

Regional-scale FWI of wide-angle OBN data from a crude initial model using graph-space optimal transport

A. Górszczyk^{1,2*}, R. Brossier¹, L. Métivier^{1,3}

¹ Univ. Grenoble Alpes, ISTerre, F-38058 Grenoble, France

² Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

³ CNRS, Univ. Grenoble Alpes, LJK, F-38058 Grenoble, France

SUMMARY

Regional-scale imaging with full-waveform inversion of sparse long-offset ocean-bottom node data is inherently difficult. In particular, the large number of wavelengths that have to propagate within the volumetric model-space makes the inversion prone to cycle-skipping. While building an accurate initial model that predict the synthetic data with half-wavelet accuracy might be problematic due to the sparsity of the ocean-bottom nodes, solutions can be found from full-waveform inversion schemes based on alternative misfit functions. The recently developed graph-space optimal transport misfit function is designed for an improved convexity with respect to shifted patterns in the seismic signals - and therefore to the kinematic errors of the velocity model. In principles, this shall make it possible to successfully invert the data starting from a crude initial model. Here we show how using full-waveform inversion based on this misfit function, combined with multiscale data-selection strategy, we are able to make inversion converge from a simple 1D starting model. Despite the fact that the kinematic inaccuracy between the observed and the synthetic first-arrivals is reaching five cycles in the initial stage, we are able to bring the data in phase and obtain geologically consistent velocity model of the complex subduction zone. The presented approach can therefore significantly relax the constraint on the kinematic accuracy of the initial FWI model which is typically derived by traveltime tomography.

INTRODUCTION

Building an accurate initial velocity models for full-waveform inversion (FWI) from sparse ultra-long offset ocean-bottom node (OBN) data remains challenging. This is because the accumulation of the kinematic error along the diving and refracted wavepaths, additionally enriched with wide-angle reflections, makes it difficult to satisfy the cycle-skipping criteria (Pratt, 2008). This is problematic for FWI applications, because in the presence of large inaccuracies of the starting model, the classical L^2 -norm misfit function drives the inversion process towards non-informative local minima. Moreover, the wide range of various arrivals, combined with the sparsity of the receivers, increases the non-linearity of the inverse problem and rises the possibility of reconstruction of geologically meaningless model. Therefore, despite the fact that long-offset stationary-receiver seismic acquisitions have ability to provide low-frequency and multi-component data, the routine processing of these data with FWI remains challenging. On the other hand, previous sparse OBN studies (e.g. Kamei et al. (2012); Górszczyk et al. (2017); Davy et al. (2017)) have demonstrated a great potential to efficiently constrain the subsurface velocity at depths which are beyond the range of typical streamer acquisitions. In particular, wavefields traveling in the deep crust and upper mantle undershoot shallower structures, and provide illumination of the target from different perspectives than short-offset surface acquisition.

Natural solution to mitigate the cycle-skipping problem during FWI of long-offset OBN data can rely on building more accurate initial models (e.g. with slope tomography, Sambolian et al. (2018)), applying signal transformations which broaden the attraction valley of the L^2 -norm misfit function (e.g. envelopes, Bozdag et al. (2011)), or designing careful multiscale workflow (e.g. Górszczyk et al. (2017)). Recent approaches based on extended search space by relaxation of the wave-equation constraint (e.g. van Leeuwen and Herrmann (2013); Aghamiry et al. (2018)) also provide promising results in mitigating cycle-skipping. On the other hand, the development of more convex misfit functions has recently led to the design of optimal transport (OT)

based misfit measurement techniques. In particular, the graph-space optimal transport (GSOT) has provided promising results in terms of velocity model reconstruction (Métivier et al., 2019). Compared with standard L^2 -norm, GSOT is convex with respect to the patterns in the waveform which can be shifted in time for more than half-period. Therefore, coupled with a proper data-selection strategy, this misfit function has potential to reduce the risk of cycle-skipping at the initial FWI stage.

To test this hypothesis, here we apply a GSOT-based time-domain acoustic FWI workflow to a 2D wide-angle sparse OBN dataset acquired in the geologically challenging environment of the eastern Nankai Trough (Figure 1). We aim at reconstructing the complex velocity structure of the subduction zone starting from a simple 1D model. We compare our results with a Laplace-Fourier FWI study (Górszczyk et al., 2017) performed using the same dataset and a L^2 -norm misfit function, where significant efforts were devoted to derive an accurate initial velocity model. In the approach we present here, despite obvious cycle-skipping in the initial FWI model, the GSOT misfit function is still able to match the corresponding data-samples and converge towards a correct solution.

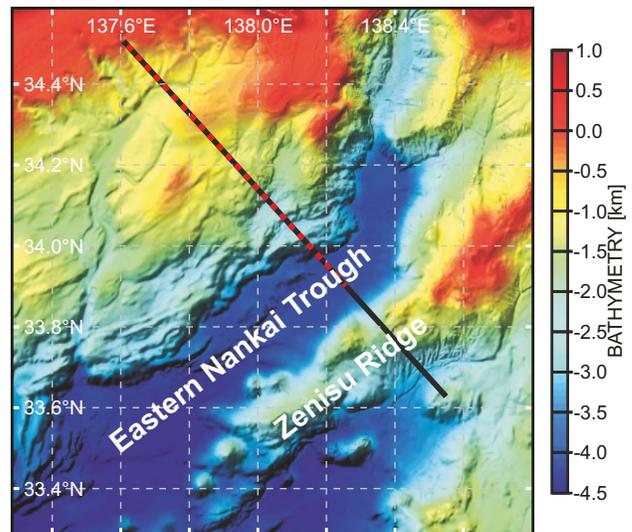


Figure 1: TKY-21 survey location in the eastern Nankai Trough. Colorscale represents bathymetry variations. The black and dotted red lines represent the shot and the receiver lines respectively.

GSOT FORMULATION

In this study we follow the GSOT strategy presented in Métivier et al. (2019). First we define a seismic trace $d(t)$ discretized as (d_1, \dots, d_N) . We denote its discrete graph by $(t, d) = ((t_1, d_1), \dots, (t_N, d_N)) \in (\mathcal{R}^2)^N$. Let now d_{cal} and d_{obs} be a calculated and observed trace respectively. With (t, d_{cal}) and (t, d_{obs}) we denote their discrete graphs consisting of N delta Dirac functions in 2D time/amplitude space. The GSOT distance between d_{cal} and d_{obs} is given by solving the linear assignment problem:

$$h_2(d_{cal}, d_{obs}) = \min_{\sigma \in \mathcal{S}(N)} \sum_{i=1}^N c_{i, \sigma(i)}(d_{cal}, d_{obs}), \quad (1)$$

Graph-Space Optimal Transport FWI of OBN data

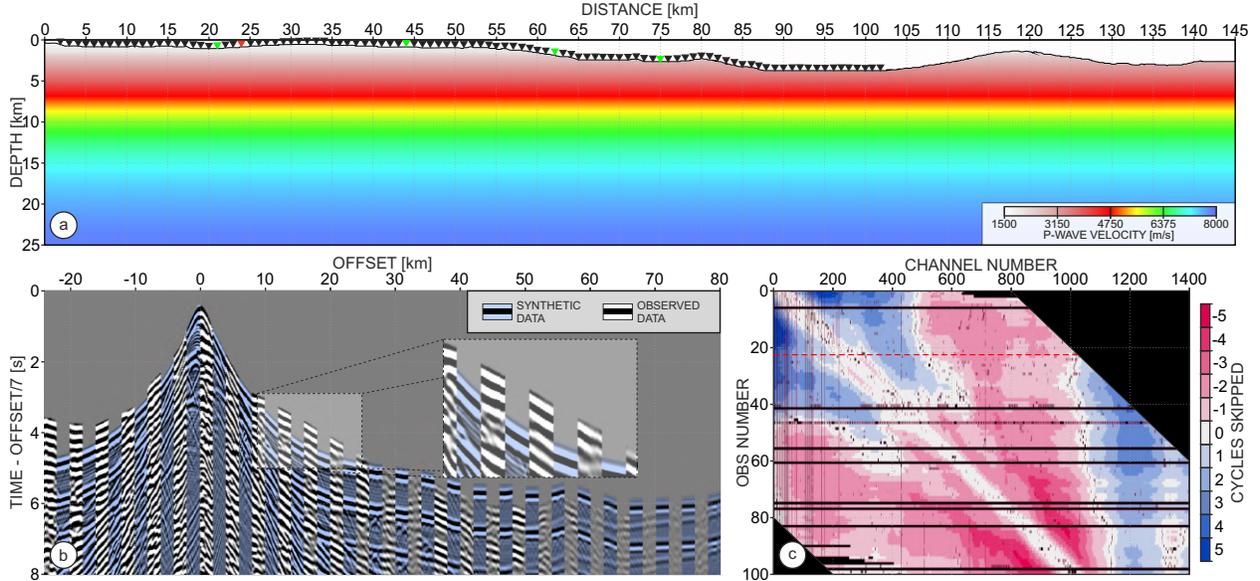


Figure 2: a) 1D initial FWI model. Triangles mark the OBN positions. Red triangle mark the position of the OBN for the gather in (b). Green triangles mark the positions of the OBNs in Figure 5. (b) Gather extracted from the OBN 23. Every 20 traces of the observed data (white and black phases) are interleaved with the following 20 traces of the synthetic data (light blue and black phases) generated in the initial FWI model from (a). Inset present the zoom on the heavily cycle-skipped data. (c) Map showing the number of skipped cycles for each trace in the dataset - arranged in the source-receiver coordinates. Blue/red color indicates that the first breaks of the synthetic data arrive later/earlier than their observed counterparts. Black color marks the traces excluded from the inversion.

	STAGE1	STAGE2	STAGE3	STAGE4
τ value	4.0 s	1.0 s	1.0 s	0.2 s
Time window	0.1 s + 0.5 s taper	0.1 s + 1.0 s taper	0.1 s + 2.0 s taper	0.1 s + 2.0 s taper
Amplitude	NO AVO	SQRT(AVO)	AVO	AVO
Smoothing	2.0×2.0	1.6×0.8	1.6×0.8	0.8×0.4
Iterations	80	40	40	140

Table 1: Summary of the inversion steps.

where $S(N)$ denotes the space of permutation of $\{1, \dots, N\}$, and c_{ij} is the distance between the points i and j of the discrete graph of d_{cal} and d_{obs} :

$$c_{ij}(d_{cal}, d_{obs}) = |t_i - t_j|^2 + \left| \frac{\tau}{A} (d_{cal,i} - d_{obs,j}) \right|^2. \quad (2)$$

In Equation 2, $\frac{\tau}{A}$ is a factor rescaling properly the amplitude values with respect to time-shift values. A is the maximum peak amplitude difference (calculated separately for each pair of the observed and calculated seismograms), while τ is the maximum estimated time shift between d_{obs} and d_{cal} . This ensures the convexity of the GSOT distance for time up to around τ . With increasing τ , the attraction valley of the misfit function becomes wider, while with the τ approaching to zero it becomes more narrow and the GSOT converges towards the L^2 -norm. This linear assignment problem is efficiently solved using the auction algorithm (Bertsekas and Castanon, 1989). The final cost-function we use for the purpose of FWI application with N_s shots containing N_r receivers is defined as:

$$\min_m f[m] = \sum_{s=1}^{N_s} \sum_{r=1}^{N_r} w^{s,r} h_2(d_{cal}^{s,r}[m], d_{obs}^{s,r}), \quad (3)$$

where $d_{obs}^{s,r}$ and $d_{cal}^{s,r}[m]$ are the observed and synthetic (calculated in model m) traces respectively associated with source s and receiver r . The coefficients $w^{s,r}$ correspond to a trace-by-trace weighting, typically used to restore the AVO trend in the data, which is removed prior to the calculation of the GSOT cost-function due to the normalization by the factor $\frac{\tau}{A}$.

FIELD DATA

The TKY-21 dataset was acquired in the by JAMSTEC in 2001 during the KY0106 cruise of R/V Kaiyo. The experiment was conducted

within the framework of the academic French-Japanese project SFJ-OBS, aiming on the deep crustal imaging of the Tokai area (eastern Nankai Trough, Figure 1). The acquisition comprises 100 OBNs (4.5-Hz three-component geophones and hydrophones) deployed with 1 km intervals and 1404 air-gun shots (array volume of 196 l) spaced 100 m apart. This shot/receiver spatial sampling of the target is sufficiently redundant for FWI application. We pre-process the raw field data starting from removing the DC component and scaling the data-samples by square root of time (3D to 2D amplitude-spreading correction). We apply a spectral whitening (due to the strong bubble effect) followed by a bandpass filter - 1.5 Hz to 3.5 Hz. Finally, we normalize the OBN gathers with their root-mean-square amplitude, to compensate for differences in the instrument response and seabed coupling.

WORKFLOW

We apply a time-domain visco-acoustic FWI workflow, relying on our visco-acoustic full-waveform modeling and inversion solver TOYX-DAC.TIME. The intrinsic attenuation mechanism is based on the generalized Maxwell body. The crude Q model consist of the fluid and solid part for which the Q -factor equals 10000 and 200 respectively. The density model is assumed to be homogeneous (1000 kg/m³). The inversion is based on the l -BFGS local optimization scheme (Nocedal, 1980). The gradient is computed in the time-domain using the checkpoint-assisted reverse forward simulation (CARFS) strategy (Yang et al., 2016).

Our 4-stages FWI scheme is summarized in Table 1. It incorporates progressive reduction of the gradient regularization and simultaneous extension of the time windows. The regularization is performed as a non-stationary Gaussian smoothing of the gradient where the correlation lengths are based on an estimation of the local wavelength. The

Graph-Space Optimal Transport FWI of OBN data

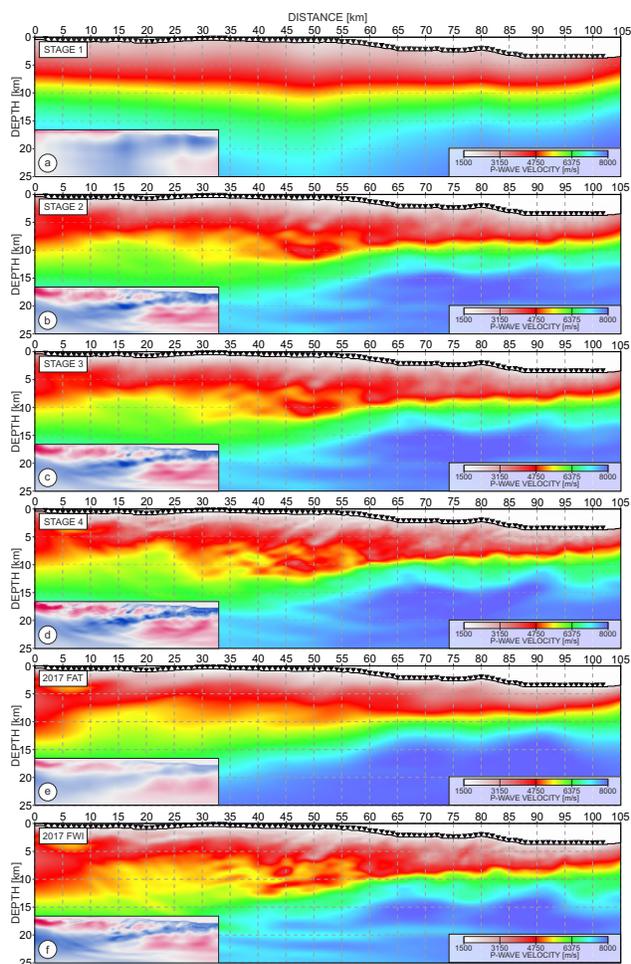


Figure 3: a)-(d) Model evolution after each FWI stage. (e)-(f) FAT and FWI models derived by Górszczyk et al. (2017). Insets show the difference between the presented and the initial model from Figure 2a. The color-scale is ± 2000 m/s.

time window of length of 0.1 s starts at the first-arrival time and is extended from one FWI stage to another with progressively longer tapers. In Table 1, the time window which we use at the initial FWI stages is narrow and focused mainly around the first arrivals. Consequently, the volume of the data which is actually compared with the GSOT misfit function is reduced. Muting of the secondary arrivals and wide-angle reflections from the observed data reduces the risk of matching them with the first arrivals from the synthetic data. Such a mismatch may likely occur taking into account the large time shifts between the observed and initial data in Figure 2c. We use a maximum time window of ~ 2 s after the first arrival on the final stage of the workflow to limit the amount of elastic energy in the data (mainly from near and intermediate offsets), as we rely on a visco-acoustic inversion.

In the GSOT misfit function, τ is the crucial parameter to tune. It controls the amount of the kinematic mismatch that might be handled by the GSOT. As the model accuracy increases from one FWI stage to another, we reduce the value of τ . This makes the slopes of the attraction valley of GSOT misfit function steeper, translating to an improved convergence. At the same time the global minimum becomes sharper reducing the null space, which might result not only from the presence of local minima, but also from various factors such as presence of noise, visco-acoustic approximation of the wave propagation, simplified subsurface parametrization, or imperfect source wavelet.

As mentioned, introducing the normalization factor $\frac{\tau}{A}$ in Equation 2 causes loss of the amplitude versus offset trend (AVO) in the misfit function. This trend can be restored with appropriate $w^{s,r}$ weighting.

At the initial FWI stage, we use constant $w^{s,r}$ equal to 1; namely, we compare trace-normalized seismograms to increase the contribution of the far-offset data in the gradient. Those data correspond to the waves penetrating in the deeper subsurface and have typically much smaller amplitudes than the near-offset data. Trace normalization at the initial FWI stage mitigates this issue at the price of AVO information loss. In stage 2, we introduce an intermediate misfit function weighting based on the square root of the AVO. This is due to the large amplitude variations between near- and far-offset traces (several orders of magnitude) observed in this dataset. In stage 3 and 4 the $w^{s,r}$ is calculated as a trace-by-trace root mean square (RMS) value of the observed seismogram to restore the AVO trend in our misfit function. To re-estimate the maximum peak amplitude difference between the observed and calculated data (A in Equation 2), we restart the inversion after every 10 FWI iterations. We run the inversion until no significant model updates are observed. We pre-estimate one source wavelet for all OBN gathers from the short-offset water wave (Pratt, 1999), and we keep it the same for all FWI stages.

RESULTS

Figure 3a-d and Figure 4 illustrate, respectively, the model and the data evolution after each FWI stage. During stage 1, we recover long-wavelength positive (red) and negative (blue) perturbations (see the inset in Figure 3a), which leads to a smooth model exhibiting the general trend of the subduction zone. However, the synthetic data in Figure 4a (blue-shaded traces) are still locally significantly cycle-skipped. This mismatch results most likely from the strong smoothing of the gradient at this stage, which hampers the intermediate- and small-scale perturbations needed to explain the data more precisely. Therefore, in stage

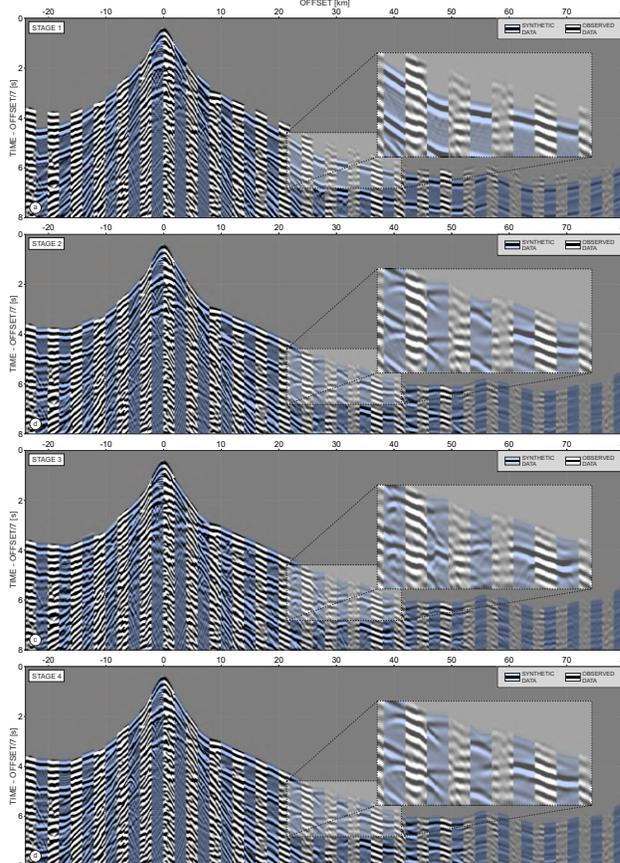


Figure 4: (a)-(d) Data fitting evolution (OBN 23, Figure 2b) after FWI stages 1-4 respectively. Every 20 traces of the observed data are interleaved with the following 20 traces of the synthetic data (blue-shaded traces). Insets show the zoom on the complex waveform package.

Graph-Space Optimal Transport FWI of OBN data

2, we use a smaller τ value to improve the convergence of inversion, and simultaneously, we increase the resolution of the introduced perturbation by means of reduced smoothing and a slightly extended time window. Importantly, we switch to a data weighting according to the square root of the AVO. In this way, we put more weight into the energetic short- and intermediate-offset data that carry rich information about the underlying geological features. It is clear that after stage 2, the reconstructed model (Figure 3b) has much higher resolution. One can observe the signature of the complex structures building the accretionary prism. The shape of the oceanic crust and the Moho is also reconstructed.

This increase in the model resolution is reflected by the improved data fitting presented in Figure 4b. There is no clear evidence of cycle-skipping within the first arrivals which indicates a significant improvement of the model with respect to the previous stage. In the inset, one can observe that the synthetic data contain a complex package of wide-angle reflections - although their kinematics and dynamics are not precisely reconstructed yet. The model after stage 3 (Figure 3c) shows mostly improvement of the resolution of the shallow and intermediate structures. This result is due to the weighting of the misfit function applied according to the restored AVO trend and extended time-window. At this stage, we further improve the continuity of the phases between synthetic and field data panels (also for the later arrivals), as seen in Figure 4c. The final model is presented in Figure 3d. A smaller τ value and a smaller smoothing make it possible to sharpen the structure within the accretionary prism. Shallow sedimentary basins, as well as sequences of thrusts of various scales are clearly visible now. The characteristic undulations of the subducting oceanic crust in the Tokai region, coupled with a wavy nature of the underlying Moho are also made evident. The final velocity perturbations presented in the inset exhibit a wide range of introduced structures varying both in terms of spatial scale and magnitude. In Figure 4d one can observe further improvement of the continuity of phases and amplitude trends between synthetic and field data traces.

In Figure 3e we show the initial model used during Laplace-Fourier FWI study of Górszczyk et al. (2017). Note that it exhibits sharper velocity structure than the results we obtained with our GSOT-based FWI after stage 1 (Figure 3a). This initial model was derived through the first arrivals tomography (FAT) conducted in few stages. After each stage the first-break picks were refined according to resulting travel-time error and the FAT was calculated with the new set of updated picks. This back-and-forward procedure yields a model which satisfies the cycle-skipping criterion for a subsequent L^2 -norm based FWI. However, the first-breaks picking requires a lot of human interaction from the interpreter - especially in the area of lack of coherence of the first arrivals and noisy intervals of traces. This subjective picking, combined with the sparsity of the OBN deployment, can lead to difficulties with obtaining a regional-scale initial model which produces synthetic wavefield with half a wavelet accuracy for all picked first arrivals. As we show here, with GSOT misfit-function, we are able to drastically reduce the efforts for initial model building. Starting from a rough 1D model, we converge to a solution which in good agreement with the FWI model from Górszczyk et al. (2017) (Figure 3f) despite notably different parametrization and modeling engines.

To obtain a better insight into the final data fitting in Figure 5, we present another 4 observed/synthetic OBN gathers recorded at different locations along the profile (green triangles in Figure 3). The presented waveforms cover a broad range of arrivals that might be recorded in the subduction zone environment. One can observe the good continuity of the phases not only in terms of the first arrivals but also in terms of later wide-angle reflections and free-surface multiples appearing at various offsets. The amplitude trends of the synthetic data also follow those from field data - although the visco-acoustic approximation of our FWI causes limited amplitude reconstruction in particular for later arrivals. It is noteworthy here that for our inversion, we use only the data from the 2 s time window starting at the first arrival. Such a time window makes it possible to incorporate the majority of the P-wave arrivals constraining the velocity model, and at the same time, it limits the amount of waveforms related to elastic effects that

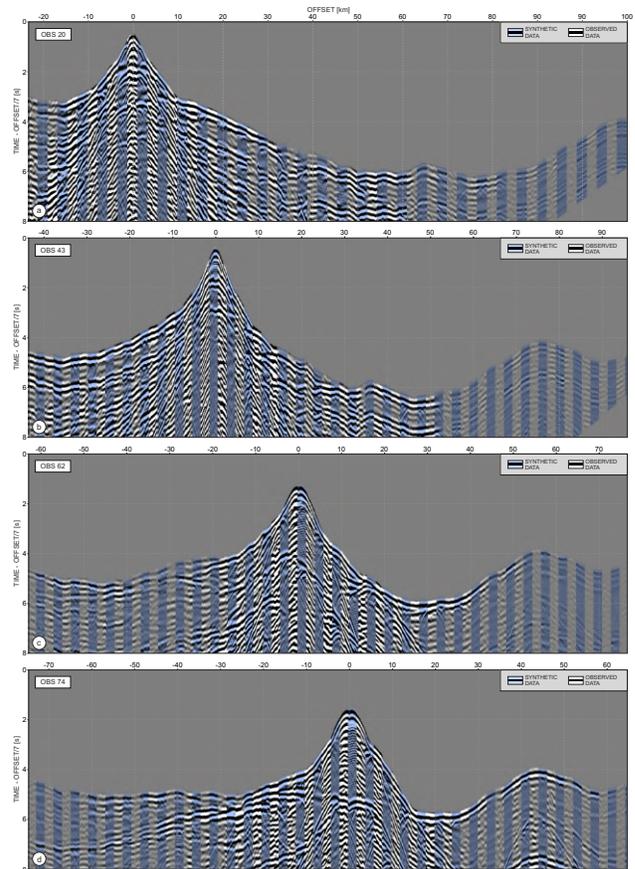


Figure 5: (a)-(d) Comparison of data fitting after stage 4 of FWI - the OBS positions are marked by the green triangles in Figure 2a. Every 20 traces of the observed data are interleaved with the following 20 traces of the synthetic data (blue-shaded traces).

we cannot handle due to the simplified approximation of the physics we use here.

CONCLUSION

The FWI workflow that we present exhibits significantly less dependency on the accuracy of the initial velocity model. The GSOT misfit function, combined with a progressive input data selection makes it possible to maintain the convergence of the inversion even when the initial data are shifted for more than few cycles with respect to the observed data. We believe that further applications of OT-based misfit functions can significantly reduce the overall risk of cycle-skipping during FWI of long-offset OBN data. Starting from a 1D model here is to be understood as a proof of concept in a “worst case” situation. We can envision future regional-scale OBN studies where one could build rapidly initial velocity models from traveltime tomography. The obtained velocity model, even if it were not to satisfy the cycle-skipping criteria, could serve as a initial velocity model for a GSOT-based FWI. In such a way, one could start FWI with a better initial model than what we presented here, still taking advantage from the improved convexity of the GSOT misfit function.

ACKNOWLEDGEMENTS

This study was partially funded: (i) by the SEISCOPE consortium (<http://seiscope2.osug.fr>), sponsored by AKERBP, CGG, CHEVRON, EQUINOR, EXXON-MOBIL, JGI, SHELL, SINOPEC, SISPROBE, and TOTAL; (ii) the Polish National Science Center, (grant no: 2019/33/B/ST10/01014). The study was granted access to the HPC PL-Grid Infrastructure (grant id: 3dwind2).

REFERENCES

- Aghamiry, H., A. Gholami, and S. Operto, 2018, Improving full-waveform inversion by wavefield reconstruction with alternating direction method of multipliers: *Geophysics*, **84**, no. 1, R125–R148.
- Bertsekas, D. P., and D. Castanon, 1989, The auction algorithm for the transportation problem: *Annals of Operations Research*, **20**, 67–96, doi: <https://doi.org/10.1007/BF02216923>.
- Bozdog, E., J. Trampert, and J. Tromp, 2011, Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements: *Geophysical Journal International*, **185**, 845–870.
- Davy, R. G., J. V. Morgan, T. A. Minshull, G. Bayrakci, J. M. Bull, D. Klaeschen, T. J. Reston, D. S. Sawyer, G. Lymer, and D. Cresswell, 2017, Resolving the fine-scale velocity structure of continental hyperextension at the Deep Galicia Margin using full-waveform inversion: *Geophysical Journal International*, **212**, 244–263, doi: <https://doi.org/10.1093/gji/ggx415>.
- Gorszczyk, A., S. Operto, and M. Malinowski, 2017, Toward a robust workflow for deep crustal imaging by FWI of OBS data: The eastern Nankai Trough revisited: *Journal of Geophysical Research: Solid Earth*, **122**, 4601–4630.
- Kamei, R., R. G. Pratt, and T. Tsuji, 2012, Waveform tomography imaging of a megasplay fault system in the seismogenic Nankai subduction zone: *Earth and Planetary Science Letters*, **317–318**, 343–353, doi: <https://doi.org/10.1016/j.epsl.2011.10.042>.
- Metivier, L., R. Brossier, Q. Merigot, and E. Oudet, 2019, A graph space optimal transport distance as a generalization of Lp distances: Application to a seismic imaging inverse problem: *Inverse Problems*, **35**, 085001.
- Nocedal, J., 1980, Updating Quasi-Newton matrices with limited storage: *Mathematics of Computation*, **35**, 773–773, doi: <https://doi.org/10.1090/S0025-5718-1980-0572855-7>.
- Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain — Part 1: Theory and verification in a physical scale model: *Geophysics*, **64**, 888–901, doi: <https://doi.org/10.1190/1.1444597>.
- Pratt, R. G., 2008, Waveform tomography — Successes, cautionary tales, and future directions: 70th Annual International Conference and Exhibition, EAGE — Full-Waveform Inversion: Current Status and Perspectives, Extended Abstracts, doi: <https://doi.org/10.3997/2214-4609.201405056>.
- Sambolian, S., S. Operto, A. Ribodetti, F. B. Tavakoli, and J. Virieux, 2018, Parsimonious slope tomography based on eikonal solvers and the adjoint-state method: 80th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: <https://doi.org/10.1093/gji/ggz150>.
- van Leeuwen, T., and F. J. Herrmann, 2013, Mitigating local minima in full-waveform inversion by expanding the search space: *Geophysical Journal International*, **195**, 661–667, doi: <https://doi.org/10.1093/gji/ggt258>.
- Yang, P., R. Brossier, L. Metivier, and J. Virieux, 2016, Wavefield reconstruction in attenuating media: A checkpointing-assisted reverse-forward simulation method: *Geophysics*, **81**, no. 6, R349–R362, doi: <https://doi.org/10.1190/geo2016-0082.1>.