Nankai Trough Velocity Structure Reconstruction using
FWI of Wide-angle OBS Data with Graph-space Optimal
Transport Misfit Function

Andrzej GÓRSZCZYK1,2,*, Romain BROSIÈR1 and Ludovic MÉTIVIER1,3

1 ISTerre, Université Grenoble Alpes, Grenoble, France
2 Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland
3 Université Grenoble Alpes, CNRS, LJK, Grenoble, France


Technical development oriented on the detailed seismic imaging and velocity model building coupled with rapid increase of the computing power available nowadays make it possible to process large volumes of seismic data using numerically intensive approaches based on the wavefield propagation. In particular, full-waveform inversion (FWI) was intensively developed during last decade and is now routinely employed in the hydrocarbon exploration industry for velocity model building with the resolution allowing for direct, high-confidence interpretation of the reconstructed velocity models (Operto et al., 2013). As a consequence of successful FWI applications to the marine streamer data, recent years brought more interest into the seismic acquisition with densely deployed ocean-bottom nodes (OBN), which allow to record long-offset low-frequency 4C data preferable for FWI (e.g., Roende et al., 2020). In particular, wavefields traveling in the deep crust and undershooting shallower structures, provide the illumination of the target at various perspectives. This in turn helps to efficiently constrain the subsurface velocity at depths which are beyond the range of typical streamer acquisitions.

In the academic community, although the crustal-scale seismic data acquisitions utilizing ocean-bottom seismometers (OBS) are conducted since decades, they are rarely coupled with FWI (e.g., Kamei et al., 2012; Davy et al., 2017; Gorskczyk et al., 2017, 2021). On the acquisition side, usually limited number of available instruments translates to the crude wavenumber coverage of the model space during FWI. To mitigate this issue, seismic surveys with maximum OBS spacing of few kilometers (<5 km) shall be considered. On the processing side, even when the receiver sampling is sufficiently dense, the wide range of various arrivals recorded by long-offset deployments, increases the non-linearity of the inverse problem and rises the possibility of cycle-skipping, which guides FWI towards the reconstruction of geologically meaningless model. In this perspective, building of an accurate initial FWI model from long-offset OBS data is fundamental for successful velocity reconstruction. This crucial task, however, remains challenging 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). In consequences, successful FWI applications to OBS data are hampered, since in the presence of large errors of the starting model the classical L2-norm misfit function drives the inversion process towards non-informative local minima.

The solution to the cycle-skipping problem during FWI of long-offset OBS data can rely on deriving more accurate initial models (e.g., with slope tomography, Sambolian et al., 2021). 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 (Metivier et al., 2019). Compared with standard L2-norm, GSOT is convex with respect to the patterns in the waveform which can be shifted in time for more than half-wavelength, and therefore, this misfit function has potential to reduce the risk of cycle-skipping resulting from the inaccurate initial FWI model.

To test this hypothesis, here we apply a GSOT-based time-domain acoustic FWI workflow to a 2D wide-angle OBS dataset from the geologically challenging environment of the eastern Nankai Trough. The TKY-21 dataset (Fig. 1a) was acquired in the by JAMSTEC in 2001 during the KY0106 cruise of R/V Kaiyo. The acquisition comprises 100 OBS (4.5-Hz three-component geophones and hydrophones) deployed with 1 km intervals and 1404 air-gun shots (array volume of 1961) spaced 100 m apart. This shot/receiver spatial sampling is sufficiently dense for FWI application.

In contrast to the Laplace-Fourier FWI study of Gorsczyk et al. (2017) performed using the same dataset and a L2-norm misfit function (where significant efforts were devoted to derive an accurate initial velocity model), here we aim at reconstructing the same velocity structure of the subduction zone starting from a simple 1D model (Fig. 1b). In the approach we present here, despite obvious cycle-skipping in the initial FWI model (Figs. 1c–d), the
The GSOT misfit function is still able to match the corresponding data-samples and converge towards a correct solution.

In this study we follow the GSOT strategy of Metivier et al., 2019 with key features described hereafter. Instead of comparing directly the seismic traces \( d(t) \) discretized as \((d_1, ..., d_N)\), we use their discrete graphs \( (t, d) = ((t_1, d_1); \ldots; (t_N, d_N)) \in (\mathbb{R})^N \). In such a case, if \( d_{\text{cal}} \) and \( d_{\text{obs}} \) are the calculated and observed traces respectively, then \((t, d_{\text{cal}})\) and \((t, d_{\text{obs}})\) are their discrete graphs consisting of \( N \) delta Dirac functions in 2D time/amplitude space. The GSOT distance between \( d_{\text{cal}} \) and \( d_{\text{obs}} \) is given by solving the linear assignment problem:

\[
h_2(d_{\text{cal}}, d_{\text{obs}}) = \min_{\pi \in \mathbb{S}(N)} \sum_{i=1}^{N} c_{\pi(i)}(d_{\text{cal}}, d_{\text{obs}})
\]

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

\[
c_{\pi(i)}(d_{\text{cal}}, d_{\text{obs}}) = |t_i - t_j|^2 + \left[ \frac{1}{A} |d_{\text{cal},i} - d_{\text{obs},j}| \right]^2
\]

In Equation 2, \( A \) and \( \tau \) are the factors rescaling the amplitude component of the misfit with respect to the traveltimes component of the misfit. \( A \) is the maximum peak amplitude difference (calculated separately for each pair of the observed and calculated seismograms), while \( \tau \) is the maximum expected time-shift between signals in \( d_{\text{cal}} \) and \( d_{\text{obs}} \). This ensures the convexity of the GSOT for time up to around \( \tau \). With increasing \( \tau \), the attraction valley of the misfit function becomes wider, while with the \( \tau \) approaching to zero it becomes more narrow and the GSOT converges towards the L2-norm.

We apply time-domain visco-acoustic FWI workflow, relying on our visco-acoustic full-waveform modeling and inversion solver TOYx-DAC TIME. We design GSOT-based FWI scheme consisting of 4 stages. For each stage we essentially manipulate 4 inversion parameters: (i) we progressive narrow the attraction valley of the misfit function by the means of decreasing \( \tau \) value; (ii) we reduce the gradient regularization to allow introduction of smaller scale perturbations; (iii) we extend the mute time-windows that control amount of the inverted input data; (iv) we manipulate the AVO trend to boost the contribution of long-offset traces.

In the GSOT misfit function, \( \tau \) 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 \( \tau \). 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.

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. By reducing the correlation lengths of the smoothing operator we allow the fine scale perturbations to update of the model.

We use the time-window of length of 0.1 s starts at the first-arrival time and we extend it from one FWI stage to
another with progressively longer tapers. Initially the time-window 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. Extension of the time-window also corresponds to the narrowing of the scattering angles, which in turns translates to increased resolution of the model perturbations. We use a maximum time window of 2 s after the first arrival on the final 4\textsuperscript{th} 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 our dataset the difference in amplitude between short- and far-offset data is roughly three order of magnitude. To boost the contribution of the weak-amplitude far-offset traces (which represent the waves traveling in the deep part of the model), we start the inversion with trace-normalised data. In the following FWI stages we use square-root of the AVO and true AVO to account also for the amplitude variations in our data.

Figs. 2a–d and Figs. 2e–h 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 Fig. 2a), which leads to a smooth model exhibiting the general trend of the subduction zone. The synthetic data in Fig. 2e (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 2, we use a smaller $\tau$ value to improve the convergence of inversion, and simultaneously, we increase the resolution of the introduced perturbations 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 (Fig. 2b) 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 Fig. 2f. 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 (Fig. 2c) 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 Fig. 2g. The final model is presented in Fig. 2d. A smaller $\tau$ 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

Fig. 2. (a–d) Model evolution after each FWI stage. Insets show the difference between the presented and the initial model from Fig 1b. The color-scale is +/- 2000 m/s; (e–h) Data fitting evolution (OBS 23, Fig. 1c) 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.
final velocity perturbations presented in the inset exhibit a wide range of introduced structures varying both in terms of spatial scale and magnitude. In Fig. 2h one can observe further improvement of the continuity of phases and amplitude trends between synthetic and field data traces.

The FWI workflow that we presented 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. 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 OBS studies where one could rapidly build initial velocity models using traveltimes. The obtained velocity model, even if it were not to satisfy the cycle-skipping criteria, could serve as an 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. We believe that further applications of OT-based misfit functions can significantly reduce the overall risk of cycle-skipping during FWI of OBS data. This in turn shall encourage and stimulate future high-resolution deep crustal-scale academic studies utilising FWI of OBS data.

Key words: time-domain full-waveform inversion, alternative misfit function, high-resolution crustal-scale imaging

Acknowledgments: 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 No. plg3dwind3).

References