer*: In this case, we believe it is a good idea to briefly remind the main steps of the DC technique, together with exposing the different domains in which each back-propagation is performed.

Page 6:

- *Referee*: Lines 5-13: I agree that the amplitudes can be affected by DC and that one has to very careful what part of the data is suitable for FWI. Thus I think it would be good that you mention this more explicitly. This also support your decision to do FWI on non-DC data.

- *Answer*: The rationale for this decision is now explained in Line 12: "As amplitudes are affected by the re-datuming process, it is not straightforward performing FWI using DC data. However, arrival times of the different phases are correct (which is our main objective) if the Vp model of the water layer is accurate enough. A better approximation of the true wavefield, and thus a better result, would be achieved using denser and larger arrays. For sparse data sets the redatumed arrivals would be less focused due to a larger amplitude loss and noise effects, nevertheless correctly located."

Page 7:

- *Referee*: (a) While you provide detailed description of re-datuming, the description on tomography is limited. (b) It is not clear how you combine refraction TTT on DC data and reflection TTT on non-DC data. Please, provide more information. (c) It would be good to indicate the events reflection and refraction by introducing an additional figure or insert them somehow in the figures 4 and 6.

- *Answer*: Thanks for the comments.

(a) Concerning the detailed description of re-datumming, we did so due to the lack of references explaining the details and steps of this technique. In contrast, the details of TOMO2D and FWI codes are already well-documented in the references provided. This is why we only summarize the description of tomographic techniques.

(b) The joint refraction and reflection TTT is performed using travel times of both refractions and reflections as input. In the data input file of TOMO2D one must introduce for each source its 2D position and the number of receivers associated to that source. After that it must contain for each associated receiver, the 2D positions, if the travel time corresponds to a refraction or reflection to trace the ray properly, the travel time pick and its error. In the case of DC data, the 2D coordinates will be the ones on the virtual surface, but for the reflections will be the location of the air gun and streamer. To clarify how refraction and reflection are combined, a sentence is now added in Page 6 Line 31: "Therefore, the data input file contains both type of travel times picks, each with their corresponding source/receiver positions. In the case of DC arrivals, the 2D coordinates will be the ones on the virtual surface, but for the reflections they will be at the actual location of the air gun source and of the individual streamer channels. Further details for our particular study area and experiment are explained in Section 4.2." Moreover, an additional explanation of how refractions and reflections are treated in the TTT is also added in Page 7 Line 12: "A 1-D floating reflector is defined along the profile. This means that it is parametrized as a node array that is independent from the velocity mesh, and both reflector array and velocity mesh can be spatially variable." and Line 14:"The partial derivatives of travel time residuals with respect to velocity (for first arrivals and reflections) and reflector depth (only for reflections) are introduced in the Fréchet matrix with additional regularization constraints (smoothing, damping). Thus, this interface depth model is inverted simultaneously with the Vp model. The inverse problem to iteratively update the Vp and depth models is solved using the sparse matrix solver of the LSQR algorithm of Paige and Saunders (1982). The difference between observed and synthetic travel times is iteratively minimized as a least-squares problem until a stopping criteria is fulfilled. Regularization constraints, smoothing and damping, are applied to stabilize the minimization process avoiding singularity of the matrix. For further details, see Korenaga et al. (2000) and Meléndez et al. (2015)".

(c) The DC seismic phases used are now included in Figure R1 (d)-e)-f)) (Figure 6 in the manuscript).We picked the TOB on common mid-point (CMP) gathers where it is a bright continuous reflection, that is also interpreted at the same time in a fully processed stack image, where the TOB is a clear event. Using the CMP gathers to pick the TOB removes the uncertainty of the reflector location in the shot gather domain. Therefore, it is not recommended to insert them in Figure 4 because the TOB displays the real laterally complex geometry. The CMP picks are then sorted back to Shot gather geometry and used for tomography. An image showing the reflector picking procedure (Fig. R2) is also added as additional material (Fig. A1). A specific comment on this topic is now included in Page 10

Line 28: "The selected reflection corresponds to the top of the basement (TOB), which is a clear event (e.g. see Figure 9). Reflection travel times are picked from MCS common mid-point gathers where it is a bright continuous reflection, which is also interpreted at the same time in a fully processed stacked image (Fig. A1)."

Page 9:

- *Referee*: Question: You mention that the number of iterations per frequency band is 10 and then you say that the stopping criteria is governed by the Arminjo rule. Does this mean that in all cases this criteria was reached within 10 iterations?

- *Answer*: We fix the maximum number of iterations to be 10. As it is shown in Figure 12, the inversion converges typically after 6 iterations. In this case, the Arminjo rule is applied as an extra stopping criteria to make the algorithm more efficient: if the Arminjo rule is not fulfilled within 10 iterations, then the inversion stops. Thus, we modified the sentence in the manuscript to be more comprehensible, using the words "An additional stopping criteria" rather than "Finally, the stopping criteria" in Line 15.

- *Referee*: **Section 4.2:** This section is a mix of methodology and results and should be rewritten. The part including lines 20-30 belongs clearly to methodology section and can be used to address the comment I raised for the content presented on Page 7.

- *Answer*: Thanks for the comment. In this paper, the methodology shows the characteristics of the codes used throughout the whole strategy. The part including lines 20-30 describes the joint refraction and reflection TTT in this particular real case scenario. All that information depends on the study area and field data of the experiment explained in Section 3, not the tomographic method itself. Therefore, we have included these parameters in Table R1. Instead of lines 20-30, now the section starts in Page 10 Line 27 and several sentences are added about specific aspects on the data set used: "The inversion parameters used in the TTT are shown in Table 1. We do not use the entire data set to reduce computational burden but all receivers are used to ensure data redundancy. The selected reflection corresponds to the top of the basement (TOB), which is a clear event (e.g. see Figure 9). Reflection travel times are picked from MCS common mid-point gathers where it is a bright continuous reflection, which is also interpreted at the same time in a fully processed stacked image (Fig. A1). The source and receiver positions were projected to a straight line defined between first and last shots, preserving their offset distance."

- *Referee*: Question: In lines 23-24 you mention that you pick refraction arrivals for all of the offsets (up to 6 km). My experience with DC and your example shot gathers in Figure 6 show that your maximum offset for picking should be <4.5 km. This is a known problem in DC. (a) How do you deal with this problem? (b) What is the effect on TTT velocity, if any?

- *Answer*: (a) We have changed Fig. 6 in the manuscript to show and compare manual picks at three DC shotgathers with those obtained using the final TTT Vp model (d)-e)-f) in Fig. R1). To prove the validity of our DC picks we have performed an additional test that consists on reproducing the DC processing but using as input synthetic streamer shotgathers that have been simulated with the final FWI Vp model (Fig. R3). First arrivals identified in the real DC shotgathers (blue dots) remarkably coincide for all offsets with the first arrival travel-times of the final TTT Vp model (red squares) and

also the ones imaged in the DC wavefields that are obtained using the synthetic data. This fact justifies the inclusion of the DC travel times even for large offsets. The new image is included as Fig. 16 together with the following explanation in Page 15 Line 20: "To check the first arrival travel-time errors introduced by the DC procedure we have reproduced the redatuming processing but using synthetic shotgathers simulated with the streamer acquisition geometry and the final FWI Vp model (Fig. 10). As it is shown in Fig. 16 (a)-b)-c), the first arrivals identified in the real DC shotgathers (blue dots) coincide remarkably well for all offsets with the first arrival travel-times of the final TTT Vp model (red squares) and also with the ones in the DC wavefields that are obtained using the synthetic data. The similar results for the first arrivals justify the inclusion of the DC travel times even for large offsets."

(b) There is no critical effect on the TTT velocity model because as can be seen in Figure R3 the travel times of the final TTT Vp model (red squares) coincide for both first arrivals of the DC shotgathers (real and synthetic). Moreover, the validity of the input picks is also shown when the TTT Vp result is converted to two-way-time and superimposed to the time migrated image in Figure 9, because of the spatial coincidence of major velocity contrasts with reflectivity changes, together with the good match of the TOB geometry resolved from a flat boundary.

Page 11:

- *Referee*: Line 22: I would rather use "gently" that "softly' in this context.

- *Answer*: We changed it as suggested (Page 11 Line 24).

Page 15:

- *Referee*: Line 22: The elastic effect may not be significant in your case, as you have sediments on top of igneous basement. You may cite Warner et al., (EAGE abstract, I believe it is 2012).

- *Answer*: We now cite Warner et al. (2013) as suggested (Page 15 Line 34).

Figure 3:

- *Referee*: The annotations are too small, as well as numbers indicating lat and long . Please, increase the font.

- *Answer*: We have modified the figure for a better visualization. Also, new references are added to provide data sources of the map and additional information on the characterization of the study area (Fig. R4).

Figure 11:

- *Referee*: (a) Correct me if I am wrong, but it seems that you have completely reversed polarity in

Initial/Resultant model with respect to Real data. (b) It would be good to show residuals or at least provide a wiggle plot with observed and synthetically calculated data superimposed to understand if the problem is due to plotting or there is really an issue with data polarity.

- *Answer*: Thanks for noticing that the figure was not the proper one. The Initial/Resultant data plotted were not filtered at the same frequency as the Real data, the figure is now corrected (Fig. R5).

(a) In several places there might be a small time shift between the observed and synthetic data due to the difficulty of simulating the exact time when the shot is produced and the exact relative distances between source/receiver positions. Moreover, the data preconditioning of the FWI selects a time window to fit the signal after the seafloor reflection arrival and its multiple. This data-driven preconditioning is applied because the seafloor reflection is one example of a discontinuity that is hard to fit by the acoustic formulation of the wave propagation.

(b) Initial (d) and final (e) residuals and wiggle plots (f) with observed and synthetically calculated data superimposed are added for a clearer visualization of the result in Fig. R5. Therefore, the text is now modified in Page 13 Line 18: "To show the model improvement in the data domain, in Fig. 11 we compare a recorded shotgather (Fig. 11c) and a synthetic one generated with the FD solver [Dagnino et al., 2014] using the initial TTT (Fig. 11a) and final FWI (Fig. 11b) velocity models. In contrast with the synthetic data generated with the TTT velocity model (Fig. 11a), the shotgather generated with the FWI velocity model (Fig. 11b) presents some near-vertical reflections. Thus, the seismogram simulated with the final model shows a larger number of seismic events compared to the initial one, which only recovered the first arrival phases and the TOB reflection from the TTT. The data-driven preconditioning of the FWI strategy targeted the energy corresponding to the near-vertical reflections. Therefore, in that region final residuals (Fig. 11e) are smaller than initial ones (Fig. 11d), but are larger around the seafloor reflection. Again, in Fig. 11f aside from wave amplitudes, the main difference between the data generated with the TTT model (orange line) and target (black line) seismogram is the presence of reflected waves. We observe a better fit between the final (blue line) and real data set (black line), except for the effects that are not modelled, such as 3D diffractions, or the arrivals not included in the data-based pre-conditioning."

Figure 16:

- *Referee*: Is there a particular reason why you chose to do your interpretation on time seismic section?

- *Answer*: We have included Figure 16 to Figure 15 and modified Figure 15 adding the interpretation on it (Fig. R6).

Bibliography:

Arnulf, A.F., Singh, S.C., Harding, A.J., Kent, G.M., and Crawford, W.C.: Strong seismic heterogeneity in layer 2A near hydrothermal vents at the Mid-Atlantic Ridge, Geophys. Res. Lett., 38, L13320, http://doi.org/ 10.1029/2011GL047753, 2011.

Arnulf, A.F., Harding, A.J., Kent, G.M., Singh, S.C., and Crawford, W.C.: Constraints on the shallow

velocity structure of the Lucky Strike Volcano, Mid-Atlantic Ridge, from downward continued multichannel streamer data, J. Geophys. Res.-Sol. EA, 119, 1119-1144, http://doi.org/10.1002/2013JB010500, 2014.

Ballesteros, M., Rivera, J., Muñoz, A., Muñoz-martín, A., Acosta, J., Carbó, A., and Uchupi, E.: Alboran Basin, southern Spain-Part II: Neogene tectonic implications for the orogenic float model, Mar. Petrol Geol., 25, 1, 75-101, http://doi.org/10.1016/j.marpetgeo.2007.05.004, 2008.

Berryhill, J.R.: Wave-equation datuming, Geophysics, 44, 1329-1344, https://doi.org/10.1190/1.1441010, 1979.

Berryhill, J.R.: Wave-equation datuming before stack, Geophysics, 49, 11, 2064-2066, https://doi.org/10.1190/1.1441620, 1984.

Cho, Y., Ha, W., Kim, Y., Shin, S., and Park, E.: Laplace-Fourier-Domain Full Waveform Inversion of Deep-Sea Seismic Data Acquired with Limited Offsets, Pure Applied Geophysics, 173, 749-773, http://doi.org/10.1007/s00024-015-1125-7, 2016.

Gómez de la Peña, L., Gràcia, E., Muñoz, A., Acosta, J., Gómez-Ballesteros, M., Ranero, C.R., and Uchupi, E.: Geomorphology and Neogene tectonic evolution of the Palomares continental margin (western Mediterranean), Tectonophysics, 689, 25-39, http://doi.org/10.1016/j.tecto.2016.03.009, 2016.

Gràcia, E., Bartolome, R., Lo Iacono, C., Moreno, X., Stich, D., Martínez-Díaz, J. J., et al.: Acoustic and seismic imaging of the Adra Fault (NE Alboran Sea): in search of the source of the 1910 Adra earthquake, Nat. Hazard Earth Sys., 12, 11, 3255-3267, http://doi.org/10.5194/nhess-12-3255-2012, 2012.

Harding, A. J., Arnulf, A. F., and Blackman, D. K.: Velocity structure near IODP Hole U1309D, Atlantis Massif, from waveform inversion of streamer data and borehole measurements. Geochem. Geophy. Geosy., 17, 6, 1990-2014, https://doi.org/10.1002/2016GC006312, 2016.

Henig, A. S., Blackman, D. K., Harding, A. J., Canales, J. P., and Kent, G. M.: Downward continued multichannel seismic refraction analysis of Atlantis massif oceanic core complex, 30N, mid-Atlantic ridge. Geochem. Geophy. Geosy., 13, Q0AG07, https://doi.org/10.1029/2012GC004059, 2012.

Korenaga, J., Holbrook, W.S., Kent, G.M., Kelemen, P.B., Detrick, R.S., Larsen, H.-C., Hopper, J.R., and Dahl-Jensen, T.: Crustal structure of the southeast Greenland margin from joint refraction and reflection seismic tomography, J. Geophys. Res., 105, B9, 21591-21614, 2000.

Meléndez, A., Korenaga, J., Sallarès, V., Miniussi, A., and Ranero, C.R.: TOMO3D: 3-D joint refraction and reflection traveltime tomography parallel code for active-source seismic data-synthetic test, Geophys. J. Int., 203, 158-174, http://doi.org/10.1093/gji/ggv292, 2015.

Paige, C.C., and Saunders, M.A.: LSQR: An algorithm for sparse linear equations and sparse least squares, ACM T. Math. Software, 8, 1, 43-71, http://doi.org/10.1145/355984.355989, 1982.

Qin, Y., and Singh, S.C.: Detailed seismic velocity of the incoming subducting sediments in the 2004

great Sumatra earthquake rupture zone from full waveform inversion of long offset seismic data, American Geophysical Union, http://doi.org/10.1002/2016GL072175, 2017.

Shtivelman, V., and Caning, A.: Datum correction by wave-equation extrapolation, Geophysics, 53, 10, 1311–1322, https://doi.org/10.1190/1.1442409, 1988.

Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Stekl, I., Guasch, Ll., Win, C., Conroy, G., and Bertrand, A.: Anisotropic 3D full-waveform inversion, Geophysics, 78, 2, R59–R80, https://doi.org/10.1190/geo2012-0338.1, 2013.

Table:

Inversion parameters	Values
Grid dimension (long x deep)	8.25 x 5 km
Range of water layer depth	1.1-1.9 km
Node spacing (dx = dz)	25 m
Number of DC shotgathers picked	121
Distance between DC shots picked	~500 m
Maximum experimental offset	~6 km
Total of first arrival travel times	58,080
Total of reflected travel times	56,920
Forward star order (x,z) (node connexions)	(7,7)
Eliminate data outliers with chi values	> 15
Tolerance level to terminate inversion	0.001
Number of iterations	10
Velocity smoothing parameter (λ_v)	75
Depth smoothing parameter (λ_z)	10
Average velocity perturbation limit (%)	15
Average depth perturbation limit (%)	15
Top velocity smoothing correlation lengths (x,z) (L_{Htop} , L_{Vtop})	(0.6,0.05) (km)
Bottom velocity smoothing correlation lengths (x,z) (L_{Hbot} , L_{Vbot})	(3,0.5) (km)
Horizontal smoothing correlation length for the reflector (x) (L_z)	1.5 km

Table R1 (Table 1 in the manuscript). Relevant inversion parameters used in the travel-time tomographic inversion.
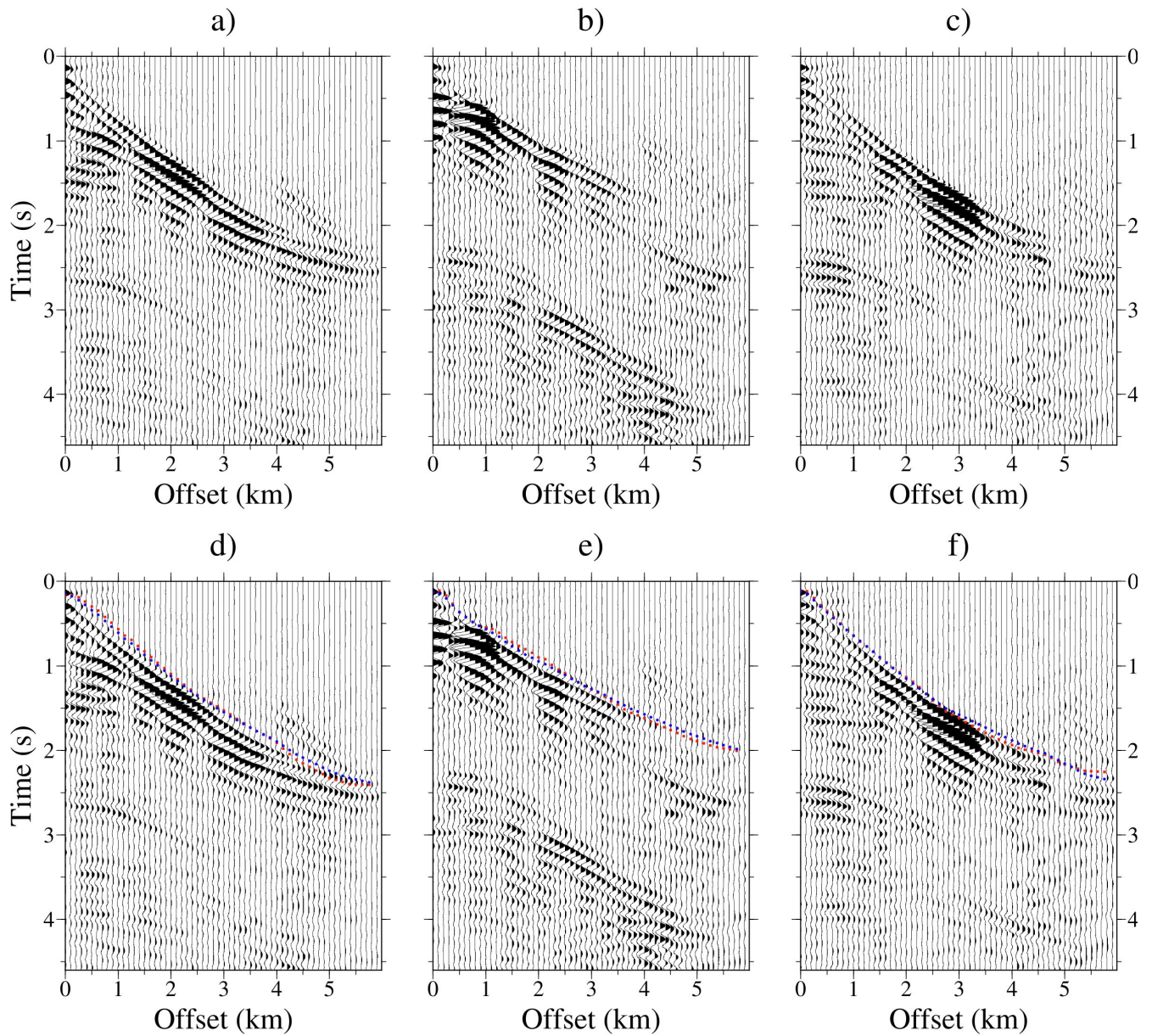
Figures:

Figure R1 (Figure 6 in the manuscript): Seismic data obtained from the DC of the streamer shots recorded in the TOPOMED cruise. Only 10 traces each km are plotted (so 1/5 of the total) for clarity. From left to right, shot locations are at 171.1, 138.7 and 112.5 km along the profile. Lower panels (d)-e)-f)) show shots (a)-b)-c)) together with first arrivals used as input for the TTT (blue dots) and the first arrivals simulated with the TTT model (red squares).
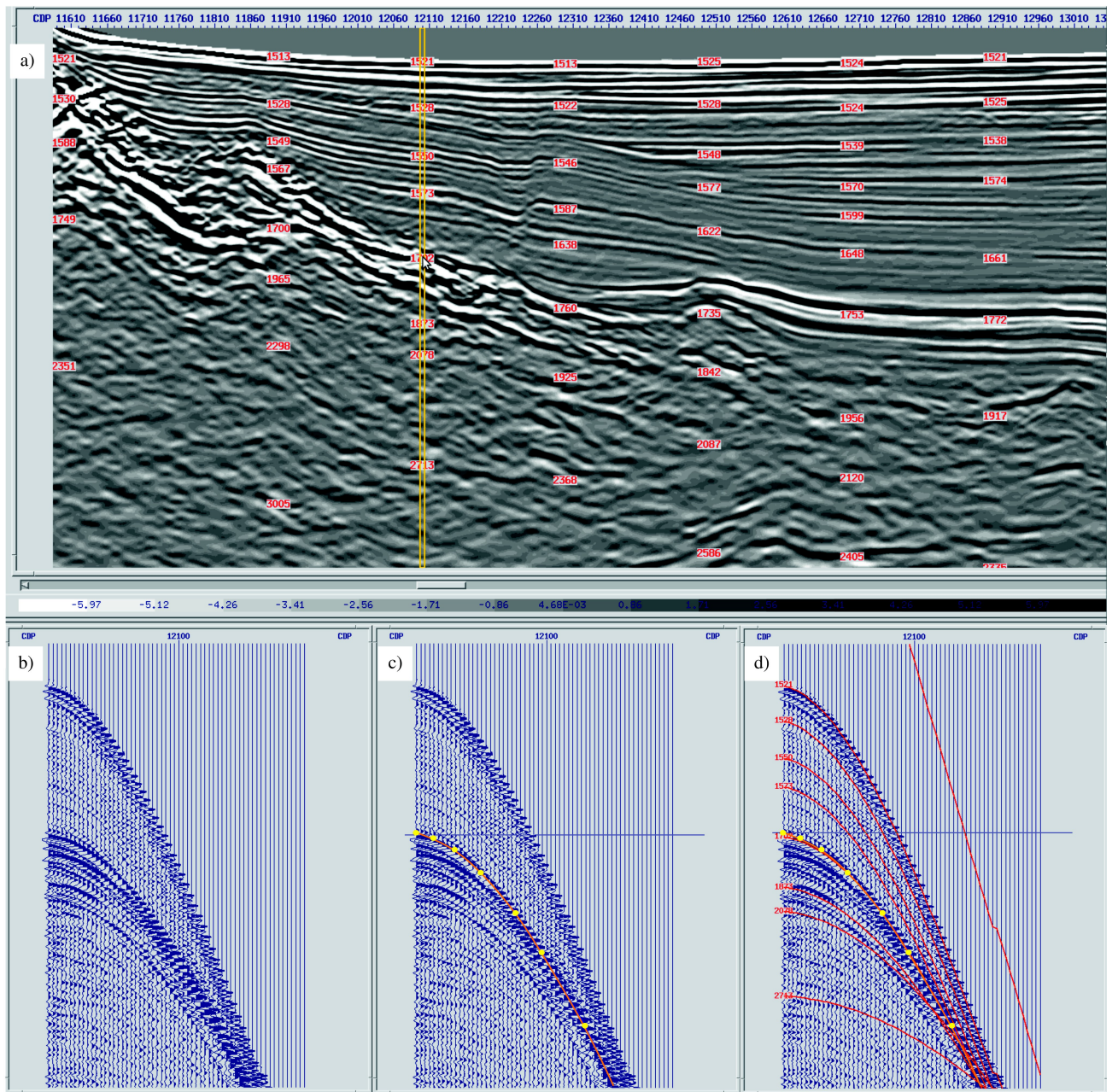
Figure R2 (Figure A1 in the manuscript): Reflector picking procedure. Processed stack image of a section of the seismic profile (a) showing the location of the CMP gather that is plotted with (c), (d) and without (b) the TOB reflection travel times (yellow circles) used as input for the TTT. The TOB is a bright continuous reflection on CMP gathers. The arrow in (a) and horizontal line in (c), (d) indicate the position of the TOB reflection in the stack image and CMP gathers. Red lines in (d) correspond to velocity parabolas that are used as a velocity control during the reflector picking.
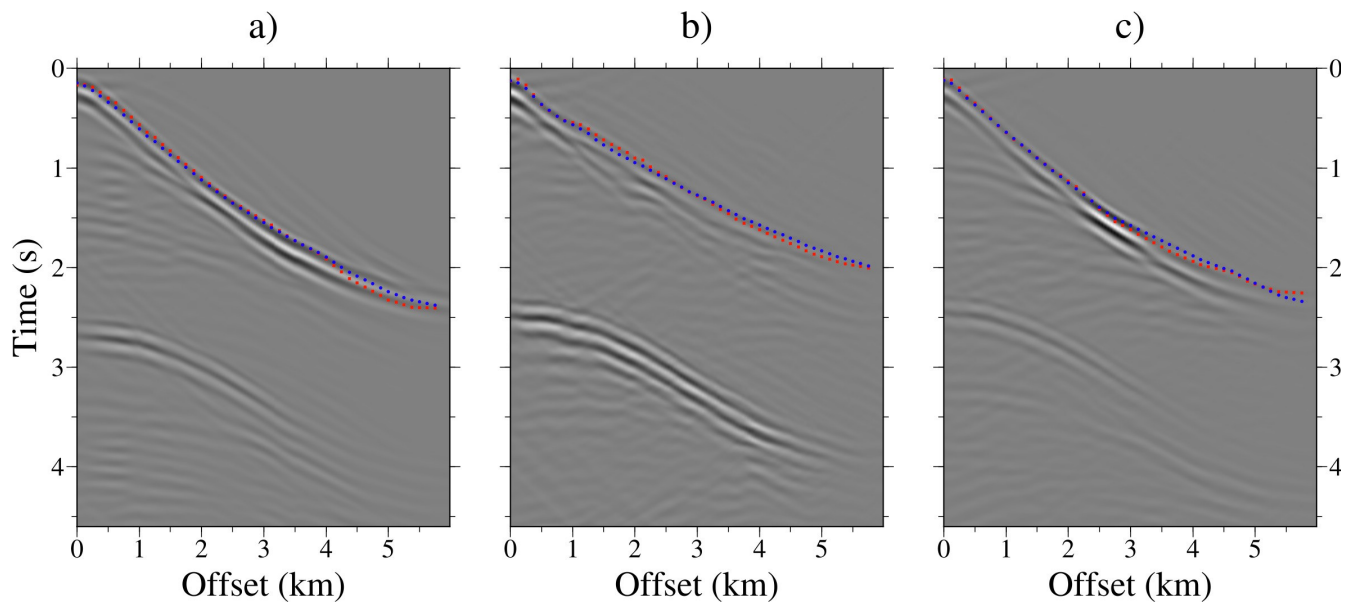
Figure R3: Panels a)-b)-c) show synthetic seismic data obtained from the DC of the streamer shots simulated using the final FWI Vp model. From left to right, shot locations are at 171.1, 138.7 and 112.5 km along the profile. First arrivals used as input for the TTT are plotted as blue dots and the ones simulated with the final TTT Vp model as red squares.
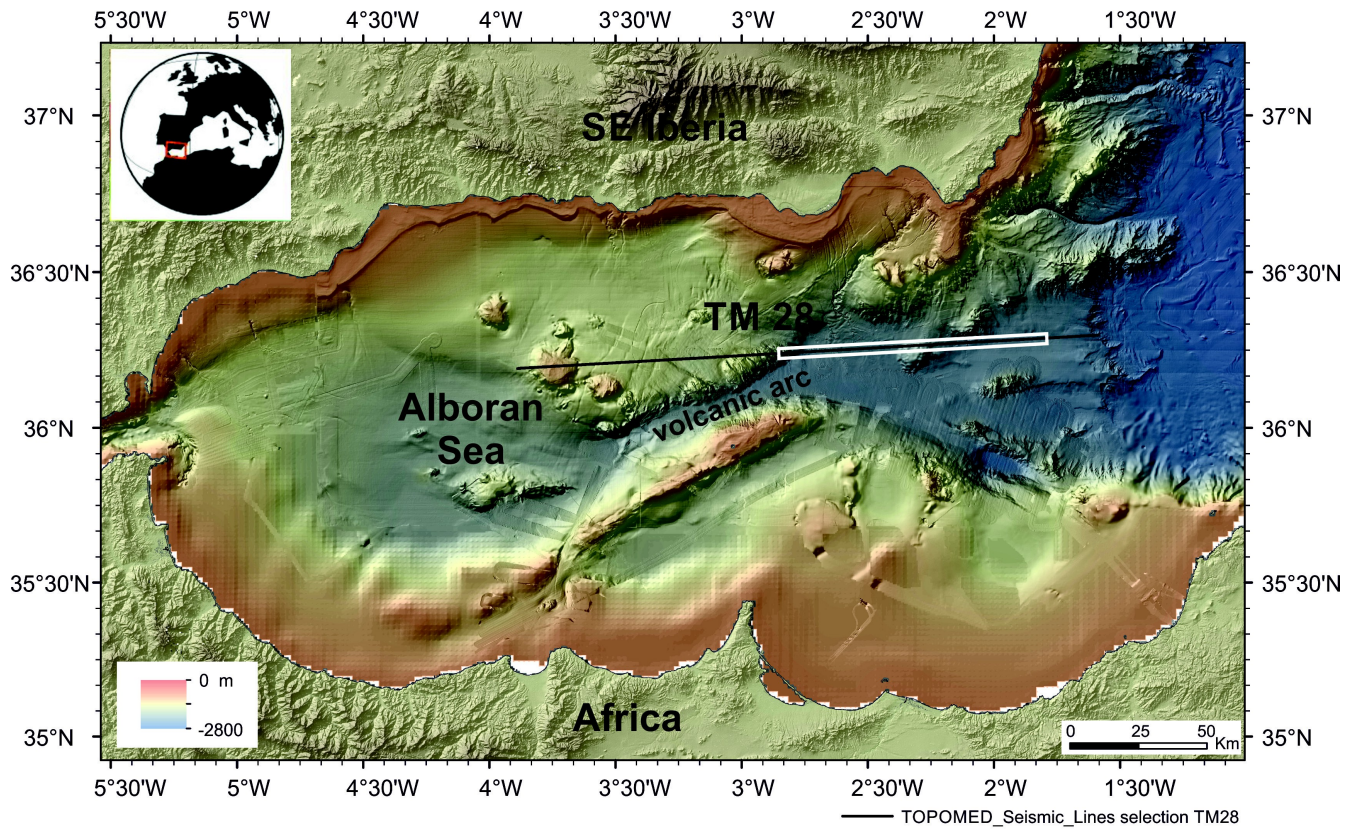
Figure R4 (Figure 3 in the manuscript). Relief map of the study area (Ballesteros et al., 2008; Gràcia et al., 2012, Gómez de la Peña et al., 2016). The field data used in the test corresponds to the white box along TM28 profile.
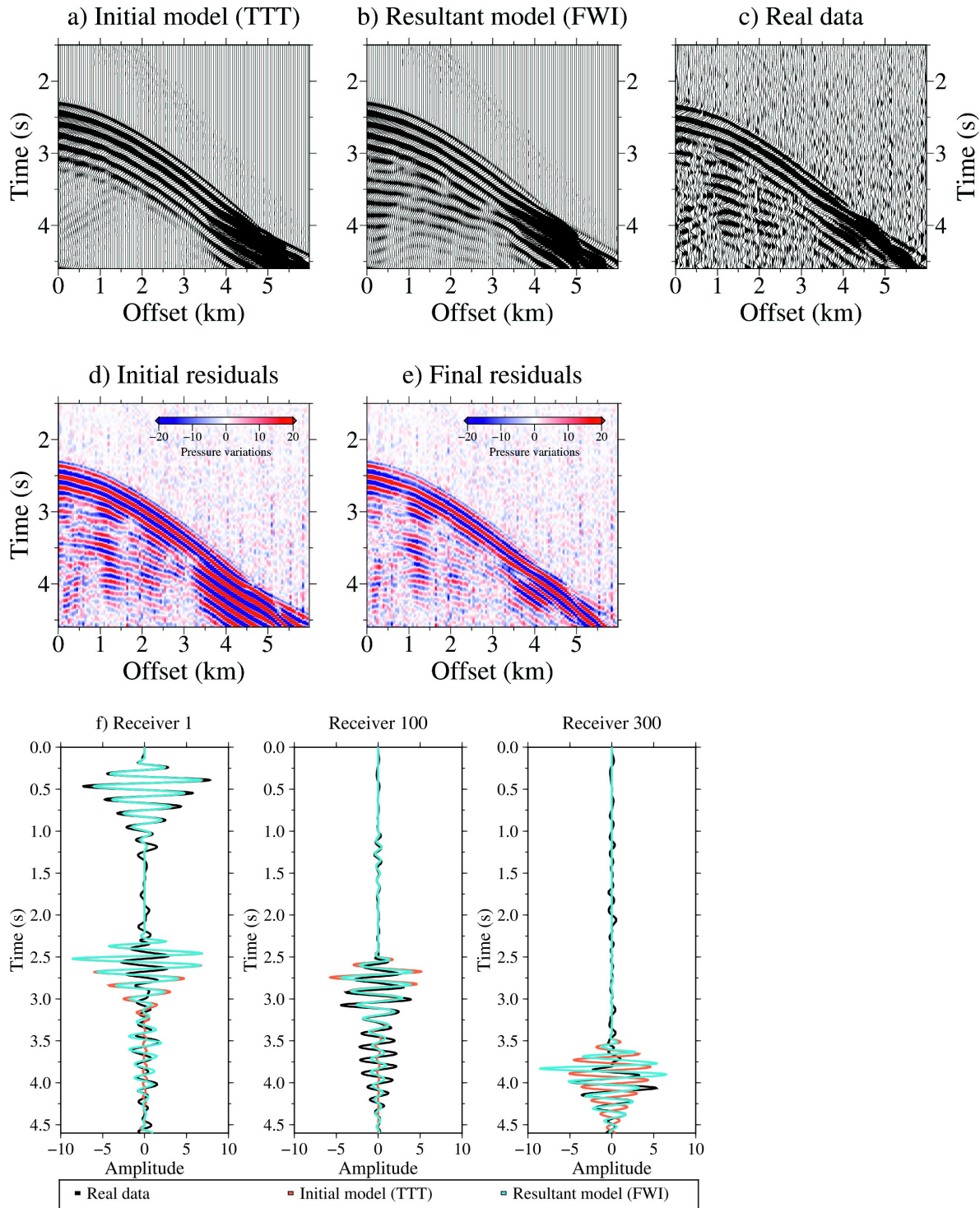
Figure R5 (Figure 11 in the manuscript): Streamer synthetic seismograms, which are computed using the initial (a) and resultant (b) velocity model presented in Fig. 7c and Fig. 10 and the FD solver of the FWI algorithm, together with the real data (c). Only every forth trace at 50 m is plotted. The location of the streamer is at a distance of ~ 101 to 107 km along profile. Initial (d) and final (e) residuals are displayed for a better comparison of the results, as well as (f) initial (orange, a)), resultant (blue, b)) and real (black, c)) seismic traces for receivers 1, 100 and 300.
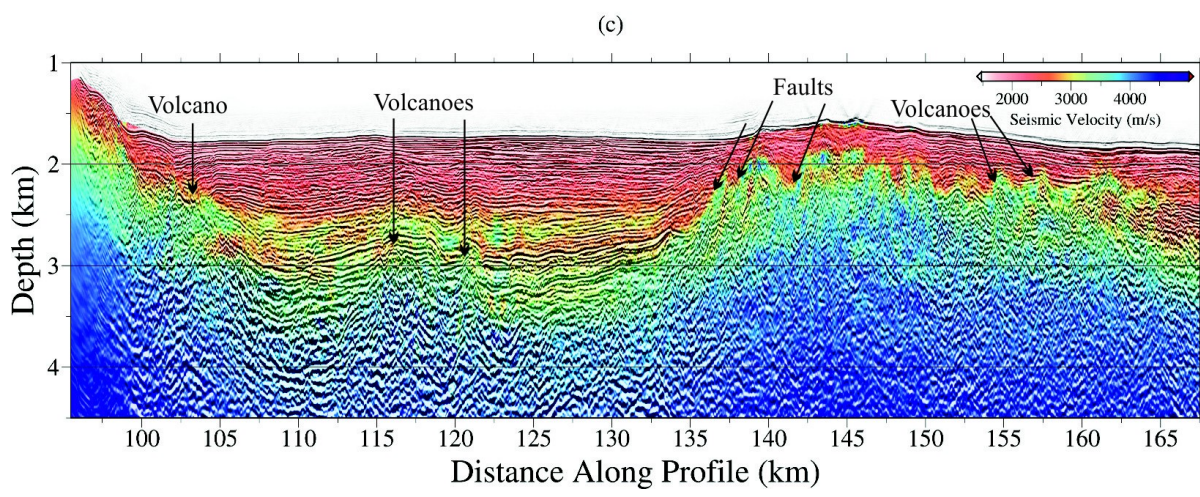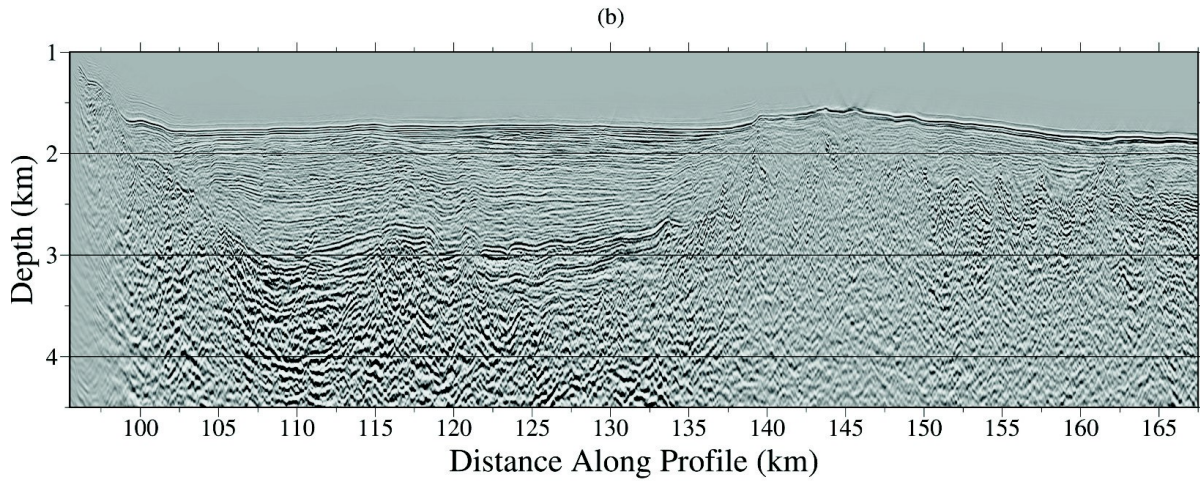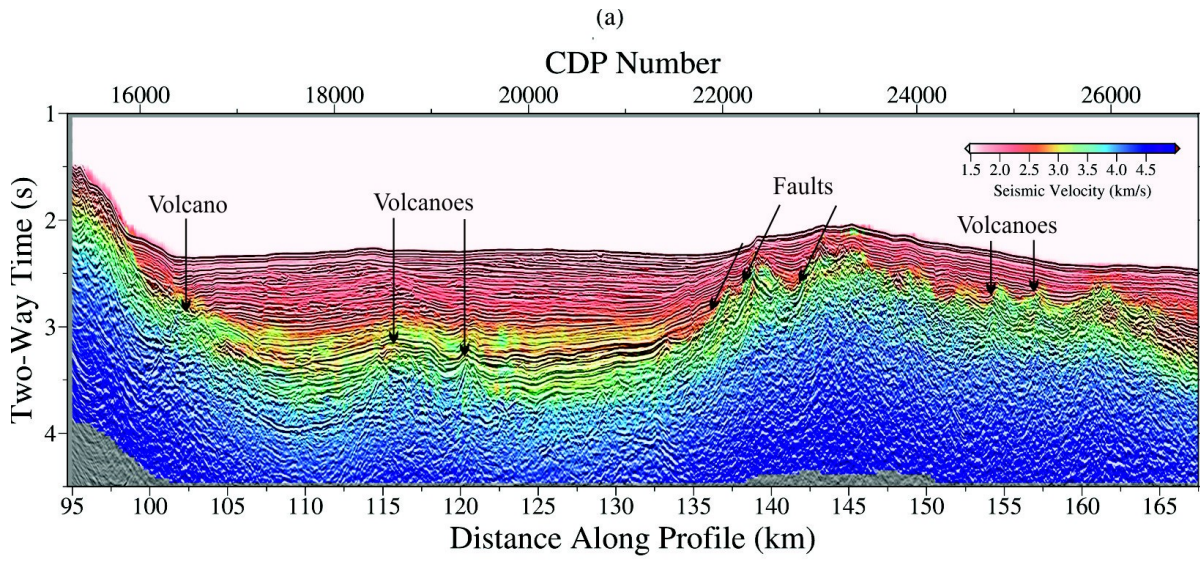
Figure R6 (Figure 16-15 in the manuscript): (a) Geological interpretation superimposed to Fig. 13. (b)

Kirchhoff depth migration result based on the velocity model shown below (c) obtained after the modeling sequence proposed on this paper (joint DC refraction and streamer reflection TTT + FWI). Geological interpretation is also superimposed to (c).

Interactive comment on "Full waveform inversion of short-offset, band-limited seismic data in the Alboran basin (SE Iberia)" by C. Gras et al.

We thank the referee for the review, comments and corrections. We must clarify that to estimate the initial velocity model for the FWI we do not perform just three iterations of a joint refraction and reflection TTT, but a two-step layer-stripping strategy, with ten iterations at each inversion step. We expect it is now clearer in the text. The rest of referee comments, bibliography and figures are addressed below (in green).

Specific Comments:

- *Referee*: One aspect in this paper is the DC of MCS used as a starting model for the FWI. The results look very good for the real data but difficult to quantify. (a) I would expect to see a synthetic shot-gather with the DC result. (b) The Fig. 2b and 2d looks as if both are syn. modeled which is ok for a general scheme picture. (c) A 2-step DC of fig.11a or 11b would be for this paper more appropriate especially for the discussion of how much offset of the diving wave show useful information.

- *Answer*: We have implemented the referee's suggestions as follows: (a) We now show in Fig. R1 and R2 the travel time picks of the DC shotgather used as input for the TTT inversion (blue dots), together with the picks obtained from the final TTT model (red squares) for three different shotgathers. Figure R2 (a)-b)-c)) compares the travel times for both input and final TTT picks with the results of applying DC to the streamer shots simulated using the final FWI Vp model. The travel-time differences between the picks made at the DC and those predicted by the TTT model are of 37 ms on average (Fig. 8a blue line in the manuscript). (b) The Figure R3 b and d (Fig 2b and d in the manuscript) are cartoons of synthetic generated shotgathers. (c) The obtained result after each step of the DC for three streamer shotgathers (Fig. R4, Fig. 4 in the manuscript), are shown in Fig. R5 and Fig. R1 (Figs. 5 and 6 in the manuscript). Figure R6 (Fig. 11a and b in the manuscript) show low-pass filtered shotgathers (< 8 Hz) generated with the real streamer geometry using the initial and final FWI model, respectively.

Individual comments/corrections:

Page 2:

- *Referee*: Line 19... near offset reflected alone.

- *Answer*: We changed this this by a new sentence in Line 18: "Due to the inherent velocity-depth trade-off and to the possible errors in the identification of the reflector boundaries, we cannot expect a precise velocity model by travel time inversion using only one or few reflectors with no lateral or vertical continuity".

- *Referee*: Lines 15-19: This is not fully correct especially with a streamer length of 6km at this target depth. Iterative prestack depth migration with the combination of reflection tomography velocity updating and the local reflector dip is a robust method if reflection continuity exists laterally as well as vertically. Please reformulate Line 15-19 with the special situation of only one TOB basement reflection event.

- *Answer*: Thanks for the clarification. A comment making clear this issue has now been added (see answer to previous question).

- *Referee*: Line 25: I think it should be: have no low energy content below 4-5 Hz.

- *Answer*: The sentence in the text has been re-written to clarify that and now Line 23 is: "The most commonly used sources in MCS systems are airguns, which are typically not capable of generating energy below 4-5 Hz, so in those cases the signal to noise ratio is very low."

- *Referee*: Line 29: Correction: . . .Singh, 2017).

- *Answer*: We removed the extra bracket.

- *Referee*: Line 33: Correction: Shah et al., 2012.

- *Answer*: We corrected the year (Line 35).

Page 6:

- *Referee*: Lines 18-19... Another effect could happen by large shot point distances. Here the receiver gathers may be spatially aliased.

- *Answer*: We agree that the quality of the DC result is highly influenced by the distances between shots and also receivers, so it gets worse and can be spatially aliased for sparse data sets. The sparse data set/aliasing issue is now specifically mentioned in Page 4 Line 24: "The proper source spacing and optimal recording time step are the ones that avoid aliasing issues, grid dispersion and reduce the effects caused by the discrete approximation of the wavefield."

Page 7:

- *Referee*: Line 20: Question: By introducing a free surface with a source / streamer depth of 10m in general a source or receiver ghost will be modelled. Did this not happen because of the limit of 20 Hz and the corresponding grid spacing?

- *Answer*: The source and receivers ghosts are modeled by the forward solver. Ghosts are reflections of the wave at the sea surface, and thus they are considered in the wave equation and solved in the propagation. The peak of the source frequency is at 20 Hz, but this is not a limit.

- *Referee*: Line 26: Question: Please explain T/6 and 6T and how do you remove it?

- *Answer*: In our data set, there is no useful signal above T/6 and below 6T, where T is the period of the highest inverted frequency. These cutoff values are determined by checking the source signature. Different sources as streamer stretching can introduce noise at these frequencies. To avoid introducing this noise in the data, source signal is zeroed outside these limits.

Page 10 :

- *Referee*: Line 23: almost 6 km? I would choose something like 3.5-4.5 km. Here a synthetic model would help (See Specific Comments above)

- *Answer*: As explained above, we have changed Fig. 6 in the manuscript to show and compare first

arrival picks at three DC shotgathers with those modeled using the final TTT Vp model (d)-e)-f) in Fig. R1).

We have also performed an additional test that consists on reproducing the DC processing but using as input synthetic streamer shotgathers that have been simulated with the final FWI Vp model (Fig. R2). The new image is included as Fig. 16 together with the following explanation in Page 15 Line 20 : "To check the first arrival travel-time errors introduced by the DC procedure we have reproduced the redatuming processing but using synthetic shotgathers simulated with the streamer acquisition geometry and the final FWI Vp model (Fig. 10). As it is shown in Fig. 16 (a)-b)-c), the first arrivals identified in the real DC shotgathers (blue dots) coincide remarkably well for all offsets with the first arrival travel-times of the final TTT Vp model (red squares) and also with the ones in the DC wavefields that are obtained using the synthetic data. The similar results for the first arrivals justify the inclusion of the DC travel times even for large offsets."

- *Referee*: Lines 25-26: Picking reflection travel times on unmigrated data especially at TOB were many diffraction originates seems dangerous. Additional a local reflector dip may influence the velocity. Was the TOB reflector premigrated before traveltime picking and dip corrected ?

- *Answer*: We picked the TOB on common mid-point (CMP) gathers where it is a bright continuous reflection, which is jointly interpreted in a fully processed stack image, where the TOB is a clear event. Using the CMP gathers to pick the TOB removes the uncertainty of the reflector location in shot gather images where the TOB displays the real laterally complex geometry. The CMP picks are then sorted back to the shot gather domain and used for tomography. An image showing the reflector picking procedure (Fig. R7) is now introduced in the additional material (Fig. A1). A specific comment on this topic is now included in Page 10 Line 28 : "The selected reflection corresponds to the top of the basement (TOB), which is a clear event (e.g. see Figure 9). Reflection travel times are picked from MCS common mid-point gathers where it is a bright continuous reflection, which is also interpreted at the same time in a fully processed stacked image (Fig. A1)."

- *Referee*: Lines 25-26: Please explain in more detail how the reflections were included in the TTT (floating reflector?) because it seems not to be a standard reflection tomography approach (see e.g.: Enhanced velocity estimation using gridded tomography in complex chalk, M. Sugrue, et al., Geophysical Prospecting, 2004, 52, 683–691.

- *Answer*: Yes, reflector depth parameters are introduced as a floating reflector. This means that it is parametrized as a node array that is independent from the velocity mesh, and both reflector array and velocity mesh can be spatially variable. Then the partial derivatives of travel time residuals with respect to velocity (for first arrivals and reflections) and reflector depth (only for reflections) are introduced in the Fréchet matrix with additional regularization constraints (smoothing, damping), and the minimization is made in a least square sense by applying the LSQR algorithm for sparse matrix inversion (Paige and Saunders, 1982). The inversion procedure is explained in detail in Korenaga et al. (2000) and Meléndez et al. (2015). We have added this clarification in the text (Page 7 Line 10).

Page 13:

- *Referee*: Lines 30-31: By Quality factor you mean the anelastic/intrinsic Attenuation Q. The inversion of Q removes time variant phase distortion (dispersion) generated by the earth's rock properties (amplitude and phase correction), it does not belong to the class of geometrical corrections. Value of 100 seems reasonable (100-200 for wet sand). Only amplitude or also phase correction for Q-Attenuation?

- *Answer*: The Quality factor corrects for phase effects and energy loss of the seismic signal caused by absorption and dispersion. The correction applied is based on Wang (2002). We balanced only amplitudes with a value of 100 found through analysis.

Page 15:

- *Referee*: Line 32: this high velocity layer within sediments can already be seen in the final TTT Fig. 7d. but not as strong and detailed as by the FWI.

- *Answer*: We have modified Line 32 because it is true that the TTT result shows a slightly increase of velocity in that part. The new sentence is in Page 16 Line 11: "In addition, the high-velocity layer within sediments on top of this structure is only clearly defined in the FWI Vp model (Fig. 10)".

Page 16:

- *Referee*: Line 25: Correction: there are not known salt deposits.

- *Answer*: We changed it as suggested (Page 17 Line 6).

Figure 2b and 2d:

- *Referee*: Correction: velocity labels difficult to read. Position the annotations at the end of the arrows.

- *Answer*: We changed the velocity labels and the position of the annotations as suggested (Fig. R3).

Figure 7:

- *Referee*: Correction: inversion step. (d)

- *Answer*: We changed it as suggested.

Figure 14a and 15a:

- *Referee*: the seismic sections could have less gain at least on a paper print.

- *Answer*: We changed it as suggested (Fig. R8 and R9).

Bibliography:

Korenaga, J., Holbrook, W.S., Kent, G.M., Kelemen, P.B., Detrick, R.S., Larsen, H.-C., Hopper, J.R., and Dahl-Jensen, T.: Crustal structure of the southeast Greenland margin from joint refraction and reflection seismic tomography, J. Geophys. Res., 105, B9, 21591-21614, 2000.

Meléndez, A., Korenaga, J., Sallarès, V., Miniussi, A., and Ranero, C.R.: TOMO3D: 3-D joint refraction and reflection traveltime tomography parallel code for active-source seismic data-synthetic test, Geophys. J. Int., 203, 158-174, http://doi.org/10.1093/gji/ggv292, 2015.

Paige, C.C., and Saunders, M.A.: LSQR: An algorithm for sparse linear equations and sparse least squares, ACM T. Math. Software, 8, 1, 43-71, http://doi.org/10.1145/355984.355989, 1982.

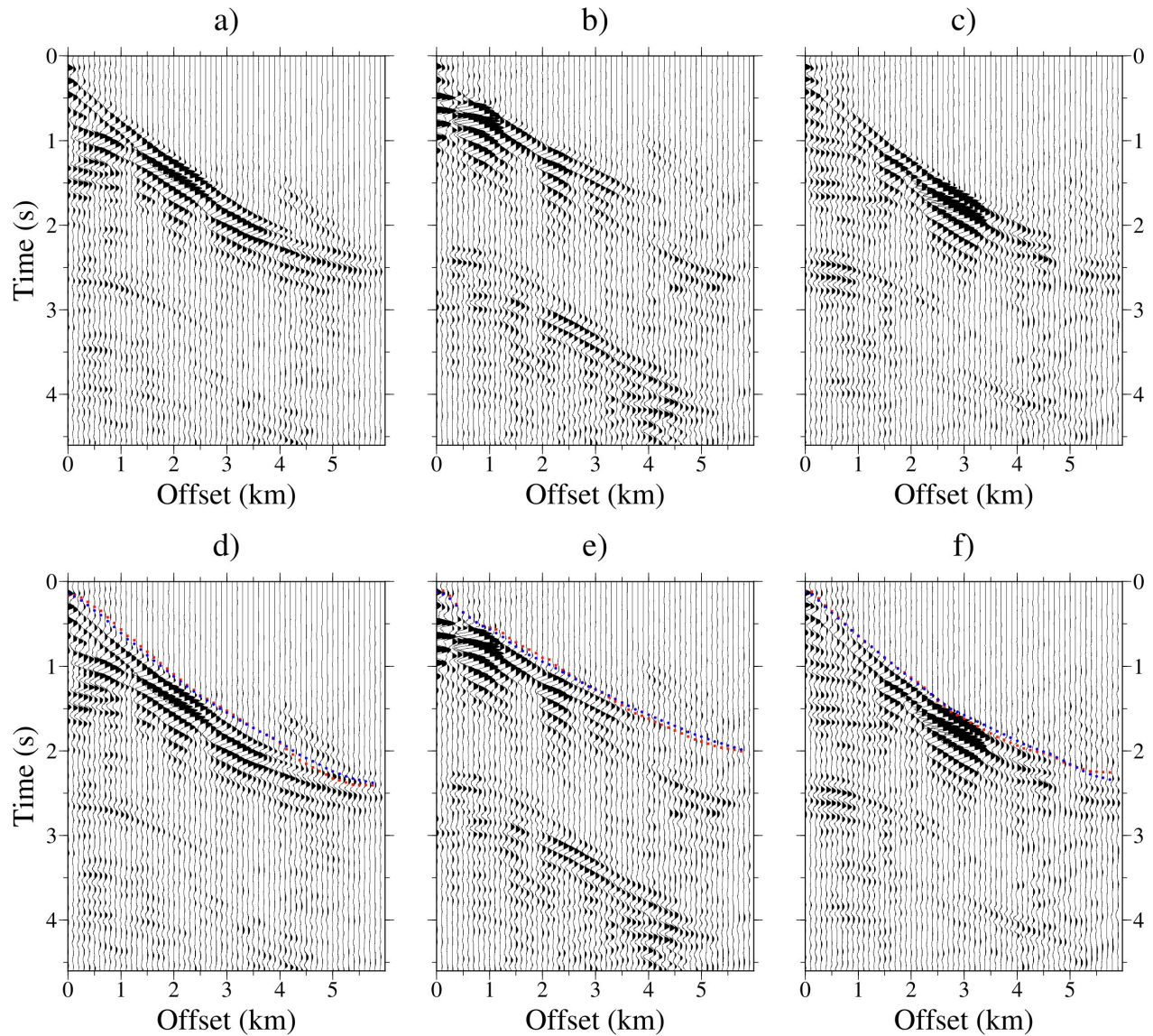Wang , Y.: A stable and efficient approach of inverse Q filtering, Geophysics, 67, 2, 657-663, https://doi.org/10.1190/1.1468627, 2002.

Figures:

Figure R1 (Figure 6 in the manuscript): Seismic data obtained from the DC of the streamer shots recorded in the TOPOMED cruise. Only 10 traces each km are plotted (so 1/5 of the total) for clarity. From left to right, shot locations are at 171.1, 138.7 and 112.5 km along the profile. Lower panels (d)-e)-f)) show shots (a)-b)-c)) together with first arrivals used as input for the TTT (blue dots) and the first arrivals simulated with the TTT model (red squares).
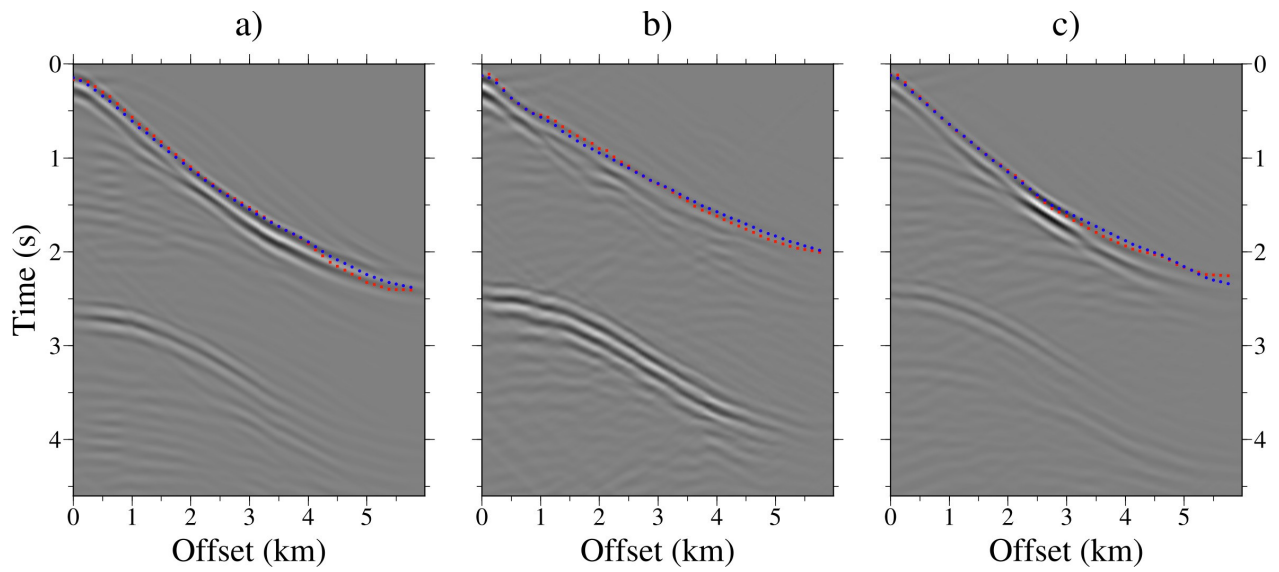
Figure R2: Panels a)-b)-c) show synthetic seismic data obtained from the DC of the streamer shots simulated using the final FWI Vp model. From left to right, shot locations are at 171.1, 138.7 and 112.5 km along the profile. First arrivals used as input for the TTT are plotted as blue dots and the ones simulated with the final TTT Vp model as red squares.
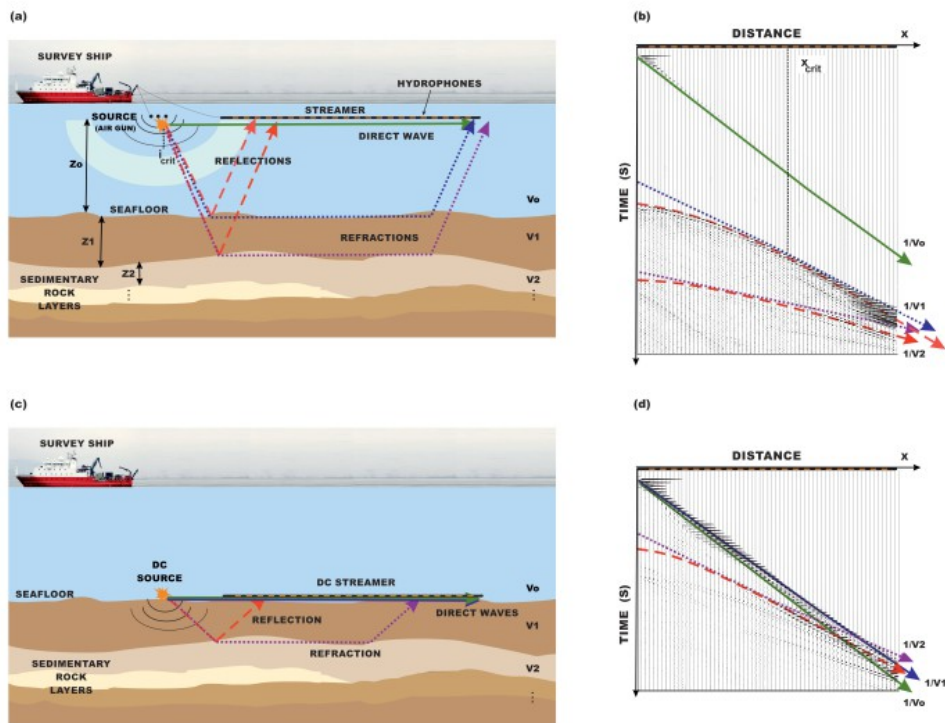

Figure R3 (Figure 2 in the manuscript): (a) Scheme of a streamer acquisition geometry of a marine MCS experiment showing the main element as well as the paths of the seismic phases that can be recorded in the streamer hydrophones. Where $z_j j = 0,1,\ldots$ defines the depth of each layer, $v_j j = 0,1,\ldots$ its Vp and $i_{crit_j} j = 0,1, \ldots$ its critical angle. Green line is the direct water wave, blue and purple dotted lines are refracted ray paths in different interfaces, and its corresponding reflected ray paths are the pink and red dashed lines respectively. (b) Example of a synthetic shotgather recorded

using a streamer-type geometry, as the one that is shown in (a). Each trace in the record section corresponds to a seismogram recorded at increasing source-receiver offset. Refractions are not visible, so they cannot be identified, and travel times cannot be used for velocity modeling. (c) Scheme of the virtual set up to be simulated after DC. In this case, both sources and receivers are located at the seafloor. Green line is the direct water wave, blue line is the head wave that travels through the seabed interface, dotted purple and red dashed lines are respectively the refracted and reflected ray paths of an interface of the subsurface. (d) Example of a synthetic shotgather recorded using the virtual geometry setup shown in (c). Refractions are visible as first arrivals from zero offset.
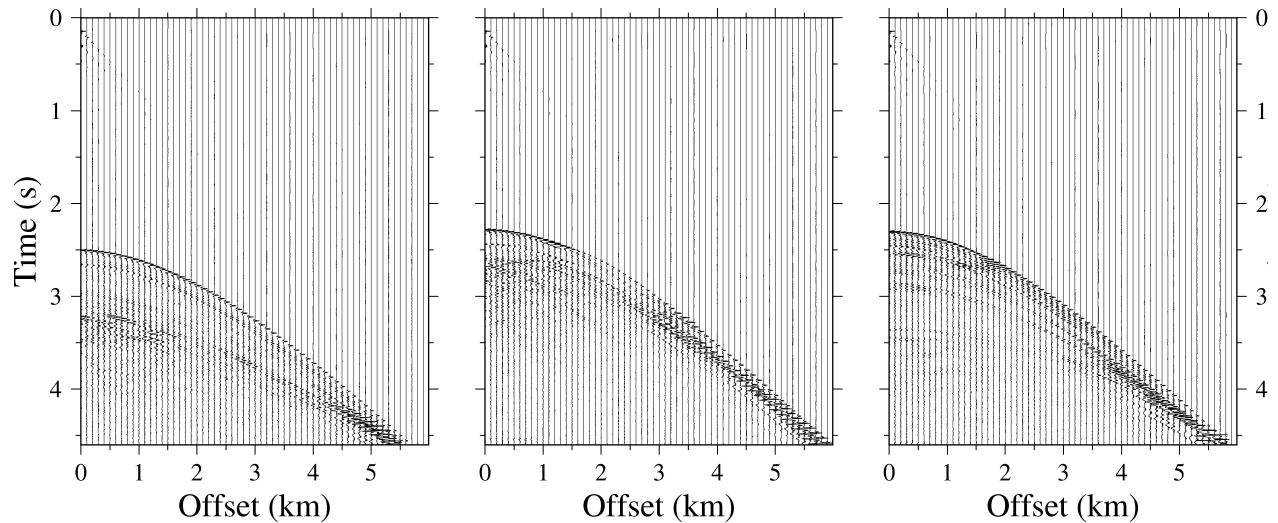

Figure R4 (Figure 4 in the manuscript): Decimated example of the field data set, only 10 traces each km are plotted (so 1/5 of the total) for clarity. From left to right, shot locations are at 171.1, 138.7 and 112.5 km along the profile.
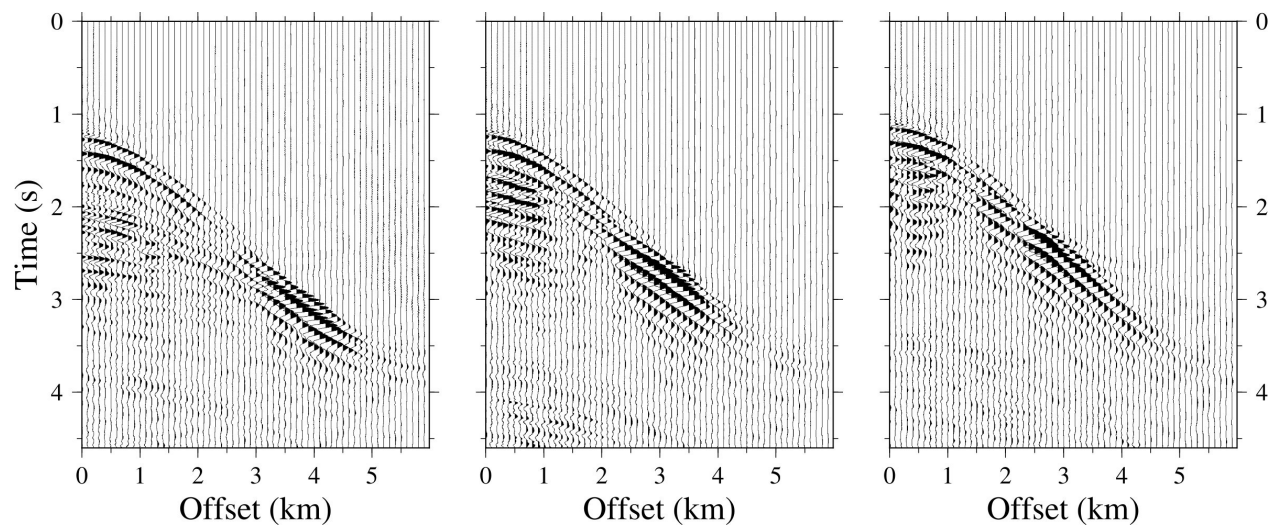

Figure R5 (Figure 5 in the manuscript): Seismic data obtained after the first DC step of the streamer shots recorded in the TOPOMED cruise. The wavefield are displayed for every fifth trace (so 10 traces each km, 1/5 of the total). From left to right, shot locations are at 171.1, 138.7 and 112.5 km along the profile.
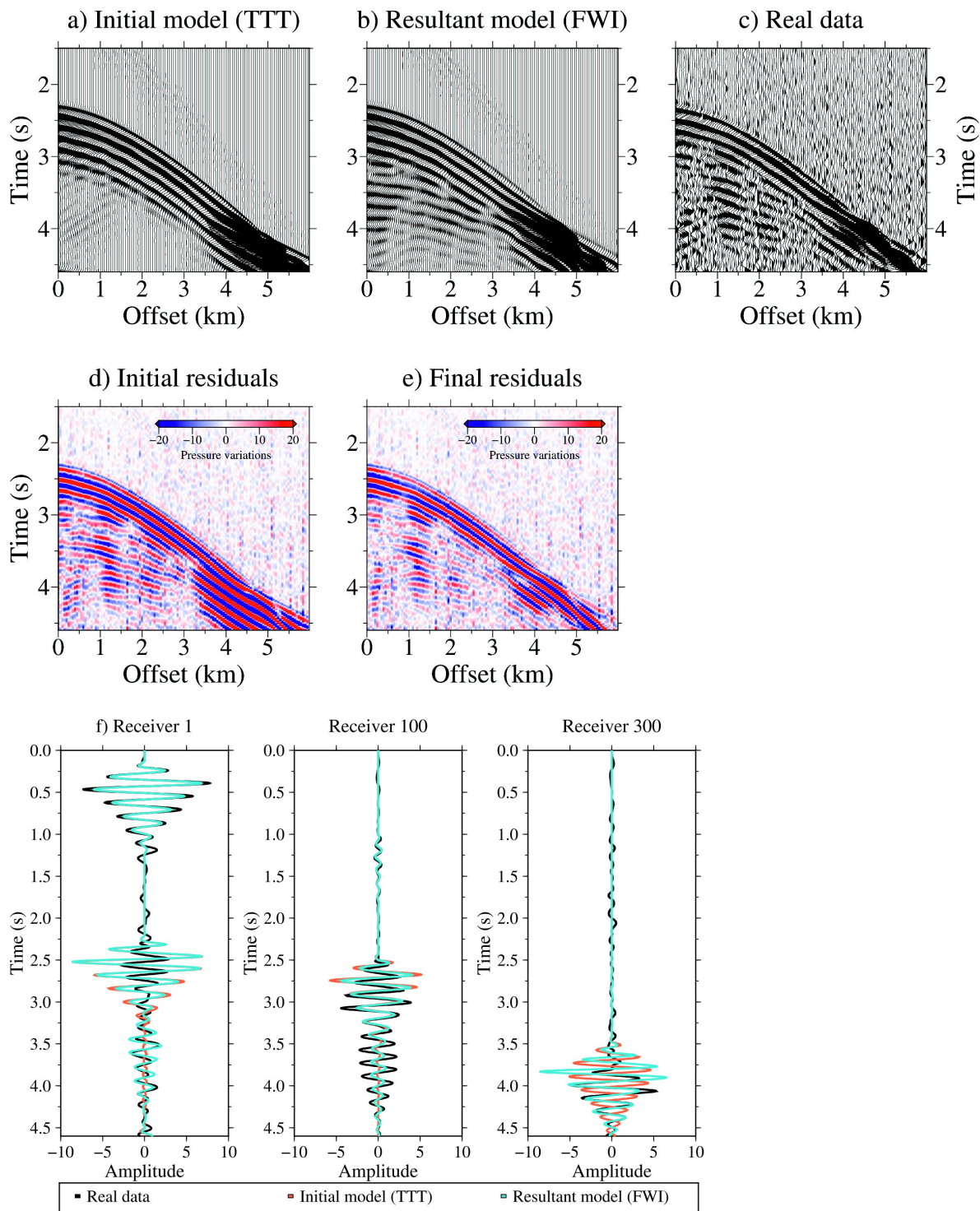
Figure R6 (Figure 11 in the manuscript): Streamer synthetic seismograms, which are computed using the initial (a) and resultant (b) velocity model presented in Fig. 7c and Fig. 10 and the FD solver of the FWI algorithm, together with the real data (c). Only every forth trace at 50 m is plotted. The location of the streamer is at a distance of ~ 101 to 107 km along profile. Initial (d) and final (e) residuals are displayed for a better comparison of the results, as well as (f) initial (orange, a)), resultant (blue, b)) and real (black, c)) seismic traces for receivers 1, 100 and 300.
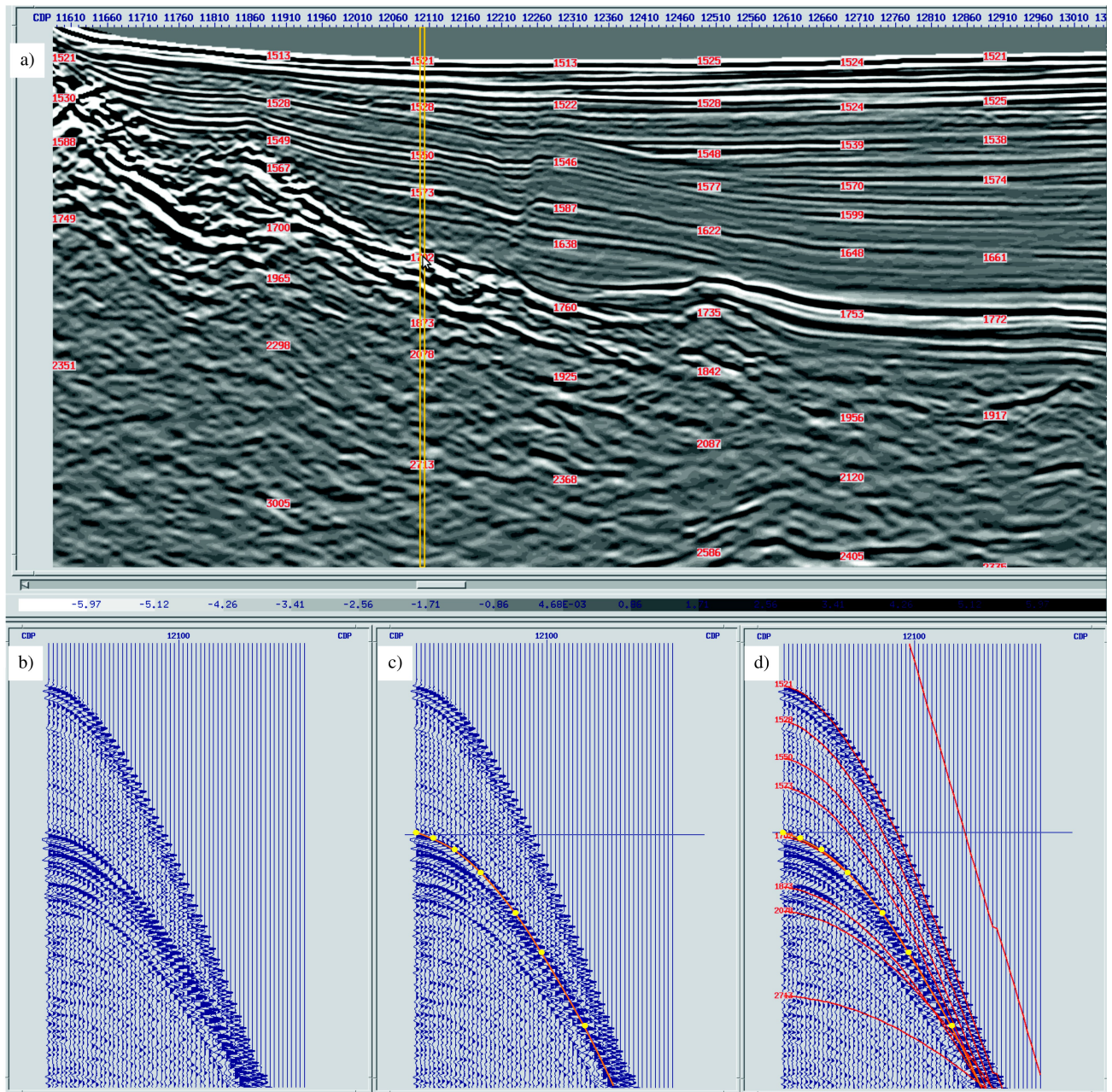
Figure R7 (Figure A1 in the manuscript): Reflector picking procedure. Processed stack image of a section of the seismic profile (a) showing the location of the CMP gather that is plotted with (c), (d) and without (b) the TOB reflection travel times (yellow circles) used as input for the TTT. The TOB is a bright continuous reflection on CMP gathers. The arrow in (a) and horizontal line in (c), (d) indicate the position of the TOB reflection in the stack image and CMP gathers. Red lines in (d) correspond to velocity parabolas that are used as a velocity control during the reflector picking.
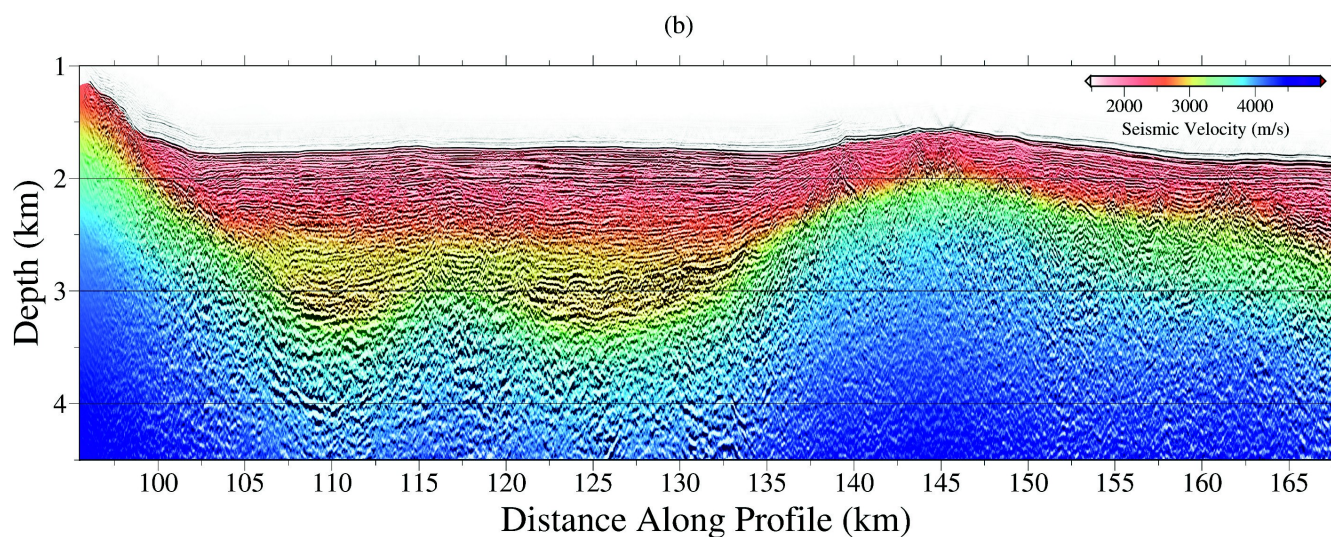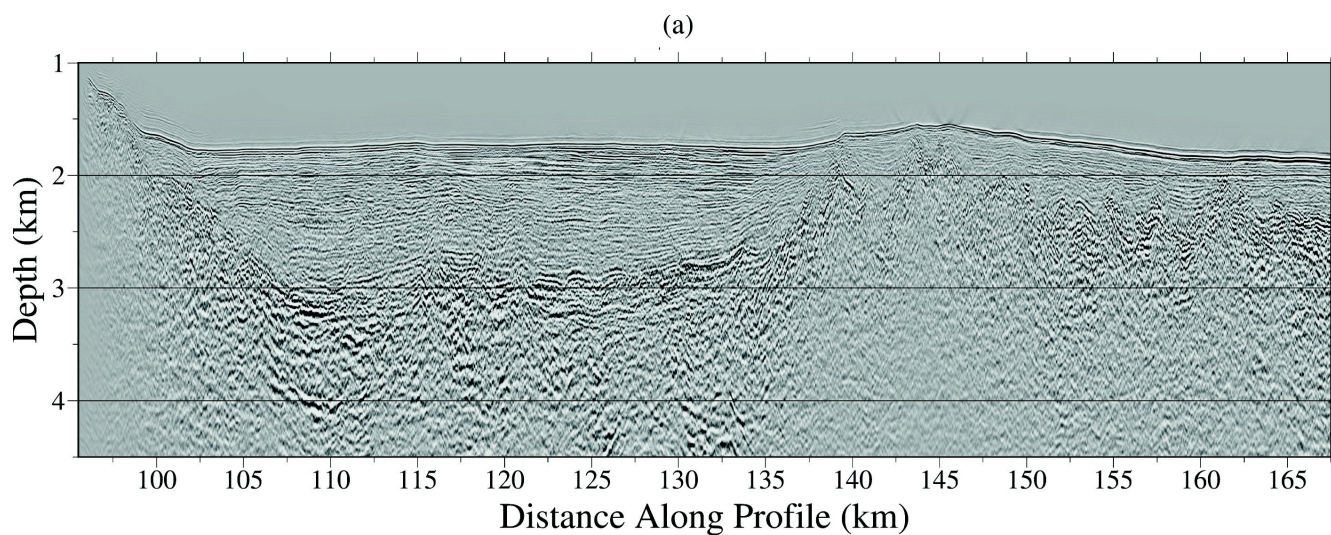
Figure R8 (Figure 14 in the manuscript): a) Kirchhoff depth migration result based on the velocity model shown below (b) obtained after the joint DC refraction and streamer reflection TTT.
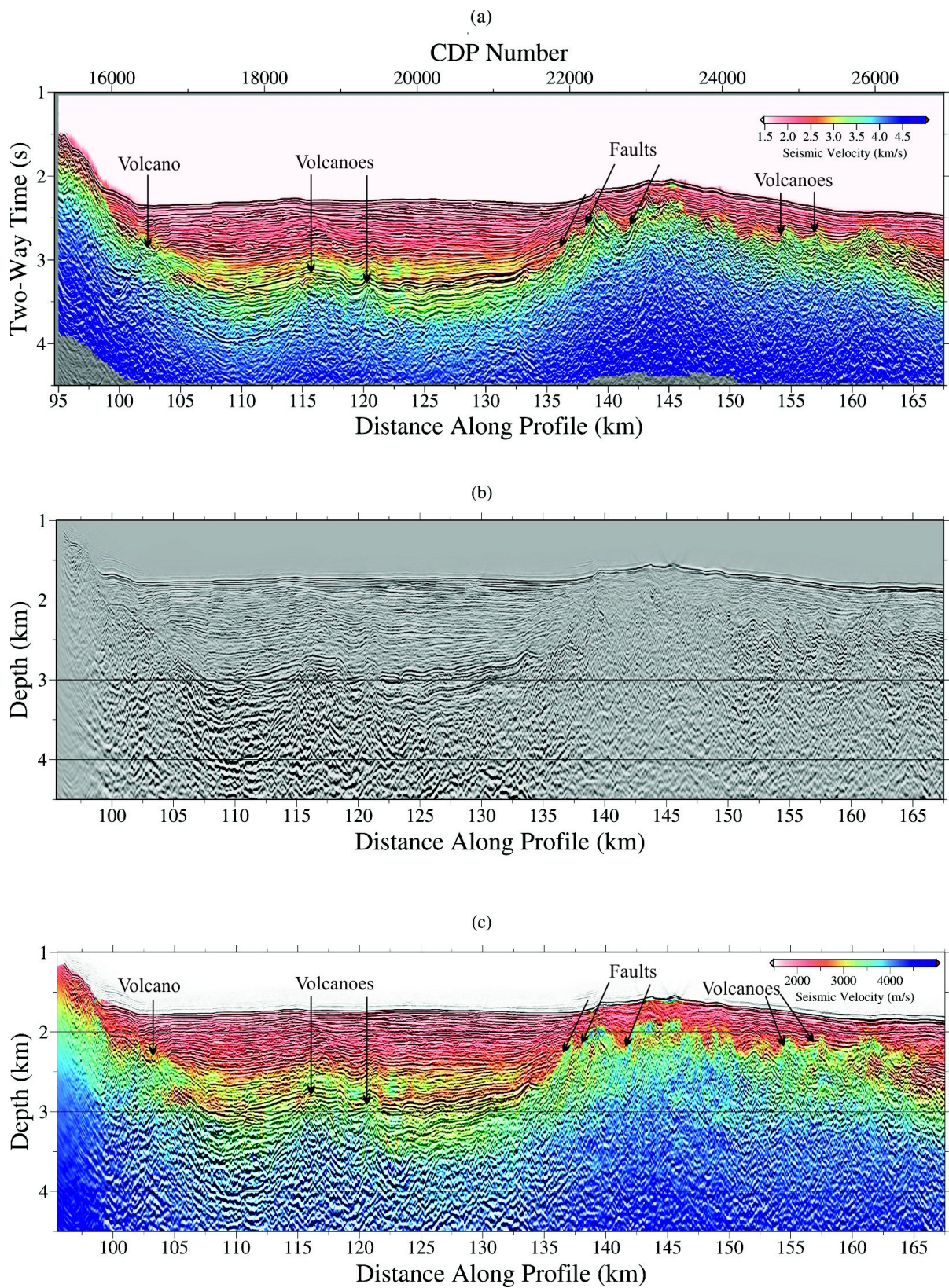
Figure R9 (Figure 16-15 in the manuscript): (a) Geological interpretation superimposed to Fig. 13. (b) Kirchhoff depth migration result based on the velocity model shown below (c) obtained after the modeling sequence proposed on this paper (joint DC refraction and streamer reflection TTT + FWI). Geological interpretation is also superimposed to (c).

Full waveform inversion of short-offset, band-limited seismic data in the Alboran basin (SE Iberia)

Clàudia Gras[1], Daniel Dagnino[1], C. Estela Jiménez-Tejero[1], Adrià Meléndez[1], Valentí Sallarès[1], and César R. Ranero[2]

[1]Barcelona Center for Subsurface Imaging, ICM, CSIC, 08003, Barcelona, Spain
[2]ICREA, Passeig de Lluís Companys, 23, 08010, Barcelona, Spain

Correspondence: Clàudia (gras@icm.csic.es)

Abstract. We present a high-resolution P-wave velocity model of the sedimentary cover and the uppermost basement until $\sim$ 3 km depth obtained by full-waveform inversion (FWI) of multichannel seismic data acquired with a 6 km-long streamer in the Alboran Sea (SE Iberia). The inherent non-linearity of the method, especially for short-offset, band-limited seismic data as this one, is circumvented by applying a data processing/modeling sequence consisting of three steps: (1) data re-datuming by back-propagation of the recorded seismograms to the seafloor; (2) joint refraction and reflection travel-time tomography combining the original and the re-datumed shot gathers; and (3) FWI of the original shot gathers using the model obtained by travel-time tomography as initial reference.

The final velocity model shows a number of geological structures that cannot be identified in the travel-time tomography models or easily interpreted from seismic reflection images alone. A sharp strong velocity contrast accurately defines the geometry of the top of the basement. Several low-velocity zones that may correspond to the abrupt velocity change across steeply dipping normal faults are observed at the flanks of the basin. A 200-300 m thick, high-velocity layer embedded within lower velocity sediment may correspond to evaporites deposited during the Messinian crisis. The results confirm that the combination of data re-datuming and joint refraction and reflection travel-time inversion provides reference models that are accurate enough to apply FWI to relatively short offset streamer data in deep water settings starting at field-data standard low frequency content of 6 Hz.

Copyright statement. TEXT

1 Introduction

Seismic methods are one of the most powerful existing geophysical tools to extract information on the structure and properties of the Earth′s subsurface. These techniques have been, and currently continue to be, widely used to obtain images of the sediments and crust and to map the variations of physical properties, particularly P-wave velocity (V_P). The V_P distribution provides information on the subsurface geology and on the type of rocks constituting the different structures. The most objective and widely-used method to retrieve V_P models is seismic tomography, using either travel-time information only, as in travel-

time tomography (TTT) (e.g., Zelt and Smith, 1992; Korenaga et al., 2000; Hobro et al., 2003), or a more complete set of waveform attributes including both phase and amplitudes as in full waveform inversion (FWI) (e.g., Virieux and Operto, 2009). Whereas TTT is a robust, moderately non-linear technique that provides coarse V_P models, FWI is computationally demanding and strongly non-linear, but it has the potential to provide higher resolution V_P models. For their characteristics, TTT and FWI are considered to be complementary, so that they are often combined and applied together. In general, TTT is applied first to get a moderate resolution V_P model, which is then used as initial model for a multi-scale FWI. The key of a successful combined inversion is to obtain an initial FWI model that is kinematically correct, in the sense that the simulated and recorded waveforms are not cycle-skipped at the lowest frequency available to perform FWI (e.g., Shipp and Singh, 2002). Long-offset refracted waves containing long-wavelength information are a key ingredient to obtain a kinematically correct reference macro-velocity model through TTT. Due to the intrinsic limited offset of near-vertical reflection acquisitions, refracted waves are typically not observed in seismic reflection recordings, so that they are barely used to perform seismic tomography. These narrow-azimuth reflection acquisitions, sample only small values of scattering angles at a diffracting point if the background model is smooth. In the framework of the single-scattering approximation, for small values of scattering angles only the high-wavenumbers of the subsurface can be recorded and thus, reconstructed from the wavefields applying conventional FWI (Brossier et al., 2014).

One example of near-vertical seismic acquisition is marine multichannel reflection seismic (MCS) systems. The setup of MCS systems is designed to record near-vertical reflections with high redundancy (large number of channels), but offset is limited to the streamer length. In deep water settings, the critical distance for refracted waves is often beyond this distance. Near-vertical reflections hold indirect information on the velocity field and on the depth of the layer where they are reflected. Due to the inherent velocity-depth trade-off and to the possible errors in the identification of the reflector boundaries, we cannot expect a precise velocity model by travel time inversion using only one or few reflectors with no lateral or vertical continuity.

Aside from the lack of long-offsets or wide-aperture acquisition systems and hence, of refracted phases and wide scattering angles needed for TTT, the ability to solve low-wavenumber information of the subsurface is also influenced by the low angular frequency content of the signal (Brossier et al., 2014). The most commonly used sources in MCS systems are airguns, which are typically not capable of generating energy below 4-5 Hz, so in those cases the signal to noise ratio is very low. The combination of (1) lack of refracted waves that make difficult to apply TTT and (2) lack of low frequency components in the dataset that aggravate the issue of cycle-skipping, makes it even more challenging to perform TTT+FWI with MCS data.

A common approach to mitigate the above-mentioned issues and apply TTT+FWI to marine MCS data in deep water settings is by re-datuming the collected data (Arnulf et al., 2011, 2014; Henig et al., 2012) to a virtual recording surface first (Qin and Singh, 2017), so that refracted waves can be identified as first arrivals in the virtual recordings to perform TTT. Then first arrival TTT is applied, and the resulting model is used as a reference for multi-scale FWI starting at the lowest available frequency. The success of this approach relies on obtaining a TTT V_P that allows overcoming cycle skipping. This issue is outstanding and difficult to overcome when data lack frequencies below 3-4 Hz (e.g., Shah et al., 2012; Jiménez Tejero et al., 2015), which is the case of most field MCS data.

Our goal here is designing and testing a data processing/ modeling strategy that allows applying FWI to limited-offset MCS data in deep water settings starting FWI at frequencies higher than 5 Hz. For this, we propose a modified version of the joint

TTT+FWI approach that combines wave equation datuming with joint refraction and reflection TTT (instead of first arrival TTT), prior to FWI. The performance of the proposed approach is demonstrated by applying it to field MCS data acquired with a 6 km streamer in the $\sim$ 2 km deep Alboran basin (W Mediterranean). The final FWI model display a number of details of geological interest that cannot be seen in the initial TTT model, including a layer that may correspond to evaporites embedded in the sedimentary basin, faults and short wavelength highs and lows of the crystalline basement. Finally we show that this V_P model allows to significantly improving pre-stack depth migrated (PSDM) images.

2 Methodology

Having a good initial model is essential to overcome the inherent non-linearity of FWI for band-limited seismic data. A key ingredient to build that model is using travel times of refracted waves, which for deep water and short-offset settings are not visible as first arrivals. The objective of our workflow is solving these challenges and applying FWI under these adverse conditions. The proposed approach consists of the following steps: 1) data re-datuming by Downward Continuation (DC) of the recorded data to the seafloor, 2) joint refraction and reflection TTT of the original and re-datumed shot gathers, 3) FWI of the original shot gathers using the model obtained in (2) as initial reference, 4) pre-stack depth migration of the original shot gathers using the FWI V_P model. In the next sections we describe the fundamental aspects of the different methodologies. The general workflow is summarized in Fig. 1.

2.1 Wave equation datuming (Downward continuation)

As stated in the introduction, refractions are essential to build a macro-velocity model with the correct low wavenumber content on it. As MCS seismic acquisition systems have a limited offset (Fig. 2a), the first arrivals in deep-water environments are typically dominated by the direct wave as well as reflections, whereas refractions are hidden behind these phases (Fig. 2b). Several methods have been proposed to recover refracted signals by changing the plane of the acquisition geometry (e.g., Schuster and Zhou, 2006; Vrolijk et al., 2012; Wapenaar et al., 2014; Cho et al., 2016). The goal of wave equation datuming and, particularly, of DC, is to extrapolate the recorded data to a virtual datum surface so that refractions can be retrieved as first arrivals. The technique used here to back-propagate the recorded data follows the steps proposed by Berryhill (1979, 1984) and the scheme of McMechan (1982, 1983). This procedure is easy to implement and can deal with spatial velocity variations within the water column and irregular datum surfaces, which are often located at the seafloor (Fig. 2c).

The idea is eliminating the effect of the water layer by simulating a sea bottom acquisition (both sources and receivers). Without the water layer, the direct-wave signal and the seafloor reflections disappear, and the critical distance is reduced to zero offset. Thus, the virtual recordings contain early refractions visible as first arrivals (e.g. Fig. 2d).

In 2D, the recorded wavefield can be expressed as,

$$S(x_s, z_s; t) * G(x_s, z_s; x_r, z_r; t) = u(x_r, z_r; t) \tag{1}$$

3

Where t represents the time, x and z the horizontal and vertical spatial coordinates, S the source signal, G the Green's function that defines a forward wavefield extrapolation process in a heterogeneous subsurface observed at (x_r, z_r) from a source at (x_s, z_s) and u is the seismic trace. In the streamer configuration (Fig. 2a), there are two contributions of the wave propagation through the water column that can be removed from the data by applying wave equation datuming. Following the matrix operator notation used by Berkhout (1981, 1997a, b), the true two dimensional (2-D) wavefield propagation from sources to receivers is also formulated here by means of matrices but divided in three equations, each considering a different part of the total trajectory of the wavefield, as

$$S^T(x_s, z_s; t) * G_d(x_s, z_s; x_d, z_{bat}; t) = u_d^T(x_d, z_{bat}; t) \tag{2}$$

$$u_d^T(x_d, z_{bat}; t) * G_e(x_d, z_{bat}; x_u, z_{bat}; t) = u_u^T(x_u, z_{bat}; t) \tag{3}$$

$$u_u^T(x_u, z_{bat}; t) * G_u(x_u, z_{bat}; x_r, z_r; t) = u^T(x_r, z_r; t) \tag{4}$$

Where Eq. 2 corresponds to the downward propagation (G_d) from the source (S) surface (x_s, z_s) to the seafloor (x_d, z_{bat}) through the water column; Eq. 3 describes the effects of the wave propagation through the Earth's subsurface (G_e) for an acquisition system deployed at the seafloor from an incoming wavefield at (x_d, z_{bat}) to an outcoming wavefield at (x_u, z_{bat}), and 4 represents the upward propagation (G_u) from the seafloor (x_u, z_{bat}) to the recording surface (x_r, z_r) through the water media. The u_d^T, u_u^T and u^T represent the pressure field expressed at its corresponding datum surface as a row vector, being T the transpose symbol. The G_d and G_u operators are equal only if sources and receivers are located at the same horizontal coordinates and thus the downward and upward vertical propagation follow the same path.

Here the inverse extrapolation follows the boundary value migration (BVM) scheme of McMechan (1982, 1983). In this case, the inverse extrapolation is described by the source, wave equation and boundary conditions used. The source function through time is formed by the observed wavefield, which is introduced in the 2-D model plane sequentially and backwards in time as a boundary value. The wavefield transferred from the seismogram to the model at each time step is taken as an equivalent source. The complete seismogram is used, thus no pre-selection and arrival identification is needed. Prior to re-datuming, we: (1) deleted noisy channels, (2) applied Butterworth 1/3-40/60 filtering and (3) muted direct arrival. We did not apply amplitude corrections. The proper source spacing and optimal recording time step are the ones that avoid aliasing issues, grid dispersion and reduce the effects caused by the discrete approximation of the wavefield.

The inserted wavefield is propagated back in time following the acoustic wave equation solved by the forward extrapolation scheme (Dagnino et al., 2016). The following 2-D acoustic differential equation,

$$\frac{1}{\kappa(x,z)} \frac{\partial^2 u(x,z;t)}{\partial t^2} + \nabla \cdot \left(\frac{1}{\rho(x,z)} \nabla u(x,z;t) \right) = s(x,z;t) \tag{5}$$

Where x is the horizontal space coordinate, z the depth or vertical space coordinate, t the time, u the pressure wavefield, s the source signal, κ the compressibility and ρ the density of the medium; is implemented in a recursive and explicit finite difference (FD) scheme of the two-way wave equation in the space-time domain (Dagnino et al., 2016). The inverse extrapolation is the reverse of the forward propagation because the recorded traces that act as sources (s) are reversed in time. This is because for a lossless medium, the wave equation is symmetric in time (reversible) and the Green's function is reciprocal. The idea is that reflectors exist in the Earth at places where the onset of the downgoing wave is time coincident with an upcoming wave (Claerbout, 1976). Thus, the true forward propagation of the wavefield is compensated by the inverse extrapolation that sends the recorded energy back through the same paths, but in the opposite direction. The progressive focusing in time and space of the wavefronts makes it possible to recover the wavefield at a later stage. Hence, the redatumed wavefield is directly imaged at the new surface depth along the extrapolation time. A multishooting technique is applied to back-propagate the wavefield from all the recorded seismograms at once, increasing the signal-to-noise ratio (SNR) and making the process more compact both in memory and computational time.

Unlike re-datuming approaches that use the Kirchhoff implementation (Berryhill, 1979, 1984; Shtivelman and Caning, 1988; Arnulf et al., 2011), the computation of extrapolated data at intermediate depth levels is needed. Despite the fact that it increases the computational complexity as compared with the Kirchhoff method (Arnulf et al., 2014), the wave equation imaging allows including laterally and vertically variable water velocity, in contrast with other techniques in which the water velocity is considered constant (Cho et al., 2016) or need a specific ray tracing (Shtivelman and Caning, 1988). Recursive extrapolation makes that spatial velocity variations can be handled properly. To define the 2-D water velocity model we use in situ oceanographic measurements of water properties obtained with expendable bathy-thermographic (XBT) probes. We have also tested the effect of the water velocity model used to DC the shotgathers versus the travel time picks. In our area of study the velocity changes in the water layer are small (<20 km/s) so considering either a realistic or an homogeneous model have a minor effect in the first arrival travel time picks, smaller than the travel time picking uncertainty ($\pm$ 0,035 s). Moreover, the new virtual positions do not necessarily have to be in a flat surface or at several meters above seafloor surface as in other studies (e.g., Harding et al., 2016), here the new virtual surface is the seafloor relieve itself extracted from bathymetric data. In contrast to Harding et al. (2016), which use the DC data to perform FWI, we only use first arrival travel times to perform TTT, so surface-related effects do not affect here the inversion results. The main drawback of this DC strategy is the computational cost, worse wavefield reconstruction in the far offsets caused by the truncation of the recorded data, and potential artefacts due to propagation effects.

In the inverse extrapolation process, complex-frequency-shifted perfectly matched layers (CFS-PML) (Zhang and Shen, 2010) are imposed as absorbing boundary conditions to avoid spurious reflections in all model boundaries that may cause interference and mask the correct seismic phases. The only exception is the time-dependent boundary values associated to the equivalent sources.

In summary, the main steps of the DC approach are the following: first, we back-propagate the shotgathers through the water layer from the receiver positions to a virtual surface located at the seafloor. Then, we sort the resulting virtual seismograms into common receiver gathers. After that, we back-propagate the receiver gathers through the water layer, this time from the

virtual surface to the source positions, in the opposite direction to its true trajectory. Finally, we resort the resultant wavefield in the original shotgather domain (Fig. 2d). The rearrangement of the data is possible because of the reciprocity principle, so that if source and receiver have identical directional characteristics, then interchanging the positions of sources and receivers yields the identical seismic trace (Berryhill, 1984).

5 The final virtually recorded wavefield with sources and receivers located at the seafloor can be expressed as,

$$S^T(x_s, z_s; t) * G_e(x_s, z_{bat}; x_r, z_{bat}; t) = u_r^T(x_r, z_{bat}; t) \tag{6}$$

Because the redatumed wavefield (u_r) at (x_r, z_{bat}) is built by using finite, discrete and single-sided recordings, it is not the pressure wavefield that would be recorded by performing the experiment with the virtual geometry setup. Moreover, the energy coming from the direct and surface waves are not included. In general, the amplitude of the resulting wavefield is reduced due

10 to the energy loss through the absorbing model boundaries during the inverse extrapolation process. Therefore, amplitude is not preserved as in other re-datuming approaches (Schuster and Zhou, 2006) and its attenuation factors are related to geometrical spreading and transmission losses. As amplitudes are affected by the re-datuming process, it is not straightforward performing FWI using DC data. However, arrival times of the different phases are correct (which is our main objective) if the V_P model of the water layer is accurate enough. A better approximation of the true wavefield, and thus a better result, would be achieved

15 using denser and larger arrays. For sparse data sets the redatumed arrivals would be less focused due to a larger amplitude loss and noise effects, nevertheless correctly located.

The accuracy of the resulting wavefield also depends on the precision of the forward solver. The numerical algorithm (Dagnino et al., 2016) uses a Runge-Kutta method of fourth order in time (Lambert, 1991) and sixth order approximation in space discretized with a straggered grid (Virieux, 1986). It employs the extrapolated output as input for the next extrapola-

20 tion step (recursively) plus four weighted average increments.

Another effect that can affect the result is the presence of artefacts such as diffraction tails located at the edges of the array and therefore where the observed wavefronts are truncated. These artefacts are due to the finite aperture of the recording system. Their effects are larger for small aperture setups, where they can interfere with the energy focusing in the new datum surface and mask the true arrivals. Moreover, the finite aperture of the acquisition array causes a non-uniform illumination of

25 the wavefronts, which results in energy mitigation near the recording limits. Finally, the resulting extrapolated wavefield is also influenced by numerical dispersion, which attenuates high frequencies as it travels backwards in time through the medium. Thus, the spectrum of the frequency bandwidth of the data is reduced due to the Earth effect that acts as a filter (Berkhout, 1997a, b).

2.2 Joint refraction and reflection travel-time tomography

30 The goal of this step is to obtain a reference V_P model by TTT of some predefined seismic phases; in particular, we combine first arrivals identified at the virtual DC recordings with reflections recorded at the original MCS data. Therefore, the data input file contains both type of travel time picks, each with their corresponding source/receiver positions. In the case of DC arrivals,

the 2D coordinates will be the ones in the virtual surface, but for the reflections they will be at the actual location of the air gun source and of the individual streamer channels. Further details for our particular study area and experiment are presented in Section 4.2. As a rule of thumb, it is considered that V_P models obtained by TTT are correct to a depth that is about a half of the maximum experimental offset.

We use a version of the TOMO2D code (Korenaga et al., 2000) modified by Begovič et al. (2017) to enable locating both sources and receivers within the water layer so as to invert streamer data. A two-dimensional V_P model is parameterized in our study as a regular mesh hanging from the seafloor, i.e. the model is limited to the subsurface medium and does not include the water layer. Propagation velocity in the water layer is fixed to 1.5 km/s. Arrival time differences using a realistic water V_P model are much smaller than the expected data misfit, typically several tens of milliseconds. A 1-D floating reflector is defined along the profile. This means that it is parametrized as a node array that is independent from the velocity mesh, and both reflector array and velocity mesh can be spatially variable. Travel times for given V_P and reflector depth models are computed using a ray tracing strategy that combines the graph method (Moser, 1991) with a subsequent ray bending refinement (Moser et al., 1992).

The partial derivatives of travel time residuals with respect to velocity (for first arrivals and reflections) and reflector depth (only for reflections) are introduced in the Fréchet matrix with additional regularization constraints (smoothing, damping). Thus, this interface depth model is inverted simultaneously with the V_P model. The inverse problem to iteratively update the V_P and depth models is solved using the sparse matrix solver of the LSQR algorithm of Paige and Saunders (1982). The difference between observed and synthetic travel times is iteratively minimized as a least-squares problem until a stopping criteria is fulfilled. Regularization constraints, smoothing and damping, are applied to stabilize the minimization process avoiding singularity of the matrix. For further details, see Korenaga et al. (2000) and Meléndez et al. (2015).

2.3 Full waveform inversion

The last inversion step of the proposed approach is using the V_P model obtained by TTT as initial one for adjoint-state FWI. In this case, we use the original seismograms recorded by the MCS system as input to refine the V_P model. As in the case of the DC, the FWI code uses the acoustic domain FD solver of Dagnino et al. (2014, 2016) for the forward- and back-propagation of the wave equation.

2.3.1 Forward problem

For the forward problem, we consider the 2-D acoustic wave equation in the time domain expressed in the form that appears at Eq. 5 to simulate the synthetic pressure wavefield. The code includes CFS-PML (Zhang and Shen, 2010) as boundary conditions, and we use 25 layers on the left, bottom and right boundaries to eliminate numerical reflections and a free surface at the top of the model that represents the water/air discontinuity. The optimal temporal step to solve the forward problem fulfills the usual Courant-Friedrich-Lewy stability condition to avoid instabilities in the PML (Dagnino et al., 2014). To estimate the source for the first iteration we use a Ricker wavelet with central frequency of 20 Hz. Then the source wavelet is inverted at each iteration as proposed by Pratt (1999) prior to velocity inversion, so the updated source signature is used in the subsequent

iterations (Dagnino et al., 2016). The source signature is calibrated and inverted for all the individual shots up to 0.5 km offset. The initial (T/6) and final (6T) oscillations of the source signal are removed to avoid the introduction of noise in the synthetic data.

2.3.2 Inverse problem

5 The goal of the inversion is to find the parameters m, in our case V_P, that minimize the discrepancy or misfit between some predefined waveform attributes of the observed (u^o) and synthetic (u^s) seismic traces at each iteration (Jiménez Tejero et al., 2015). We use the classical L2-waveform objective function (χ) that is built as the least-squares norm of the corresponding misfit (Tarantola, 1987) for a source as follows,

$$\chi = \sum_{r,t} \frac{(u^s(x_r,t) - c_r u^o(x_r,t))^2}{2} \qquad (7)$$

10 where x_r is the receiver position, t the time and c_r the calibration term. The calibration term is calculated independently for each source-receiver as in Dagnino et al. (2016), but fixing the sea bottom reflection instead of the direct water wave. This term is useful to correct the source signature changes due to an irregular bathymetry and the inhomogeneous source directivity. The wavefield of each shotgather is resampled to fit on the finite difference grid. The synthetic data for the misfit calculation is simulated solving the forward problem previously explained in a reference model and its updates.

15 The acoustic 2-D forward modeling code compensates the actual 3-D amplitude decay of the field data multiplying the field data by $\sqrt{t}$ (Hicks and Pratt, 2001; Dagnino et al., 2016) before the calculation of the residuals. Moreover, the energy of the shots is normalized dividing it by the total source energy to overcome problems due to elastic effects. The density model of the subsurface used for the inversion also affects the reflectivity part of the data set, so it is important to interpret the reflectors not only as velocity contrasts. In our case, it is updated after the inversion of each frequency band using the V_P model obtained at

20 the previous iteration and the Nafe-Drake (Ludwig) experimental relationship (Mavko et al., 1998).

The following step consists on updating the model in the direction where the misfit decreases. For this purpose, we calculate the gradient of the misfit following the adjoint-state method proposed by Lailly (1983) and Tarantola (1984). The adjoint field is obtained back-propagating the residuals in time for each receiver position. The adjoint source is calculated as $f(x,z;t) = -\nabla_{u^s}\chi$.

25 The gradient of the objective function with respect to the model, $\nabla_\kappa \chi$, is computed by convolution of the forward propagated wavefield of the source term and the adjoint or back-propagated wavefield of the residuals from the receiver location (Tarantola, 1984).

$$\nabla_\kappa \chi = \sum_s \frac{1}{\kappa(x,z)} \int_t \frac{\partial u(x,z)}{\partial t} \frac{\partial f(x,z)}{\partial t} \qquad (8)$$

Because of non-linearity, the gradient can point toward a local minimum. In the FWI code, a gradient-based preconditioning

30 concentrates model updates in the regions where the gradient is more reliable. Moreover, it also adds a weighting to compensate

the amplitude decay of the signal away from the source. Furthermore, the gradient is balanced to reduce its sensitivity to the presence of 'spikes' with a smoothing regularization based on a 2-D low-pass zero-phase Butterworth filter. Details on the gradient pre-conditioning techniques can be found in Dagnino et al. (2014, 2016).

We use the steepest descendent (SD) approach in the optimization problem because it is particularly little sensitive to noise. The step used to find the direction in the model space where the misfit function locally decreases, is just the opposite of the gradient in the SD approach, $p_i = -\nabla\chi_i$. In the FWI algorithm, we use a normalized search direction. Then, the optimal step length (α_i) is determined by imposing $\alpha_i = \min_{c>0}\chi(m_i + cp_i)$. The final step is obtained after acquire the minimum of a polynomial approximation over three steps calculation in the search direction of the misfit function, p_i. Then, the model can be updated as $m_{i+1} = m_i + \alpha_i p_i$.

A multi-scale approach where frequency bands are inverted sequentially, from low to high frequencies (Bunks et al., 1995), is applied to reduce the risk of falling into a local minima and to add details to the model progressively. In particular a low-pass Butterworth filter is applied to the data, so that let solve different frequency bands of the same data at each iteration, restricting the inversion to a limited bandwidth. Implementing this strategy mitigates the non-linearity, because the objective function is smoothed when the data is filtered. High wavenumbers are then incorporated sequentially into the model successive iterations. The maximum number of iterations per frequency step is fixed to 10. An additional stopping criteria that we use in the inversion is the Arminjo rule (Nocedal and Stephen, 2006) when changes are < 0.01.

2.4 Pre-Stack Depth Migration

We apply a 2-D Kirchhoff depth migration to image the subsurface structures directly in depth by using the V_P models obtained and the high frequency content of the data set. The migration module applied is part of the Seismic Unix platform. We use the sukdmig2d command for the data migration. The travel-time table required for the depth migration is obtained with the rayt2d command, which calculates the 2-D ray tracing along the V_P model.

3 Study area and data set

The field data used to test the described approach correspond to an experiment conducted in the Alboran Sea, a complex basin located in the Westernmost Mediterranean (Fig. 3) (Booth-Rea et al., 2007, 2018). The target of this survey was to image the structure and properties of the sediment units and the nature of the uppermost basement.

The seismic records used in this study were collected in 2011 on board Spanish R/V Sarmiento de Gamboa, as part of the TOPOMED cruise. We select an 80 km-long section of the TM28 profile that crosses the central and deepest part of the basin, across a volcanic arc (Gómez de la Peña et al., 2018), to test the approach. The input data consist of series of seismograms (traces) recorded with a 6 km-long streamer (Fig. 4). The streamer has 480 channels with a group interval of 12.5 m. The length of the trace record used is 8 s and the time sampling is 2 ms. In total, we use 1517 air-guns shots with an average spacing of 50 m. The source power was 4600 c.i., producing a central frequency of ~ 20 Hz. The system was towed at a depth of 10 m, and the nearest-offset distance was 203.7 m. The early arrivals in the streamer recordings are dominated by the shallow

near-vertical sediment reflections, as it can be seen in Fig. 4. Hence, data pre-conditioning is essential to identify first arrivals and make the data set appropriate for TTT.

4 Results

Here we show the results after applying each step of the modeling strategy described in the previous sections (see Fig. 1).
5   We work under the premise that we have no a priori information on the structure and properties of the subsurface, so the goal is recovering all the possible information on the V_P model from the MCS data alone. Before starting our processing and modeling workflow, we apply a 2-D band-pass minimum-phase Butterworth filter to the dataset for swell noise removal. The low-cut and high-cut of the band-pass filter are 2 and 60 Hz, respectively. The DC result of the streamer recordings is first shown to make the early refractions visible as first arrivals. Then, a macro-velocity model is obtained by joint refraction and
10   reflection TTT, which is then used as initial model to perform FWI starting at $\sim$ 6 Hz, the lowest signal frequency available in data set. Afterwards, we perform a PSDM of the recorded data using the 2-D V_P models obtained by TTT and by FWI.

4.1 Downward continuation results

We compute the data re-datuming by wave equation DC of the streamer recordings to the seafloor surface. Figure 5 shows the result after the first step of the DC for 3 shotgathers derived using the full wavefield of each shot and a heterogeneous water
15   layer model built as explained in the methodological section. The grid size in both x and z direction is 12.5 m. We image the wavefronts of all the seismic events recorded in the streamer traces reconstructed by the solver after its back-propagation to the virtual receiver positions in the correct reversed time (i.e. the OBS-type acquisition setting). In this way, we recover several first arrivals that were obscured in the original recordings. However, the main first arrival along the recordings is still the seafloor reflection.

20   In the second step, we resort the virtual OBS-type wavefield in receiver gathers, and then, the data from all shotgathers are combined to construct the final virtual shotgather. Figure 6 shows the final result for the 3 shotgathers also shown in Fig. 5. The DC has collapsed the seafloor reflection towards a single point, so refractions from shallow subsurface can now be identified and tracked as first arrivals from zero offset. First arrivals are more difficult to identify at long offsets, because of amplitude attenuation and the appearance of diffraction tails. Lower panels (d)-e)-f)) in Fig. 6 show the first arrivals used as input for the
25   TTT (blue dots).

4.2 Joint refraction and reflection travel-time tomography results

The inversion parameters used in the TTT are shown in Table 1. We do not use the entire data set to reduce computational burden but all receivers are used to ensure data redundancy. The selected reflection corresponds to the top of the basement (TOB), which is a clear event (e.g. see Figure 9). Reflection travel times are picked from MCS common mid-point gathers
30   where it is a bright continuous reflection, which is also interpreted at the same time in a fully processed stacked image (Fig. A1). The source and receiver positions were projected to a straight line defined between first and last shots, preserving their offset

10

distance. Correlation lengths for the velocity model are set at the top and bottom nodes of the grid and are linearly interpolated for the intermediate nodes. Velocity gradient is stronger in the vertical direction, so the vertical correlation lengths are selected to be shorter than the ones defined in the horizontal direction. Our inversion process follows the layer-stripping strategy as described in Meléndez et al. (2015). In the first inversion step, the initial TOB depth model is a flat boundary located at 2.25

5 km depth, and the initial V_P model follows the linear function of depth from the seafloor $V_P(z)[km/s] = 1.61 + 0.72 \cdot z[km]$, increasing from 1.61 km/s at the seafloor to 4.13 km/s at a depth of 3.5 km below the seafloor (Fig. 7a). In the second inversion step we use the inverted TOB from the first step (Fig. 7b) as initial TOB depth model, and the same data set and inversion parameters. Regarding the V_P model, we build a new initial model that is equal to the V_P model above the TOB obtained in the first step, whereas below it V_P is defined as $V_P(z) = V_{TOB} + (z - z_{TOB}) \cdot \frac{(V_{z_f} - V_{TOB})}{(z_f - z_{TOB})}$, varying from 3.6 km/s at the TOB

10 interface (V_{TOB}) to 6 km/s (V_{z_f}) at a depth of 8.6 km below the seafloor (z_f) (Fig. 7c). The goal of the layer-stripping strategy is to recover the sharp velocity contrast at the sediment-basement reflecting boundary, that might otherwise appear as a smooth velocity gradient, contributing also to the improvement of the deepest part of the model.

The final macro-velocity model is presented in Fig. 7d. The joint refraction and reflection TTT allows recovering the long-wavelength geometry of the sharp sediment-basement boundary. The ray coverage of the model inversion is quantified by

15 the derivative weight sum (DWS) (Toomey et al., 1994) (Fig. 7), and it is influenced by the geometry of the experiment and the subsurface velocity distribution. Thus, ray coverage decreases to the edges of the model and with depth, and it is denser beneath the source locations. The model is generally well constrained in the central part that is covered by both refractions and reflections, whereas the lateral areas mapped only by reflections are subject to a higher degree of velocity-depth ambiguity. In the central area, the overall trend is an increase of velocity with depth, from ~ 1.6 km/s at the seafloor to ~ 4.0 km/s at

20 the bottom. The results indicate that a high velocity anomaly (>3.5 km/s) is located at ~ 115 km distance and ~ 3 km depth. Slower velocities are obtained at shallower depths below the TOB. The velocity contrast that accurately follows the geometry of the TOB is delimiting steeply dipping discontinuities. As an example, we show that on the left hand side of the profile in Fig. 7d the basement is located just below the seafloor (~ 200 m respect to the seafloor), whereas ~ 15 km further along the profile the basement position is at a depth of ~ 3.5 km. On the other side, the geometry of the TOB is gently waved around an average

25 depth of 2 km. The TOB boundary marks strong velocity changes with a velocity jump of ~ 0.5 km/s between sediment and basement velocities (>2.7 km/s).

Figure 8 shows histograms of travel time residuals for three data groups after the first and last iterations of the two inversion steps of the layer-stripping strategy. The distributions of data misfit with the initial model (red lines) are asymmetric and wide for both refractions and reflections. After the first inversion step (orange lines), the distributions are symmetric and

30 narrower with the highest pick around zero. At the initial stage of the second inversion step of the layer-stripping (green lines), an increase of positive travel-time residuals is shown respect to the distribution after the first step (orange lines) due to the insertion of higher velocities below the TOB to recover the sharp velocity contrast. Thus, the distributions (green lines) are slightly asymmetric and wider than after the first step of the layer-stripping (orange lines). The final distributions of travel time residuals (blue lines) are narrower than the previous ones and approach to a Gaussian centered at zero with the largest

35 counts. This fact evidences the improvement of the velocity model, and therefore the corresponding travel time fitting during

the inversion. Lower panels (d)-e)-f)) in Fig. 6 show for 3 DC shotgathers the first arrivals used as input for the TTT (blue dots) and the first arrivals recovered after TTT (red squares). The final velocity model in Fig. 7d shows a good convergence with root mean square (RMS) residual travel times of $\sim$ 20 ms, which is of the order of the picking uncertainty ($\chi^2 \approx 1$). In contrast, the initial RMS before the inversion (Fig. 7a) is $\sim$ 420 ms.

5 Moreover, to assess the reliability of the inversion results, we convert the velocity and TOB models to two-way-time and superimpose the results on top of the time migrated image (TMI) (Fig. 9). The figure shows that the TTT V_P and TOB models and the MCS image are both consistent, since both the velocity contrasts and the depth of the TOB are located in the same two-way-time where the TMI displays important reflectivity changes and the TOB discontinuity. The good match validates our travel time picks, given that neither the interpolation of the reflected picks or the inclusion of the far-offset first arrival DC

10 travel times appear to introduce substantial artifacts or errors in the model. However, the TTT model lack resolution compared to the time migrated MCS image. It does not reproduce the sharp geometry and changes in amplitude of the anomalies in detail, as the highs and lows of the TOB, the faults at the flanks, or the sediment layering seen in the TMI.

4.3 Full waveform inversion results

After TTT, we perform FWI of the original streamer shotgathers starting from a smoothed version of the TTT V_P model of

15 Fig. 7d. A total of 1,512 shotgathers (>720,000 seismograms) are used for the inversion. The length of the traces is 4.6 s. In order to reduce computational costs, we invert sequentially 4 different sets of 378 shotgathers for each frequency band. Source and receiver positions are also projected to a straight line defined between first and last shots, preserving their offset distance. As in TTT, no decimation is applied to receiver sampling.

  We apply a data preconditioning to emphasize reflections between the first arrival, typically the seafloor reflection, and its

20 first multiple, using a data-windowing. We define a time window centred at the seafloor reflection. We identify this phase by using a maximum kurtosis and k-statistics criteria (Saragiotis et al., 2002). We calculate the travel time of the seafloor reflection approximately using a water velocity of 1.5 km/s, the seafloor depth and the source-receiver offset distance. When the first arrival is not found, for example due to a noisy channel, or the observed and simulated seismograms are cycle-skipped for the inverted frequency, then the trace is eliminated. The trace value is set to zero before the first arrival travel time, and after

25 that it is balanced by a function defined as $\sqrt{t}$ to compensate the amplitude decrease in depth. Finally, the trace is also set to zero after the travel time of the first multiple. Therefore, the gradient calculation focuses on information of the data set coming from the near-vertical reflections. Additionally, when the absolute difference between the maximum amplitudes of a field and its corresponding synthetic traces is higher than three times the maximum amplitude of the synthetic trace, it is considered to be a noisy channel and the trace is eliminated. Another issue to consider when comparing synthetic and field data is having a

30 high signal to noise ratio (SNR), so each trace is stacked with its two neighbors at frequencies lower than 8 Hz.

  We start the multi-scale FWI at 6 Hz. The maximum frequency inverted is 16 Hz. A frequency step of 0.5 Hz is applied at the beginning and 1 Hz at the final stages of the multi-scaling inversion. A total of 14 different frequencies in the multi-scaling strategy are inverted. The size of the space grid for the inversion ($dx < \frac{\lambda_{inv}}{5}$) depends on the frequency step and the minimum

velocity of the model, being smaller at high frequencies. The minimum and maximum velocities were constrained as gradient preconditioning and set at 1.4 and 5.0 km/s.

Figure 10 displays the final FWI velocity model. This model has higher resolution than the TTT one (Fig. 7d), showing velocity contrasts of intermediate and short wavelength and a number of geologically meaningful structures that cannot be identified in the TTT model. The improvement due to the higher inverted frequencies is clearer in the shallow part of the model. The velocity contour ~ 3.25 km/s, which corresponds to the sediment-basement boundary, is better focused and detailed highs and lows are recovered. Thus, in the right hand side of the profile the velocity variations reproduce the TOB geometry accurately. Dipping low-velocity features are shown with high resolution. The velocity differences between these anomalies and the surrounding media are of 0.5-1.0 km/s at ~ 2.25 km of depth between 137 km and 165 km along the profile. The sedimentary package shows velocities from ~ 1.7 km/s near the seafloor increasing to ~ 3.0 km/s just above the basement. The blue colour found in the deeper parts of the model, corresponds to high velocities of more consolidated rocks. However, there are areas where the high velocities are shallower than in others, due to the action of normal faults, as in the left edge of the model.

A basin from 95 km to 137 km distance is displayed between of two sloping faults. In addition, a 200-300 m thick, ~ 3 km/s high-velocity layer of ~ 25 km long embedded within the sedimentary package is shown up to a volcano-like structure at ~ 2.6 km depth. The key point is that all these features have been imaged with a streamer of only 6 km long thanks to the workflow followed here.

To show the model improvement in the data domain, in Fig. 11 we compare a recorded shotgather (Fig. 11c) and a synthetic one generated with the FD solver (Dagnino et al., 2014) using the initial TTT (Fig. 11a) and final FWI (Fig. 11b) velocity models. In contrast with the synthetic data generated with the TTT velocity model (Fig. 11a), the shotgather generated with the FWI velocity model (Fig. 11b) presents some near-vertical reflections. Thus, the seismogram simulated with the final model shows a larger number of seismic events compared to the initial one, which only recovered the first arrival phases and the TOB reflection from the TTT. The data-driven preconditioning of the FWI strategy targeted the energy corresponding to the near-vertical reflections. Therefore, in that region final residuals (Fig. 11e) are smaller than initial ones (Fig. 11d), but are larger around the seafloor reflection. Again, in Fig. 11f aside from wave amplitudes, the main difference between the data generated with the TTT model (orange line) and target (black line) seismogram is the presence of reflected waves. We observe a better fit between the final (blue line) and real data set (black line), except for the effects that are not modelled, such as 3D diffractions, or the arrivals not included in the data-based pre-conditioning.

The misfit reduction for three steps of the multi-scale strategy is shown in Fig. 12. Normalized least-squared misfit is reduced after the whole inversion, and the final residual tends to zero typically after 5 iterations.

As in the case of the TTT model, we have converted the FWI V_P model to two-way-time and we have superimposed it to the TMI (Fig. 13). Velocity contrasts present a remarkable match with the reflectors of high amplitude of the TMI. The geometry of the TOB interface reflector, for example, is recovered with great detail. Here, velocity differences are of short wavelength, so many of the details that were previously not displayed can now be properly identified (Fig. 15a).

4.4 Pre-Stack Depth Migration results

The seismic data used for the PSDM are not the same used for FWI, but data that have been processed to attenuate coherent and incoherent noise. Streamer field data were processed using Wiener and a surface consistent deconvolution to increase the vertical resolution attenuating the ringing of the source. Data were sorted in CDP domain. An amplitude balance (Quality factor of 100) was applied to recover the energy lost by geometrical spreading.

We interpolate the FWI result to a 3.125 m grid interval in the horizontal and 3 m in the vertical for the ray tracing. The receivers were spaced 12.5 m, so the horizontal midpoint distance was 6.25 m for the migration. First, we perform a 2-D Kirchhoff depth migration using the processed data set and the TTT velocity model (Fig. 14a). The same PSDM image is also shown in Fig. 14b superimposed with the TTT velocity model, providing additional information on the rock properties. The resolution of the velocity models obtained through TTT is similar to the typical velocity models built for PSDM, although the latter are based fundamentally on reflections. Thus, the PSDM result shown in Fig. 14 should be comparable to the one that would be obtained with conventional velocity analysis and PSDM. Although the main features are imaged, i.e. the basin geometry, some structures and interfaces are not as well defined as when the PSDM is performed using a high resolution V_P model, as it is shown below.

Finally, we repeat the PSDM but using instead the final FWI V_P model, obtained after the processing+ modeling sequence proposed in this work (Fig. 15 b and c). An accurate image of the real structure of the shallow subsurface is displayed directly in depth in Fig. 15b. The sediment layer is thicker on the left side of the profile (~ 1.5 km), where we find the basin, than on the right (~ 0.5 km), where the basement gets shallower. As it can be seen in both the FWI velocity model as well as the PSDM image, the basement is severely folded and cut by clear fault structures along the whole profile. The strong lateral velocity changes produced by the faults have caused some migration smiles at the TOB. The deepest part and edges of the profile where FWI and TTT have the largest uncertainties have likewise the largest uncertainty in the PSDM result. On the whole, the vertical section in Fig. 15b shows the correct geometry of the structures in depth with a high resolution, clearer than in Fig. 14 when a TTT model was used. As an example, the depth and location of the TOB boundary is now comparatively more precise.

A good fit is obtained between the shape of the geological features and the isovelocity contours in Fig. 15c. The volcano-like structure and dipping faults, coincide with well-defined velocity anomalies. Furthermore, a high velocity layer, which was not visible with TTT in Fig. 14b, is also clearly imaged within the sediment package. The combined interpretation helps us to better define the geological structures and to a proper characterization of the nature of the features.

5 Discussion

Information on the structure and properties of the subsurface in the Alboran margin is retrieved by applying a combination of TTT and FWI, particularly the V_P of the sedimentary basin and the geometry of the TOB discontinuity. The area has a heterogeneous vertical and horizontal velocity gradient, with a V_P distribution showing velocity contrasts delineating boundaries of a geologically complex area.

Using the TTT information alone, a coarse V_P model is obtained from no a priori information. No refractions are observed as first arrivals in the original data, so the shotgather recordings are first re-datumed to the seafloor by wave equation downward continuation. We simulate a virtual acquisition system with both sources and receivers located at the seafloor for the TTT, so that refractions can be tracked as first arrivals from zero offset. These near-offset crustal refractions provide excellent structural detail just in the upper portion of our model. But, by improving the velocity detail in the upper region of the model (Fig. 7b), then we obtain better resolution of the deeper structure (as shown in the second TTT inversion, Fig. 7d). In our case, some deep regions of the model are fitted only by the TOB reflection. Under these circumstances, the inversion result suffers from velocity-depth ambiguity. Streamer TTT exploits the dense and even spatial distribution of MCS data, and is further improved upon by DC, which allows for the inclusion of shallower refractions that were only previously obtainable using either seafloor receivers and sources (Henig, 2013). It is important to notice that these highlighted phases would correspond to raypaths in the upper region of the subseafloor section. Moreover, Fig. 9 validates the TTT model obtained, and thereby the travel time picks used, because of the matching of TWT converted velocity and TOB models with the TMI.

While some studies use waveform-modeling in downward continued data (Qin and Singh, 2018), here only the travel time information is used for inversion. The wavefield during the back-propagation is altered due to several factors. The wavefield at the virtual positions obtained is affected by grid dispersion and clearly shows amplitude attenuation. Moreover, the field data used to reconstruct the virtual seismogram is a discrete single-sided recording and, therefore, part of the signal is missing during the back-propagation hindering complete reproduction of the wavefield. This justifies using travel time information from the DC shot gathers instead of the whole waveform.

The data re-datuming allows recovering and identifying the refracted phases as first arrivals, but not all the energy collapses at its corresponding point, making even more difficult the first arrival picking (see Fig. 6). To check the first arrival travel-time errors introduced by the DC procedure we have reproduced the redatuming processing but using synthetic shotgathers simulated with the streamer acquisition geometry and the final FWI V_P model (Fig. 10). As it is shown in Fig. 16 (a)-b)-c), the first arrivals identified in the real DC shotgathers (blue dots) coincide remarkably well for all offsets with the first arrival travel-times of the final TTT V_P model (red squares) and also with the ones in the DC wavefields that are obtained using the synthetic data. The similar results for the first arrivals justify the inclusion of the DC travel times even for large offsets.

The target of applying FWI is to obtain a high resolution model of the properties of the subsurface. However, due to its non-linear behaviour typical FWI applications use data with low frequency content or good kinematical starting models from a prior information (e.g., Sirgue and Pratt, 2004; Brossier et al., 2009, 2014; Morgan et al., 2013, 2016). Here, we have implemented and tested a methodological solution (Fig. 1) to overcome this challenge. In our case, the TTT model (Fig. 7d), is essential to perform FWI from data which lack of low frequencies (<6 Hz). As state in Introduction, TTT is a very robust technique and highly used to build good background initial models to FWI (e.g., Shipp and Singh, 2002; Dagnino et al., 2014; Qin and Singh, 2017, 2018). The TTT result is improved here during the FWI by introducing smaller wavenumbers into the V_P model especially from near-vertical reflections (Fig. 11). Despite the higher level of detail, not all the observed signal is matched by the acoustic approach of FWI (Fig. 11). Possible mismatches most likely originate by the presence of noise, elastic (Warner et

al., 2013; Marjanovic et al., 2019), anisotropy or attenuation effects, which were not taken into account. The seafloor reflection is one example of a discontinuity that is hard to fit by the acoustic formulation of the wave propagation.

In both TTT and FWI V_P models we clearly identify the TOB as an abrupt velocity change. However, the shape of this discontinuity and its velocity contrast differs between the two models in some areas. The FWI result shows sharper boundaries with pronounced dipping contacts at $\sim$ 137-167 km along the profile, that are not displayed in the model obtained by TTT. Those dipping contacts, which may represent fault planes in some cases, are of short wavelength and separate zones of abrupt velocity changes that might be associated to lithological variations. This difference becomes more evident at $\sim$ 115 km along the profile, where a volcano-like structure is imaged. This feature might be a volcano due to the location of the profile, which crosses an area where there is supposed to be a volcanic arc. Moreover, the geometry of the volcano-like structure and irregular TOB boundary are only clearly imaged in depth when the PSDM is performed with the high-resolution V_P model provided by FWI (compare Figs. 15 and 14). In addition, the high-velocity layer within sediments on top of this structure is only clearly defined in the FWI V_P model (Fig. 10). This layer is likely to be an evaporite layer deposited during the Messinian crisis.

To assess the match of the velocity model to the seismic boundaries obtained with more tradition processing and imaging, the V_P model is converted to two-way-time and compared with the TMI. As state above, the two-way-time transformed velocity model has an excellent match with MCS time migration; velocity changes nicely follow major reflectivity contrasts and fault locations (Fig. 13), which further support the quality of the inversion result and workflow.

6 Geological structure

The seismic line was collected to image the structure of the eastern Alboran Basin in a region where basement is interpreted to be made of magmatic arc rocks (Booth-Rea et al., 2007, 2018; Gómez de la Peña et al., 2018). The time migration and PSDM show a sediment basin with basement highs on either end of the image. The large basement highs display in the flank some indications of extensional faulting and block tilting, particularly in the NE. These faults worked during the formation of the basin and may have some syntectonic sediment tilted in between rotated fault block, but the faults were not active during the deposit of most of the sediment cover.

The basement top under the basin and also along the southern ridge, displays several small-scale highs with steep flanks and a triangular shape with fairly symmetric flanks that do not support tectonic rotation and may indicate small-scale volcanoes. So that the image supports that the basement formed by the interplay between magmatic and extensional processes as expected in a magmatic arc.

The sediment cover displays several units bound by unconformities and cut by small sub-vertical, currently active faults. Based on regional geology, we infer that most of the sediment sequence is possibly Pliocene but the oldest layers may be late Miocene, although no drill hole in the vicinity provides a calibration. The seismic velocity distribution in the basin infill may provide some further clues on the age and nature of the sediment. The V_P model of the largest basin shows a gradual increase in velocity with depth to $V_P \sim 3$ km/s (change reddish to orange in Fig. 15a,c) that abruptly in underlain by ~ 0.5 km/s slower velocity body (change from orange to reddish in Fig. 15a,c). This high to low velocity change is marked by a high

amplitude reflection that marks an unconformity -possibly erosional- with the underlying unit. The high velocity is anomalous because clastic material -turbidites in mots of the basin- typically increases in V_P with compaction driven decrease in porosity. The high velocity layer may indicate the presence evaporites from the end Messinian period characterized across much of the Mediterranean by a relatively large drop down of the water layer and deposit of oversaturated deposits. In the deep basin further

5   to the east there are considerably thick salt deposits that be easily identified in seismic reflection images due to the mobility and formation of diapiric structures (Booth-Rea et al., 2007). Here, in the region of this study there are not known salt deposits and the presence of Messinian deposits without drilling information is unconstrained. A potential interpretation is that the high velocity layer represents an anhydrite rich layer that has relatively high velocities (like salt layers) but it is not as mobile.

7   Conclusions

10   We show that the proposed workflow including DC of MCS data to the seafloor, joint refraction and reflection TTT and FWI, allows obtaining a detailed, geologically meaningful V_P model from a non-optimal dataset. Although streamer length is of only 6 km, water depth is considerably large ($\sim$ 2 km), and seismic records lack low frequencies (<6 Hz), the inversion converged successfully providing a detailed velocity model to a depth of 3-4 km. The excellent match between the two-way-time transformed velocity model and the time-migrated MCS image validates the results. The PSDM images together with the

15   V_P model reproduce the structures and properties of the sedimentary layers and the uppermost part of the basement with high resolution directly in the depth domain. Aside from the irregular TOB, a volcano-like structure and steeply dipping anomalies at the flanks of the basin that may correspond to faults are clearly imaged. Moreover, the model allows identifying the location, geometry, thickness and velocity of a high-velocity layer (possibly evaporites on top of the Messinian unconformity) embedded within sediments. We conclude that the proposed method is valid and can be applied under these circumstances. Applying it

20   to longer streamer data should improve V_P models and allow extending the models to greater depths. Here, we recover all the possible subsurface information from the MCS recordings alone. Future research can be directed to apply this strategy to longer offset data from wide-angle acquisition geometries to increase the coverage and to better constrain the model properties.

References

Arnulf, A.F., Singh, S.C., Harding, A.J., Kent, G.M., and Crawford, W.C.: Strong seismic heterogeneity in layer 2A near hydrothermal vents at the Mid-Atlantic Ridge, Geophys. Res. Lett., 38, L13320, http://doi.org/ 10.1029/2011GL047753, 2011.

Arnulf, A.F., Harding, A.J., Kent, G.M., Singh, S.C., and Crawford, W.C.: Constraints on the shallow velocity structure of the Lucky Strike Volcano, Mid-Atlantic Ridge, from downward continued multichannel streamer data, J. Geophys. Res.-Sol. EA, 119, 1119-1144, http://doi.org/ 10.1002/2013JB010500, 2014.

Ballesteros, M., Rivera, J., Muñoz, A., Muñoz-martín, A., Acosta, J., Carbó, A., and Uchupi, E.: Alboran Basin, southern Spain-Part II: Neogene tectonic implications for the orogenic float model, Mar. Petrol Geol., 25, 1, 75-101, http://doi.org/10.1016/j.marpetgeo.2007.05.004, 2008.

Begovič, S., Meléndez, A., Ranero, C.R., and Sallarès, V.: Joint refraction and reflection travel-time tomography of multichannel and wide-angle seismic data, EGU General Assembly, Vienna, Austria, 23-28 April 2017, EGU2017-17231, 2017.

Berkhout, A.J.: Pushing the limits of seismic imaging. Part I: prestack migration in terms of double dynamic focusing, Geophysics, Soc. Expl. Geophys., 62, 3, 937-953, https://doi.org/10.1190/1.1444201, 1997a.

Berkhout, A.J.: Pushing the limits of seismic imaging. Part II: integration of prestack migration, velocity estimation, and AVO analysis, Geophysics, Soc. Expl. Geophys., 62, 3, 954-969, https://doi.org/10.1190/1.1444202, 1997b.

Berkhout, A.J.: Wave field extrapolation techniques in seismic migration, a tutorial, Geophysics, 46, 12, 1638-1656, https://doi.org/10.1190/1.1441172, 1981.

Berryhill, J.R.: Wave-equation datuming, Geophysics, 44, 1329-1344, https://doi.org/10.1190/1.1441010, 1979.

Berryhill, J.R.: Wave-equation datuming before stack, Geophysics, 49, 11, 2064-2066, https://doi.org/10.1190/1.1441620, 1984.

Booth-Rea, G., Ranero, C.R., Martínez-Martínez, J.M., and Grevemeyer, I.: Crustal types and Tertiary tectonic evolution of the Alborán sea, western Mediterranean, Geochem. Geophy. Geosy., 8, 10, 1525-2027, https://doi.org/10.1029/2007GC001639, 2007.

Booth-Rea, G., Ranero, C.R., and Grevemeyer, I.: The Alboran volcanic-arc modulated the Messinian faunal exchange and salinity crisis, Sci. Rep.-UK, 8, 13015, https://doi.org/10.1038/s41598-018-31307-7, 2018.

Brossier, R., Operto, S., and Virieux, J.: Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion, Geophysics, 74, 6, WCC105-WCC118, https://doi.org/10.1190/1.3215771, 2009.

Brossier, R., Operto, S., and Virieux, J.: Velocity model building from seismic reflection data by full-waveform inversion, Geophysical Prospecting EAGE, 1-14, http://doi.org/10.1111/1365-2478.12190, 2014.

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

Cho, Y., Ha, W., Kim, Y., Shin, S., and Park, E.: Laplace-Fourier-Domain Full Waveform Inversion of Deep-Sea Seismic Data Acquired with Limited Offsets, Pure Applied Geophysics, 173, 749-773, http://doi.org/10.1007/s00024-015-1125-7, 2016.

Claerbout, J.F.: Fundamentals of Geophysical Data Processing, McGraw-Hill Inc., New York, 1976.

Dagnino, D., Sallarès, V., and Ranero, C.R.: Scale- and parameter-adaptive model-based gradient pre-conditioner for elastic full-waveform inversion, Geophys. J. Int., 198, 1130-1142, http://doi.org/10.1093/gji/ggu175, 2014.

Dagnino, D., Sallarès, V., Biescas, B., and Ranero, C.R.: Fine-scale thermohaline ocean structure retrieved with 2-D prestack full-waveform inversion of multichannel seismic data: Application to the Gulf of Cadiz (SW Iberia). J. Geophys. Res.-Oceans, 121, 5452-5469, http://doi.org/10.1002/2016JC011844, 2016.

Gómez de la Peña, L., Gràcia, E., Muñoz, A., Acosta, J., Gómez-Ballesteros, M., Ranero, C.R., and Uchupi, E.: Geomorphology and Neogene tectonic evolution of the Palomares continental margin (western Mediterranean), Tectonophysics, 689, 25-39, http://doi.org/10.1016/j.tecto.2016.03.009, 2016.

Gómez de la Peña, L., Ranero, C., and Gràcia, E.: The crustal domains of the Alboran Basin (Western Mediterranean), Tectonics, 37, 3352-3377, https://doi.org/10.1029/2017TC004946, 2018.

Gràcia, E., Bartolome, R., Lo Iacono, C., Moreno, X., Stich, D., Martínez-Díaz, J. J., et al.: Acoustic and seismic imaging of the Adra Fault (NE Alboran Sea): in search of the source of the 1910 Adra earthquake, Nat. Hazard Earth Sys., 12, 11, 3255-3267, http://doi.org/10.5194/nhess-12-3255-2012, 2012.

Harding, A. J., Arnulf, A. F., and Blackman, D. K.: Velocity structure near IODP Hole U1309D, Atlantis Massif, from waveform inversion of streamer data and borehole measurements. Geochem. Geophy. Geosy., 17, 6, 1990-2014, https://doi.org/10.1002/2016GC006312, 2016.

Henig, A. S., Blackman, D. K., Harding, A. J., Canales, J. P., and Kent, G. M.: Downward continued multichannel seismic refraction analysis of Atlantis massif oceanic core complex, 30N, mid-Atlantic ridge. Geochem. Geophy. Geosy., 13, Q0AG07, https://doi.org/10.1029/2012GC004059, 2012.

Henig, A.S.: Seismic structure of shallow lithosphere at locations of distinctive seafloor spreading, ProQuest Dissertations And Theses, Ph.D. Thesis, University of California, San Diego, Publication Number: AAT 3558093, ISBN: 9781303023583, Source: Dissertation Abstracts International, 74-08(E), Section: B., 141, 2013.

Hicks, G.J., and Pratt, R.G.: Reflection waveform inversion using local descent methods: Estimating attenuation and velocity over a gas-sand deposit, Geophysics, 66, 2, 598-612, 2001.

Hobro, J.W.D., Singh, S.C., and Minshull, T.A.: Three-dimensional tomographic inversion of combined reflection and refraction seismic traveltime data, Geophys. J. Int., 152, 1, 79-93, https://doi.org/10.1046/j.1365-246X.2003.01822.x, 2003.

Jiménez Tejero, C., Dagnino, D., Sallarès, V., and Ranero, R.C.: Comparative study of objective functions to overcome noise and band-width limitations in full waveform inversion, Geophysical, J. Int., 203, 1, 632-645, http://doi.org/10.1093/gji/ggv288, 2015.

Korenaga, J., Holbrook, W.S., Kent, G.M., Kelemen, P.B., Detrick, R.S., Larsen, H.-C., Hopper, J.R., and Dahl-Jensen, T.: Crustal structure of the southeast Greenland margin from joint refraction and reflection seismic tomography, J. Geophys. Res., 105, B9, 21591-21614, 2000.

Lailly, P.: The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering, Theory and Application, Society for Industrial and Applied Mathematics, in edited by J. Bednar et al., 206-220, SIAM, Philadelphia, Penn, 1983.

Lambert, J.D.: Numerical Methods for Ordinary Differential Systems: The Initial Value Problem, John Wiley, N. Y., 1991.

Marjanovic, M., Plessix, R.-E., Stopin, A., and Singh, S.: Elastic versus acoustic 3-D Full Waveform Inversion at the East Pacific Rise 9°50'N, Geophys. J. Int., 216, 1497-1506, http://doi.org/10.1093/gji/ggy503, 2019.

Mavko, G., Mukerji, T., and Dvorkin, J.: The Rock Physics handbook, Cambridge University Press, 1998.

McMechan, G.A.: Determination of source parameters by wavefield extrapolation, Geophys. J. Roy. Astr. S., 71, 613-628, https://doi.org/10.1111/j.1365-246X.1982.tb02788.x, 1982.

McMechan, G.A.: Migration by extrapolation of time-dependent boundary values, Geophys. Prospect., 31, 413-420, https://doi.org/10.1111/j.1365-2478.1983.tb01060.x, 1983.

Meléndez, A., Korenaga, J., Sallarès, V., Miniussi, A., and Ranero, C.R.: TOMO3D: 3-D joint refraction and reflection traveltime tomography parallel code for active-source seismic data-synthetic test, Geophys. J. Int., 203, 158-174, http://doi.org/10.1093/gji/ggv292, 2015.

Morgan, J., Warner, M., Bell, R., Ashley, J., Barnes, D., Little, R., Roele, K., and Jones, C.: Next-generation seismic experiments: wide-angle, multi-azimuth, three-dimensional, full-waveform inversion, Geophys. J. Int., 195, 1657-1678, http://doi.org/10.1093/gji/ggt345, 2013.

Morgan, J., Warner, M., Arnoux, G., Hooft, E., Toomey, D., Brandon, V., and Wilcock, W.: Next-generation seismic experiments-II: wide-angle, multi-azimuth, 3-D, full-waveform inversion of sparse field data, Geophys. J. Int., 204, 1342-1363, http://doi.org/10.1093/gji/ggv513, 2016.

Moser, T.J.: Shortest path calculation of seismic rays, Geophysics, 56, 1, 59-67, https://doi.org/10.1190/1.1442958, 1991.

5   Moser, T.J., Nolet, G., and Snieder, R.: Ray bending revisited, Bull. Seismol. Soc. Am, 82, 1, 259-288, 1992.

Nocedal, J., and Stephen, J.W.: Numerical Optimization, American Mathematical Society, Springer, Providence, Rhode Island, USA, http://doi.org/10.1007/978-0-387-40065-5, 2006.

Paige, C.C., and Saunders, M.A.: LSQR: An algorithm for sparse linear equations and sparse least squares, ACM T. Math. Software, 8, 1, 43-71, http://doi.org/10.1145/355984.355989, 1982.

10 Pratt, R.G.: Seismic waveform inversion in the frequency domain, Part I: theory and verification in a physical scale model, Geophysics, 64, 3, 888-901, http://doi.org/10.1190/1.1444597, 1999.

Qin, Y., and Singh, S.C.: Detailed seismic velocity of the incoming subducting sediments in the 2004 great Sumatra earthquake rupture zone from full waveform inversion of long offset seismic data, American Geophysical Union, http://doi.org/10.1002/2016GL072175, 2017.

Qin, Y., and Singh, S.C.: Insight into frontal seismogenic zone in the Mentawai locked region from seismic full waveform inversion of
15 ultra-long offset streamer data, American Geophysical Union, http://doi.org/10.1029/2018GC007787, 2018.

Saragiotis, C.D., Hadjileontiadis, L.J., and Panas, S.M.: PAI-S/K: A robust automatic seismic P phase arrival identification scheme, IEEE Transactions on Geoscience and Remote Sensing, 40, 6, 1395-1404, https://doi.org/10.1109/TGRS.2002.800438, 2002.

Schuster, G.T., and Zhou, M.: A theoretical overview of model-based and correlation-based redatuming methods, Geophysics, 71, 4, SI103-SI110, http://doi.org/10.1190/1.2208967, 2006.

20 Shah, N., Warner, M., Nangoo, T., Umpleby, A., Stekl, I., Morgan, J., and Guasch, L.: Quality assured full-waveform inversion: Ensuring starting model adequacy, SEG Annual Meeting, Las Vegas, Nevada, 4-9 November 2012, https://dx.doi.org/10.1190/segam2012-1228.1, 2012.

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

25 Shtivelman, V., and Caning, A.: Datum correction by wave-equation extrapolation, Geophysics, 53, 10, 1311-1322, https://doi.org/10.1190/1.1442409, 1988.

Sirgue, L., and Pratt, R.G.: Efficient waveform inversion and imaging : a strategy for selecting temporal frequencies, Geophysics, 69, 1, 231-248, http://doi.org/10.1190/1.1649391, 2004.

Tarantola, A: Inversion of seismic reflection data in the acoustic approximation, Geophysics, 49, 8, 1259-1266,
30 http://doi.org/10.1190/1.1441754, 1984.

Tarantola, A.: Inverse problem theory: Methods for data fitting and model parameter estimation, Elsevier Science Publ. Co., Inc, New York, 1987.

Toomey, D.R., Solomon, S.C., and Purdy, G.M.: Tomographic imaging of the shallow crustal structure of the East Pacific Rise at 9°30'N, J. Geophys. Res., 99, 24135-24157, https://doi.org/10.1029/94JB01942, 1994.

35 Virieux, J.: P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method, Geophysics, 51, 4, 889-901, http://doi.org/10.1190/1.1442147, 1986.

Virieux, J., and Operto, S.: An overview of full-waveform inversion in exploration geophysics, Geophysics, 74, 6, WCC1-WCC26, http://doi.org/10.1190/1.3238367, 2009.

Vrolijk, J.-W., Haffinger, P., and Verschuur, E.: Multi-datum based estimation of near-surface full-waveform redatuming operators, J. Appl. Geophys., 82, 1, 30-45, http://doi.org/10.1016/j.jappgeo.2012.02.004, 2012.

Wapenaar, K., Thorbecke, J., van der Neut, J., Broggini, F., Slob, E., and Snieder, R.: Marchenko imaging, Geophysics, 79, 3, WA39-WA57, http://doi.org/10.1190/GEO2013-0302.1, 2014.

5 Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Stekl, I., Guasch, Ll., Win, C., Conroy, G., and Bertrand, A.: Anisotropic 3D full-waveform inversion, Geophysics, 78, 2, R59-R80, https://doi.org/10.1190/geo2012-0338.1, 2013.

Wessel, P., and Smith, W.: New version of the generic mapping tools, EOS, 76, 33, 329-336, https://doi.org/10.1029/95EO00198, 1995.

Zelt, C.A., and Smith, R. B.: Seismic traveltime inversion for 2-D crustal velocity structure, Geophys. J. Int., 108, 1, 16-34, https://doi.org/10.1111/j.1365-246X.1992.tb00836.x, 1992.

10 Zhang, W., and Shen, Y.: Unsplit complex frequency-shifted PML implementation using auxiliary differential equations for seismic wave modeling, Geophysics, 75, 4, T141-T154, http://doi.org/10.1190/1.3463431, 2010.

Table 1. Relevant inversion parameters used in the TTT.

Inversion parameters	Values
Grid dimension (long x deep)	8.25 x 5 km
Range of water layer depth	1.1-1.9 km
Node spacing (dx = dz)	25 m
Number of DC shotgathers picked	121
Distance between DC shots picked	~ 500 m
Maximum experimental offset	~ 6 km
Total of first arrival travel times	58,080
Total of reflected travel time	56,920
Forward star order (x,z) (node connexions)	(7,7)
Eliminate data outliers with chi values	> 15
Tolerance level to terminate inversion	0.001
Number of iterations	10
Velocity smoothing parameter (λ_v)	75
Depth smoothing parameter (λ_z)	10
Average velocity perturbation limit (%)	15
Average depth perturbation limit (%)	15
Top velocity smoothing correlation lengths (x,z) ($L_{H_{top}}$, $L_{V_{top}}$)	(0.6,0.05) (km)
Bottom velocity smoothing correlation lengths (x,z) ($L_{H_{bot}}$, $L_{V_{bot}}$)	(3,0.5) (km)
Horizontal smoothing correlation length for the reflector (x) (L_z)	1.5 km

Figure 1. General scheme of the workflow. We build the V_P model sequentially, from long- to small-scales, taking advantage of all the information contained in our data. We apply the different techniques referred in the text (DC + TTT + FWI + PSDM) to obtain a high-resolution image of the subsurface.
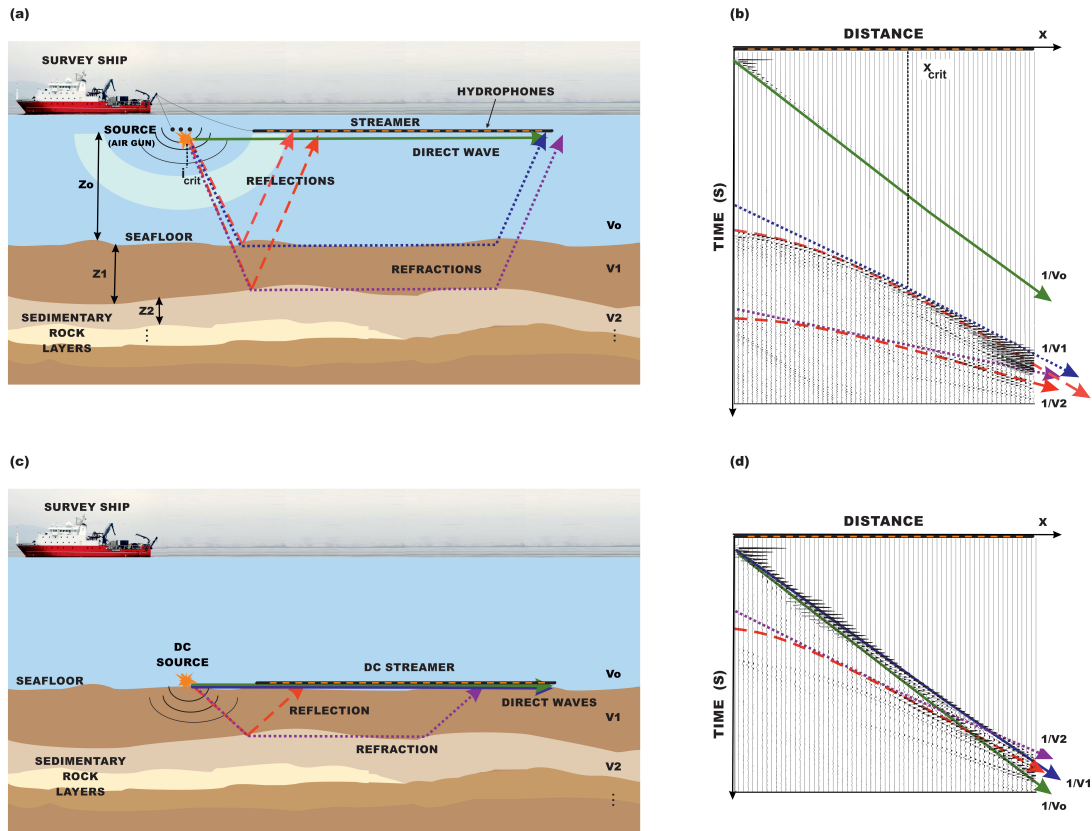
Figure 2. (a) Scheme of a streamer acquisition geometry of a marine MCS experiment showing the main element as well as the paths of the seismic phases that can be recorded in the streamer hydrophones. Where z_j $j = 0, 1, ...$ defines the depth of each layer, v_j $j = 0, 1, ...$ its V_P and i_{crit_j} $j = 0, 1, ...$ its critical angle. Green line is the direct water wave, blue and purple dotted lines are refracted ray paths in different interfaces, and its corresponding reflected ray paths are the pink and red dashed lines respectively. (b) Example of a synthetic shotgather recorded using a streamer-type geometry, as the one that is shown in (a). Each trace in the record section corresponds to a seismogram recorded at increasing source-receiver offset. Refractions are not visible, so they cannot be identified, and travel times cannot be used for velocity modeling. (c) Scheme of the virtual set up to be simulated after DC. In this case, both sources and receivers are located at the seafloor. Green line is the direct water wave, blue line is the head wave that travels through the seabed interface, dotted purple and red dashed lines are respectively the refracted and reflected ray paths of an interface of the subsurface. (d) Example of a synthetic shotgather recorded using the virtual geometry setup shown in (c). Refractions are visible as first arrivals from zero offset.

25

Figure 3. Relief map of the study area (Ballesteros et al., 2008; Gràcia et al., 2012; Gómez de la Peña et al., 2016). The field data used in the test corresponds to the white box along TM28 profile.
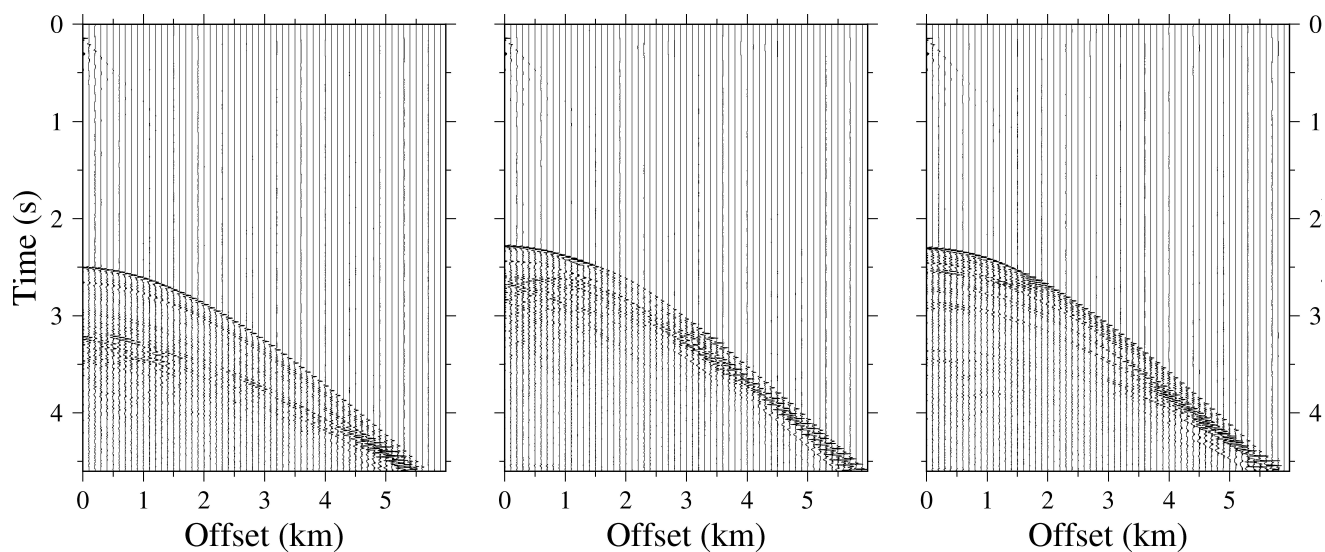
Figure 4. Decimated example of the field data set, only 10 traces each km are plotted (so 1/5 of the total) for clarity. From left to right, shot locations are at 171.1, 138.7 and 112.5 km along the profile.
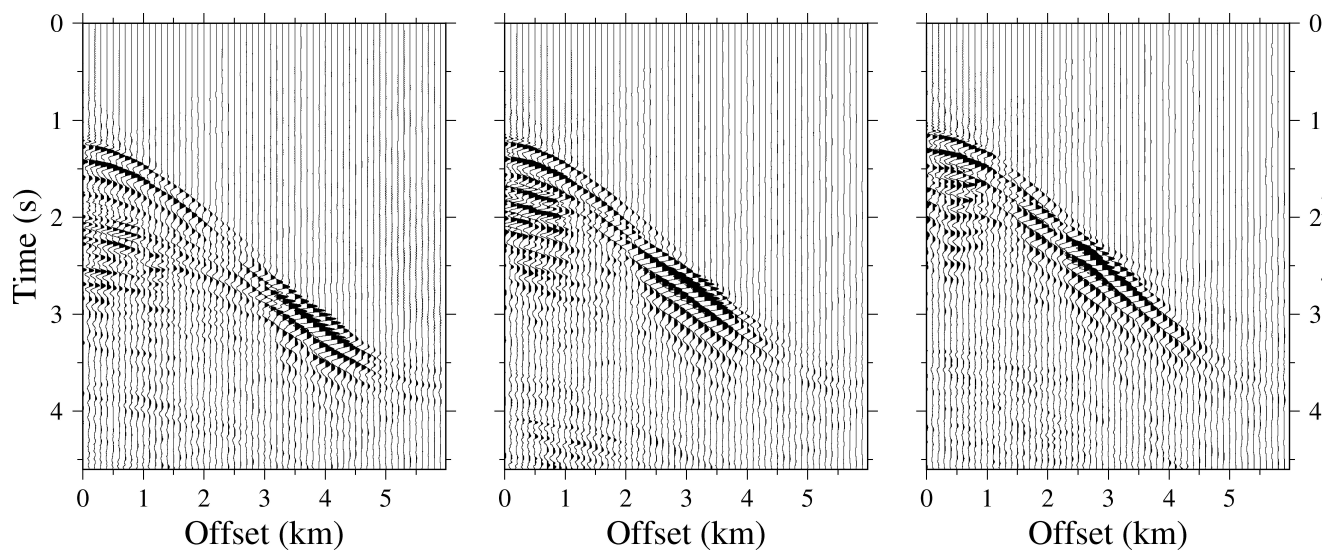

Figure 5. Seismic data obtained after the first DC step of the streamer shots recorded in the TOPOMED cruise. The wavefield are displayed for every fifth trace (so 10 traces each km, 1/5 of the total). From left to right, shot locations are at 171.1, 138.7 and 112.5 km along the profile.
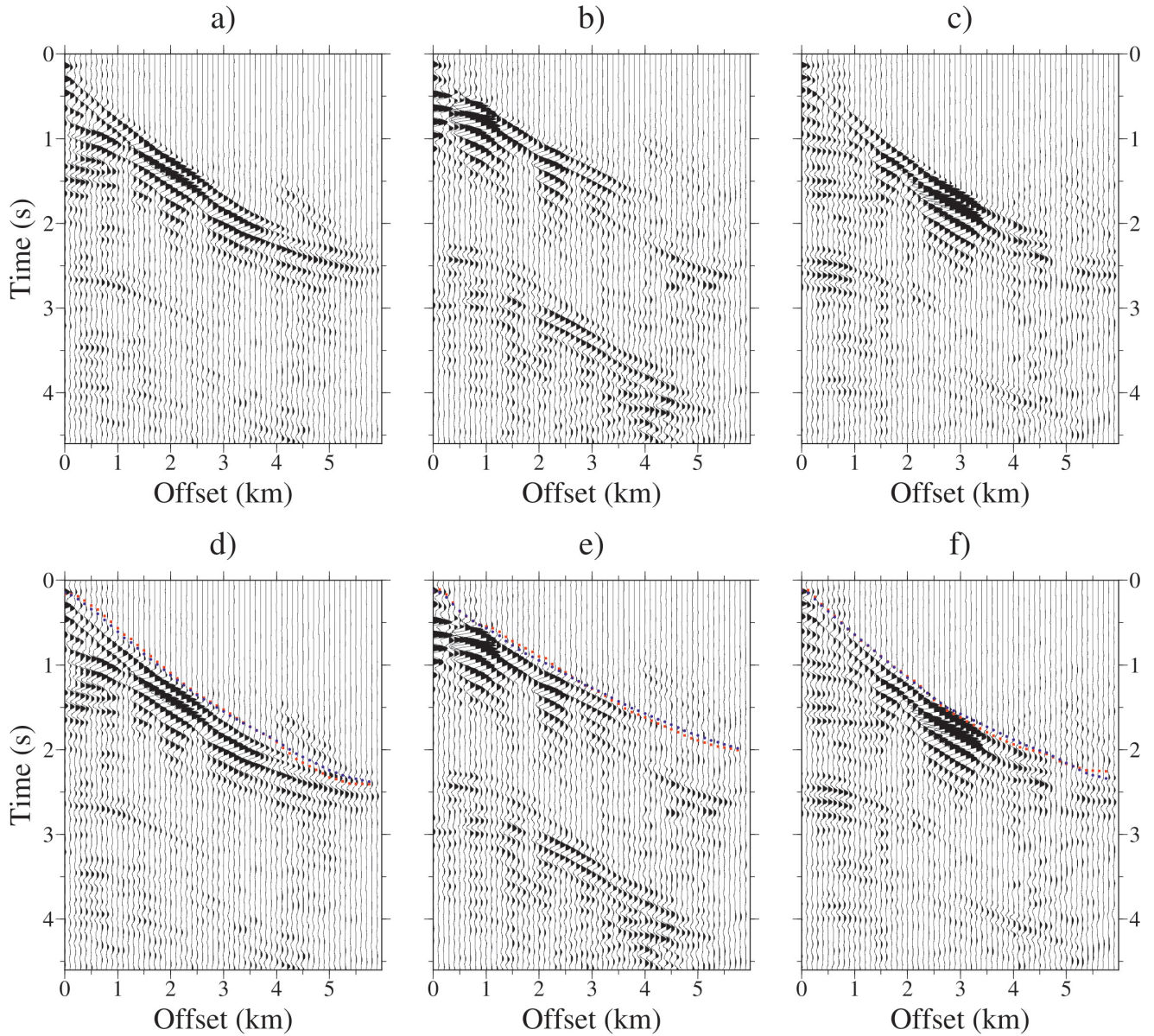
Figure 6. Seismic data obtained from the DC of the streamer shots recorded in the TOPOMED cruise. Only 10 traces each km are plotted (so 1/5 of the total) for clarity. From left to right, shot locations are at 171.1, 138.7 and 112.5 km along the profile. Lower panels (d)-e)-f)) show shots (a)-b)-c)) together with first arrivals used as input for the TTT (blue dots) and the first arrivals simulated with the TTT model (red squares).
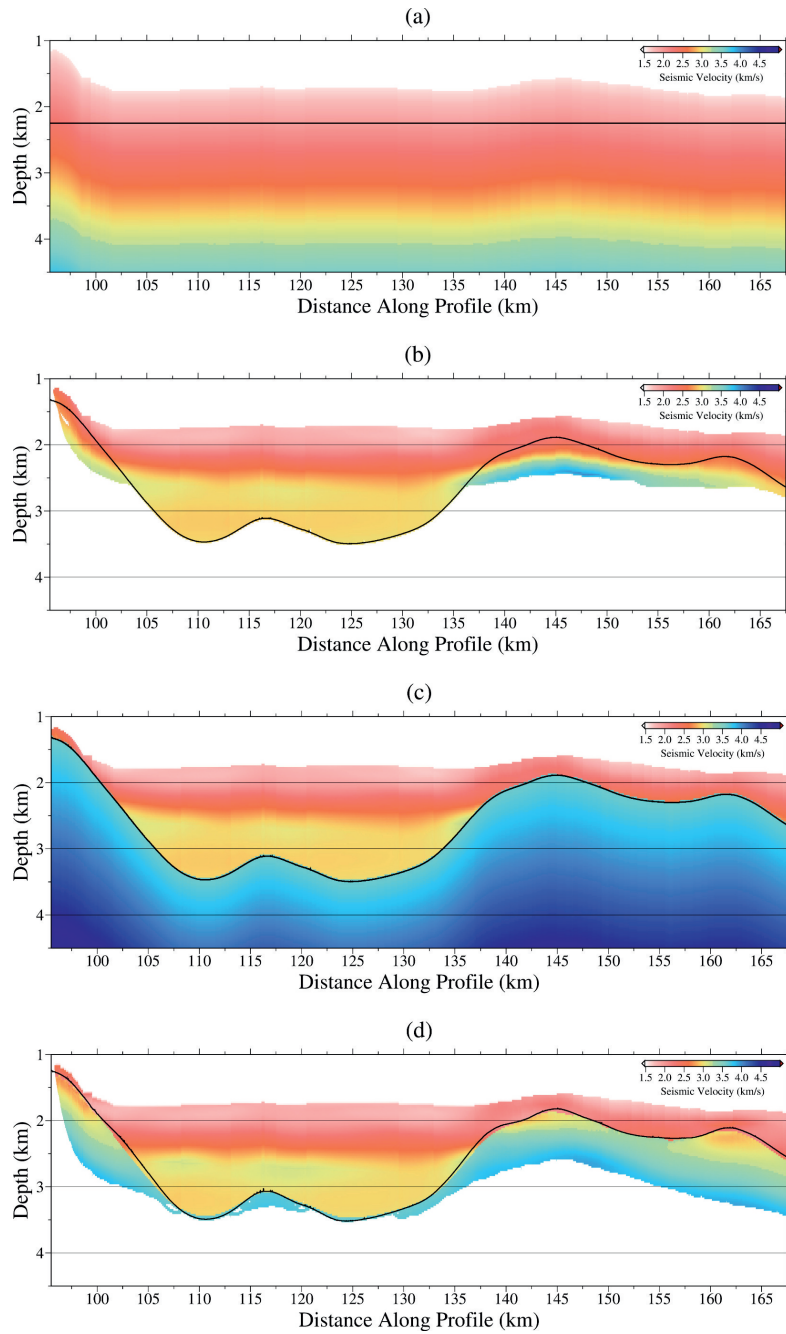
Figure 7. (a) Initial V_P and TOB models for the first inversion step. (b) Inverted V_P and TOB models after the first inversion step using (a) as initial ones. (c) Initial V_P and TOB models for the second inversion step. (d) Inverted V_P and TOB models after the second and final inversion step using (c) as initial ones. Both inversion steps use the same inversion parameters in Table 1, and the data set is made of the first arrivals of the DC shotgathers and the reflected phase of the TOB discontinuity identified in the original streamer records. Only the ray-covered areas of the model, where the derivative weight sum (DWS) is not zero, are plotted. The thick black line corresponds to the location of the TOB.
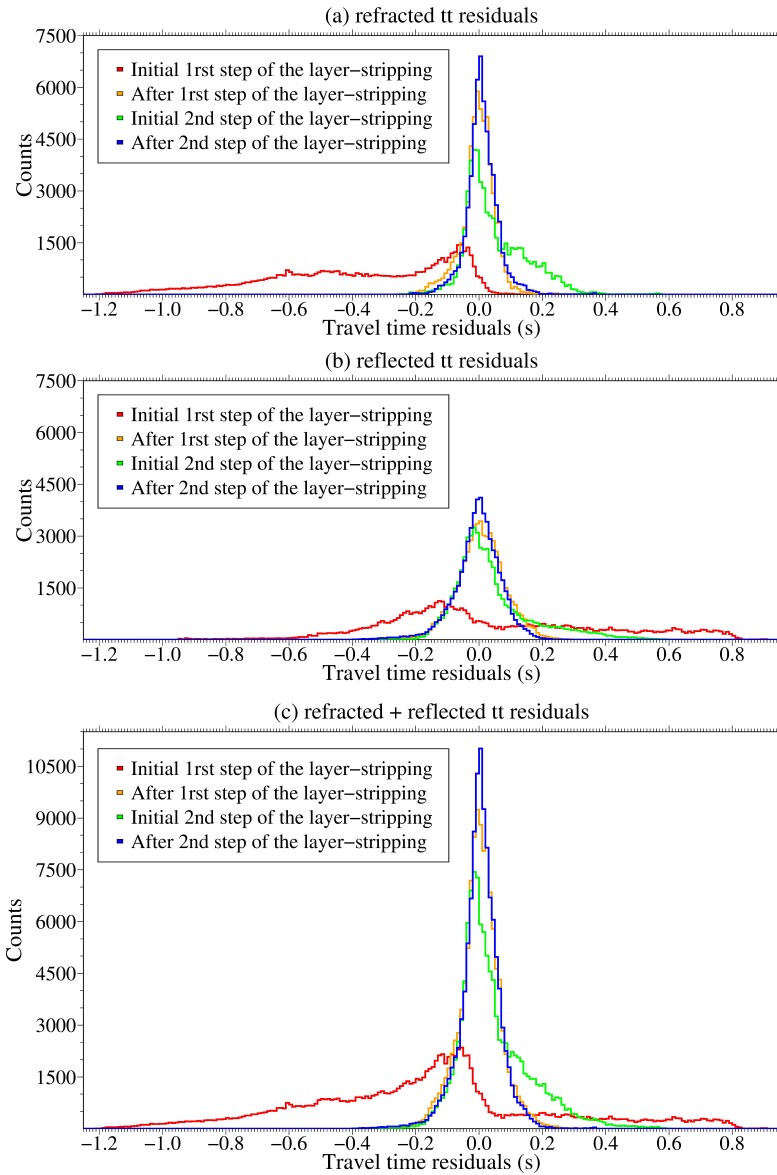
Figure 8. Histogram of travel-time residuals. (a) Upper panel shows only the residuals of the refractions obtained with the initial (red), after the first inversion step (orange), initial for the second inversion step (green) and final (blue) models. (b) Middle panel shows only the residuals of the reflections obtained with the initial (red), after the first inversion step (orange), initial fo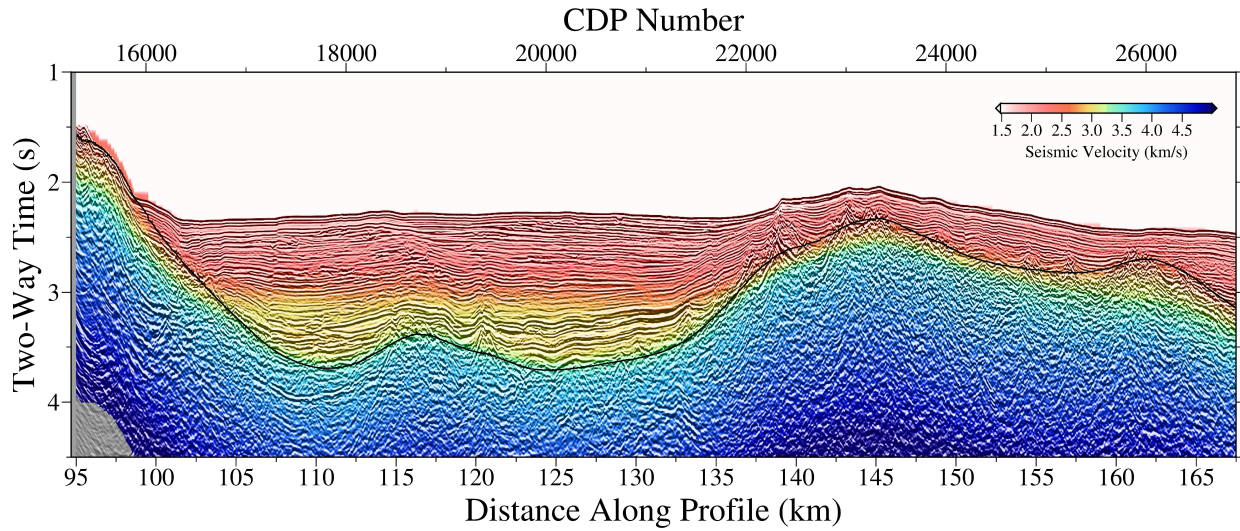r the second inversion step (green) and final (blue) models. (c) Lower panel shows the residuals of both phases (refractions and reflections) for the initial (red), after the first inversion step (orange), initial for the second inversion step (green) and final (blue) models.

Figure 9. Two-way-time transformed V_P and TOB models obtained after the joint DC refraction and streamer reflection TTT are shown superimposed on the time migrated MCS image.
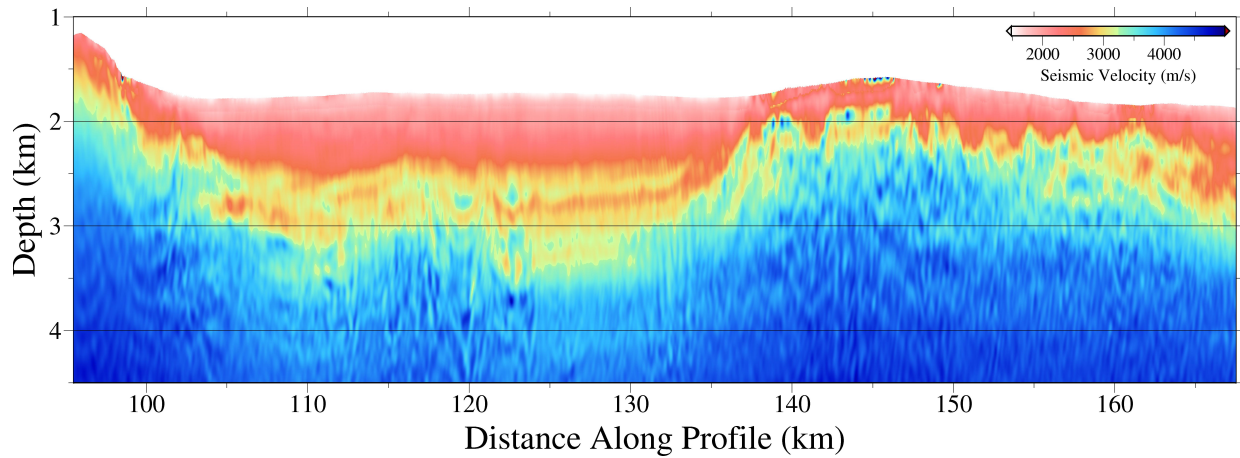
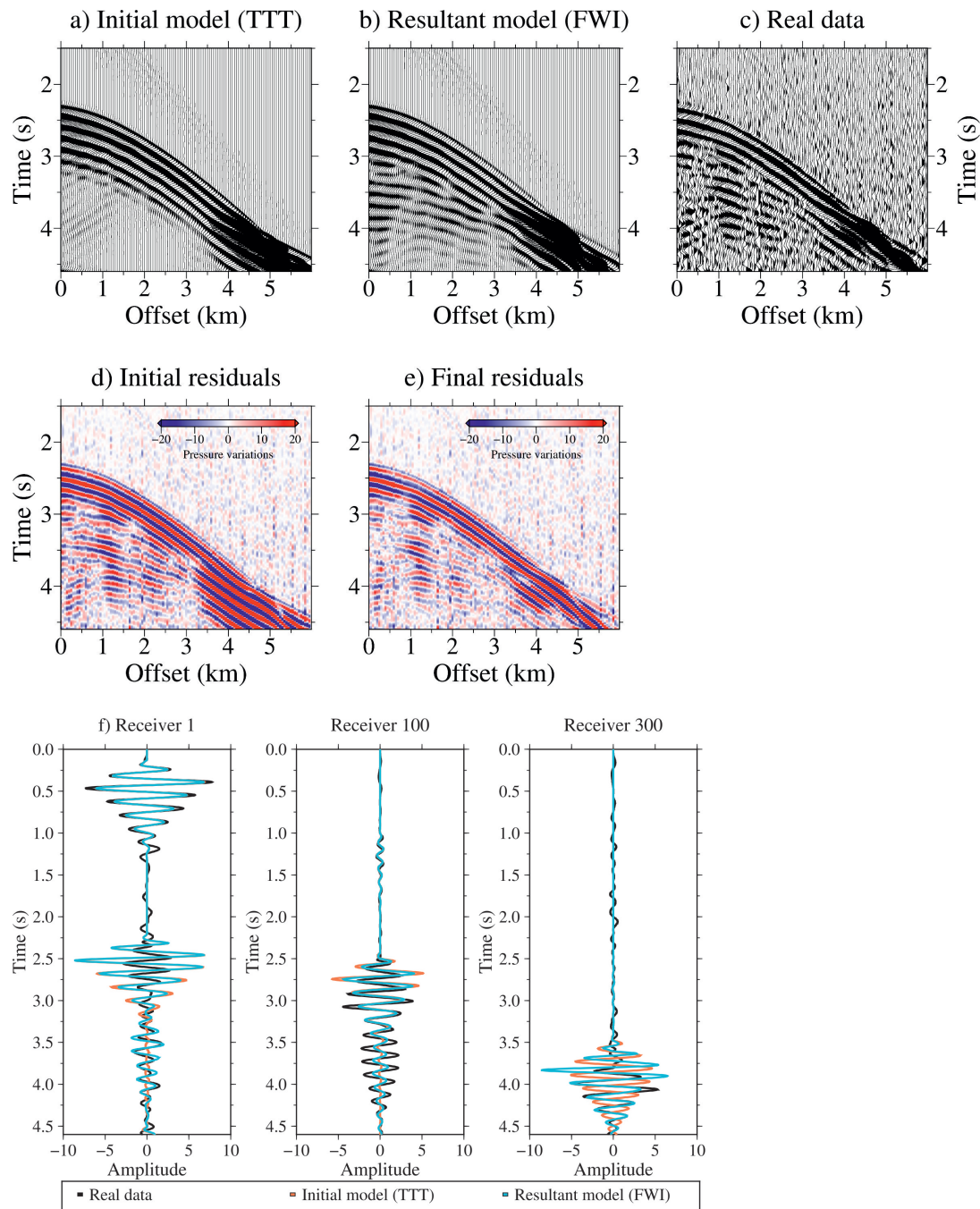**Figure 10.** Final velocity model after the FWI using Fig. 7d as initial model.

Figure 11. Streamer synthetic seismograms, which are computed using the initial (a) and resultant (b) velocity model presented in Fig. 7d and Fig. 10 and the FD solver of the FWI algorithm, together with the real data (c). Only every forth trace at 50 m is plotted. The location of the streamer is at a distance of ~ 101 to 107 km along profile. Initial (d) and final (e) residuals are displayed for a better comparison of the results, as well as (f) initial (orange, a)), resultant (blue, b)) and real (black, c)) seismic traces for receivers 1, 100 and 300.
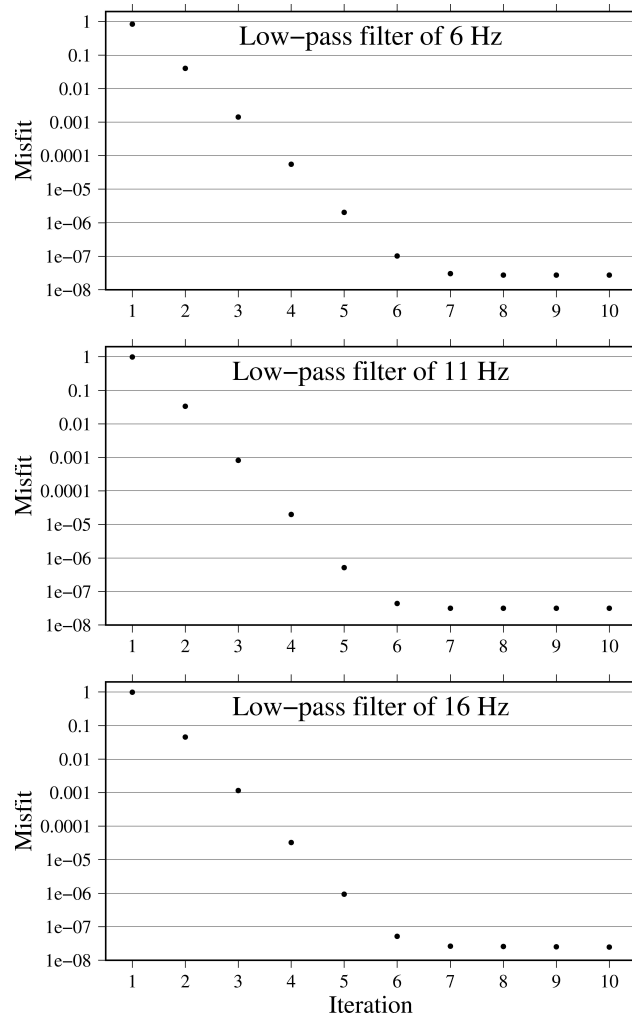
Figure 12. Misfit decrease plotted in a log axis along the iterations. The upper panel corresponds to the misfit reduction of the seismic information contained in the first multi-scaling step (low-pass filter of 6 Hz), the middle one for an intermediate step (low-pass filter of 11 Hz), and the lower panel for the final frequency band inverted (low-pass filter of 16 Hz).
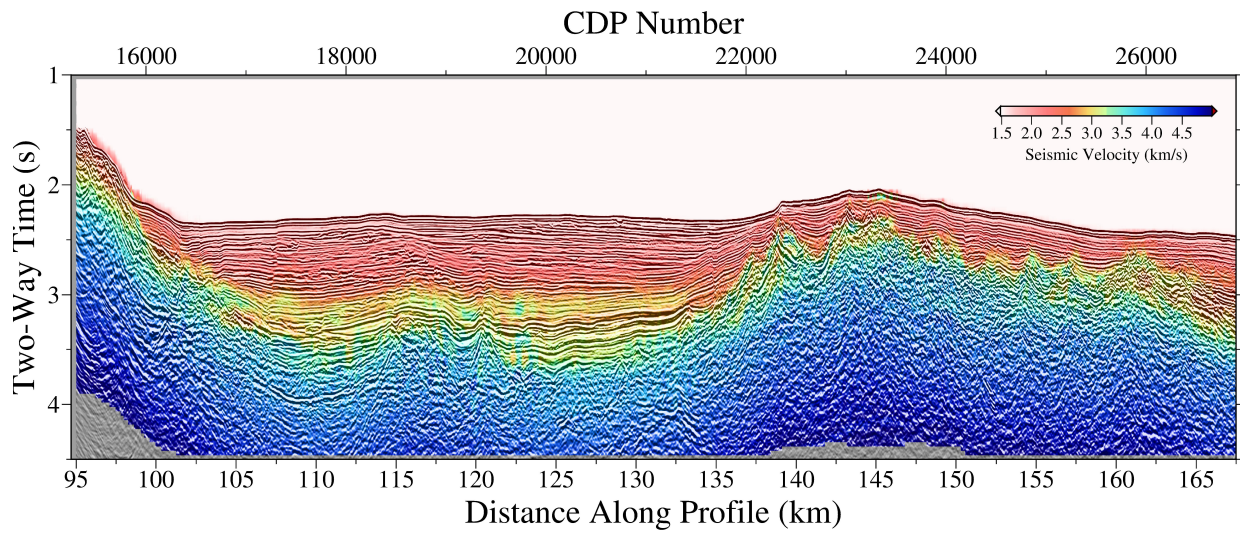
Figure 13. Two-way-time transformed V_P model obtained after the modeling sequence proposed on this paper (joint DC refraction and streamer reflection TTT + FWI) is shown superimposed to the time migrated MCS image.
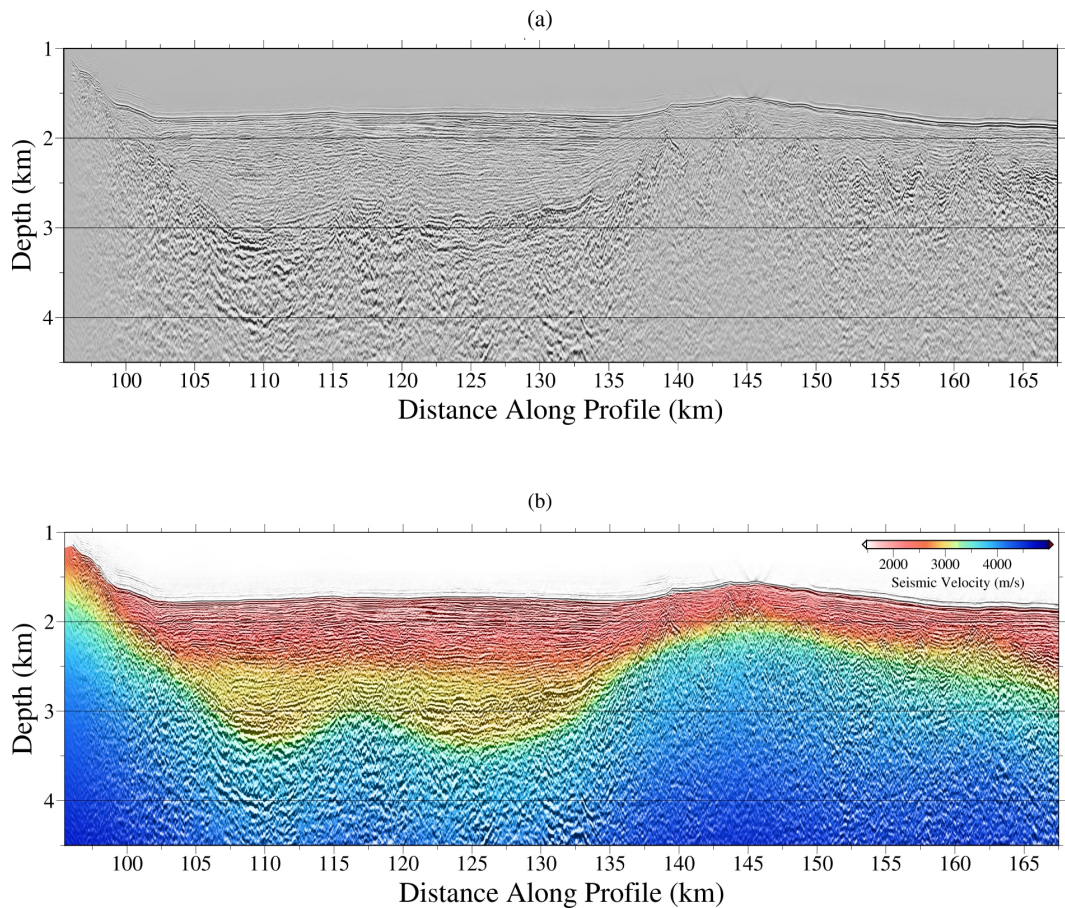
Figure 14. (a) Kirchhoff depth migration result based on the velocity model shown below (b) obtained after the joint DC refraction and streamer reflection TTT.
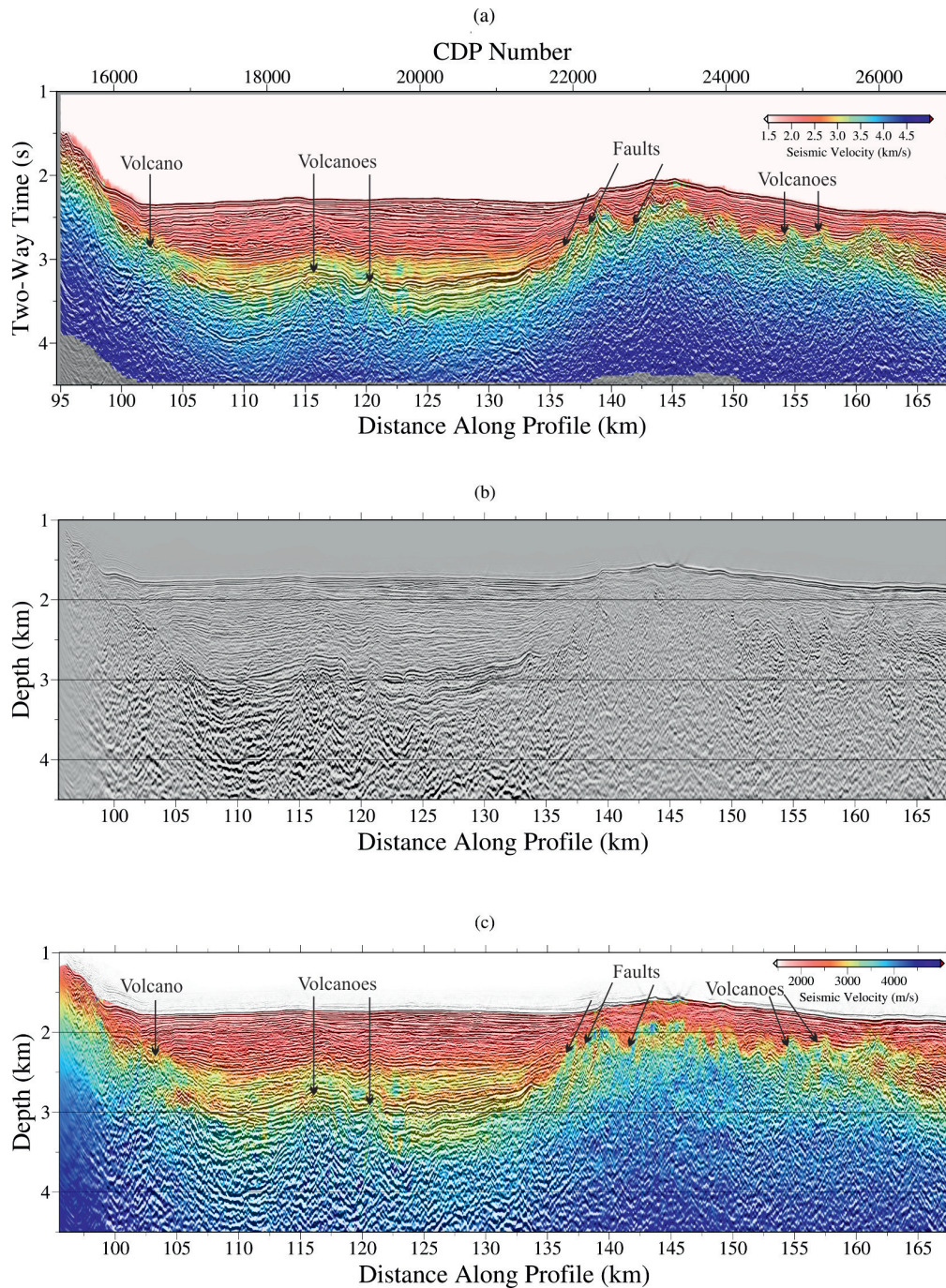
Figure 15. (a) Geological interpretation superimposed to Fig. 13. (b) Kirchhoff depth migration result based on the velocity model shown below (c) obtained after the modeling sequence proposed on this paper (joint DC refraction and streamer reflection TTT + FWI). Geological interpretation is also superimposed to (c).
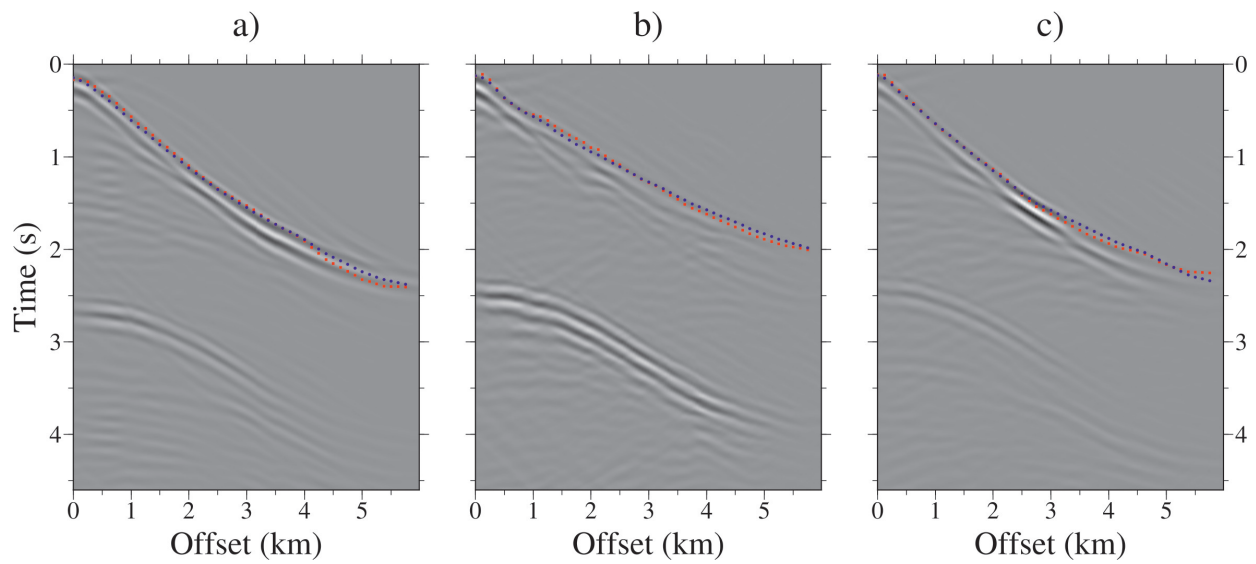
Figure 16. Panels a)-b)-c) show synthetic seismic data obtained from the DC of the streamer shots simulated using the final FWI V_P model. From left to right, shot locations are at 171.1, 138.7 and 112.5 km along the profile. First arrivals used as input for the TTT are plotted as blue dots and the ones simulated with the final TTT V_P model as red squares.