



HAL
open science

Localized adaptive waveform inversion: regularizations for Gabor deconvolution and 3-D field data application

Peng Yong, Romain Brossier, Ludovic Métivier, Jean Virieux

► To cite this version:

Peng Yong, Romain Brossier, Ludovic Métivier, Jean Virieux. Localized adaptive waveform inversion: regularizations for Gabor deconvolution and 3-D field data application. *Geophysical Journal International*, 2023, 235 (1), pp.448-467. 10.1093/gji/ggad225 . hal-04278807

HAL Id: hal-04278807

<https://hal.science/hal-04278807>

Submitted on 10 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Localized adaptive waveform inversion: regularizations for Gabor deconvolution and 3-D field data application

Peng Yong¹,^{ORCID} Romain Brossier,¹ Ludovic Métivier^{1,2} and Jean Virieux¹

¹Univ. Grenoble Alpes, ISTERre, F-38000 Grenoble, France. E-mail: peng.yong@univ-grenoble-alpes.fr

²Univ. Grenoble Alpes, CNRS, LJK, F-38000 Grenoble, France

Accepted 2023 May 24. Received 2023 May 19; in original form 2023 March 4

SUMMARY

Recently, we have developed a localized adaptive waveform inversion (LAWI) method to tackle the cycle-skipping issue in velocity reconstruction through seismic waveform inversion. The LAWI method employs a local matching filter, computed using Gabor deconvolution, to measure the instantaneous time-shift between observed and calculated data. Unlike the adaptive waveform inversion (AWI) approach, the LAWI method can take the non-stationarity between observed and calculated data into account. In this work, we investigate two types of regularization based on prior information about the expected filter, which could be a minimum-norm filter or a delta-shape filter, with regard to their effects on the robustness and resolution of inversion. We demonstrate on synthetic data the advantages and disadvantages of these two types of prior information, where the delta-type LAWI may handle multiple observed phases not initially predicted by the starting velocity model. Therefore, we apply the delta-type LAWI to a high-quality 3-D field data set in the North Sea, eliminating the need for data-windowing tuning, which can be tedious and time-consuming for 3-D data. Under different workflows with varying reliable initial models and frequency bands of the pressure data considered, we show that the LAWI approach is robust, effective and efficient for reconstructing the *P*-wave velocity, while other approaches such as AWI and graph-space optimal-transport method may require meticulous data-tuning strategies to converge to the correct model. Well logs and data fits, primarily from early arrivals, give us confidence that this LAWI approach could be applied to various acquisitions and subsurface targets, thanks to its phase-driven principle.

Key words: Waveform inversion; Inverse theory; Fourier analysis; Computational seismology.

1 INTRODUCTION

Seismic imaging by one integrated step that merges velocity model building and migration has been a long-term goal for numerous exploration geophysicists. Full-waveform inversion (FWI), proposed by Lailly (1983) and Tarantola (1984), theoretically, is able to automatically utilize the whole recorded data (e.g. transmitted waves, primary reflections, multiples, and ghosts) to produce high-quality images of broad wavenumber spectra. However, the journey of bringing FWI from theory to 3-D field data application is full of challenges. Thanks to full-azimuth and wide-offset data of high quality and the progress in high-performance computing, a spectacular FWI result on an industrial-scale 3-D seismic field data was obtained by Sirgue *et al.* (2010). Afterwards several successful exploration applications were reported on marine data (Warner *et al.* 2013; Operto *et al.* 2015; Shen *et al.* 2018; Huang *et al.* 2021) and land data (Plessix *et al.* 2012; Pan *et al.* 2018; Sedova *et al.* 2019).

These successful field data applications have greatly boosted the research interest for FWI. Let us mention that velocity reconstruction by FWI is a highly nonlinear inverse problem (Tarantola 1984, 1986). When observed data and predicted data generated by starting models differ by more than half-wave cycle, FWI can be trapped into local minima, failing to produce geologically meaningful velocity models (Virieux & Operto 2009). Building phase-compatible initial velocity model is a remedy, and such a model is easier to design for low-frequency seismic data (Sirgue *et al.* 2010). However, in seismic exploration, initial velocity models provided by tomographic methods rely on kinematic information that ignores the limited finite-frequency seismic content. These provided velocities as initial models may not make FWI free of local minima issues. On the other hand, high-fidelity low-frequency data require low-frequency source injection and high-quality geophones, together with a long-offset and full-azimuth acquisition (Dellinger *et al.* 2016; Vigh *et al.* 2021). Such an acquisition system is often costly,

especially for high-density acquisition. Alternative strategies should be investigated and we shall focus on this issue in this paper.

Enhancing phase information in misfit design can mitigate the nonlinearity of waveform inversion and relax the accuracy requirement on initial models and the quality requirement on low-frequency data (Luo & Schuster 1991; Luo *et al.* 2016; Zhang *et al.* 2018). Measuring time-shift between observed and calculated data is a difficult and long-standing task in exploration seismology. In the early work of wave-equation traveltimes tomography (Luo & Schuster 1991), time-shift is taken as the lag that maximizes the cross-correlation function. In principle, its applicability is limited to single-phase data, such as identifiable first-arrivals. Besides, its adjoint source derivation relies on the assumption that observed and calculated data only differ by a time-shift (identical waveforms), while this assumption is not in agreement with realistic situations. Inspired by the differential semblance optimization for migration velocity analysis (Symes & Carazzone 1991; Shen & Symes 2008), van Leeuwen & Mulder (2010) propose a penalization-based cross-correlation objective function in the data domain, which enables the derivation of adjoint source without the need for this identical waveform assumption. Following, Luo & Sava (2011) suggest using deconvolution instead of cross-correlation in order to handle oscillatory seismic signals. Deconvolution operation can make the matching filter more focused and confined at zero lag with model update. Warner & Guasch (2016) design a normalized form of penalization-based deconvolution misfit, named adaptive waveform inversion (AWI), which shows a great superiority over penalization-based methods in dealing with cycle-skipping issues.

In the previous work (Yong *et al.* 2023), we have highlighted that the misfit in AWI can be considered as an estimation of the centroid time of the matching filter, with the assumption that the global matching filter in AWI is invariant for all events in seismograms. This suggests that AWI is effective in mitigating cycle-skipping issue when processing data that is dominated by single events, such as non-triplicated diving waves. However, seismic phases often have different time-shifts in exploration data, making the global matching filter vulnerable to interference between different events, especially in the presence of multiple events or phases with significant energy. This can lead to the AWI misfit function becoming non-convex with respect to a local time-shift. See Figs. 6 and 7 in Yong *et al.* (2023) for further details. In realistic scenarios, multiple phases can arise in a data trace through various means, including multiple transmission ray paths, primary or multiple reflection, and elastic conversion, all of which can cause AWI to fail. To address this issue, we have proposed a non-stationary extension of AWI, referred to as localized AWI (LAWI), that assesses phase mismatch between observed and predicted data locally. This is achieved through a time–frequency analysis of signals using a Gabor transform (Gabor 1946), treating each event independently, so that the same phase detection mechanism that makes AWI succeed with single-event data works locally, everywhere.

Although a variety of methods related to misfit function modification have been proposed to handle cycle-skipping issues during the last decade, they enhance phase information or low-frequency component in different ways. Some of them are computationally challenging for 3-D applications. Some may lack robustness when proceeding with complex or noisy field data (Guasch *et al.* 2019; Pladys *et al.* 2021). For applications to a wide range of 3-D field data sets with confidence, the method requires to be sufficiently robust, efficient, and effective. In this work, following the non-stationary strategy promoted by Yong *et al.* (2023), we investigate two regularization strategies used for Gabor deconvolution in LAWI, with a

focus on the effect of injected prior information on the inversion results. The robust regularization one illustrated with synthetic examples is applied to a 3-D ocean-bottom-cable (OBC) data set from the North Sea. For challenging workflows where cycle-skipping issues are expected, we compare the LAWI approach with the delta-type regularization to other methods, such as the AWI strategy (Warner & Guasch 2016) and the graph-space optimal-transport (GSOT) strategy (Métivier *et al.* 2018, 2019) on this available real data.

The structure of the paper is organized as follows: We begin by outlining the mathematics of the LAWI method and discuss numerical implementations of Gabor deconvolution from a point of view of regularization. Then, we study the characteristics of two implementations of Gabor deconvolution through simple signal analysis and 2-D synthetic inversion tests. In Section 4, we describe the field data set and its preparation for inversion, and then present a sequence of inversion results under increasingly challenging situations, as well as comparison with available well logs and data fitting. The results and possible further improvement are addressed in Section 5, and finally we draw conclusions about the LAWI method from this study.

2 THEORY

In the LAWI approach, the connection between one-trace observed and predicted data is built by a Gabor convolutional model (Margrave 1998), and the phase difference is implicitly estimated with the corresponding local matching filter (Yong *et al.* 2023). We first provide a brief recapitulation of the LAWI method, then discuss the prior information used for regularizing the Gabor deconvolution problem and propose a new type of regularization, which follows the fact that the local matching filter converges towards a delta function when the model turns into the true one.

2.1 Localized adaptive waveform inversion

With one-trace observed data $d(t)$ and predicted data $p(t)$, a local matching filter $\hat{w}(t, \omega)$ used in the LAWI approach, can be defined by a non-stationary convolutional model in the time–frequency domain (Yong *et al.* 2023) as

$$\hat{w}(t, \omega) \hat{d}(t, \omega) = \hat{p}(t, \omega), \quad (1)$$

where $\hat{d}(t, \omega)$ and $\hat{p}(t, \omega)$ are the time–frequency spectra of observed and predicted data. The Gabor transform (Gabor 1946) is applied to obtain the time–frequency spectra. The Gabor transform pair can be given by

$$\hat{f}(t, \omega) = \mathcal{G}[f](t, \omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(\xi) h_{\sigma}^{\dagger}(\xi - t) e^{-i\omega\xi} d\xi, \quad (2)$$

$$f(t) = \mathcal{G}^{-1}[\hat{f}](t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^2} \hat{f}(\xi, \omega) h_{\sigma}(t - \xi) e^{i\omega t} d\xi d\omega. \quad (3)$$

Here, symbol \dagger denotes the complex conjugate and $h_{\sigma}(t) = (\pi\sigma^2)^{-\frac{1}{4}} e^{-\frac{t^2}{2\sigma^2}}$ denotes the window function, in which σ controls the radius (Strang & Nguyen 1996; Fichtner *et al.* 2008). With the time-varying local matching filter $\hat{w}(t, \omega)$ in the frequency domain, we can obtain its time domain form $w(t, \tau)$ through

$$w(t, \tau) = \mathcal{F}_{\tau}^{-1}[\hat{w}(t, \omega)], \quad (4)$$

where \mathcal{F}_{τ}^{-1} denotes the inverse Fourier transform for the variable τ . The instantaneous centroid time-shift is estimated by

$$T(t) = \frac{\int_{\mathbb{R}} |\tau| w^2(t, \tau) d\tau}{\int_{\mathbb{R}} w^2(t, \tau) d\tau}. \quad (5)$$

All instantaneous time-shifts can be integrated under the L^n norm, with n typically chosen as 1 or 2. The misfit function of LAWI, with $n = 2$, can be expressed as

$$\mathbb{J}_{\text{LAWI}} = \frac{1}{2} \int_{\mathbb{R}} T^2(t) dt. \quad (6)$$

Let us mention that the Gabor deconvolution itself is ill-posed for noisy data (Margrave *et al.* 2011). In practice, Tikhonov regularization (Tikhonov *et al.* 2013) is often used to stabilize the deconvolution (Claerbout & Fomel 2014). We shall discuss the importance of prior information in the regularization in the following subsection.

2.2 Two possible regularizations for Gabor deconvolution

The local matching filter, defined by the Gabor deconvolution (1), is commonly computed by the equation

$$\hat{w}_z(t, \omega) = \frac{\hat{d}^\dagger(t, \omega) \hat{p}(t, \omega)}{\hat{d}^\dagger(t, \omega) \hat{d}(t, \omega) + \epsilon}, \quad (7)$$

where a small positive number ϵ prevents the division by zero (Margrave *et al.* 2011). From a point of view of regularization, the local matching filter, calculated by eq. (7), is the minimum-norm solution of the regularized least-squares problem defined by

$$\min_{\hat{w}_z(t, \omega)} \frac{1}{2} \int_{\Omega} \left| \hat{p}(t, \omega) - \hat{d}(t, \omega) \hat{w}_z(t, \omega) \right|^2 + \epsilon |\hat{w}_z(t, \omega)|^2 d\omega, \quad (8)$$

where symbol Ω denotes the frequency band of observed and predicted signals, and the local matching filter. Each time component (7) of the local matching filter will be used when considering the misfit function (6). The value of the regularization parameter ϵ (assumed to be time-independent) limits the admissible solution space of the local matching filter $\hat{w}(t, \omega)$. This commonly used regularization is designed for numerical stability. In practice, the regularization parameter ϵ selection depends on the SNR of observed data. When dealing with low SNR data, it is often necessary to choose a larger value for ϵ . Such a zero-type regularization has been used by Guasch *et al.* (2019) in the field-data application of AWI. However, the local matching filter is supposed to converge into a delta function with iteration in LAWI, while this regularization attempts to restrict it to a zero norm, which constitutes a certain inconsistency between the expected convergence of the filter and the applied prior information (Yong *et al.* 2022).

Following the same conceptual strategy as Warner & Guasch (2016) (eq. 15 in their article), the local matching filter should tend toward a delta function with model update. Thus the expected behavior expressed by

$$p(t) \rightarrow d(t), \quad \hat{w}(t, \omega) \rightarrow 1, \quad (9)$$

can be related to the following least-squares problem

$$\min_{\hat{w}_d(t, \omega)} \frac{1}{2} \int_{\Omega} \left| \hat{p}(t, \omega) - \hat{d}(t, \omega) \hat{w}_d(t, \omega) \right|^2 + \epsilon |\hat{w}_d(t, \omega) - 1|^2 d\omega, \quad (10)$$

leading to another local matching filter given by

$$\hat{w}_d(t, \omega) = \frac{\hat{d}^\dagger(t, \omega) \hat{p}(t, \omega) + \epsilon}{\hat{d}^\dagger(t, \omega) \hat{d}(t, \omega) + \epsilon} = \hat{w}_z(t, \omega) + \frac{\epsilon}{\hat{d}^\dagger(t, \omega) \hat{d}(t, \omega) + \epsilon}. \quad (11)$$

Such an expression (11) differs from the one (7) when applying the zero-type regularization. We name this new regularization delta-type regularization to make it distinct from the standard zero-type regularization.

For these two implementations of Gabor deconvolution, the adjoint sources in LAWI share the same following expression

$$r_{\text{LAWI}} = 2\mathcal{G}^{-1} \left[\mathcal{F}_\tau \left[\frac{T(t)(|\tau| - T(t))w(t, \tau)}{\int_{\mathbb{R}} w^2(t, \tau) d\tau} \right] \frac{\hat{d}(t, \omega)}{\hat{d}(t, \omega) \hat{d}^\dagger(t, \omega) + \epsilon} \right], \quad (12)$$

where the local matching filter could be either the zero-type filter w_z or the delta-type filter w_d , leading to different numerical adjoint fields. The detailed derivation of adjoint sources in LAWI can be found in Yong *et al.* (2023). In addition, the wave equation and the model-gradient construction used in the following numerical experiments can be found in Appendix A. Let us explore on synthetic examples the influence of these two different regularization taken as prior information.

3 ANALYSIS ON SYNTHETIC DATA

Before considering a realistic 2-D inversion example, let us illustrate on one trace the influence of single and multiple observed phases.

3.1 Simple signal analysis

We start with one-event signal experiment to show how resolution could be slightly downgraded with the delta-type regularization. Fig. 1(a) presents two 5 Hz Ricker wavelets with a 0.5 s time-shift. The estimated instantaneous time-shifts by the two implementations of Gabor deconvolution are shown in Figs. 1(b) and (c), and both estimations are close to the true value (0.5 s). The delta-type regularization generates a non-zero signal along the zero-lag axis (see eq. 11), coming from the injected prior information of a delta function (Fig. 1c), while the zero-type regularization does not (Fig. 1b). Such non-zero signal decreases the estimated time-shift between the calculated and observed data (Fig. 1d). Let us underline that adjoint sources are different (Fig. 1e).

For the delta-type regularization, these non-zero energies along the zero-lag axis induce a slight loss of resolution in the inversion. Let us estimate the objective function variation with respect to two parameters, the amplitude scaling and the time-shift through the relation

$$d_{\text{obs}}(t) = A d_{\text{cal}}(t + \tau). \quad (13)$$

The objective function with the delta-type regularization becomes flat when approaching the global minimum, reducing the expected resolution of the final result (Fig. 2). As expected, both misfit functions are nearly agnostic to amplitude scaling and mainly sensitive to time-shifts (Yong *et al.* 2023).

In addition, the zero-type regularization makes use of events in observed seismogram that are locally coherent to the events occurring in predicted data: an expression of the minimum-norm solution. The new regularization does not. Let us illustrate it with a case where observed data contains two events while only one event is present in the predicted data, as shown in Fig. 3(a). The time-shifts and the adjoint sources presented in Figs. 3(d) and (e) clearly show that the delta-type regularization accounts for the two events at 1.25 and 3.25 s using only one predicted event with a different behavior. For the zero-type regularization, the local matching is zero when the predicted data is zero, no matter whether observed data is zero or not. Consequently, the zero-type regularization cannot take the observed second event into account (Fig. 3d). Conversely, for the delta-type regularization, the local matching filter has a non-zero energy pattern diffusely distributed at 1.5 and of 3.25 s along

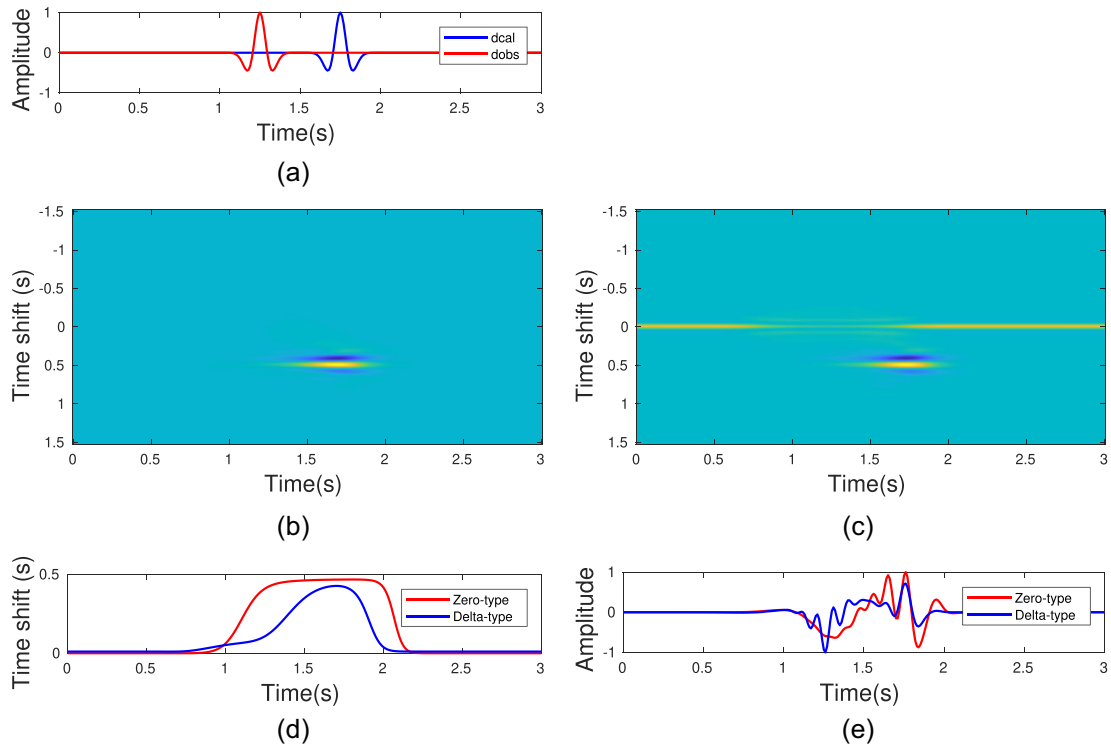


Figure 1. Single event study: observed and calculated data (a), the local matching filters $w(t, \tau)$ with zero-type regularization (b) and delta-type regularization (c), the estimated time-shifts (d), and the associated adjoint sources (e). Please note the slight delay in the expected time-shift for delta-type regularization.

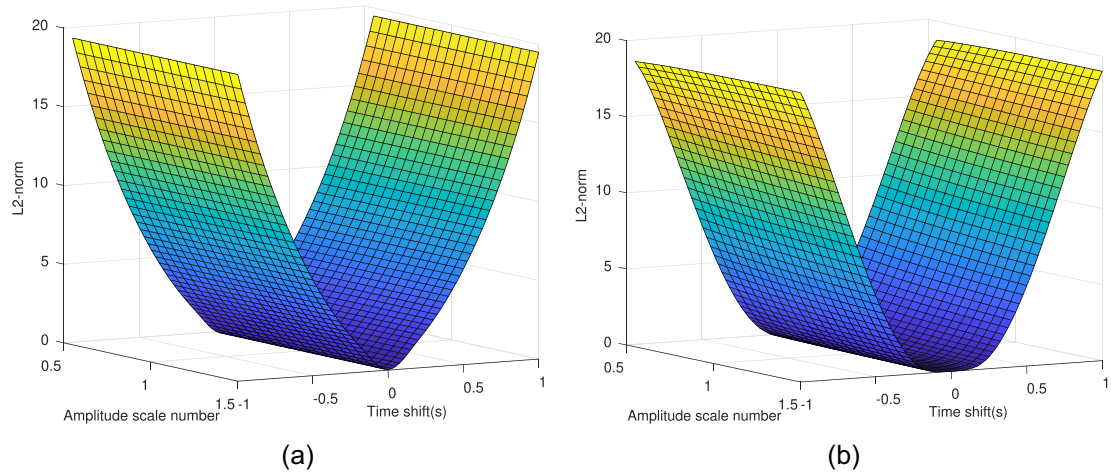


Figure 2. Objective functions with respect to two parameters of time-shift and amplitude scaling: zero-type regularization (a) and delta-type regularization (b). Please note the flatness of the delta-type objective function.

the time axis (Fig. 3c). Such a diffusive pattern underestimates the time-shift estimation for the observed event at the time 1.5 s but creates a time-shift for the second observed event at the time 3.25 s from the single predicted event. The delta-type regularization enables LAWI to take into account some information of the second observed event, even though there is no locally associated event in the predicted data. The adjoint source has a corresponding signal which cannot be observed for the zero-type regularization (Fig. 3e).

In seismic exploration, the observed data usually contains direct wave, diving wave, reflected wave, etc. When inversion starts from a smooth velocity model, the calculated data may only contains direct and diving waves. Thus the number of events in calculated

data and observed data are not the same. LAWI with the zero-type regularization focuses on the time-shifts between diving and direct waves only, which might not provide sufficient information to match the observed reflections to the predicted data. When a significant part of the energy in the observed data is related to reflections, a non-convergence issue can arise for LAWI with the zero-type regularization. This is because, with model update, reflections will appear in the predicted data, which leads to time-shifts of the new reflections (not existing at the beginning) and may make the objective function increase (Yong *et al.* 2023). As the delta-type regularization can account for the whole information in observed data from the starting iteration of inversion, it does not suffer from such a non-convergence issue.

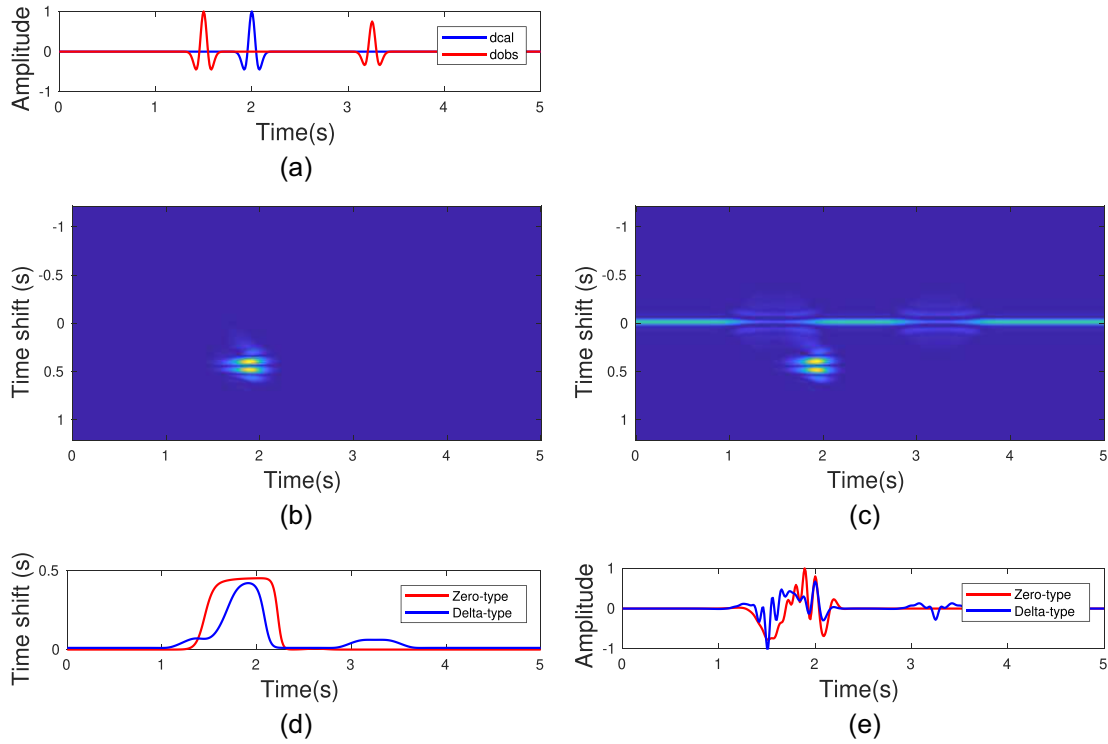


Figure 3. Multiple-event observed trace: the observed and calculated data (a), the local matching filter $|w(t, \tau)|$ with zero-type regularization (b) and delta-type regularization (c), the related time-shifts (d), and the adjoint sources (e).

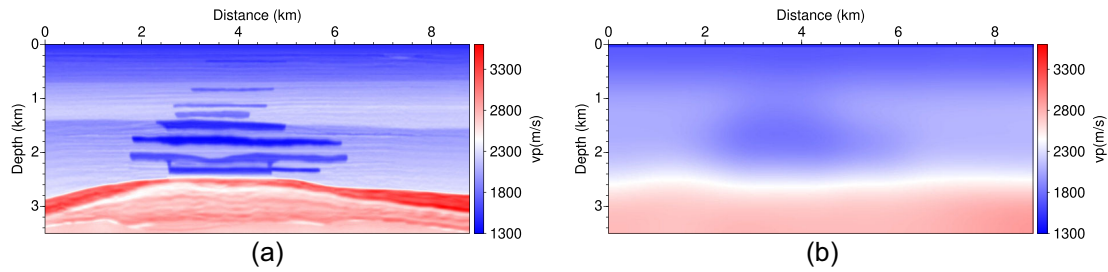


Figure 4. 2-D Valhall models: true velocity (a) and initial velocity (b).

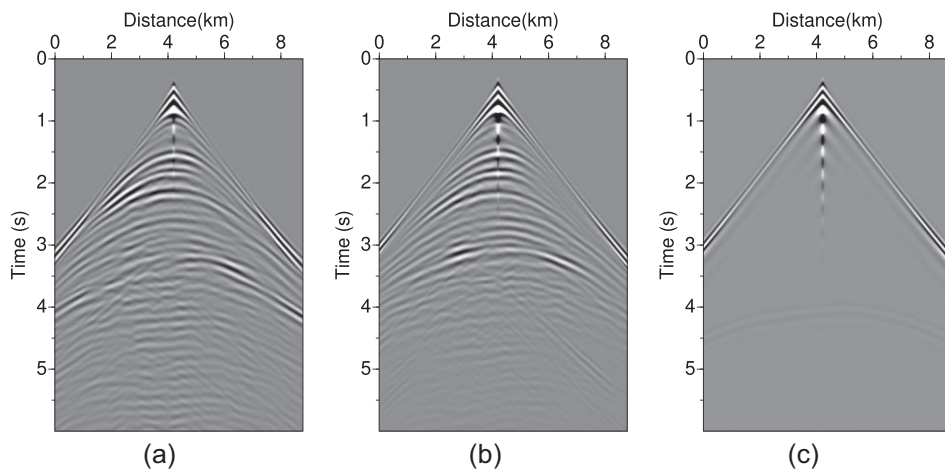


Figure 5. The observed data (a) generated with the true velocity and density model and the calculated data using the initial velocity model with the true density (b) and with the smooth density model (c).

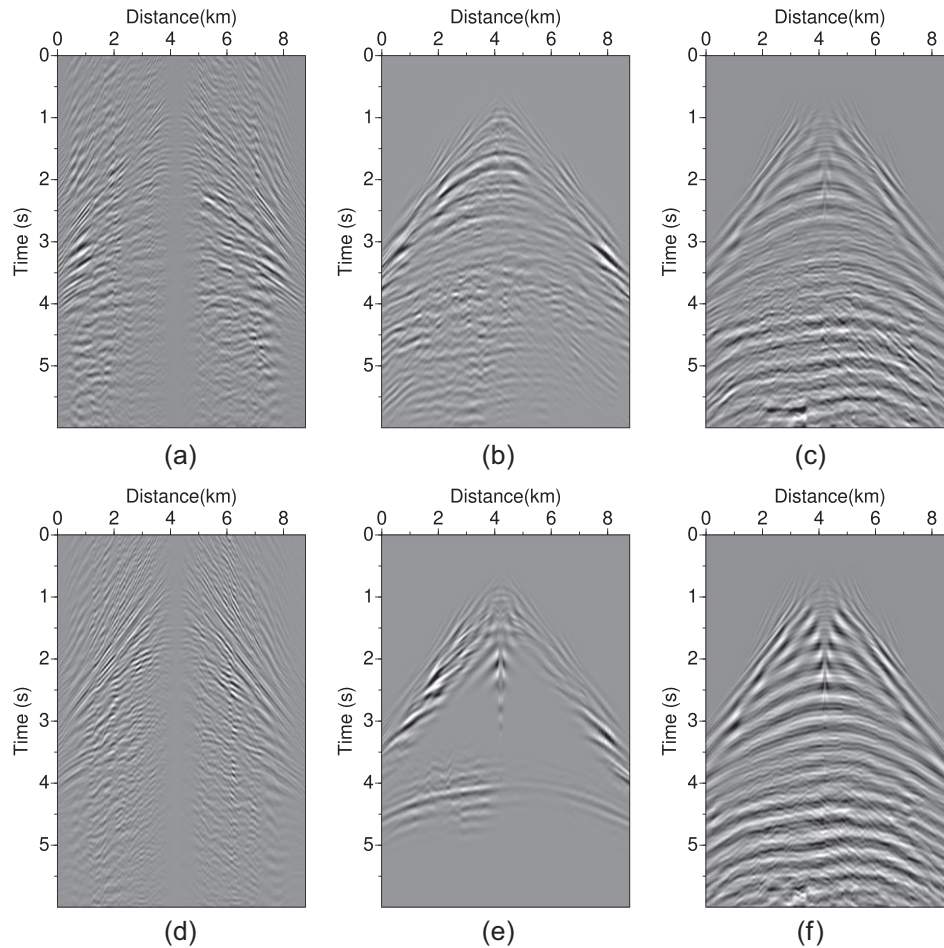


Figure 6. The adjoint sources: the true density model (a–c) and the smooth density model (d–f). AWI with the zero-type regularization (a and d), LAWI with the zero-type regularization (b and e) and LAWI with the delta-type regularization (c and f).

3.2 2-D synthetics example

Let us now consider a 2-D realistic model in a shallow water environment similar to the 3-D field data we shall consider later, to further illustrate the two identified features of LAWI. Fig. 4 presents the true and initial velocity models. The density model, used to generate the observed data, is calculated following the Gardner’s relation (Gardner *et al.* 1974)

$$\rho = 309.6V^{0.25}. \quad (14)$$

The source function is a 5 Hz Ricker wavelet with a low-cut filter to remove energy below 2 Hz. A fixed-spread acquisition is used, with 32 equally spaced sources and 352 equally spaced receivers with interval of 25 m placed on the surface. The initial model is built by using the `sfsmooth` command in the open-source Madagascar software (Fomel *et al.* 2013): repeatedly (15 times) applying a triangle filter with a radius of 250 m to the true model. Standard least-squares FWI suffers from cycle-skipping issues with this initial model. The ℓ -BFGS method ($\ell = 5$) is applied for the model update (Métivier & Brossier 2016). The maximum number of iteration is set to 50.

The observed data contains both transmitted and reflected phases (Fig. 5a). An initial model with the true density model gives synthetic data which include reflected phases (Fig. 5b). Another initial model is built with a smooth density model generated by applying

the Gardner’s law to the smooth initial velocity model. Seismic traces mainly contain transmitted diving and direct phases with one weak reflection from localized fast velocity variation at a depth around 2.5 km, arriving at a time around 4 s (Fig. 5c). From the adjoint sources shown in Fig. 6, LAWI with the zero-type regularization only accounts for events occurring in the predicted data (Figs. 5b and e), while LAWI with the delta-type regularization can bring the whole information contained in observed data to the model gradient (Figs. 5c and f). The conventional AWI displays more complex adjoint sources mainly driven by transmitted phases (Figs. 5a and d). Moreover, strong non-causal artefacts occur in the AWI adjoint sources, coming from the stationary assumption in AWI (Yong *et al.* 2023).

Let us consider the inversion results starting with the smooth velocity model and with the true density using the AWI approach and the two LAWI strategies. AWI has difficulties to reconstruct the low-velocity layers (Fig. 7a), whereas LAWI with zero-type regularization converges to the true velocity model (Fig. 7c). The result provided by LAWI with delta-type regularization has a slightly lower resolution as expected (Fig. 7e). When inversion begins with the same smooth velocity model complemented by the smooth density model, AWI evolves to a wrong model (Fig. 7b). LAWI with the zero-type regularization, which suffers from the aforementioned shortcomings, cannot update the given initial model (Fig. 7d). Interestingly, the delta-type regularization can take all reflections into

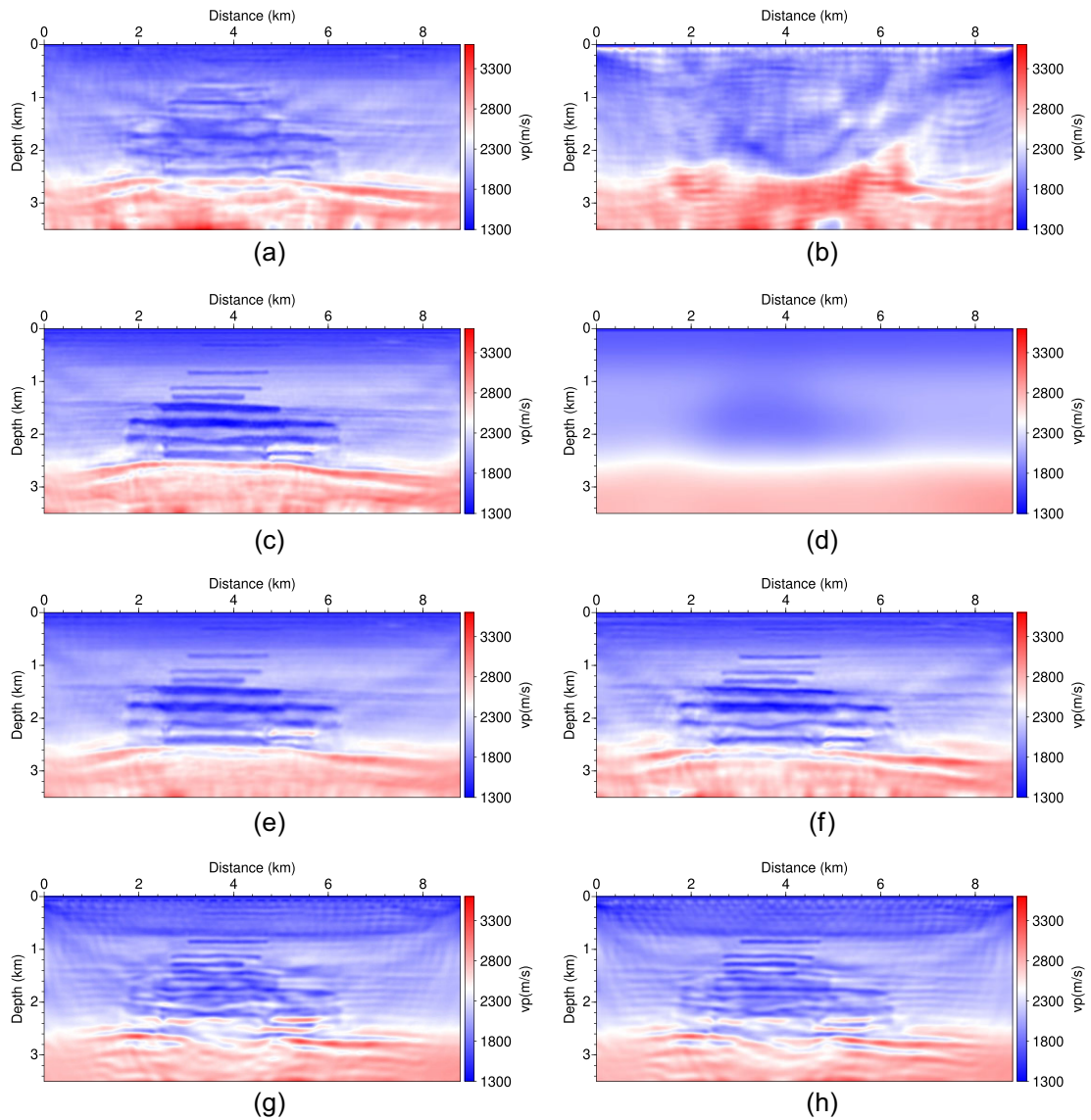


Figure 7. Inversion results with (left) and without (right) the true density model: AWI (a and b), LAWI with zero-type regularization (c and d), LAWI with delta-type regularization (e and f), standard FWI (g and h). Optimization failure is observed for the zero-type regularization (d): the displayed model is the input model.

account at the beginning of inversion, leading to a model very similar to the one obtained when considering true density model. One may notice that the result starting with the smooth density model (Fig. 7f) has a little more high-wavenumber component than when starting with the true density model (Fig. 7e). This is because the density model is not updated: related data differences are accounted in the velocity updating. Standard least-squares FWI results with the two starting models are shown in Figs. 7(g) and (h): these velocity models are very similar with the small high-wavenumber velocity component when considering the smooth density model which is not updated in this illustration.

It is necessary to mention that a weighting-window strategy, following the data-domain layer stripping principle, can help LAWI with the zero-type regularization to avoid the non-convergence issue. However, it requires manually adjusting compared areas in seismograms during inversion, which could be difficult when considering 3-D applications. The weighting-window strategy is also very helpful for AWI, as it can mitigate the nonlinearity between

observed and calculated data (see more illustrations in Yong *et al.* 2023). Let us mention that, to guide the matching filters towards a delta function, we use a monotonically increasing penalization function with respect to the time lag in this work, while alternative forms of penalty function is possible, which may make the AWI approach amenable to more robustly interpret complex signals (Warner & Guasch 2016; Sun & Alkhalifah 2020).

4 3-D FIELD DATA APPLICATIONS

The LAWI approach with delta-type regularization is applied to a field data set for which the zero-type regularization may require more tuning efforts from the user.

4.1 Data set

We consider the pressure component recorded during active experiments for a 3-D field target from the North Sea. The acquisition

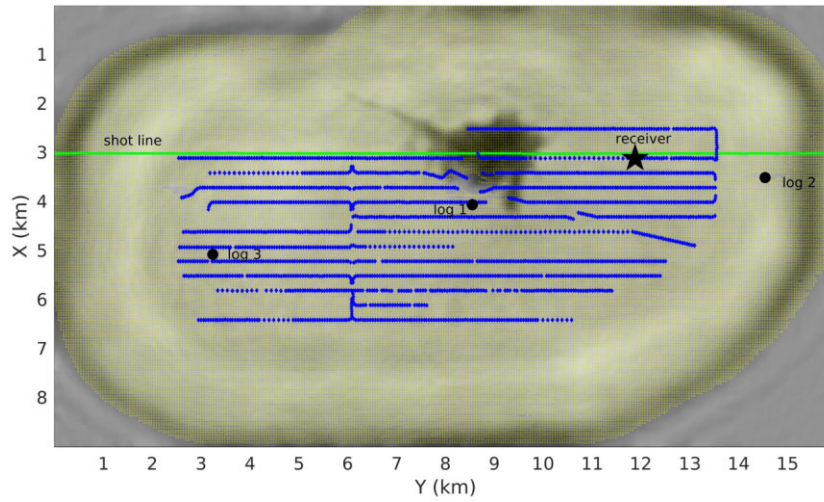


Figure 8. The acquisition of 3-D OBC data overlaid on a depth slice ($z = 1$ km) of the velocity reconstructed from LAWI: locations of sources (yellow dots) and receivers (blue diamonds), three well logs. The data corresponding to the green line of shots and the receiver (black star) is extracted for data processing illustration and quality check of the reconstructed models.

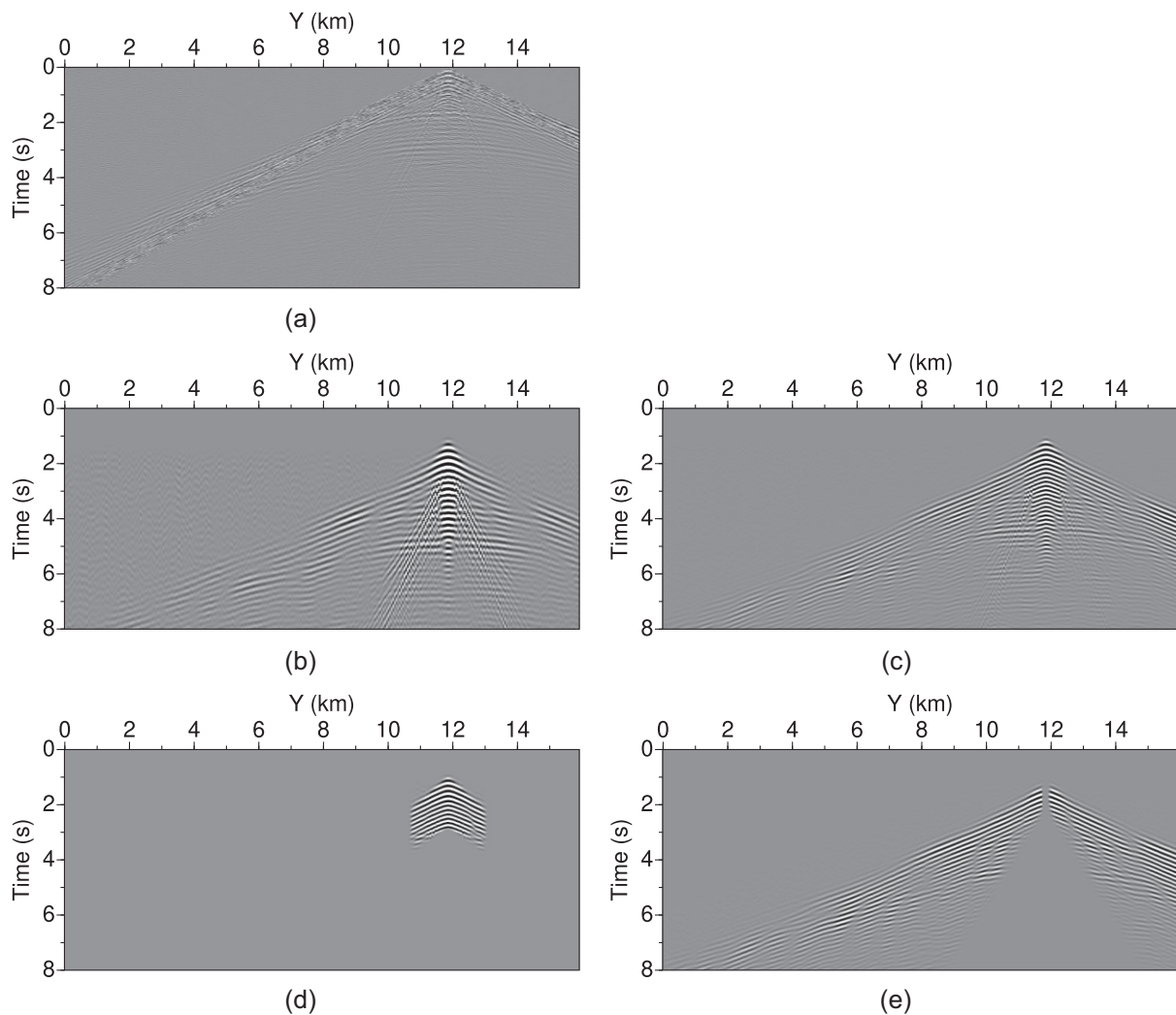


Figure 9. Hydrophone shot record: (a) full bandwidth raw data, (b) 2.5–5.0 Hz data, (c) 2.5–7.0 Hz data. The data weighting strategy is used to select suitable data for source estimation (d) and velocity inversion (e) on 2.5–7.0 Hz data.

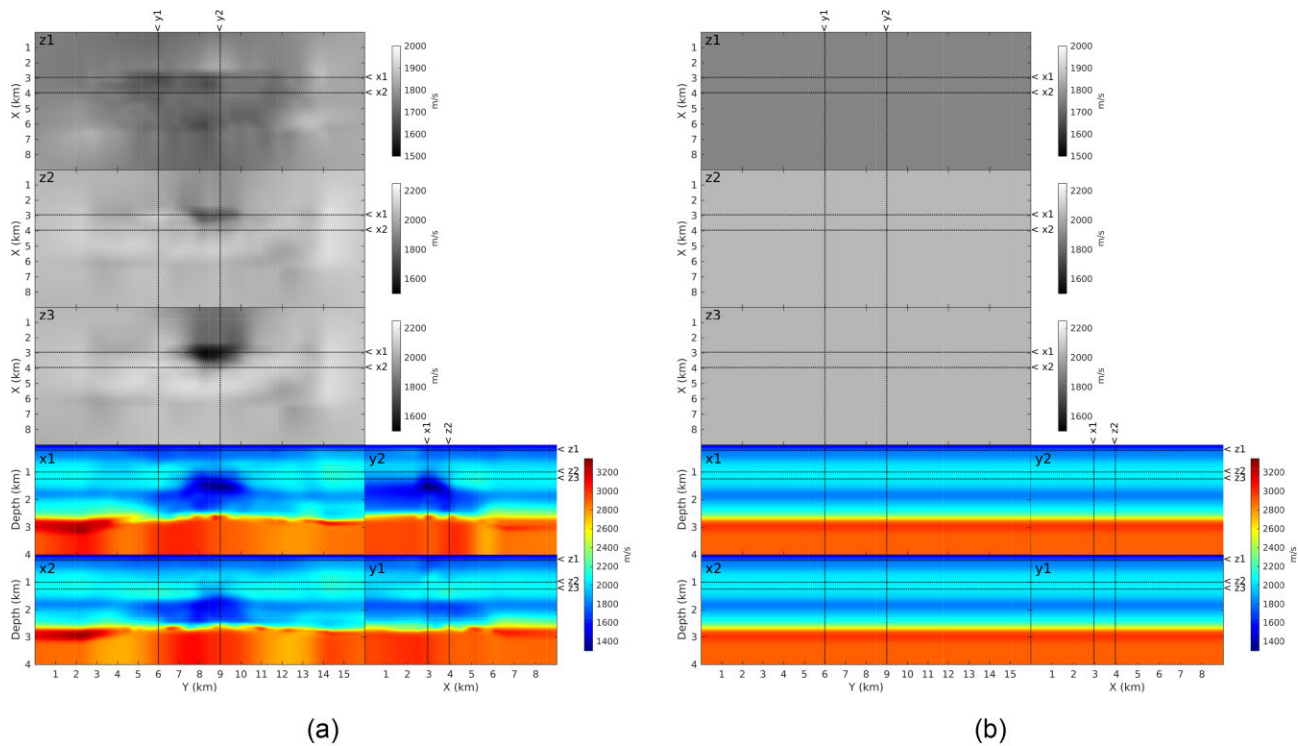


Figure 10. Initial velocity models: (a) tomographic model and (b) simple stratified model. Grey-scale horizontal slices are extracted from the depths of 0.2 km (z_1), 1.0 km (z_2) and 1.25 km (z_3), respectively. Inline vertical slices are from $x = 2.95$ km (x_1) and $x = 3.95$ km (x_2). Crossline vertical slices are at $y = 6$ km (y_1) and $y = 9$ km (y_2).

configuration consists of 2044 receivers, shown as blue diamonds in Fig. 8, and 50 824 shots, fired every 50 m inline. The area covered by the survey is approximately 145 km². For saving computational cost of waveform inversion, we exchange the source and receiver positions with the spatial reciprocity theorem for waveform prediction. In other words, we have now 2044 shot gathers.

The source-subsampling strategy promoted by Warner *et al.* (2013) is applied. The 2044 gathers are divided into 16 batches without data repetition to make sure that all shots are used during inversion. The first 15 batches contain 128 random gathers while the last batch contains only 124 gathers. For model gradient estimation, we run three iterations of ℓ -BFGS on each batch, meaning that all data are reached 48 times. For each data set, we run 96 iterations in total, meaning that each shot is used six times in the inversion. The source-subsampling strategy makes it possible to run large-scale 3-D waveform inversion problems with a relatively small HPC resource (Kamath *et al.* 2021; Pladys *et al.* 2022).

Fig. 9(a) displays the raw data along the green line ($x = 3$ km) in Fig. 8, recorded at the location of the black star in Fig. 8 ($x = 3.12$ km, $y = 11.88$ km). The bandpass filtered data in the frequency range of 2.5–5.0 Hz and 2.5–7.0 Hz are shown in Figs. 9(b) and (c), respectively. Considering narrower lower-frequency range than the range 2.5–5.0 Hz during the inversion for overcoming possible cycle-skipping issues is not effective because the SNR for these low frequencies is too low. A data weighting strategy is applied for data selection (Kamath *et al.* 2021; Pladys *et al.* 2022): the near-offset direct waves (Fig. 9d) is extracted for wavelet estimation in the frequency domain (Pratt 1999), and a short-offset mute is applied on

bandwidth data for removing low-velocity/low-frequency Scholte waves. The same mute is applied to synthetic data when doing the data comparison. No other processing such as deghosting and demultiple is applied to the data for inversion.

4.2 Inversion tests

Three starting configurations are considered for respective performance of LAWI (using delta-type regularization) and FWI with respect to cycle-skipping issues:

- (1) Starting frequency band (2.5–5.0 Hz) and a tomographic starting model (frequency band (2.5–7.0 Hz) data used subsequently)
- (2) Starting frequency band (2.5–5.0 Hz) and a simple stratified starting model (frequency band (2.5–7.0 Hz) data used subsequently)
- (3) Starting frequency band (2.5–7.0 Hz) and a simple stratified starting model.

Local minima could be met more often from the first setting to the third one. Fig. 10 displays the two initial velocity models: a tomographic model and a simple stratified model. The simple stratified model is generated by extrapolating horizontally the vertical velocity profile of the tomographic model at $y = 0$. The stratified model contains major extended stratified structures from the tomographic model. The significant difference comes from the focused low-velocity anomaly in the central part of the target. With the first inversion configuration, we expect that LAWI and FWI provide similar high-quality results. For the later two configurations, we shall

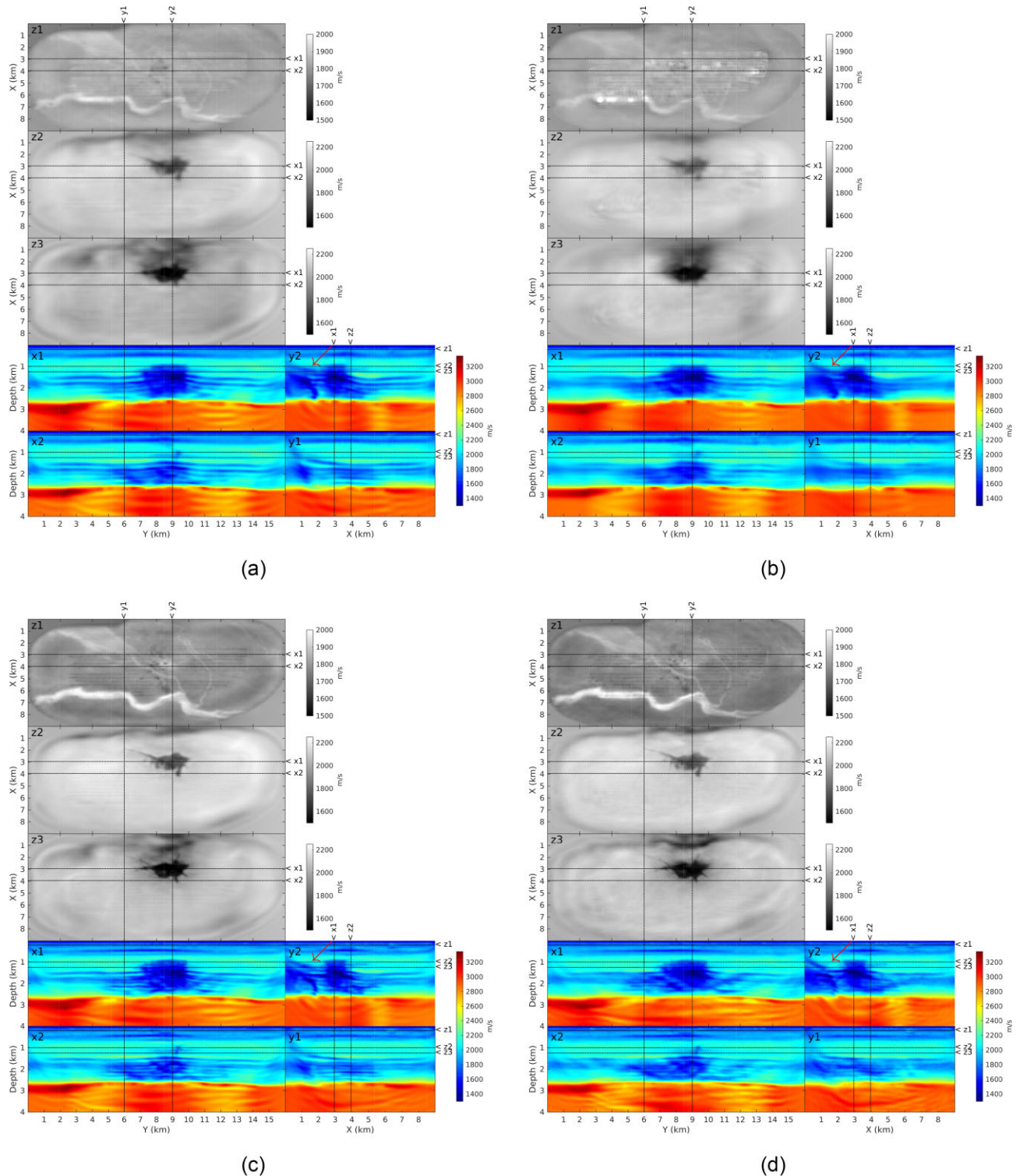


Figure 11. Inversion starting with the tomographic model: reconstructed velocity models with 2.5–5.0 Hz data using FWI (a) and LAWI (b), and subsequently reconstructed velocity models with 2.5–7.0 Hz data using FWI (c) and LAWI (d). Please note the increase in resolution when increasing the frequency range.

focus on the low-velocity anomaly for assessing the reconstruction quality as well as the local minima possibilities.

Because we consider pressure data, P -wave velocity reconstruction is performed. It is enough to consider a 3-D viscoacoustic vertical transversely isotropic (VTI) time-domain wave equation (Yang *et al.* 2018), solved by a 4th-order staggered grid finite-difference

method (Levander 1988). The gradient computation and the related forward and adjoint equations can be found in Appendix A. All inversions use the same VTI anisotropy model described in recent publications (Kamath *et al.* 2021; Plady *et al.* 2022), and such anisotropy is not updated during the inversion. The importance of including anisotropy is well described in the work of Prieux *et al.*

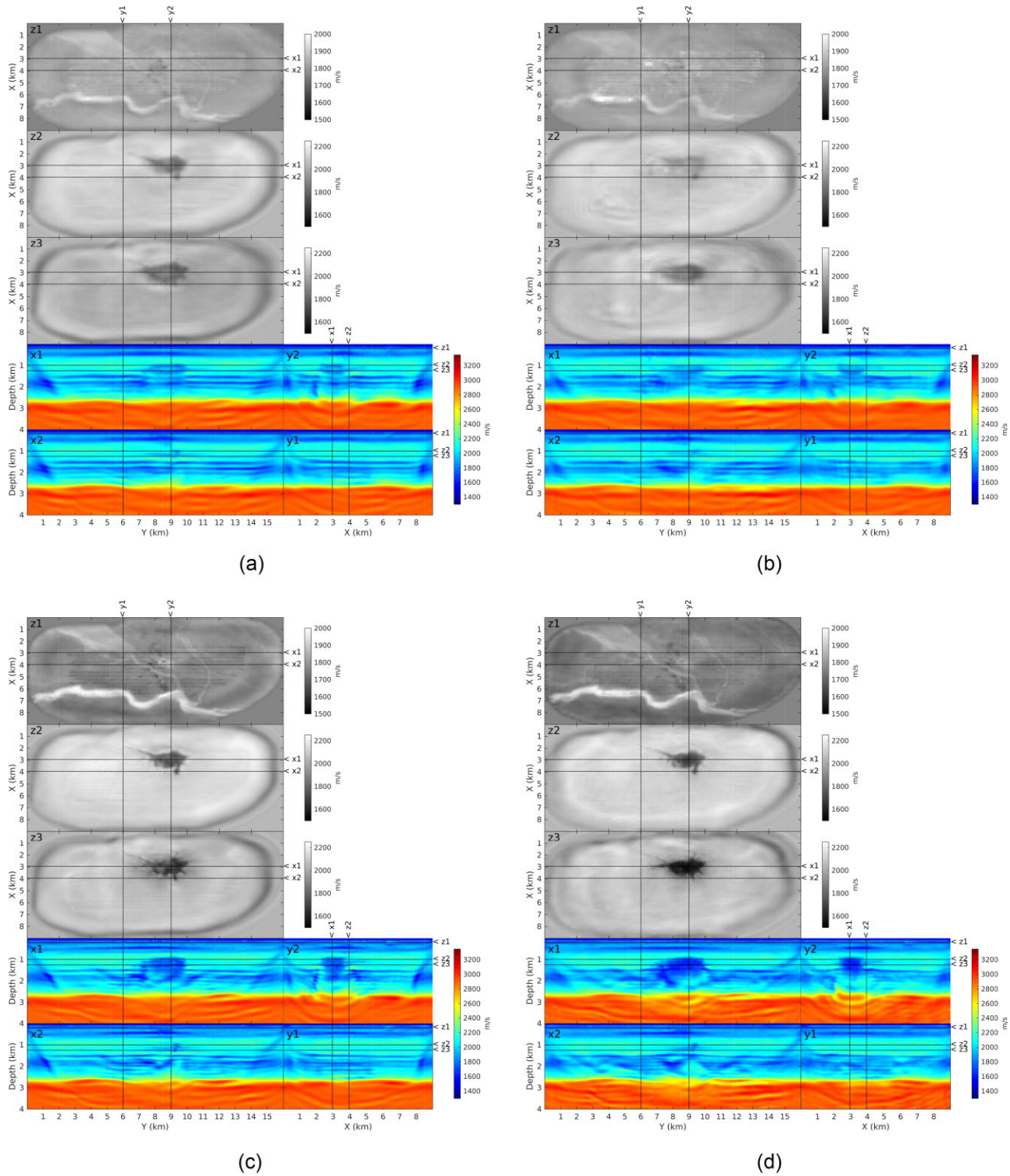


Figure 12. Inversion starting with the simple stratified model: reconstructed velocity models with 2.5–5.0 Hz data using FWI (a) and LAWI (b), and subsequently reconstructed velocity models with 2.5–7.0 Hz data using FWI (c) and LAWI (d). The low-velocity anomaly reconstructed by FWI at the depth of 1.25 km is not as good as the one obtained with the tomographic starting model.

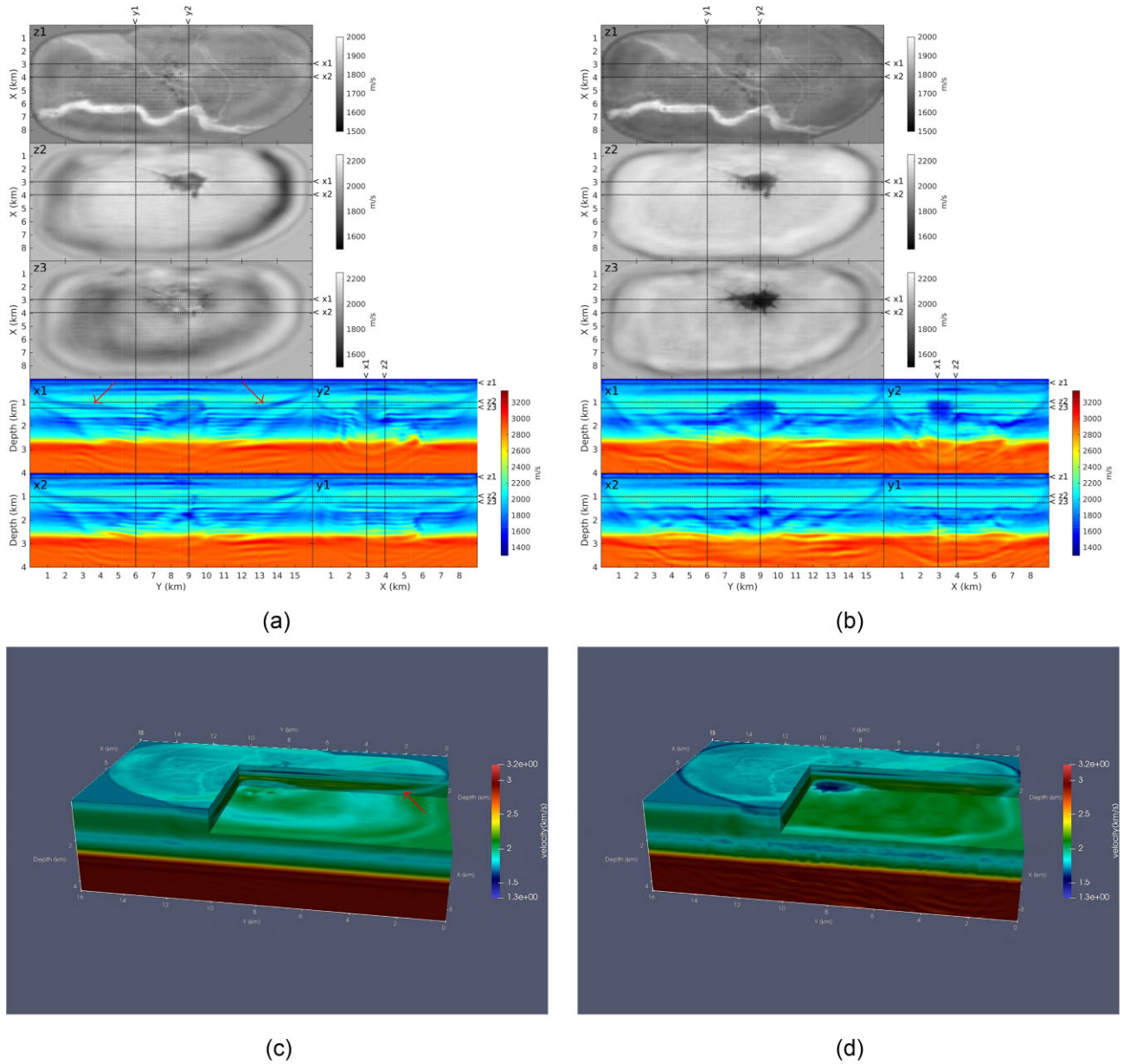


Figure 13. Inversion starting with the simple stratified model after 96 iterations: reconstructed velocity models directly using 2.5–7.0 Hz data by FWI (a) and LAWI (b). 3-D view of the reconstructed velocity models with FWI (c) and LAWI (d). For the FWI approach, the spurious bend low-velocity layers between 1.0–1.5 km depth are caused by the cycle-skipping issue (red arrows). 3-D view clearly shows that FWI is unable to recover the central low-velocity anomaly.

(2011). The attenuation effect is also taken into account with a fixed attenuation model of $Q = 1000$ in the water column and $Q = 200$ in the sediments is generated according to the root-mean square amplitudes of early arrivals (Operto *et al.* 2015; Kamath *et al.* 2021). In addition, the density models are derived from selected starting velocity models with the Gardner’s law (Gardner *et al.* 1974).

4.2.1 Workflow using the first frequency-band data with a tomographic starting model

This workflow should serve as a reference because the initial model provides in-phase predictions with the low-frequency data content. Selecting such an initial model with such a frequency band should

avoid cycle-skipping issue: FWI and LAWI approaches are supposed to be equally successful. Following the multiscale strategy (Bunks *et al.* 1995), we first run inversion on 2.5–5.0 Hz data, then on 2.5–7.0 Hz data using the previously updated velocity model. For each frequency band, 96 iterations in total are performed, namely six iterations per shot. This number of iterations represents a practical compromise between computational cost and accuracy of inversion results.

Figs. 11(a) and (b) show the inversion results of FWI and LAWI on 2.5–5.0 Hz data, and the subsequent inversion results with 2.5–7.0 Hz data are displayed in Figs. 11(c) and (d). The two methods produce similar high-quality inversion results. From the depth slices, we can clearly observe the resolution improvement compared

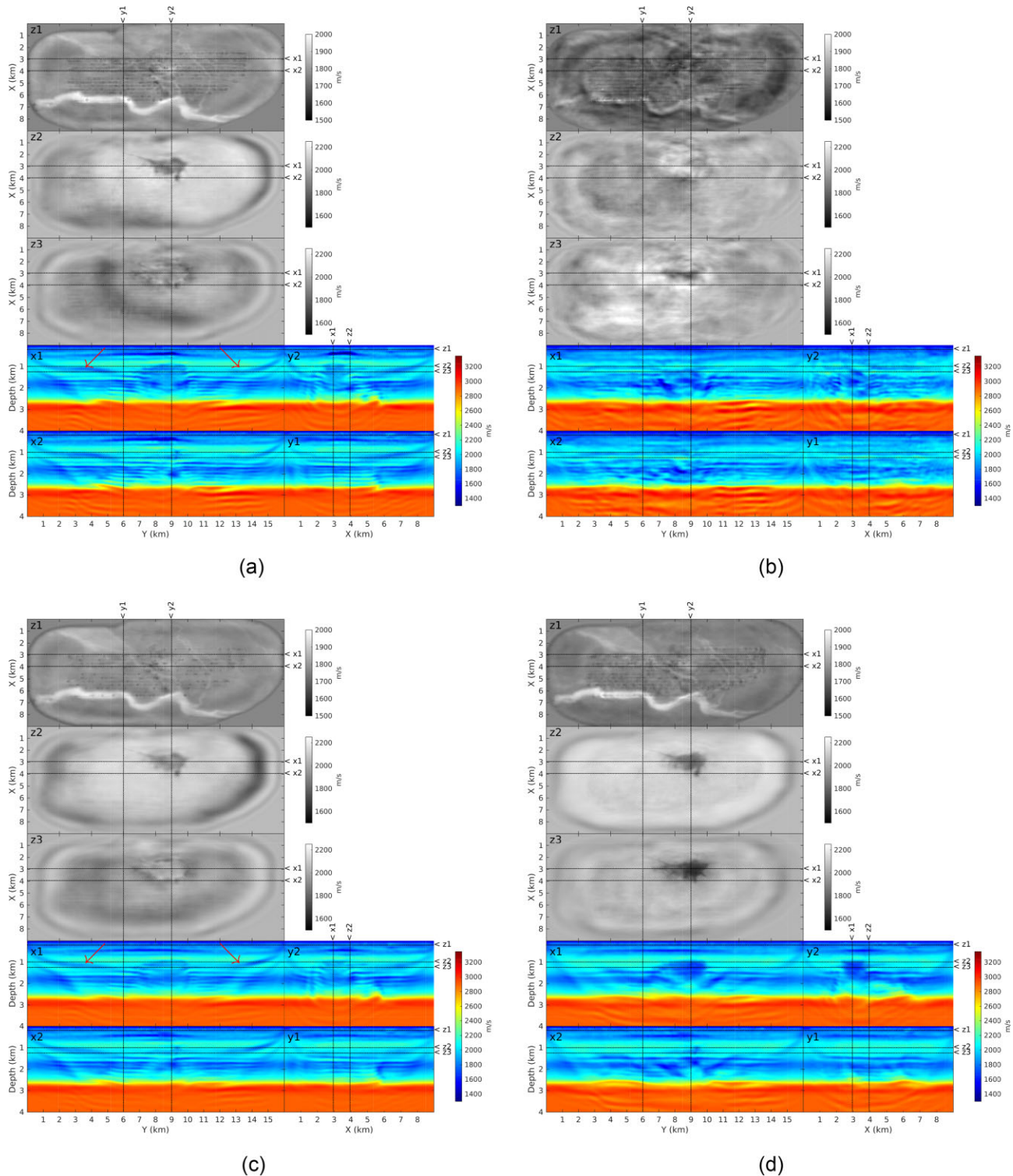


Figure 14. Inversion starting with the simple stratified model after 48 iterations: reconstructed velocity models with 2.5–7.0 Hz data using the GSOT objective function (a), the AWI method (b), the FWI method (c) and the LAWI method (d).

to the tomographic initial model: the channel-shaped high-velocity structure from the 0.2 km depth slice, and the top and middle of the low-velocity anomaly from the 1.0 and 1.25 km depth slices. Besides, we can see a vertical structure from the profile of $x = 3.95$ km. The difference between the inversion results by two meth-

ods is highlighted by red arrows. One can notice that the velocity of the corresponding part in the tomographic initial model is obtained by horizontal extrapolation from the low-velocity anomaly to the model edge. Therefore, this part in the initial tomography model could be inaccurate, and the inversion results indicates that this area

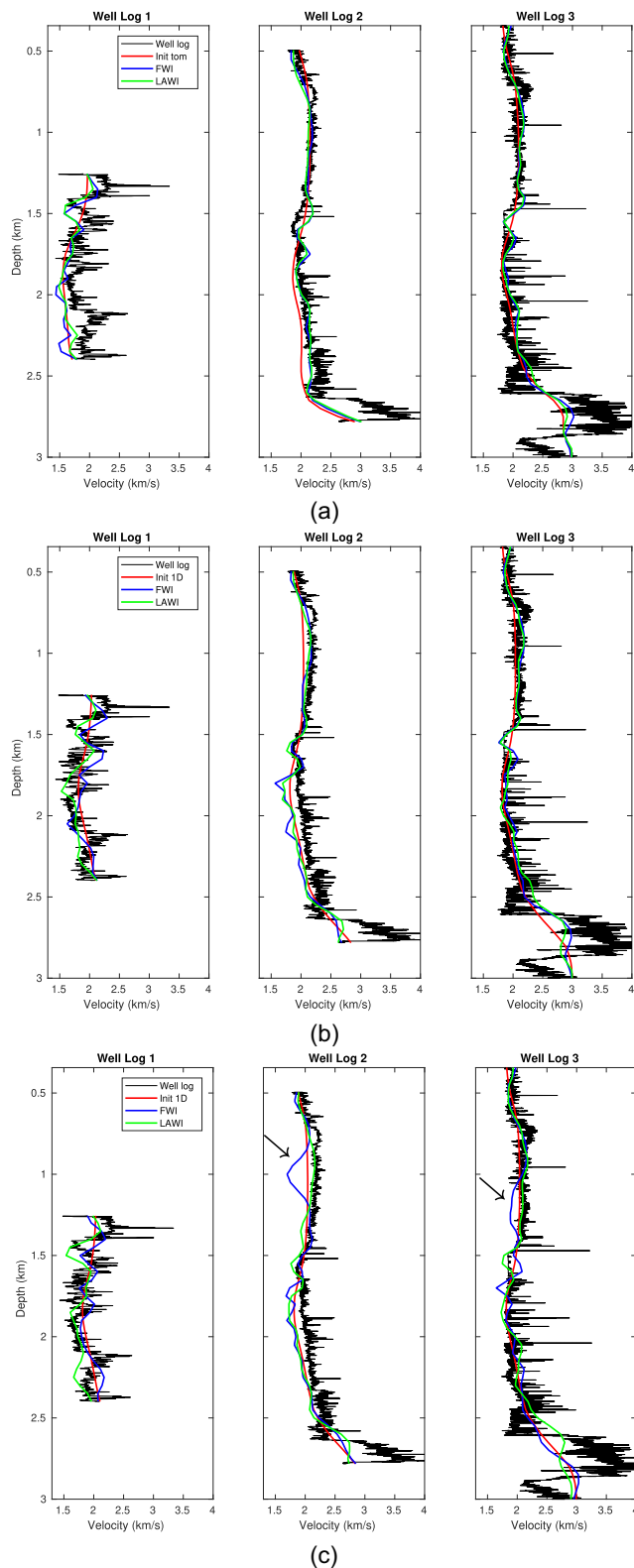


Figure 15. Well-log comparison: (a) starting with low frequency and the tomography model, (b) starting with low frequency and the stratified model and (c) starting with higher frequency and the stratified model.

should have a higher velocity, which is confirmed in the latter two tests. LAWI has a better capability to correct this intermediate-scale velocity deviation than FWI does.

In addition, one can observe some artefacts shows at the illumination boundary. When the starting velocity, close to the illumination boundary, is inaccurate, strong artefacts occur in this area, clearly found in the depth slices at 1.25 km. Both FWI and LAWI are mainly driven by diving and direct waves, the maximum depth that diving waves can reach is about 2 km. The reconstructed models below 2 km depth could be less reliable than the one above it.

4.2.2 Workflow starting with the first frequency-band data and a simple stratified starting model

The only difference with respect to the previous workflow is the stratified starting model. This starting model does not contain the low-velocity anomaly at the central location, while the background layer information is reasonably well described. The same 96 iterations are run on both 2.5–5.0 Hz and subsequently 2.5–7.0 Hz data. We are especially interested to investigate whether the central low-velocity anomaly can be reconstructed or not.

Fig. 12 presents the inversion results in this new workflow. The low-velocity anomaly can be roughly recovered from the first frequency band data (2.5–5.0 Hz). Resolution improvement is observed after using the data with the frequency-band 2.5–7.0 Hz (Figs. 12c and d). The FWI approach generates a higher value for the middle portion of the low-velocity anomaly compared to the first workflow. In contrast, the LAWI approach produces a superior reconstruction that is more similar to the outcome achieved with the previous workflow, as evidenced by the depth slice at $z = 1.25$ km. The trend of having low-velocity anomaly in the left-top area in the profile of $y = 9.0$ km is now missing: such a trend comes from an intermediate-scale velocity feature of the tomographic model. Starting from a stratified model highlights that this artificial trend is very difficult to remove, especially when it occurs at the illumination boundary.

Overall, the background information in the stratified simple model is less accurate than the one in the tomographic model. Differences of background velocity in the two starting models can be more clearly observed in the well-log plots (Fig. 15). We suspect this is the reason why stronger artefacts occur at the illumination boundary in Fig. 12. The velocity beyond the illumination boundary can hardly be updated. This would generate reflections due to the inconsistent velocity update inside and outside of the illuminated area. This poor illumination prevents the correction of this artefacts through iterations.

4.2.3 Workflow starting with the full frequency-band data and a simple stratified starting model

Inversion are performed with the full 2.5–7.0 Hz data as initial data. Moreover, the starting model is the simple stratified velocity model, making this workflow a rather challenging one. The final results after 96 iterations are presented in Fig. 13. From the depth slices, one can observe that FWI fails to reconstruct the low-velocity anomaly at the depth of 1.25 km. In fact, the velocity of the top and bottom of the low-velocity anomaly is much higher than that obtained

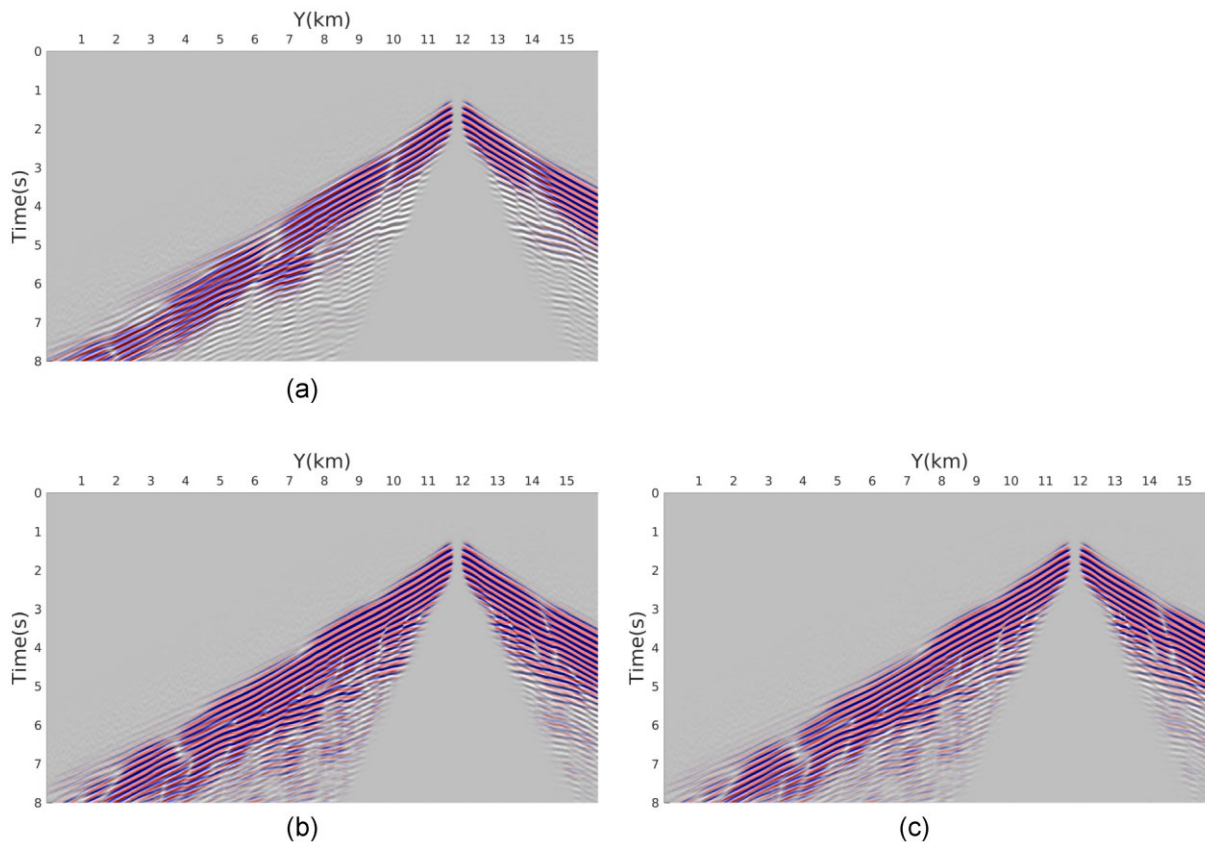


Figure 16. Data comparison for inversion starting with low frequency and the tomography model. Synthetic data are displayed in a blue-white-red colour scale, and field data (2.5–7.0 Hz) are overlapped in grey-scale with transparency. When synthetic data matches the observed data, blue pixels will be covered by black pixels. The compared data are generated with (a) the tomography model and the final reconstructed models by (b) standard FWI and (c) LAWI.

from the first workflow. On the contrary, the low-velocity anomaly reconstructed by LAWI does not depart from those obtained with the two previous workflows: an illustration of the robustness of the delta-type regularization. Besides, from Fig. 13(a), one can notice that low-velocity layers (red arrows) occurs in 1.0–1.5 km depth for the FWI approach. These bend shape low-velocity artefacts in the vertical slices are typical of cycle-skipping, and they do not appear with the LAWI approach.

For a comparison, we give the results generated by the AWI approach and the GSOT method (Métivier *et al.* 2018, 2019), with the same workflow, in this cycle-skipped circumstance. Fig. 14 displays inversion results after only 48 iterations that ensure inversion touch all data. Strong artefacts caused by cycle-skipping are observed in both velocity models built by AWI and GSOT methods (Figs. 14a and b) with similar structures of the FWI approach (Fig. 14c). The LAWI approach seems to provide superior result when considering this workflow (Fig. 14d): the main body of the central low-velocity anomaly is successfully reconstructed after 48 iterations with a reasonable results. Of course, the velocity model is reasonably improved with more iterations as shown in Fig. 13(b). No significant updates by FWI are observed between 48 iterations (Fig. 14c) and 96 iterations (Fig. 13a) because the model is stuck into a local minimum.

The effectiveness of the GSOT method for alleviating the cycle-skipping issue has been verified by the same field data application (Pladys *et al.* 2022): together with a hierarchical time-windowing strategy, the GSOT method can reconstruct similar correct velocity structures from a even cruder 1D initial velocity model (linear

increase of velocity with depth). Nevertheless, in this challenging case, without the time-windowing strategy, the GSOT method does not make inversion get rid of the local minima issue, and the quality of the GSOT result is close from the FWI result. We believe that the GSOT method is more sensitive to amplitude than the LAWI approach driven strongly by phases, thanks to amplitude normalization (Yong *et al.* 2023). Although a lot of studies show that AWI can mitigate cycle-skipping issues, it also fails in this challenging case. Furthermore, AWI could produce a very bad results when it fails to converge into a correct model, especially when predicted data lacks the reflections that occur in the observed data, because the non-stationary relationships between observed and predicted data is not taken into consideration in AWI. This observation is consistent with the findings of the 2-D synthetic study when starting with a smooth density model. It is worth noting that a similar issue related to multiple phases has been observed in the automatic cross-well tomography that relies on differential semblance optimization, as noted by Plessix *et al.* (2000). Various approaches, including the use of different data types, annihilators, and normalizations, have been attempted to mitigate this issue with field data (Plessix 2000).

4.3 Well logs

We further assess the inversion results above through well logs. There are three well logs in the survey available for quality assessment: The Log 1 is located at the centre of the acquisition survey and close to the low-velocity anomaly; the second one is at the eastern edge of the survey; the Log 3 is situated at the western of

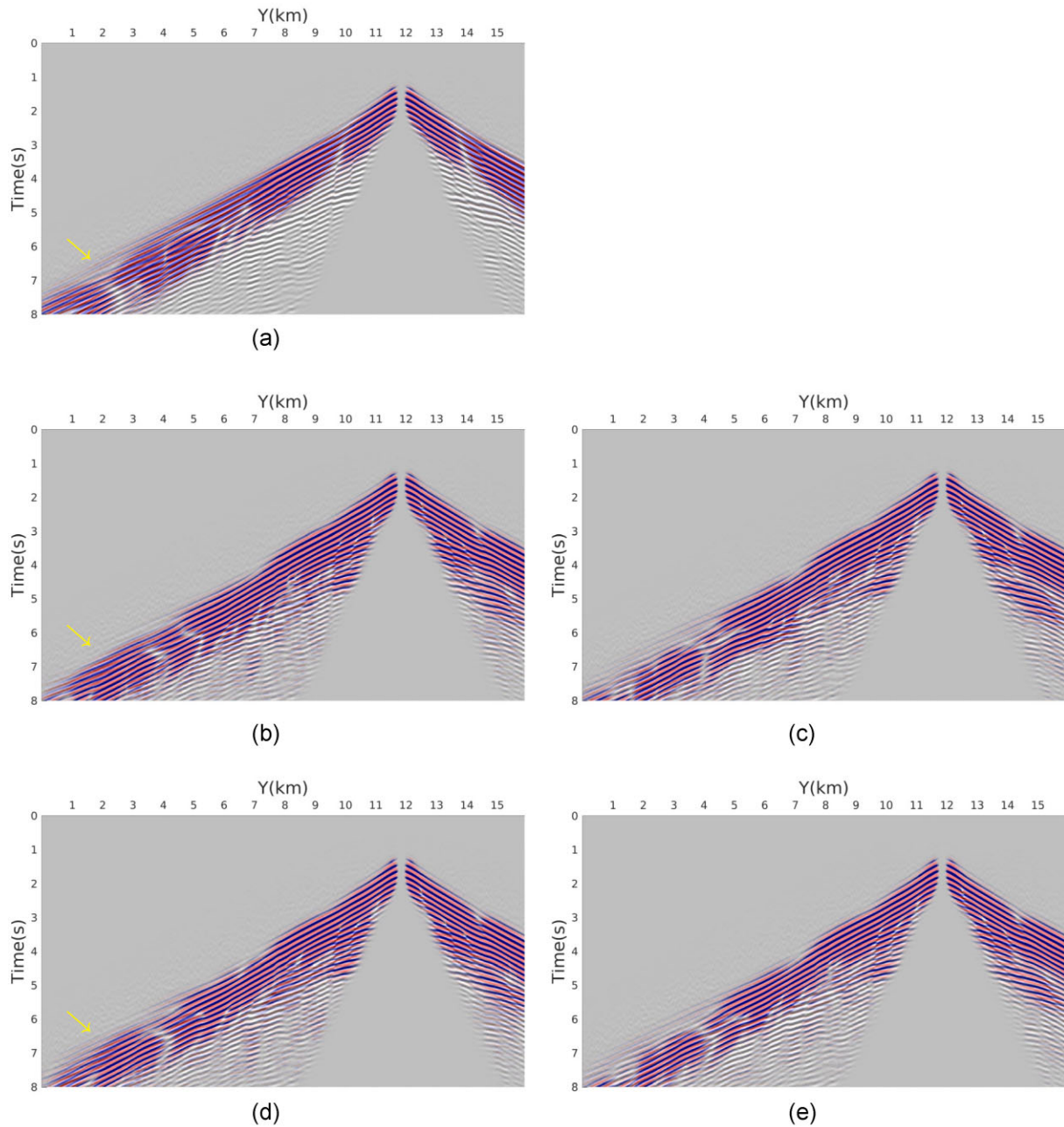


Figure 17. Data comparison for the inversion starting with the stratified initial model. The compared data are generated with (a) the stratified model, the reconstructed models by (b) FWI and (c) LAWI with low starting frequency data, and the final reconstructed models by (d) FWI and (e) LAWI with higher starting frequency data.

the survey. The details of the location of three wells can be found in Fig. 8.

Fig. 15 gives the well-log data comparison for the three workflows above. Overall, the reconstructed velocities in the first workflow improves the matching with the log data. One may notice that the initial and reconstructed velocities at the Log 1 are lower than the estimation by acoustic logging tools. The sonic log represents a velocity at a much higher frequency (about 10 kHz). However, the velocities obtained by waveform inversion is of seismic scale and influenced by the attenuative medium, which can result in slower values, due to the strong dispersion effect in an attenuative medium (Carcione 2014; Yong *et al.* 2021). A more accurate Q estimation in

the low-velocity anomaly area may improve the velocity matching at the Log 1 (Kamath *et al.* 2021).

The simple stratified starting model, used in the second and third workflows, is generated by horizontally extrapolating the velocity of the $y = 0$ km profile, therefore the velocity of the stratified model at the Log 3 is close to that in the tomographic model, and the differences can be noticed from the Log 1 and Log 2 data. The low-velocity artefacts shown in Fig. 13(a), caused by cycle-slipping issues, is highlighted by the black arrows in Fig. 15(c). Compared with two velocities reconstructed by FWI at the Log 1, one can notice there is an opposite velocity update at the depth of 1.75 km. In contrast, velocities reconstructed by LAWI have a

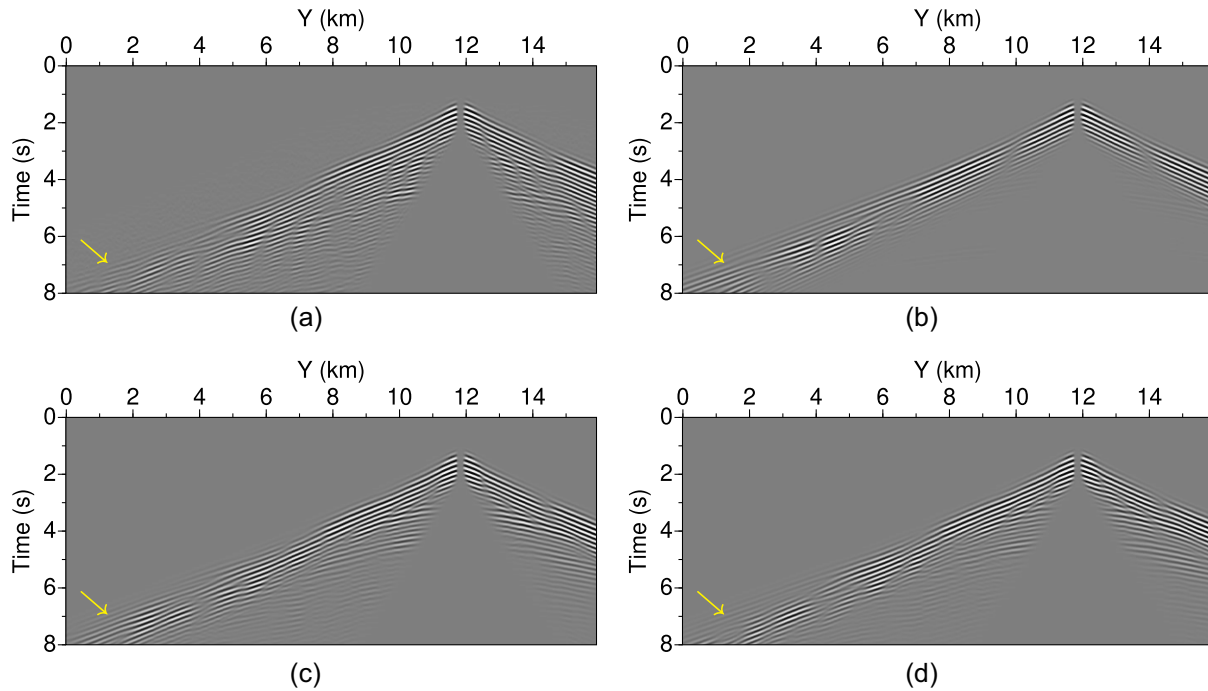


Figure 18. Data comparison for the third inversion run: (a) the observed data in 2.5–7.0 Hz, (b) predicted data by the stratified starting model, and the predicted data with the reconstructed models in the third workflow by (c) FWI and (d) LAWI, respectively. The data are that presented in Figs. 17(d) and (e).

consistent update. Besides, in Fig. 15(b), one may notice that there is a deviation at a 1.7–2.3 km depth, and the deviation is smaller in Fig. 15(c). This is related to an illumination boundary issue: the location of Log 2 is close from the edge of the acquisition, and the effective illumination area decreases with depth. As frequency and wavelength are inversely proportional to each other, waves in 2.5–5.0 Hz have a larger wavelength, which results in a larger velocity update around the illumination boundary. When moving to 2.5–7.0 Hz data, it is not able to correct velocity at the illumination boundary.

Comparing the log data from the three workflows, one can notice that LAWI generates more consistent results, which somehow demonstrates LAWI is a more stable and robust algorithm to interpret the difference between modelled and recorded data.

4.4 Data comparison

One common way to assess the quality of inversion results is to compare the predicted data generated by the reconstructed model with the observed data. Fig. 16 displays the data comparison in the inversion starting with the tomographic model. In Fig. 16(a), one can notice the phase mismatch (blue pixels) between the 2.5–7.0 Hz field data and the synthetic data from the tomographic model. These phase mismatch of diving and direct waves get corrected in Figs. 16(b) and (c): blue pixels are replaced by black pixels. Besides, more reflections are generated by the reconstructed model, while the data match of these reflections is not as good as the diving and direct waves. Overall, the reflection data of FWI result has a slight better amplitude match than that of LAWI result, because LAWI is mainly driven by phase matching and a few iterations are used in this workflow.

Fig. 17 presents the data fittings of the second and third workflow. The synthetic data generated by all of the reconstructed velocities appear similar, and all show a reasonable match to the observed

data. From previous analysis, we already know that the FWI result in the third workflow is trapped into local minima. However, the quality of the data fitting in Fig. 17(d) is quite similar to others from the previous workflow, because waveform inversion is designed to improve the data fit. The reconstructed models generally produce synthetics that resemble the observed data. Therefore, it must be careful when using data comparison to assess the quality of reconstructed models. Compared to reflections, early observed arrivals have less chance to be spuriously matched by predicted arrival than reflection phases. Thereby, they are more suitable for a quality assessment. The yellow arrows in Fig. 17 highlight the mismatch (blue parts) of early arrivals at far offsets. One can see that this mismatch occurs in the initial synthetic data, which does not exist in Fig. 16(a). FWI, in two different starting settings, is unable to reduce this phase mismatch. Conversely, LAWI results successfully correct this mismatch. This mismatch indicates that the FWI results in the second and third workflows are stuck at local minima. Although FWI in the second workflow recovers the low-velocity anomaly, its values are not as low as expected. To better observe this mismatch, we plot all synthetic and observed data involved in the third workflow in Fig. 18. In addition, we also observe that the LAWI results provide a better data fit from the cross-correlations between synthetic and field data in all three workflows as described in Supporting Information Section S1.

5 DISCUSSION

A new implementation of Gabor deconvolution is proposed in this work, which is able to bring all information of observed data in a more natural and elegant way, compared to data-weighting strategy used in our previous work (Yong et al. 2023). In addition, it can also stabilize LAWI for low-SNR data, as the proposed delta-type regularization follows the fact that the local matching filter should converge towards a delta function with model update in LAWI. For

more detailed information, please refer to Supporting Information Section S2. A slight drawback is the small resolution loss which may be overcome by taking the LAWI approach with the zero-type regularization at a later stage of the inversion, when considering higher frequencies.

The first workflow with 3-D field data demonstrates that LAWI and FWI can generate similar results when starting from an initial tomographic model and with a frequency content avoiding cycle-skipping issues. Nevertheless, the second and third workflows show that, when starting models are less accurate and/or when sufficiently low frequencies are not available, the LAWI can recover a velocity model which is similar to the one obtained by the first workflow while the FWI method cannot. Other methods, such as AWI and GSOT methods may require fine tuning for being successful, especially by data-weighting strategies. These strategies respectively reveal the importance of taking non-stationarity into account on one side and of reducing amplitude effect on the other side when designing misfit function to tackle local minima problems.

The overhead cost induced by time–frequency analysis in the LAWI approach is about 15 per cent of the computation complexity of gradient construction in the field data tests. Moreover, LAWI could need less iterations to reconstruct velocity models (see Fig. 14), which is also noticed in Yong *et al.* (2023; see Section 5.1). This means that the LAWI approach can be more efficient than the FWI method in terms of total computational time for velocity reconstruction. In short, the LAWI approach with the delta-type regularization seems to be robust, efficient, and effective in practice, therefore it should be applicable to 3-D field data over a wide range of circumstances.

High-quality velocities above 2 km depth (for instance the low-velocity anomaly) can be obtained from the field data. The deep part of the model are not sampled by diving waves: velocity variations are less reconstructed as those in the shallow part. Besides, observed reflection data are less well predicted. With more iterations, the quality of deep velocity reconstruction may be improved. However, from the wavenumber coverage analysis (Wu & Toksöz 1987; Mora 1989; Sirgue & Pratt 2004), this reconstruction will be quite difficult when considering broad-band spectra of seismic data: this challenge is also reported by several field applications (Sirgue *et al.* 2010; Messud *et al.* 2021; Pladys *et al.* 2022). Combining the LAWI approach with the so-called reflection FWI method (Xu *et al.* 2012; Zhou *et al.* 2015; Yao *et al.* 2020; Provenzano *et al.* 2023) is expected to bring a better recovery of the deep velocity model: an identified near-future investigation. Besides, when cycle-skipping issues are well-resolved, full-wavefield imaging can be realized with the FWI method (Huang *et al.* 2021). One challenge for high-frequency FWI formulation comes from computational burden, as the expense rapidly increases above 7 Hz. Furthermore, further investigations are needed to explore the impacts of shaping the source spectra, as well as the challenges posed by multiple parameters when considering anisotropy, attenuation, and elasticity.

6 CONCLUSION

FWI is a powerful tool to reconstruct high-quality subsurface images. However, due to the nonlinear nature between waveform and velocity, successful applications have a strict requirement on starting velocity models and usable lowest frequency data, which may be difficult to meet in practice. Based on the Gabor transform, the

proposed LAWI method with the delta-type and zero-type regularizations, which is mainly phase-driven, can significantly overcome the nonlinearity of velocity reconstruction by waveform inversion. Consequently, its applicability can be considered for a wider range of seismic acquisition configurations for different subsurface targets. Both synthetic and real data illustrate such potentiality during this study.

DATA AND MATERIALS AVAILABILITY

Data and code associated with this research belong to the SEISCOPE consortium. Any request related to the availability of the materials needs to be addressed to Romain Brossier (romain.brossier@univ-grenoble-alpes.fr).

ACKNOWLEDGMENTS

We thank AKERBP ASA and their partner Pandion Energy for providing the data set and permission to present this work. Peng Yong would like to thank Giuseppe Provenzano and Alexandre Hoffmann for many useful discussions on the field data processing. This study was partially funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>), sponsored by AKERBP, CGG, EXXON-MOBIL, GEOLINKS, JGI, PETROBRAS, SHELL, SINOPEC and TOTALENERGIES. This study was granted access to the HPC resources provided by the GRICAD infrastructure (<https://gricad.univ-grenoble-alpes.fr>), which is supported by Grenoble research communities, the HPC resources of Cray Marketing Partner Network (<https://partners.cray.com>) and those of IDRIS/TGCC under the allocation 046091 made by GENCI. We acknowledge careful reviews and helpful comments provided by the editor, Carl Tape, as well as the reviewer Hejun Zhu and an anonymous reviewer.

REFERENCES

- Bunks, C., Salek, F.M., Zaleski, S. & Chavent, G., 1995. Multiscale seismic waveform inversion, *Geophysics*, **60**(5), 1457–1473.
- Carcione, J.M., 2014. *Wave Fields in Real Media: Theory and Numerical Simulation of Wave Propagation in Anisotropic, Anelastic, Porous and Electromagnetic Media*, Elsevier.
- Claerbout, J.F. & Fomel, S., 2014. *Geophysical Image Estimation by Example*, Lulu. com.
- Dellinger, J. *et al.*, 2016. Wolfspar®, an “FWI-friendly” ultralow-frequency marine seismic source, in *SEG Technical Program Expanded Abstracts 2016*, pp. 4891–4895, Society of Exploration Geophysicists.
- Duvencack, E. & Bakker, P.M., 2011. Stable *P*-wave modeling for reverse-time migration in tilted TI media, *Geophysics*, **76**(2), S65–S75.
- Fichtner, A., Kennett, B. L.N., Igel, H. & Bunge, H.P., 2008. Theoretical background for continental- and global-scale full-waveform inversion in the time–frequency domain, *Geophys. J. Int.*, **175**, 665–685.
- Fomel, S., Sava, P., Vlad, I., Liu, Y. & Bashkardin, V., 2013. Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments, *J. Open Res. Softw.*, **1**(1), e8.
- Gabor, D., 1946. Theory of communication. part 1: The analysis of information, *J. Inst. Electr. Eng.-Part III: Radio Commun. Eng.*, **93**(26), 429–441.
- Gardner, G.F., Gardner, L. & Gregory, A., 1974. Formation velocity and density—the diagnostic basics for stratigraphic traps, *Geophysics*, **39**(6), 770–780.
- Guasch, L., Warner, M. & Ravaut, C., 2019. Adaptive waveform inversion: practice, *Geophysics*, **84**(3), R447–R461.
- Huang, R., Zhang, Z., Wu, Z., Wei, Z., Mei, J. & Wang, P., 2021. Full-waveform inversion for full-wavefield imaging: Decades in the making, *Leading Edge*, **40**(5), 324–334.

- Kamath, N., Brossier, R., Métivier, L., Pladys, A. & Yang, P., 2021. Multi-parameter full-waveform inversion of 3D ocean-bottom cable data from the Valhall field, *Geophysics*, **86**(1), B15–B35.
- Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migrations, in *Conference on Inverse Scattering, Theory and Application*, pp. 206–220, Society for Industrial and Applied Mathematics, Philadelphia.
- Levander, A.R., 1988. Fourth-order finite-difference P - SV seismograms, *Geophysics*, **53**(11), 1425–1436.
- Luo, S. & Sava, P., 2011. A deconvolution-based objective function for wave-equation inversion, *SEG Technical Program Expanded Abstracts 2011*, pp. 2788–2792. Society of Exploration Geophysicists.
- Luo, Y. & Schuster, G.T., 1991. Wave-equation traveltime inversion, *Geophysics*, **56**(5), 645–653.
- Luo, Y., Ma, Y., Wu, Y., Liu, H. & Cao, L., 2016. Full-traveltime inversion, *Geophysics*, **81**(5), R261–R274.
- Margrave, G.F., 1998. Theory of nonstationary linear filtering in the Fourier domain with application to time-variant filtering, *Geophysics*, **63**(1), 244–259.
- Margrave, G.F., Lamoureux, M.P. & Henley, D.C., 2011. Gabor deconvolution: Estimating reflectivity by nonstationary deconvolution of seismic data, *Geophysics*, **76**(3), W15–W30.
- Messud, J., Carotti, D., Hermant, O., Sedova, A. & Lambaré, G., 2021. Optimal transport full-waveform inversion: from theory to industrial applications with examples from the sultanate of oman, *First Break*, **39**(12), 45–53.
- Métivier, L. & Brossier, R., 2016. The SEISCOPE optimization toolbox: a large-scale nonlinear optimization library based on reverse communication, *Geophysics*, **81**(2), F11–F25.
- Métivier, L., Allain, A., Brossier, R., Mérigot, Q., Oudet, E. & Virieux, J., 2018. Optimal transport for mitigating cycle skipping in full waveform inversion: a graph space transform approach, *Geophysics*, **83**(5), R515–R540.
- Métivier, L., Brossier, R., Mérigot, Q. & Oudet, E., 2019. A graph space optimal transport distance as a generalization of L^p distances: application to a seismic imaging inverse problem, *Inverse Probl.*, **35**(8), 085001.
- Mocco, P. & Kristek, J., 2005. On the rheological models used for time-domain methods of seismic wave propagation, *Geophys. Res. Lett.*, **32**(1), L01306.
- Mora, P.R., 1989. Inversion = migration + tomography, *Geophysics*, **54**(12), 1575–1586.
- Operto, S., Miniussi, A., Brossier, R., Combe, L., Métivier, L., Monteiller, V., Ribodetti, A. & Virieux, J., 2015. Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation, *Geophys. J. Int.*, **202**(2), 1362–1391.
- Pan, W., Geng, Y. & Innanen, K.A., 2018. Interparameter trade-off quantification and reduction in isotropic-elastic full-waveform inversion: synthetic experiments and hussar land data set application, *Geophys. J. Int.*, **213**(2), 1305–1333.
- Pladys, A., Brossier, R., Li, Y. & Métivier, L., 2021. On cycle-skipping and misfit function modification for full-wave inversion: comparison of five recent approaches, *Geophysics*, **86**(4), R563–R587.
- Pladys, A., Brossier, R., Kamath, N. & Métivier, L., 2022. Robust FWI with graph space optimal transport: application to 3D OBC Valhall data, *Geophysics*, **87**(3), 1–76.
- Plessix, R.-E., 2000. Automatic cross-well tomography: an application of the differential semblance optimization to two real examples, *Geophys. Prospect.*, **48**(5), 937–951.
- Plessix, R.E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.*, **167**(2), 495–503.
- Plessix, R.E., Mulder, W.A. & ten Kroode, A. P.E., 2000. Automatic cross-well tomography by semblance and differential semblance optimization: theory and gradient computation, *Geophys. Prospect.*, **48**(5), 913–935.
- Plessix, R.-E., Baeten, G., de Maag, J.W. & ten Kroode, F., 2012. Full waveform inversion and distance separated simultaneous sweeping: a study with a land seismic data set, *Geophys. Prospect.*, **60**, 733–747.
- Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain, part I: theory and verification in a physical scale model, *Geophysics*, **64**, 888–901.
- Prieux, V., Brossier, R., Gholami, Y., Operto, S., Virieux, J., Barkved, O. & Kommedal, J., 2011. On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study, *Geophys. J. Int.*, **187**, 1495–1515.
- Provenzano, G., Brossier, R. & Métivier, L., 2023. Robust and efficient waveform-based velocity-model-building by optimal-transport in the pseudotime domain: methodology, *Geophysics*, **88**(2), U49–U70.
- Sedova, A., Messud, J., Prigent, H., Royle, G. & Lambaré, G., 2019. Acoustic land full waveform inversion on a broadband land dataset: the impact of optimal transport, in *Expanded Abstracts, 81th Annual EAGE Meeting (London)*, EAGE.
- Shen, P. & Symes, W.W., 2008. Automatic velocity analysis via shot profile migration, *Geophysics*, **73**(5), VE49–VE59.
- Shen, X., Ahmed, I., Brenders, A., Dellinger, J., Etgen, J. & Michell, S., 2018. Full-waveform inversion: the next leap forward in subsalt imaging, *Leading Edge*, **37**(1), 67b1–67b6.
- Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics*, **69**(1), 231–248.
- Sirgue, L., Barkved, O.I., Dellinger, J., Etgen, J., Albertin, U. & Kommedal, J.H., 2010. Full waveform inversion: the next leap forward in imaging at Valhall, *First Break*, **28**, 65–70.
- Strang, G. & Nguyen, T., 1996. *Wavelets and Filter Banks*, SIAM.
- Sun, B. & Alkhalifah, T.A., 2020. Joint minimization of the mean and information entropy of the matching filter distribution for a robust misfit function in full-waveform inversion, *IEEE Trans. Geosci. Remote Sens.*, **58**(7), 4704–4720.
- Symes, W.W. & Carazzone, J.J., 1991. Velocity inversion by differential semblance optimization, *Geophysics*, **56**, 654–663.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**(8), 1259–1266.
- Tarantola, A., 1986. A strategy for non linear inversion of seismic reflection data, *Geophysics*, **51**(10), 1893–1903.
- Thomsen, L.A., 1986. Weak elastic anisotropy, *Geophysics*, **51**, 1954–1966.
- Tikhonov, A., Goncharky, A., Stepanov, V. & Yagola, A., 2013. *Numerical Methods for the Solution of Ill-posed Problems*, Springer Science & Business Media.
- van Leeuwen, T. & Mulder, W.A., 2010. A correlation-based misfit criterion for wave-equation traveltime tomography, *Geophys. J. Int.*, **182**(3), 1383–1394.
- Vigh, D., Cheng, X., Jiao, K., Kang, W. & Brand, N., 2021. The impact of acquisition geometry on full-waveform inversion updates, *Leading Edge*, **40**(5), 335–341.
- Virieux, J. & Operto, S., 2009. An overview of full waveform inversion in exploration geophysics, *Geophysics*, **74**(6), WCC1–WCC26.
- Warner, M. & Guasch, L., 2016. Adaptive waveform inversion: Theory, *Geophysics*, **81**(6), R429–R445.
- Warner, M. et al., 2013. Anisotropic 3D full-waveform inversion, *Geophysics*, **78**(2), R59–R80.
- Wu, R.S. & Toksöz, M.N., 1987. Diffraction tomography and multisource holography applied to seismic imaging, *Geophysics*, **52**, 11–25.
- Xu, S., Wang, D., Chen, F., Lambaré, G. & Zhang, Y., 2012. Inversion on reflected seismic wave, *SEG Technical Program Expanded Abstracts 2012*, pp. 1–7, Society of Exploration Geophysicists.
- Yang, P., Brossier, R., Métivier, L., Virieux, J. & Zhou, W., 2018. A time-domain preconditioned truncated Newton approach to multiparameter visco-acoustic full waveform inversion, *SIAM J. Sci. Comput.*, **40**(4), B1101–B1130.
- Yao, G., Wu, D. & Wang, S.-X., 2020. A review on reflection-waveform inversion, *Pet. Sci.*, **17**(2), 334–351.
- Yong, P., Brossier, R., Métivier, L. & Virieux, J., 2021. Q estimation by full-waveform inversion: analysis and misfit functions comparison, in *First International Meeting for Applied Geoscience & Energy*, pp. 827–831, Society of Exploration Geophysicists.
- Yong, P., Brossier, R. & Métivier, L., 2022. Robust localized adaptive waveform inversion: a new regularization for gabor deconvolution, in *Second*

International Meeting for Applied Geoscience & Energy, pp. 972–976, Society of Exploration Geophysicists.

Yong, P., Brossier, R., Métivier, L. & Virieux, J., 2023. Localized adaptive waveform inversion: theory and numerical verification, *Geophys. J. Int.*, **233**(2), 1055–1080.

Zhang, Z., Mei, J., Lin, F., Huang, R. & Wang, P., 2018. Correcting for salt misinterpretation with full-waveform inversion, in *SEG Technical Program Expanded Abstracts 2018*, pp. 1143–1147, Society of Exploration Geophysicists.

Zhou, W., Brossier, R., Operto, S. & Virieux, J., 2015. Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation, *Geophys. J. Int.*, **202**(3), 1535–1554.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. Zero-lag cross-correlation analysis between field data and synthetic data (2.5–7.0 Hz) in the inversion starting with the tomography model: the synthetic data are generated with (a) the tomography model and the final reconstructed models by (b) standard FWI and (c) LAWI.

Figure S2. Zero-lag cross-correlation analysis between field data and synthetic data (2.5–7.0 Hz) in the inversion starting with the stratified model: the synthetic data are generated with (a) the stratified model and the reconstructed models by (b) FWI and (c) LAWI with low starting frequency data, and the final reconstructed models by (d) FWI and (e) LAWI with higher starting frequency data.

Figure S3. Chevron benchmark test in 0–3 Hz frequency band: observed data (a), predicted data (b), and the adjoint sources of LAWI with zero-type regularization (c) and delta-type regularization (d).

Figure S4. LAWI using zero-type regularization: initial model (a), reconstructed models after the first iteration (b), second iteration (c) and third iteration (d).

Figure S5. Reconstructed models by LAWI using delta-type regularization (a) and AWI using zero-type regularization (b) after the 15th iteration. The resolution loss is not obvious due to the limited low-frequency band.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX A: GRADIENT CONSTRUCTION BY THE ADJOINT-STATE METHOD

We give the acoustic wave equation used in this work, in which seismic attenuation is expressed by the generalized Zener body rheology (Moczo & Kristek 2005) and anisotropy effect is addressed by the stable first-order wave equation in transverse isotropy (TI) with a vertical symmetry axis (VTI), proposed by Duveneck & Bakker (2011). With the same symbols used in the previous study (Yang *et al.* 2018), the VTI viscoacoustic wave equation is writes

as

$$\begin{cases} \rho \partial_t \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} \partial_x & 0 \\ \partial_y & 0 \\ 0 & \partial_z \end{bmatrix} \begin{bmatrix} g \\ q \end{bmatrix} + \begin{bmatrix} f_{v_x} \\ f_{v_y} \\ f_{v_z} \end{bmatrix} \\ \partial_t \begin{bmatrix} g \\ q \end{bmatrix} = \begin{bmatrix} c_{11} & c_{13} \\ c_{13} & c_{33} \end{bmatrix} \begin{bmatrix} \partial_x & \partial_y & 0 \\ 0 & 0 & \partial_z \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} - \sum_{\ell=1}^L Y_{\ell} \begin{bmatrix} c_{11} & c_{13} \\ c_{13} & c_{33} \end{bmatrix} \begin{bmatrix} \xi_{\ell}^g \\ \xi_{\ell}^q \end{bmatrix} + \begin{bmatrix} f_g \\ f_q \end{bmatrix} \\ \partial_t \begin{bmatrix} \xi_{\ell}^g \\ \xi_{\ell}^q \end{bmatrix} = -\omega_{\ell} \begin{bmatrix} \xi_{\ell}^g \\ \xi_{\ell}^q \end{bmatrix} + \omega_{\ell} \begin{bmatrix} \partial_x & \partial_y & 0 \\ 0 & 0 & \partial_z \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}. \end{cases} \quad (\text{A1})$$

Here, $\mathbf{v} = (v_x, v_y, v_z)^T$ are particle velocities. g and q are horizontal and vertical stresses. ξ_{ℓ} are known as memory variables, each one associated with one reference frequency ω_{ℓ} . $Y_{\ell} \approx y_{\ell} Q_{\text{inv}}$ ($Q_{\text{inv}} = 1/Q$) with the separable approximation for the anelastic coefficients, and y_{ℓ} is computed before forward simulation by solving a least-squares problem

$$\min_{y_{\ell}} \frac{1}{2} \int_{\omega \in \bar{\Omega}} \left(\sum_{\ell=1}^n y_{\ell} \frac{\omega \omega_{\ell}}{\omega_{\ell}^2 + \omega^2} - 1 \right)^2 d\omega, \quad (\text{A2})$$

where $\bar{\Omega} = [\omega_{\min}, \omega_{\max}]$, and for three relaxation mechanisms, ω_{ℓ} are often chosen as ω_{\min} , ω_{\max} and $\sqrt{\omega_{\min} \omega_{\max}}$. The element of matrix C is related to the elastic coefficients

$$\begin{aligned} c_{11} &= \rho V_p^2 (1 + 2\epsilon) = \rho V_h^2 & c_{33} &= \rho V_p^2 = \kappa \\ c_{13} &= c_{33} \sqrt{1 + 2\delta} = \rho V_p^2 \sqrt{1 + 2\delta}, \end{aligned} \quad (\text{A3})$$

where V_p and V_h denote vertical and horizontal velocities of wave propagation. ϵ and δ are Thomsen's anisotropy parameters (Thomsen 1986). When considering isotropic media, they are simply set as zero in this study.

For the sake of simplicity, we rewrite the forward modelling equation in a compact form as

$$w : \begin{cases} \rho \partial_t v = -D^{\dagger} \sigma + f_v \\ C^{-1} \partial_t \sigma = Dv - \sum_{\ell=1}^L Y_{\ell} \xi_{\ell} + C^{-1} f_{\sigma} \\ \frac{1}{\omega_{\ell}} \partial_t \xi_{\ell} = -\xi_{\ell} + Dv, \ell = 1, \dots, L \end{cases} \quad (\text{A4})$$

With the adjoint-state method (Plessix 2006), the adjoint equation can be given by

$$\lambda : \begin{cases} \rho \partial_t \bar{v} + D^{\dagger} \bar{\sigma} + \sum_{\ell=1}^L D^{\dagger} \bar{\xi}_{\ell} = 0 \\ C^{-1} \partial_t \bar{\sigma} - D \bar{v} = r \\ \frac{1}{\omega_{\ell}} \partial_t \bar{\xi}_{\ell} - \bar{\xi}_{\ell} - Y_{\ell} \bar{\sigma} = 0, \ell = 1, \dots, L. \end{cases} \quad (\text{A5})$$

$r = \frac{\partial \chi}{\partial \sigma}$ denotes the adjoint source at the receiver locations. For standard FWI, it is the subtraction between calculated and observed data, whereas, in LAWI approach, it is given by the formula (12). The gradient of the vertical velocity V_p , of interest in this work, can be written as

$$\frac{\partial \chi}{\partial V_p} = \frac{\partial \kappa_{\text{inv}}}{\partial V_p} \frac{\partial \chi}{\partial \kappa_{\text{inv}}} = \frac{-2\rho V_p}{\kappa^2} \int_0^T dt \bar{\sigma}^{\dagger} \frac{\partial C^{-1}}{\partial \kappa_{\text{inv}}} \partial_t \sigma, \quad (\text{A6})$$

where the explicit formula of $\frac{\partial \chi}{\partial \kappa_{\text{inv}}}$ can be expressed as (Yang *et al.* 2018)

$$\frac{\partial \chi}{\partial \kappa_{\text{inv}}} = \begin{cases} \int_0^T dt \bar{p} \partial_t p, (\epsilon = \delta = 0; p = -g = -q), \\ \frac{1}{2(\epsilon - \delta)} \int_0^T dt (\bar{g} - \sqrt{1 + 2\delta} \bar{q}) \partial_t g + (-\sqrt{1 + 2\delta} \bar{g} + (1 + 2\epsilon) \bar{q}) \partial_t q (\epsilon \neq \delta). \end{cases} \quad (\text{A7})$$