3-D multicomponent full waveform inversion for shallow-seismic target: Ettlingen Line case study

Theodosius M Irnaka, Romain Brossier, Ludovic Métivier, Thomas Bohlen, Yudi Pan

Geophysical Journal International

Geophys. J. Int. (2022) **229,** 1017–1040
Advance Access publication 2021 December 17
GJI Marine Geosciences and Applied Geophysics

https://doi.org/10.1093/gji/ggab512

3-D multicomponent full waveform inversion for shallow-seismic target: Ettlingen Line case study

T.M. Irnaka [1,2,*] R. Brossier,[1] L. Métivier,[3] T. Bohlen[4] and Y. Pan [4]

[1] *Université Grenoble Alpes, ISTerre, F-*38000 *Grenoble, France. E-mail: marwan.irnaka@ugm.ac.id*
[2] *Universitas Gadjah Mada, Yogyakarta* 55281*, Indonesia*
[3] *Université Grenoble Alpes, CNRS, LJK,* 38000 *Grenoble, France*
[4] *Geophysical Institute, Karlsruhe Institute of Technology,* 76049 *Karlsruhe, Germany*

Accepted 2021 December 15. Received 2021 December 8; in original form 2020 December 1

SUMMARY

In this work, we investigate methodological development and application of viscoelastic full waveform inversion to a multicomponent source and receiver data set at the near-surface scale. The target is located in Germany and lies in an ancient war trench from the War of the Spanish Succession that was buried at approximately 1 m depth. We present the pre-processing of the data, including a matching filter correction to compensate for different source and receiver coupling conditions during the acquisition, as well as a dedicated multistep workflow for the reconstruction of both *P*-wave and *S*-wave velocities. Our implementation is based on viscoelastic modelling using a spectral element discretization to accurately account for the wave propagation complexity in this shallow region. We illustrate the inversion stability by starting from different initial models, either based on dispersion curve analysis or homogeneous models consistent with first arrivals. We recover similar results in both cases. We also illustrate the importance of accounting for the attenuation by comparing elastic and viscoelastic results. The 3-D results make it possible to recover and locate the trench precisely in terms of interpretation. They also exhibit another trench structure, in a direction forming an angle at 45° with the direction of the targeted trench. This new structure had been previously interpreted as an artifact in former 2-D inversion results. The archaeological interpretation of this new structure is still a matter of discussion.

Key words: Waveform inversion; Body waves; Surface waves and free oscillations; Wave propagation.

1 INTRODUCTION

Accurate and efficient seismic imaging techniques are crucial for near surface applications. Geotechnical applications (Kramer 1996; Stokoe *et al.* 2000), ground characterization (Foti *et al.* 2003; Roberts & Asten 2004; Chapman *et al.* 2006), infrastructure planning (Stewart *et al.* 1997; Martínez & Mendoza 2011; Pegah & Liu 2016), subsurface feature detection (Cardarelli *et al.* 2010), agriculture (Allred *et al.* 2008; Weil *et al.* 2012), archaeological studies (Wynn 1986; Vafidis *et al.* 2003; Signanini & Torrese 2004) and shallow-seismic hazard assessment (Göktürkler *et al.* 2008; Samyn *et al.* 2012) are instances of applications which focus on shallow targets from a few metres to several tens of metres deep.

Conventional methods for this type of application encompass seismic refraction (Palmer 1980), first arrival traveltime tomography (FATT, Aki *et al.* 1974) and multichannel analysis of surface waves (MASW, Park *et al.* 1999). All these methods suffer from several limitations. The seismic refraction technique uses a very crude assumption by assuming a layered subsurface, with homogeneous layers, and the velocity in each layer increases with depth. Interpreting first arrivals using Snell's law aims to infer the velocity of each layer. FATT, which only considers the first arrival signal (from *P* or *S* wave), is only sensitive to the direct and diving waves. In the context of shallow-seismic, FATT is limited because the maximum offset for such application is often too short to sample the subsurface at sufficient depth. In addition, FATT requires picking of *P* or *S* waves, and the resolution depends on the distribution of the ray path. MASW focuses on the surface waves

*ISTerre, France

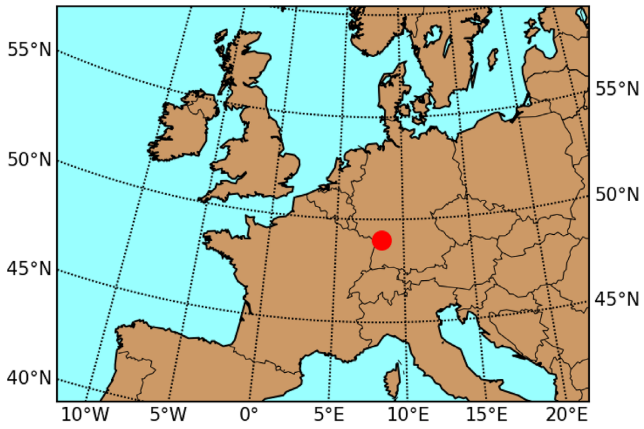

Figure 1. Rheinstetten is located at the southwest part of Germany, close to the border of France. Red circle indicates the location of the seismic acquisition.

instead of interpreting body waves. It takes advantage of the dispersive behavior of the surface waves. MASW is generally limited to the reconstruction of V_S, based on the assumption of a local 1-D model, and relies on a pre-processing stage which transfers the seismic data to the frequency–velocity domain that makes possible to pick the phase or group velocity.

Another seismic imaging method called full waveform inversion (FWI, Tarantola 1984; Virieux & Operto 2009) has started to gain popularity for shallow-seismic applications. FWI has already been successfully applied to many crustal-scale targets (Sirgue *et al.* 2010; Etienne *et al.* 2012; Prieux *et al.* 2013; Vigh *et al.* 2014; Górszczyk *et al.* 2017; He *et al.* 2019; Trinh *et al.* 2019a), as well as in seismology for lithospheric scale study (Tape *et al.* 2010; Beller *et al.* 2018; Lu *et al.* 2018) or regional and global tomography (Fichtner *et al.* 2008; Modrak & Tromp 2016). FWI is based on the iterative minimization of the difference between the observed data and the calculated data through the solution of wave propagation equations. Thus, it aims at interpreting the whole signal, contrary to the previously mentioned seismic imaging methods.

The interest for FWI comes from its high-resolution power up to half of the wavelength, the fact that it does not rely on prior assumptions regarding subsurface geometry/configuration, and the possibility of estimating 3-D quantitative subsurface models. Conventional difficulties for FWI applications are related to the design of sufficiently accurate initial subsurface models, and the risk of converging to local minima (also referred to as cycle skipping issue in the FWI community as discussed in Virieux & Operto 2009), as well as the need for a sufficiently accurate modelling engine to take into account the complexity of wave propagation and the consequently higher computational cost of the whole procedure.

Even if the lack of low frequency may not be so strong for shallow targets, the application of 3-D FWI in the near-surface context usually suffers from two specific difficulties. The first is related to the seismic acquisition, the targets are often sparsely covered due to a limited number of available sources and receivers. 2-D profiles are often used rather than full 3-D acquisition. The signal-to-noise ratio (SNR) is often poor due to weak seismic sources (sledgehammer, small vibroseis), the sources generally lack repeatability, and variable receiver-soil coupling effects can be observed. The second difficulty is due to the high complexity of near-surface elastic wave propagation, where highly dispersive surface waves dominate the recorded signal, and attenuation effects can be strong. Despite these difficulties, several 2-D FWI studies in shallow-seismic have been performed on synthetic and field data (Bretaudeau *et al.* 2013; Köhn *et al.* 2016, 2019; Wittkamp *et al.* 2018; Lamert & Friederich 2019; Wang *et al.* 2019). The extension to the 3-D case has also been proven to be feasible (Fathi *et al.* 2016; Nguyen & Tran 2018; Smith *et al.* 2019). One of the most recent examples, Smith *et al.* (2019) performed 3-D elastic FWI for tunnel detection on synthetic and field data. They use the SPECFEM package as the waveform modelling tool (Peter *et al.* 2011) and SeisFlows as the inversion engine (Modrak *et al.* 2018). They use multicomponent geophones with vertical force sources. Their initial models are obtained from traveltime tomography for V_P and MASW for V_S. Up to 20 Hz, they can reconstruct a 3-D model of the underground tunnel using field data. They also emphasize the use of horizontal source components to improve the deep reconstruction in their synthetic test.

Our study focuses on a specific 3-D data set that has been acquired to investigate a very shallow target, the Ettlingen Line, Germany (Figs 1 and 2). This target is a defensive trench which was firstly built in 1707 during the War of the Spanish Succession (Lang *et al.* 1907). The expected investigation depth reaches only 6 m. Compared with more conventional acquisitions for this target type, we have the advantage of a rather dense coverage, implying a total number of 36 source and 888 receiver positions, for a maximum offset of 41 m. In addition, the acquisition uses 3-component (3C) receivers and 3C sources. In this context, the recorded seismic signal mainly contains surface waves.

Contrary to crustal exploration case-studies, the frequency content of the data is sufficiently low to avoid the cycle skipping issues, as only one to two seismic wavelengths propagates across the target at the lowest available frequency, relaxing the constraint of accurate starting model requirement. Therefore, we show how we can perform FWI from homogeneous V_P and V_S initial models. Starting from such simple models considerably reduces the complexity of the FWI workflow, which conventionally relies on a substantial pre-processing step to build reasonable accurate initial velocity models, using tomography or MASW techniques. Another difficulty is in accurately inverting the surface waves to extract the information about V_P and V_S. We show how, based on an accurate spectral element modelling and multisteps inversion strategies, we can set up an efficient multiparameter inversion workflow, making it possible to retrieve high-resolution 3-D models for both

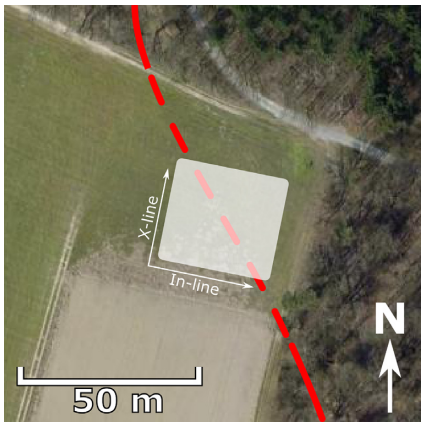

Figure 2. Basemap of the acquisition at Rheinstetten. Red solid line represents the existing location of the trench, red dashed line represents the possible buried location of the trench line, and the white rectangle represents seismic acquisition location.

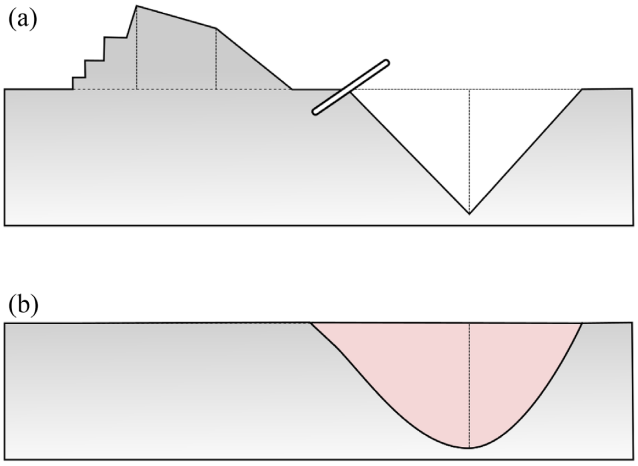

Figure 3. The original shape of the trench (a) and the current shape of buried trench (b) redrawn from Lang *et al.* (1907). In general, the current shape is smoother due to erosion. The wood palisade wall also do not exist anymore. Our target is the buried trench in which we do not know anymore the boundary and the geometry of the original trench.

parameters. The obtained results reveal a detailed 3-D structure of the Ettlingen Line, compatible with the results obtained in previous studies. The reconstructed models also exhibit a new low-velocity trench, which was previously considered to be an artifact in a 2-D FWI inversion (Wittkamp *et al.* 2018), and which was not visible on a previous MASW inversion. The 3-D coherency of this new trench makes it a very plausible structure that was previously overlooked on this site. We also discuss the inversion stability regarding the integration of attenuation in the numerical modelling engine through different numerical experiments.

2 DATA ACQUISITION AND PRE-PROCESSING

2.1 Presentation of the Ettlingen Line near Rheinstetten

The target of this study is a historical defensive trench located in Rheinstetten, Germany. The military forces that fought for the Habsburg emperor Joseph I built it up during the War of the Spanish Succession in 1707 (Lang *et al.* 1907). It had been leveled to ground in the area of our investigation (Fig. 3). The trench is surrounded by Pleistocene fluviatile sediment deposits from the Rhine River, west of the trench. Several experiments had been performed in this area to uncover the geometry of the buried trench. Wegscheider (2017) revealed the trench 2-D shape by performing a ground penetrating radar (GPR) investigation. Wittkamp *et al.* (2018) estimated a 2-D V_S model using 2-D elastic FWI focused on Love and Rayleigh waves. Pan *et al.* (2018) performed a detailed 3-D MASW, producing a 3-D V_S cube. From these studies, the Ettlingen Line emerged as an inverted triangular trench with a lower velocity than the surrounding area (consistent with a weaker consolidation of the subsurface in the trench due to the excavation prior to backfill). The width of the trench is approximately 6 m, while the depth is about 3 m.

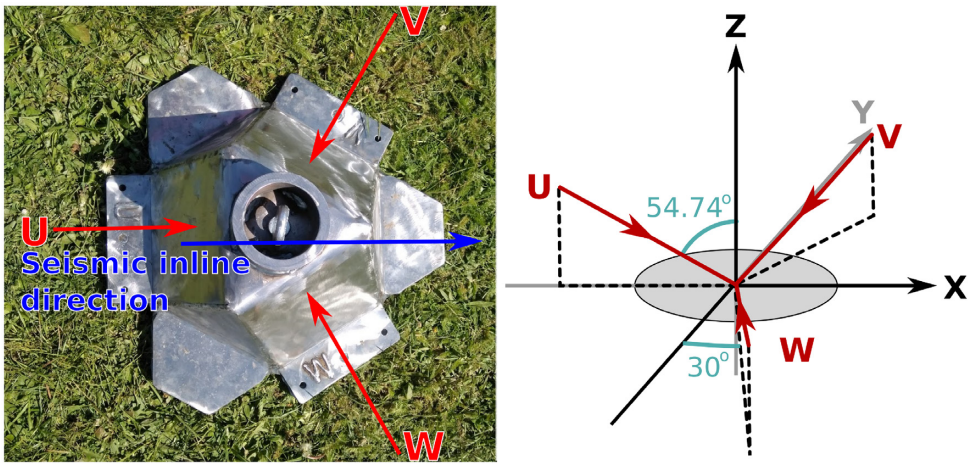

Figure 4. The Galperin source used in 9C seismic acquisition (a), and its schematic explaining the three principal orthogonal direction of the source (b) (Häusler *et al.* 2018).

Table 1. Cartesian point source weights to represent a directional point source.

Source direction	F_Z	F_X	F_Y
U	$F\cos\theta$	$F\sin\theta$	0
V	$F\cos\theta$	$-F\sin\theta\cos\phi$	$-F\sin\theta\sin\phi$
W	$F\cos\theta$	$-F\sin\theta\cos\phi$	$F\sin\theta\sin\phi$

2.2 Acquisition

The seismic data acquisition was organized by GPI Karlsruhe Institute of Technology (KIT), in collaboration with GFZ Potsdam, ETH Zurich and Univ. Grenoble Alpes (UGA) in April 2017. A 3-D seismic acquisition with 3C Galperin sources and 3C receivers was performed, resulting in a 9C seismic data (Pan, 2021).

A Galperin source is a multicomponent source designed for shallow-seismic applications (Häusler *et al.* 2018). In practice, the Galperin source acts as the base plate and needs to be used together with a sledgehammer. It has three different source directions (U, V and W), which are not aligned with the Cartesian coordinate axis. An illustration of this particular type of source is given in Fig. 4(a). It is built from iron and filled with wood inside. The wood inside the Galperin source helps to reduce the reverberation during seismic acquisition. The orientation of each source direction is depicted in Fig. 4(b). The source directions (U, V and W) form 120° angles with each other, with respect to the horizontal plane. Furthermore, all source directions form an angle θ of 54.74° with the vertical axis (Z). Source directions V and W form an angle ϕ of 30° with the Y-axis. The force distribution along Cartesian coordinates is given in Table 1, assuming a point force F on each source direction.

The 3C geophones record seismic vibrations along the vertical and horizontal components parallel to the Cartesian Z, X and Y axis. Two different types of receivers with the same eigenfrequency of 4.5 Hz have been used. For each source location and direction, three seismic stacks have been performed. The total recording time is $T = 1$ s with a sampling frequency of 4 kHz.

We have performed two sequential seismic acquisitions: each with different receiver geometries. Both acquisitions used 36 3C Galperin sources at the same positions. They are indicated by black circles in Fig. 5. The sources have been installed following a staggered pattern resulting in a minimum source distance of about 5.66 m. The first acquisition used a coarse repartition of receivers with 2 m inline and 4 m crossline spacing, giving 128 3C geophone positions in total (Fig. 5a). The second used a dense repartition of receivers, with 1 m inline and crossline spacing, giving 888 3C geophone positions in total (Fig. 5b). Because of the limited number of available equipment (160 3C geophones), the dense acquisition has been divided into six acquisition subsets (color-coded on Fig. 5b) acquired over 5 d in total, repeating all source locations for each subset.

2.3 Data evaluation

The noise level is estimated by comparing the windows where we can separate signal and noise. Signal windows are taken from far offset data (>25 m) with two windows. Noise windows are taken from ambient noise data at the same trace before the seismic wave first arrival. The length of the windows is 40 and 100 ms for noise and signal, respectively. SNR is then calculated as follows:

$$\mathrm{SNR}_{dB} = 10\log_{10}\left(\frac{S_{\mathrm{signal}}}{S_{\mathrm{noise}}}\right)^2, \tag{1}$$

where S is the spectral amplitude of the given window and SNR_{dB} is the signal-to-noise ratio given in dB.

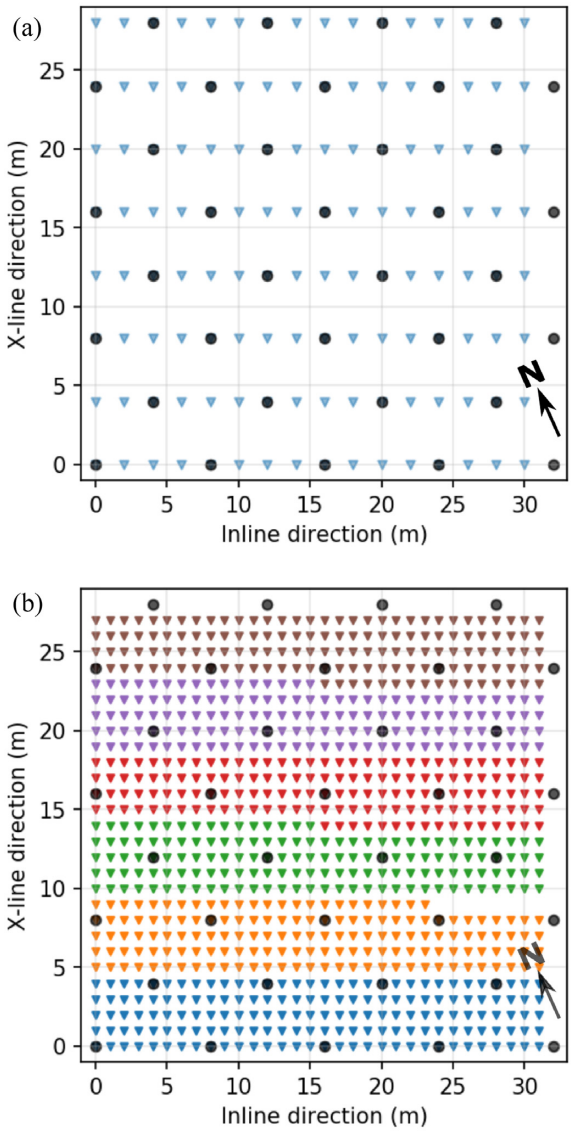

Figure 5. Coarse grid acquisition (a) and dense grid acquisition (b). Triangles represent receiver locations, whereas circles represent source locations. Different receiver colour shows different acquisition time; 1st day = blue, 2nd day = orange, 3rd day = green, 4th day = red, 5th day = purple and brown.

Fig. 6(a) presents the distribution of signal (orange line ± standard deviation) and the ambient noise (blue line ± standard deviation). We can see high SNR from 15 Hz (green line) to 145 Hz (red line). In practice, we define 3 Hz (orange line) as the lowest bound of the signal, which will be used for inversion.

2.4 Data correction using matching filter

Splitting the acquisition into six patches for the dense acquisition requires source repetition at the same location. The acquisition separation raises difficulties in source repeatability because of potential differences in source–ground coupling, trigger time, and surface condition. The latter is particularly problematic for an acquisition spread over 5 d with rain occurring at night. We illustrate this issue in Figs 7 and 8 where the vertical red lines represent subset boundaries. Fig. 7(a) shows rough seismic data before any correction. Half of the seismic data (the left part) has been acquired on April 28th, 2017, whereas the other half (the right part) has been acquired on May 2nd, 2017. Both subsets generally have the same amplitude level, but there is a visible kinematic offset or phase change. Another example of such kind of mismatch can be seen in Fig. 8(a), where a crossline section of the data shows amplitude differences from one subset to another.

In practice, those observations of phase and amplitude mismatch between the six subsets of data prevent prevents us from directly combining them in a full single single data set. One option could have been to consider the six subsets of data independently, hence increasing the total computational cost of the inversion by a factor of 6. To avoid this additional cost, we choose to correct the data before inversion, thanks to the common receiver locations between the coarse and the dense acquisition geometries. Taking advantage of having several common receiver locations between the coarse and the 6 subset of dense grid acquisitions, we can compute a matching filter aiming at transforming

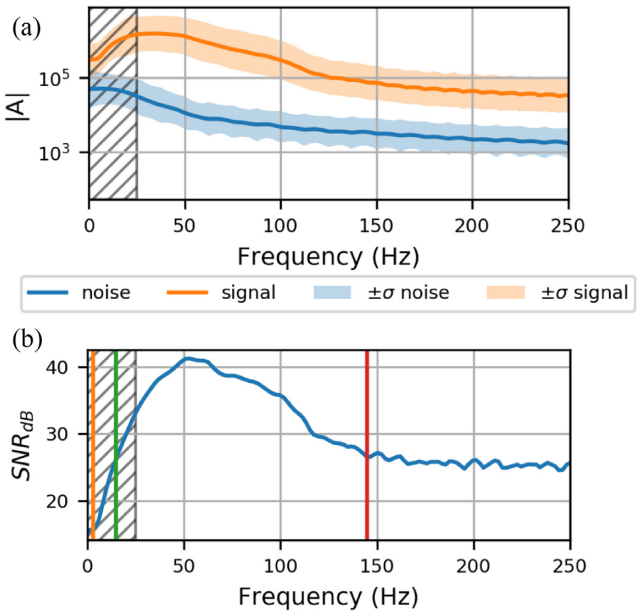

Figure 6. Signal (orange) and noise (blue) spectrum together with its standard deviation of Ettlingen Line 9C seismic data (a). SNR_{dB} of the seismic data (b). The good signal ($SNR_{dB} > 27$) ranging from 15 Hz (green line) to 145 Hz (red line). We started our FWI from 3 Hz (orange line) based on the strong coherent signal. They grey hatches represents unreliable SNR estimation.

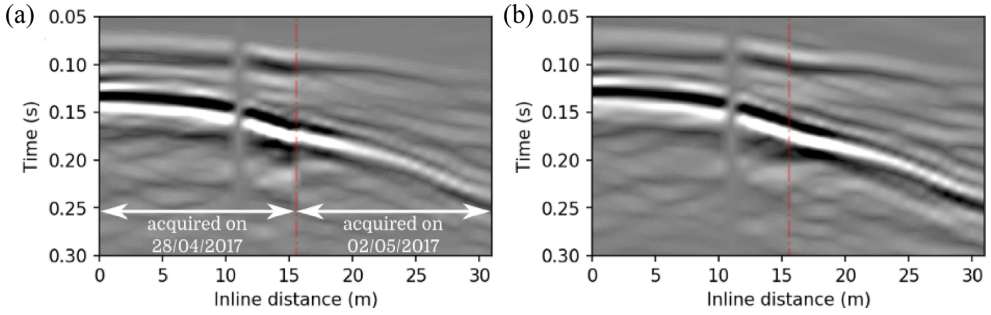

Figure 7. Seismic data before data correction (a) shows traveltime shift between acquisition subsets, and after data correction (b) where the data is more consistent.

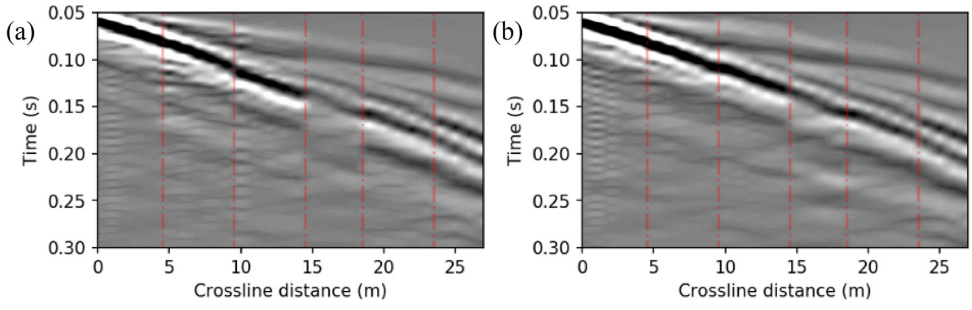

Figure 8. Seismic data before data correction (a), and after data correction (b). This seismic line is taken from the first crossline on the first source location on UZ component. Red vertical lines represent acquisition subset separation. Consistency improvements can be seen on the corrected data.

our dense grid data to match the coarse grid data. The matching filter technique relies in the solution of a linear deconvolution problem, similar to the source–time function estimation in FWI proposed by Pratt (1999) and elastic correction towards acoustic data by Agudo *et al.* (2018). The detailed formulation of the matching filter technique used for the data correction is shown in Appendix A. The main idea lies in correcting dense grid data $d_{i,\,sc,\,s,\,j,\,sr}(t)$ to match the coarse grid data $c_{i,\,sc,\,j,\,sr}(t)$ through the convolution with an optimal filter $f_{i,\,sc,\,s}(t)$

$$c_{i,sc,j,sr}(t) = d_{i,sc,s,j,sr}(t) * f_{i,sc,s}(t), \tag{2}$$

where $*$ is the temporal convolution operator, i is the source location index from 1 to N_s, sc is the source component index with $sc \in \{U, V, W\}$, j is the receiver location index from 1 to N_r, sr is the receiver component index with $sr \in \{X, Y, Z\}$, and s is the acquisition patch

index from 1 to N_a. In these notations, N_a is the number of acquisition patches, N_s is the number of source locations and N_r is the number of receiver locations.

This matching filter approach corrects both time and amplitude discrepancies. Fig. 7(b) shows a time shift correction from the matching filter, Fig. 8(b) shows an amplitude correction between each patch of acquisition. Although this strategy improves the data consistency, we also observe that it increases the noise level, especially on the data with lower SNR. However, the noise before the first arrival can be muted during the inversion.

After this correction, we can gather the six subsets of data in a full consistent data set and therefore avoid any computational overhead during the inversion.

Now that we have introduced the target, the acquisition details, and the pre-processing strategy, we present an overview of the applied methodology to invert the Ettlingen data.

3 FWI METHODOLOGY

The modelling and inversion tool this study is based on follows the development and implementation of Trinh *et al.* (2019b). We present in the following a short version of this theory and implementation, before discussing more in details about the specific points required in this shallow-target study.

3.1 Modelling of elastic and viscoelastic waves with spectral elements

The general viscoelastic wave equations are

$$\begin{aligned} \rho \partial_{tt} u_i &= \partial_j \sigma_{ij} + f_i, \\ \sigma_{ij} &= M_{ijkl} *_t \varepsilon_{kl} + \mathcal{T}_{ij}, \end{aligned} \tag{3}$$

where ρ is the density, u is the displacement field, σ and ε are the second order stress and strain tensors, f is the external force, and $\mathcal{T}$ is the external stress. The effect of the attenuation is denoted by a relaxation rate M_{ijkl}, and symbol $*_t$ represents a convolution operator in time domain between the fourth order tensor and second order tensor. We follow Einstein convention (summation over repeated indices) for these equations.

In a purely elastic wave propagation we have $M_{ijkl}(x, t) = c_{ijkl}(x)\delta(t)$, and eq. (3) simplifies into

$$\begin{aligned} \rho \partial_{tt} u_i &= \partial_j \sigma_{ij} + f_i, \\ \sigma_{ij} &= c_{ijkl}\varepsilon_{kl} + \mathcal{T}_{ij}, \end{aligned} \tag{4}$$

where c_{ijkl} is the elastic (unrelaxed) stiffness-coefficient.

In our application, isotropic medium are considered, leading to only two independent parameters in the c_{ijkl} tensor (for instance the Lamé parameters, or P-wave and S-wave velocities). Two quality factor Q_P and Q_S are used for P and S body waves when viscous media are considered. We approximate this attenuation by classically introducing a set of L standard linear solid (SLS) mechanisms, adding ordinary differential equations (ODE) and the memory variables $\psi_{s;ij}$ in eq. (3) resulting in

$$\begin{aligned} &\rho \partial_{tt} u_i = \partial_j \sigma_{ij} + f_i, \\ &\sigma_{ij} = c_{ijkl}\varepsilon_{kl} - c^R_{ijkl} \sum_{s=1}^{L} \psi_{s;ij} + \mathcal{T}_{ij}, \\ &\partial_t \psi_{s;ij} + \omega_s \psi_{s;ij} = \omega_s y_s \varepsilon_{kl}, \end{aligned} \tag{5}$$

where the scalar y_s is the dimensionless anelastic coefficient (Yang *et al.* 2016b).

Eq. (5) is discretized and solved using a finite element method, so-called spectral element method (SEM, Patera 1984). The SEM relies on weak formulation of the second-order wave equation. The seismic wavefield discretization is based on Lagrange interpolants, and the integration over elements is based on Gauss–Lobatto–Legendre (GLL) integration points. The medium is discretized using hexahedral elements. After discretization, the elastic wave equation can be represented as

$$\mathbf{M}\partial_{tt}\mathbf{u} = -\mathbf{K}\mathbf{u} + \mathbf{F}, \tag{6}$$

where $\mathbf{M}$ and $\mathbf{K}$ are the mass and stiffness matrices, and $\mathbf{F}$ is the source term. In SEM, $\mathbf{M}$ is diagonal by construction, which makes it possible to use an explicit time discretization scheme efficiently. The second-order Newmark time integration method is used in our implementation. The computationally demanding step is the product between the displacement vector and the stiffness matrix ($\mathbf{Ku}$, Komatitsch *et al.* 2000; Trinh *et al.* 2019b). This product can be calculated efficiently using a matrix-free implementation based on the factorization

$$\mathbf{K} = \mathcal{D}^w \mathbf{C} \mathcal{D}, \tag{7}$$

where $\mathcal{D}$ stands for the spatial derivatives of a vector in the Cartesian space, $\mathcal{D}^w$ is the spatial derivative weighted by GLL weights (Deville *et al.* 2002; Trinh *et al.* 2019b) and C the discrete stiffness tensor (making use of the Voigt notation to reduce the fourth-order tensor to a second-order tensor).

When considering the viscoelastic second-order wave equation, the SEM discrete formulation can be written as

$$\begin{aligned} \mathbf{M}\partial_{tt}\mathbf{u} &= -\mathbf{K}\mathbf{u} + \mathcal{D}^w C^R \sum_{s=1}^{L} \psi_s + \mathbf{F}, \\ \partial_t \psi_s + w_s \psi_s &= w_s y_s \varepsilon. \end{aligned} \tag{8}$$

where C^R is a relaxed discrete stiffness tensor which involves the quality factor coefficients Q_P and Q_S. See Trinh *et al.* (2019b) for more details.

3.2 Inverse problem

FWI is formulated as a minimization problem, which uses a least-squares misfit function $f(m)$ for m measuring the discrepancy between the calculated data $d_{\text{cal}}(m)$ and the observed data d_{obs} as

$$f(m) = \frac{1}{2} \left\| d_{\text{cal}}(m) - d_{\text{obs}} \right\|^2 . \tag{9}$$

This inversion minimizes $f(m)$ by updating the model parameters m, which gather the density ρ and any combination of the stiffness tensor coefficients C_{ij}. The calculated data $d_{\text{cal}}(m)$ is obtained through the solution of the viscoelastic equations (8) and the extraction of the wavefield values at the receiver locations. The inversion is conventionally solved using Newton-based methods, which starts from an initial model m_0 and iteratively update m as

$$m_{k+1} = m_k + \alpha_k \Delta m_k, \tag{10}$$

where k is the iteration number, α_k is the step length obtained using a line search method (Nocedal & Wright 2006) and Δm_k is the model parameter update. The quasi-Newton limited-memory Broyden–Fletcher–Goldfarb–Shanno (l-BFGS) approach is used in this study to compute Δm_k.

An adjoint-state method is used to compute the gradient $\nabla f(m)$ (Plessix 2006), leading to a zero-lag cross-correlation between the adjoint displacement field $l(m)$ (solution of the adjoint viscoelastic wave equations) and the incident acceleration field $\ddot{u}(m)$ plus an additional term related to the memory variables ψ_s as

$$\nabla f(m) = \frac{\partial \chi(m)}{\partial m} = \left\langle l(m), \frac{\partial C}{\partial C_{ij}} \ddot{u}(m) \right\rangle - \left\langle l(m), \sum_{s=1}^{L} \frac{\partial C^R}{\partial C_{ij}} \psi_s \right\rangle. \tag{11}$$

For purely elastic media, the second term of the right-hand side of course vanishes.

The gradient with respect to any parameter p, function of the density and the stiffness matrix coefficients C_{ij} can be determined following the chain-rule as

$$\frac{\partial f(m)}{\partial p} = \sum_{I=1}^{6} \sum_{J=1}^{6} \frac{\partial f}{\partial C_{ij}} \frac{\partial C_{ij}}{\partial p} + \frac{\partial f}{\partial \rho} \frac{\partial \rho}{\partial p}, \tag{12}$$

In our implementation, the incident wavefield is first stored on disk at the Nyquist sampling frequency while it is computed. It is then read from disk to build the gradient while computing the adjoint field. This strategy, which takes benefit from high-speed disks, appears to be the best compromise for our application compared to various flavors of check-pointing strategies that requires some recomputations (Griewank & Walther 2000; Symes 2007; Anderson *et al.* 2012; Komatitsch *et al.* 2016; Yang *et al.* 2016a).

The inversion generally requires regularization strategies in practice. Our implementation relies on a gradient smoothing through anisotropic Bessel filter (Trinh *et al.* 2017) implemented directly with a SEM formulation.

3.3 Multiparameter inversion strategy

Inverting shallow-seismic data with limited offset comes with inherent challenges. Even in a simple model, the surface wave can exhibit a complex waveform due to its dispersive behavior. The surface wave dispersion implies the medium properties at different depths depending on the data frequency content. Low-frequency surface wave has a deeper penetration depth and brings low wavenumber content. High-frequency surface wave brings higher-resolution information, however, at a shallower depth. This complex behavior can quickly drive the inversion towards local minima (cycle skipping at a higher frequency). To avoid this problem, we follow a standard multiscale approach, where the inversion is carried out from the lowest frequency band to the broader frequency band (Bunks *et al.* 1995). The multiscale approach is a safer strategy compared to layer-stripping strategy. The layer-stripping strategy by Liu & Peter (2019) requires a careful combination between the offset and the frequency.

In this shallow-seismic imaging context, the frequency range of a typical 4.5 Hz geophone is sufficiently low to avoid cycle-skipping. Considering a rough approximation of the target with an homogeneous model ($V_P = 260\ \mathrm{m\,s^{-1}}$ and $V_S = 150\ \mathrm{m\,s^{-1}}$) and the lowest available frequency at 4.5 Hz, only one half *P*-wavelength and one *S*-wavelength propagate across the the whole target. With such rich low-frequency content, with respect to the target, the multiscale approach is sufficient to avoid any cycle skipping issues throughout the inversion (which is much different from exploration scale targets). However, this strategy does not solve the problem of the unbalanced sensitivity between V_P and V_S on the surface wave dominated data. On the exploration scale, with a sufficiently large offset, one can perform a dynamic windowing strategy, in which the body and the surface waves can be separated (He *et al.* 2019; Trinh *et al.* 2019a). This approach is not feasible for our case study due to the strong response of surface wave and a direct multiparameter inversion is not ideal because it would led to an under-constrained V_P inversion.

For this reason, we design a specific two steps strategy. In the first step, we invert only for V_S, however, updating also for V_P by strictly enforcing an assumed linear relationship between V_P and V_S. This first step is denoted as the parameter binding strategy in the following. In the second step, we perform a true multiparameter inversion, updating simultaneously both V_P and V_S, while incorporating non-linear constraints to bound not only V_P and V_S values, but also the ratio of V_P over V_S.

More in details, the parameter binding strategy assumes the linear relationship

$$V_P(x, y, z) = \gamma(x, y, z) V_S(x, y, z) \tag{13}$$

If we consider that we invert for V_P and V_S we can denote the FWI misfit function $f(m)$ as $f(V_P, V_S)$. The parameter binding strategy consists in considering

$$g(V_S) = f(\gamma V_S, V_S) \tag{14}$$

as our misfit function, depending only on V_S.

The gradient of $g(V_S)$ can thus be written as

$$\nabla g(V_S) = \gamma\, \partial_{V_P} f(V_P, V_S) + \partial_{V_S} f(V_P, V_S). \tag{15}$$

We see that it can be directly computed from the gradient of f with respect to V_P and V_S. Implementing this strategy makes it possible to perform FWI that simultaneously updates V_P and V_S based on the scaling parameter γ.

This parameter binding strategy makes it possible to perform stable FWI in shallow-seismic case, but the relation between V_P and V_S (γ) remains fixed during the inversion. With the limited knowledge of the true γ, FWI with parameter binding can lead to an incorrect model update. The second step of multiparameter inversion is used to mitigate this issue. To stabilize this multiparameter inversion, we apply bound constraints on V_P, V_S and V_P/V_S. The bound for each of these quantities parameters can be deduced from the prior geological knowledge of the area. Enforcing these bounds amounts to define a set of non-linear constraints, as is done in Peters & Herrmann (2017) and Trinh *et al.* (2019a).

The non-linear constraints inversion is formulated as a minimization of the misfit function $f(m)$ over a restricted model space Ω which simultaneously satisfies:

(i) Bounds constraint $\mathcal{C}_1$: V_P and V_S should vary within a pre-defined range

$$\begin{aligned}\mathcal{C}_1 = \{&V_P, V_S,\\ &V_{P_{min}} \le V_P \le V_{P_{max}} \quad \text{and}\\ &V_{S_{min}} \le V_S \le V_{S_{max}}\},\end{aligned} \tag{16}$$

(ii) Ratio constraint $\mathcal{C}_2$: The value of V_P/V_S should vary within a pre-defined range

$$\mathcal{C}_2 = \{V_P, V_S, \quad r_1 \le V_P/V_S \le r_2\}, \tag{17}$$

where $V_P, V_S \in \mathcal{C}_1 \cap \mathcal{C}_2$.

In terms of the non-linear constraint, the projected parameter of V_P and V_S can be determined by solving the following equation

$$\begin{aligned}\min_{V_P, V_S} f(V_P, V_S) = &\frac{(\overline{V_P} - V_P)^2}{V_P^2} + \frac{(\overline{V_S} - V_S)^2}{V_S^2},\\ &\text{subject to} \quad r_1 \le V_P/V_S \le r_2,\end{aligned} \tag{18}$$

where $\overline{V_P}$ and $\overline{V_S}$ is the updated V_P and V_S value given by the optimization for each iteration, V_P and V_S represent a projected value with respect to the r_1 and r_2 boundary.

Trinh (2018) uses the Dykstra's algorithm to solve this non-linear model constraint (Boyle & Dykstra 1986). This approach is used in our experiment to obtain a reasonable multiparameter inversion for the shallow-seismic application. Appendix B contains more detailed information regarding the non-linear model constraint and the Dykstra's algorithm.

This non-linear constraint strategy is crucial because the V_S gradient is still strong, even after the first inversion with parameter binding strategy. Another reason is the limitation of V_P information inside the data itself (weak early body wave arrivals are recorded). Although V_P can be recovered from the surface wave, as Irnaka *et al.* (2018) had demonstrated it on a synthetic case, with the influence of the noise level

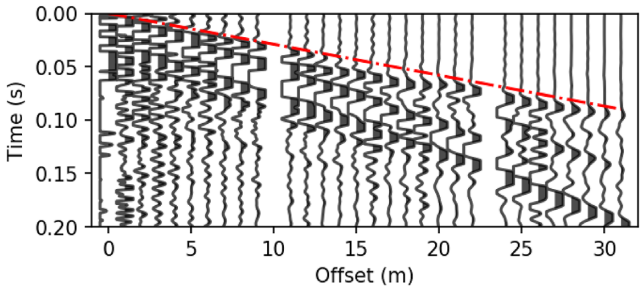

Figure 9. Seismic section on the first source location with U source direction. This seismic section is used to estimate the V_P for the homogeneous initial model. Red dashed line represents the direct P-wave arrival from the estimated homogeneous model.

and the accuracy of the acquisition, the information about V_P from the gradient remains weak compared to the information about V_S. Another source of V_P update is the body wave or P wave, which can be seen on the data. This body wave feature is masked by the strong surface wave at low frequency, and this information is difficult to retrieve. At higher frequency, the body waves are well separated from the surface wave (see Fig. 9). Without proper estimation of the V_P at the low frequency, the multiparameter inversion has difficulty in estimating the correct V_P at the higher frequency cycles.

Our approach tries to find a balance between FWI with parameter binding and FWI with multiple model constraints. We want to ensure that we would get the update of both parameters V_P, V_S and V_P/V_S for each cycle. The new value of V_P/V_S is used as the input parameter for the next frequency band of FWI with parameter binding. It is essential to note that for both FWI steps, we only use top window muting to reduce the matching filter artifacts before the first arrival. With this approach, we take all available information of both V_P and V_S from each frequency band.

3.4 Implementation: SEM46 code

The methodology that we have described has been implemented in the SEM46 package. This is a 3-D viscoelastic full waveform modelling and inversion tool based on the spectral element method and the SEISCOPE optimization toolbox (Métivier & Brossier 2016) written in Fortran (Trinh *et al.* 2019b). SEM46 is based on a simple Cartesian-based mesh design, well suited for efficient parallel computing. The specific multiparameter inversion strategy described previously has been implemented for this study based on the SEM46 package.

4 FWI APPLICATION TO THE ETTLINGEN LINE

We use a hierarchical approach based on frequency continuation, and for each frequency band, the two steps approach based on parameter binding, then non-linear model constraints are applied (Fig. 11).

4.1 Initial model building

The Ettlingen Line has a simple geology and flat topography. Our goal is to have a simple but consistent starting homogeneous velocity model.

To design it, we pick the first P-wave arrival at the longest offset on inline seismic section (Fig. 9). From this arrival, we estimate a homogeneous V_P model directly, resulting in a value of 345 m s^{-1}. We estimate the other elastic parameters (V_S and density ρ) using a fixed Poisson's ratio $\nu = 0.25$ and the Gardner relationship

$$\rho = 0.31 V_P^{0.25}, \tag{19}$$

resulting in $V_S = 199$ m s^{-1} and $\rho = 1336$ kg m^{-3}. This fixed Poisson's ratio of $\nu = 0.25$ might be a bit low for shallow sediment, but as will be shown in the following, this does not affect the inversion result reliability. For the density, even if Gardner relationship is established for sedimentary rocks, we do not reconstruct the density in our scheme and the constant value remains fixed along the process. It is worth noting that in the case of constant density, the value of density only affects the amplitude of the seismic signal, which is corrected by the source–time function inversion.

An initial attenuation model is obtained based on the previous experiment of Gao *et al.* (2020) on the same target, who have implemented multimode surface waves Q estimation (Gao *et al.* 2018). The Q_S model is 12 for the first 1 m depth. Then it linearly increases up to 80 until 5 m depth and remains constant from this depth. A constant ratio of 1.5 is chosen to estimate Q_P from Q_S based on Hauksson & Shearer (2006). Even if this ratio has been obtained from a crustal scale study, it remains consistent with the amplitude decay trend of the P wave in the data. In Fig. 10, we show a scatter plot of the P-wave amplitude versus offset (after geometrical spreading correction and a narrow bandpass filter between 50 and 80 Hz), obtained from our data, with the theoretical amplitude decay given a constant Q_P model based on Kjartansson (1979) model (using a constant V_P of 170 m s^{-1} at 65 Hz). This figure clearly shows that a Q_P value between 10 and 20 is required to match observations, validating our choice of Q_P model derived from Q_S.

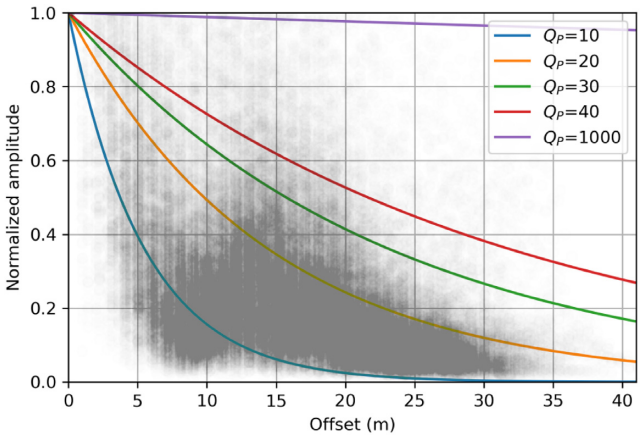

Figure 10. Normalized amplitude of P wave with respect to the offset drawn as the scatter plot on a narrow band-passed data centred at 65 Hz. The line plots represents theoretical amplitude decay given a constant Q_P model based on Kjartansson (1979) at 65 Hz. A geometrical spreading correction is applied prior to normalization.

During FWI, only V_P and V_S are reconstructed; the density ρ, Q_P and Q_S models has been considered as passive parameters. These passive parameters are used for the modelling, but their values are not updated.

4.2 Source wavelet estimation

We assume to have directional point sources (U, V, W) with known and fixed source direction for each location (Table 1). We assume to have equivalent source energy and fixed source location. Note that we kept the original direction of the Galperin source (U, V, W) in the whole workflow to avoid the additional step of source-component rotation that would require a specific inversion problem. To estimate the source–time functions, we follow a strategy with three steps: raw source–time function estimation for each source location and direction using a standard deconvolution operation (Pratt 1999), amplitude normalization to equalize the amplitude contribution of each source, and spatial weighting average to reduce the local effect and overestimation of the source. The technical details of the source estimation are given in Appendix C. These steps produce a single-source–time functions for each location and direction of the source, making a total of 108 source signals.

4.3 Multiscale strategy with two steps inversion

The lowest frequency band we start with is 3–15 Hz. We increase the frequency band by 10 Hz increments at each FWI stage. In total, we perform six FWI stages. The highest frequency band in this experiment is thus 3–65 Hz. This upper frequency limit is associated to the computational cost of the forward problem, which scales to the power 4 of frequency. This increase of computational demand with frequency limited our ability to go to higher frequencies given our available computing resources.

For each frequency band, we apply the multiparameter inversion strategy described in the previous section (Fig. 11). The first step is an inversion with parameter binding, and the second step is non-linear model constraint multiparameter inversion. We specify the initial γ as a fixed value of 1.732 (which corresponds to the constant Poisson's ratio $\nu = 0.25$). The parameter γ, which constraints V_P with respect to V_S with the relation of V_P/V_S, is updated each time the non-linear model constraints multiparameter inversion is carried out. The lower and upper boundary values we use are 100 and 2000 m s^{-1} for V_P, 50 and 800 m s^{-1} for V_S, respectively. The lower and upper boundary values for the ratio between V_P and V_S are 1.633 and 10, respectively. A short offset data muting up to 1.5 m offset is applied to remove signal with clipped amplitude. An additional top window muting is used to remove a strong noise on some part of the data, which might be increased by the matching filter correction.

In terms of numerical optimization parameters, we set the maximum line search trials to 20, and the number of stored gradients for l-BFGS to 10. The Bessel filter-based gradient smoothing uses coherent lengths of 0.1λ, 0.7λ and 0.7λ in vertical, inline, and crossline direction, respectively, where λ is the local wavelength computed from the local velocity and dominant frequency. The non-linear l-BFGS is preconditioned with a diagonal matrix whose entries are the depth to the power of two, in order to mimic the inverse Hessian operator effects.

4.4 Model reconstruction

Fig. 12 shows the evolution of the reconstructed models. We present V_S on the top row, V_P on the middle row, and Poisson's ratio ν on the bottom row. The Poisson's ratio is calculated from the reconstructed V_P and V_S. The columns represent the updated model throughout the

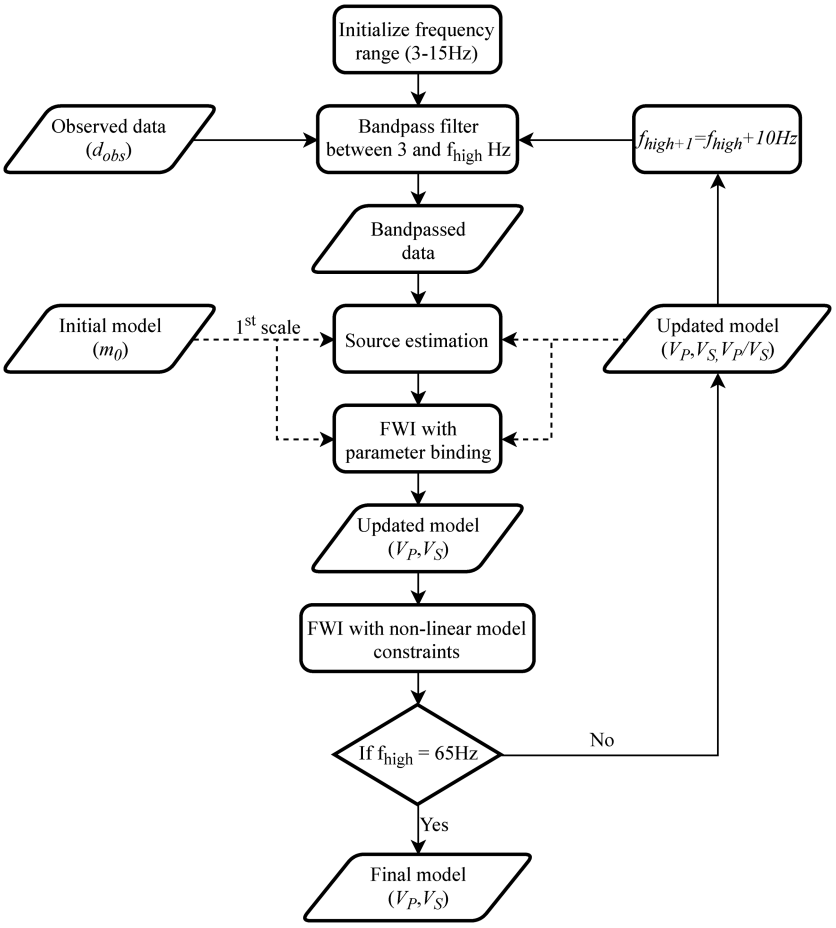

Figure 11. The proposed FWI Workflow for shallow-seismic application. We test this workflow with 9C data on Ettlingen Line case.

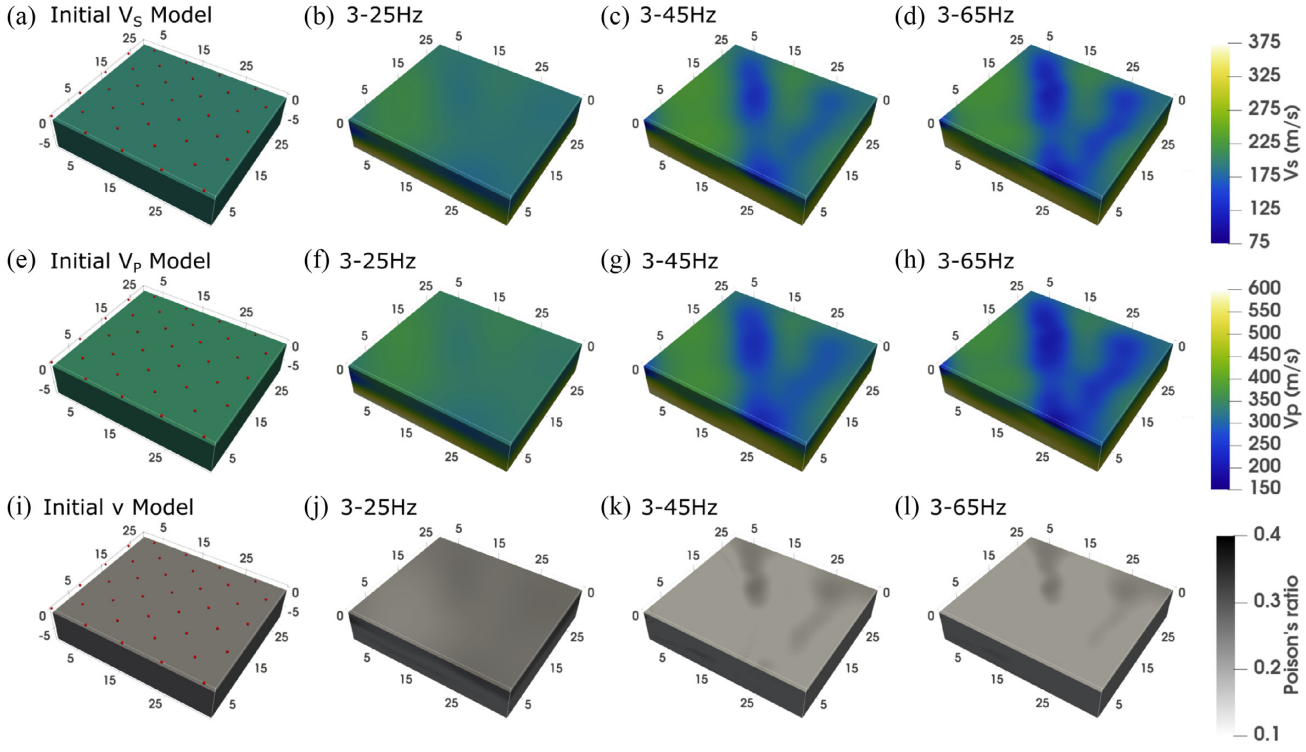

Figure 12. Reconstructed parameters (V_S at the top row, V_P at the middle row, and ν at the bottom row) at 1 m depth section. Each column is associated with a specific FWI stage (initial model on the 1st column, 3–25 Hz band on the 2nd column, 3–45 Hz band on the 3rd column, 3–65 Hz band on the 4th column).

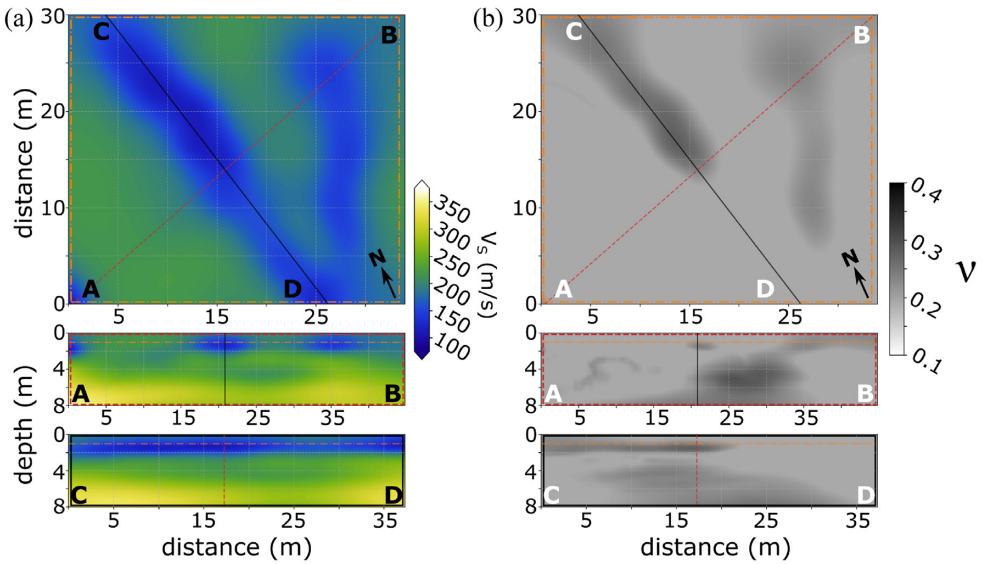

Figure 13. A horizontal slice of reconstructed V_S (a) and Poisson's ratio ν (b) at 1 m depth shows clear presence of the Ettlingen Line trench as well as a trench-like structure with north–south direction (top panel). Vertical slice of reconstructed V_S perpendicular (middle panel) and parallel (bottom panel) with the direction of Ettlingen Line.

different FWI stages, starting from the initial model on the left, 3–25 Hz band on the 2nd column, 3–45 Hz band on the 3rd column, and the final 3–65 Hz band on the 4th column. All 3-D views are presented with a 1 m depth section. The locations of the sources are indicated as red points on the initial models.

We observe a gradual reconstruction of the Ettlingen line structure through the frequency continuation strategy. On lower frequency bands (up to 25 Hz), the updates are mainly dominated by the background velocity update, which corresponds to a vertical velocity gradient. This vertical velocity gradient can be observed in the final vertical section, as shown in Fig. 13.

The Ettlingen Line, which corresponds to the low V_S and V_P velocity zone appearing at 1 m depth, oriented along the diagonal of the model, starts to be visible at 45 Hz. We also observe the appearance of a second low-velocity zone aligned with the crossline direction at this frequency at the eastern part of the model.

As expected, the Poisson's ratio is constant at the beginning of the inversion, while the multiparameter workflow we design allows us to relax the linear relationship initially imposed between V_P and V_S. Higher values of the Poisson's ratio (close to 0.4) along the two trench structures are consistent with their interpretation as less consolidated zones.

In overall, the update is rather weak at very shallow depth (Fig. 13). This observation can be associated with the smoothing and pre-conditioner effect. The current pre-conditioner choice might not be strong enough to drive a significant model update in this region.

4.5 Quality control: data fit

The inversion use the original UVW source component without any data rotation. In order to analyse the data fit, we do not perform any additional data rotation that would introduce another inverse problem with its own bias. In Fig. 14(a), we present a comparison between the 9C observed data (grey traces) and the 9C synthetic data calculated in the initial model (red traces) on the last frequency band (3–65 Hz). Both data are recorded using the source at the location inline 0 m and crossline 16 m for all nine components. The displayed seismic data is the seismic data with inline direction. The source orientations are represented in the row (U, V, W), and the receiver components are represented in columns (Z, X, Y). We perform a trace by trace normalization in order to better see the fit at a longer offset.

In this figure, we see clearly that the strong surface wave dominates the waveform. Although the homogeneous model produces a simple waveform propagation, we still observe a good surface wave matching and no evidence of cycle skipping. This good match is an indication that starting from a homogeneous model is feasible in this context.

In Fig. 14(b), we perform the same comparison, however, the synthetic is calculated using the final reconstructed V_P and V_S models. We can observe a significant improvement in the data fit in all components. The vertical components (first column), in which the waveforms are the least complex, have a better match than the horizontal components. In this seismic section, the Ettlingen Line is located approximately between traces number 10 and 20. Between those traces, we can observe that the FWI can reconstruct the small perturbartions in the recorded data. Starting from a homogeneous model, we observe the possibility to obtain a satisfactory data fit.

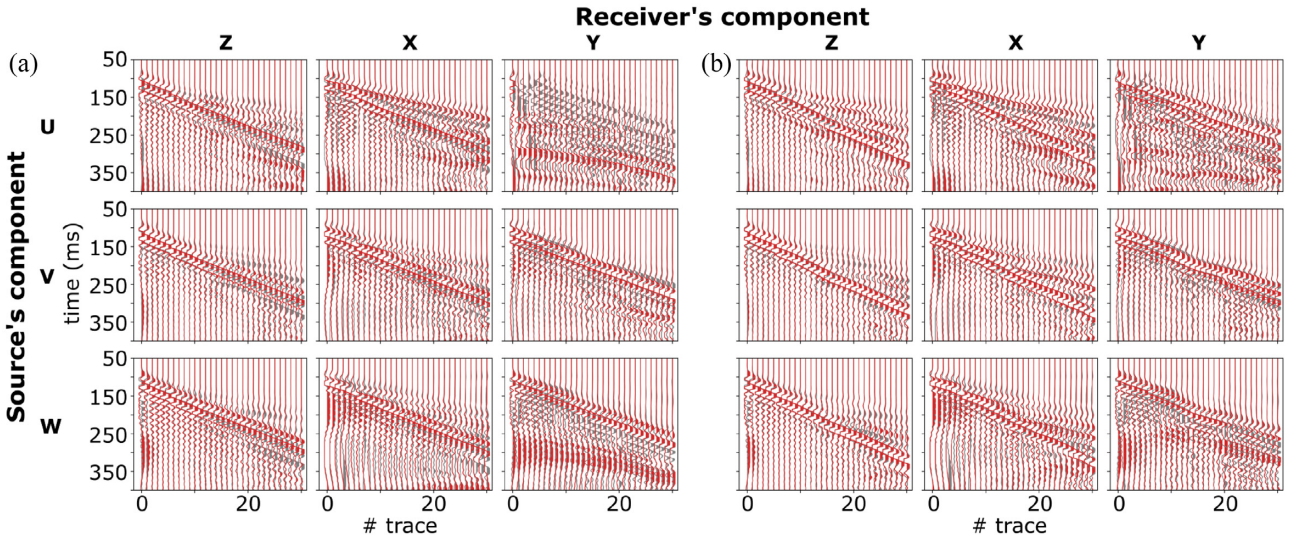

Figure 14. Comparison between 9C observed data (grey) and 9C calculated data (red) in (a) the initial model (b) the reconstructed model in inline direction at $y = 16$ m. The source is location at $x = 0$ and $y = 16$. A trace by trace normalization is used to plot the data. The larger overlap between red and grey in the final model shows a better data match.

5 DISCUSSION AND INTERPRETATION

5.1 Model interpretation

Several experiments at the Ettlingen Line have been performed with different approaches and methods. A 3-D GPR experiment has been performed by Wegscheider (2017), and 2-D elastic FWI from 2-D seismic data has been led by Wittkamp *et al.* (2018). A 3-D MASW has also been performed by Pan *et al.* (2018) using the same data.

Fig. 15 shows a comparison between these results. In Fig. 15(a), an overlay between the GPR results and the 2-D V_S model obtained by Wittkamp *et al.* (2018) is presented. GPR measurement shows an inverted triangle shape in the middle of the section, interpreted as the Ettlingen Line. The 2-D FWI was carried out using joint inversion from both Rayleigh and Love waves up until 130 Hz. The low-velocity anomaly represents the remnant of the Ettlingen Line in the middle of the model. The sharp boundary and shallow model reconstruction can be seen thanks to the higher frequency inversion. In Fig. 15(b), the corresponding 2-D slice extracted from the 3-D MASW V_S model is presented. In Fig. 15(c), we present the corresponding slice of the V_S model that have been reconstructed through 3-D multiparameter viscoelastic FWI.

Although we do not recover the high resolution coming from both GPR and high-frequency surface waves from Wittkamp *et al.* (2018), we can still locate the V_S anomaly at a similar location. This result is also in agreement with the MASW result, with, in our case, a significantly higher resolution, as expected. It is important to be noted that the previous 2-D FWI experiment presents a sharper results due to the higher frequency content which is considered compared to our experiment.

However, our approach added value relies on the fact that we estimate 3-D models, both for V_P and V_S. Both our reconstructed V_P and V_S show a distinct presence of the low-velocity anomaly. Analysing the model from the final frequency band, we can infer the geometry of the trench. It has a consistent northwest–southeast direction, with a width of around 5 and 2 m depth.

Interestingly, we can also observe an additional low-velocity anomaly of an orientation parallel to the crossline direction, with a similar trench-like structure. It is shallower than The Ettlingen Line, around 1.5 m deep, with 2–3 m width. This low-velocity anomaly is visible on the previous 2-D FWI (Fig. 15a) at a distance of around 30–35 m. However, given the location of the structure close to the edge of the acquisition, it had previously been interpreted as a probable artifact. Our 3-D reconstruction shows that this low-velocity zone has a consistent structure along the north-south direction, making it more likely to be an actual feature of subsurface structure. We do not know what the historical significance of this eastern trench-like structure is.

On the MASW result, the indication of this trench-like structure is much weaker due to the lower resolution of the 3-D estimation. However, its presence can be guessed from the smearing of the low-velocity anomaly in the southeast part of the model (Fig. 16a).

Our multiparameter reconstruction also makes it possible to better describe the physical properties of the target. In geotechnics, V_S is often used as a proxy of the soil compactness of the medium. Given the similar type of soil material in this experiment, lower V_S means a lower compaction degree. Thanks to our multiparameter inversion, we can also calculate the value of the Poisson's ratio ν from the relation between the reconstructed V_P and V_S. The Poisson's ratio can give an estimation of water saturation tendency beside the compaction level. Within the trenches, we can see a variation in Poisson's ratio. Assuming similar material was used to level the trench, the higher Poisson's ratio might correlate with higher water saturation. It can be observed at the northwest part of the Ettlingen Line (Fig. 12 l). The variation

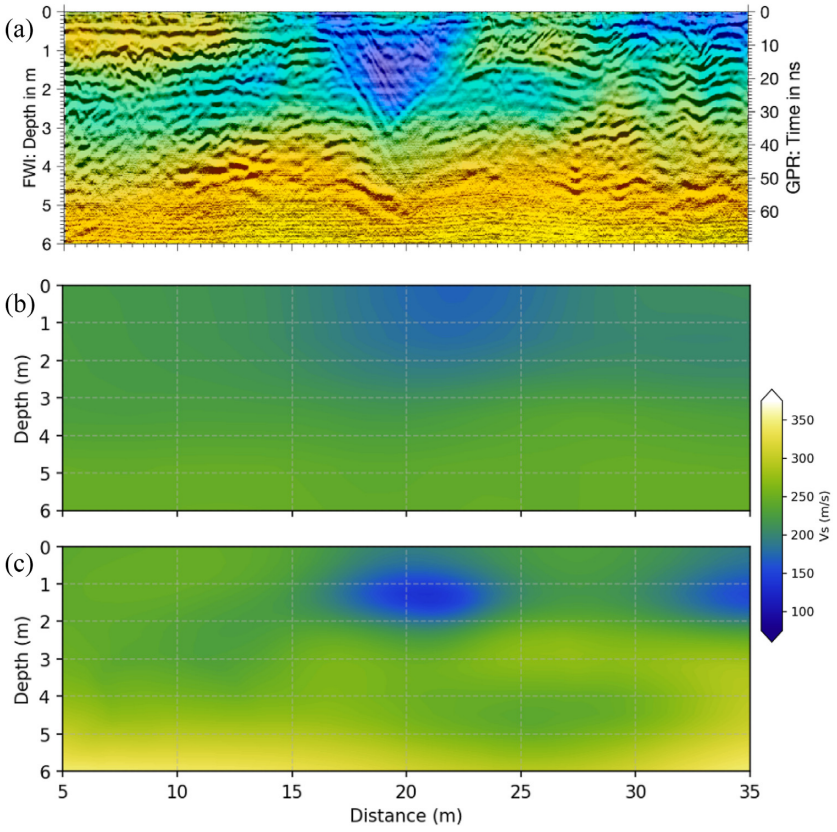

Figure 15. Vertical section of 3-D GPR and 2-D elastic joint inversion of Love and Rayleigh wave result taken from Wegscheider (2017) and Wittkamp *et al.* (2018) (a), 3-D MASW by Pan *et al.* (2018) (c) and 3-D elastic FWI starting from homogeneous model (c). The colour-scale on (a) is unscaled, blue means lower velocity and yellow means higher velocity.

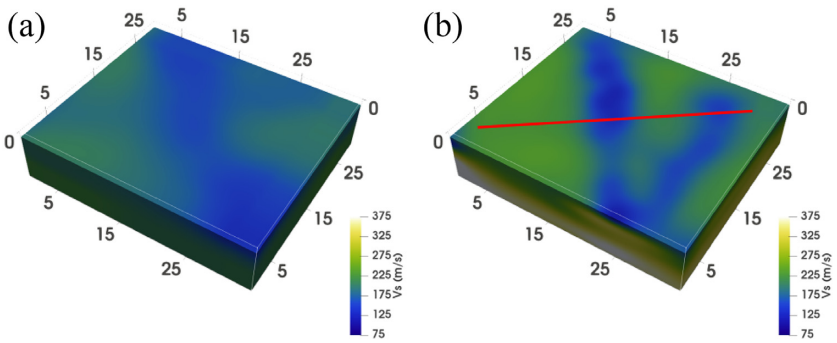

Figure 16. 3-D V_S from previous MASW study by Pan *et al.* (2018) (left-hand panel) and this study starting from homogeneous initial model (right-hand panel). Red line represents the location of the GPR line on Fig. 15.

in Poisson's ratio can also be interpreted as denoting the presence of fine heterogeneities inside the trench. The second eastern trench-like structure does not show similar Poisson's ratio variations, which indicates a probably more similar soil material within the trench.

In Fig. 13, we present horizontal and vertical sections of the 3-D model of the Ettlingen Line. The figures on the left (Fig. 13a) represent V_S, and the figures on the right (Fig. 13b) represent ν. These figures give a better view of the geometry of the trenches. The Ettlingen Line has a width ranging from 4 to 6 m, and a depth ranging from 2 to 2.5 m. This geometry is in accordance with the current existing Ettlingen Line section and illustration in Lang *et al.* (1907). The shape of this trench is also comparable to the trench during the First World War (Lynn 2013; Bull 2014).The V_S profile of the inline vertical section gives information that the Ettlingen Line structure is continuous throughout the model. The V_S of the Ettlingen is lower than the surrounding, below 140 m s^{-1}. The other trench-like structure is also continuous on the final reconstructed V_S. Comparing to the Ettlingen Line, the structure is not as straight and the width is narrower and shallower. It is around 2.5 and 1.5 m for width and depth, respectively. The V_S on the other trench-like structure is slightly higher than the one inside the Ettlingen Line.

Fig. 13(b) shows the final reconstructed Poisson's ratio at 65 Hz. In addition, higher Poisson's ratio at the trenches' location due to the different materials, compactness, or water saturation level, we observe an interesting behavior at a depth of 4–5 m. There is a Poisson's ratio

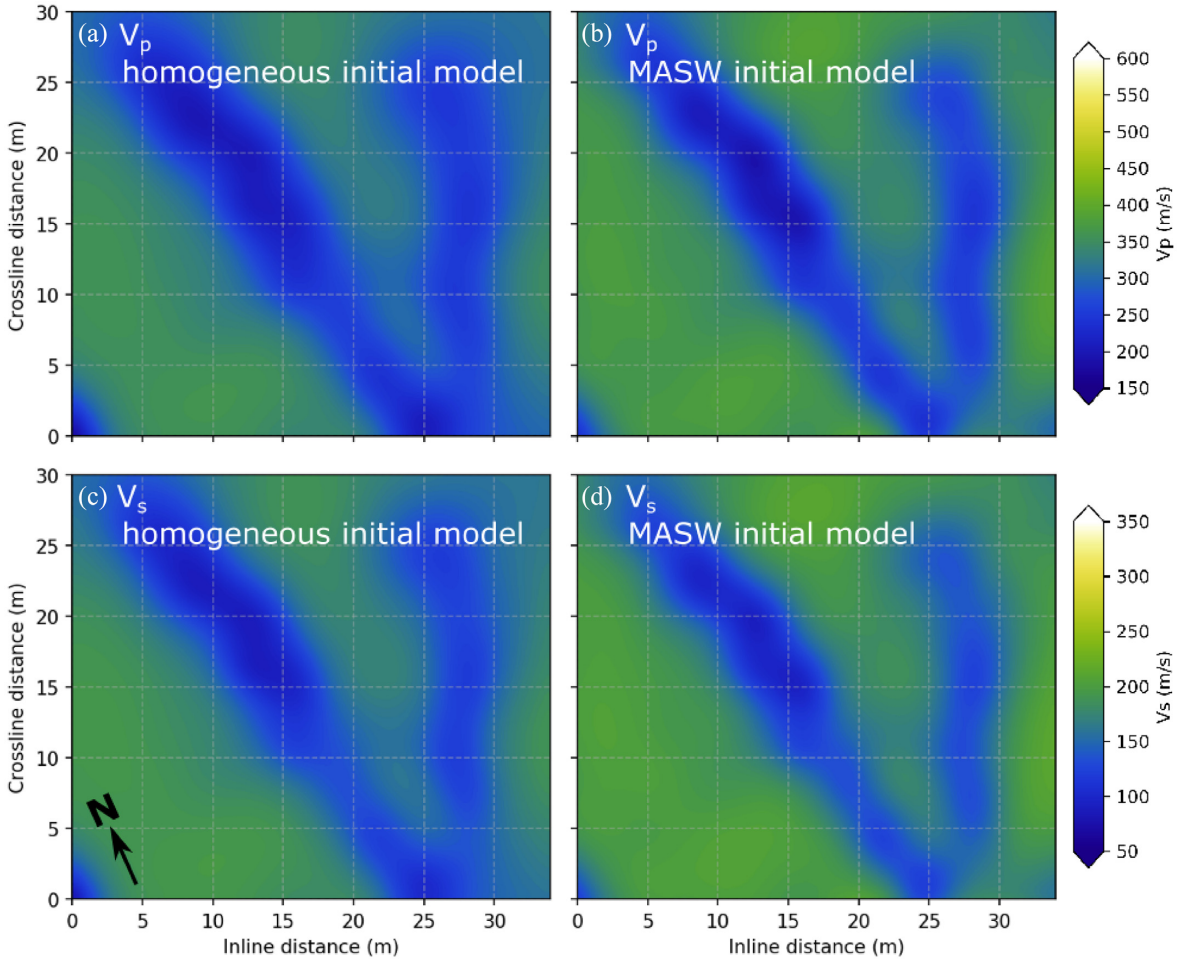

Figure 17. Horizontal slice of the inversion result at 1 m depth for both V_P (top row) and V_S (bottom row) use different initial models. Left slices represent final velocity model at 65 Hz starting from homogeneous model; whereas right slices represent final velocity model at 65 Hz starting from MASW V_S model.

contrast at this depth, from low to high value. This feature probably depicts a possible higher water saturation or even the ground-water level in the area, which was mentioned by Pan *et al.* (2018) at around 7.5 m depth.

5.2 Is it reasonable to start from a homogeneous model?

Initial model building is one of the key aspects of ensuring the FWI efficiency and can be a highly time-consuming step of the FWI workflow. We used an initial homogeneous model, computed from very crude assumptions, advocating that the usable frequency content of the signal compared to the target's size was preventing in this case from any cycle skipping effects.

We perform an additional experiment to validate this assumption, where we use the 3-D V_S MASW model from (Pan *et al.* 2018) as an initial model. We cut the model below 7 m depth, where a strong interface exists from the MASW result. Then we extrapolate the velocity value from the depth of 6.5 m. A fixed Poisson's ratio of 0.25 is used to estimate the V_P. As we used previously, the same Gardner relationship is taken to estimate the density from the initial model.

We then apply the same workflow we used using the homogeneous initial model. The reconstructed V_P and V_S from both starting models are presented in Fig. 17. The difference between both reconstructed model can be estimated by calculating the average difference values between two models divided by the average value of the mean model. We obtain 7 per cent differences between both models, which confirms the fact that considering this type of target, with a sufficiently good quality data, a complex initial model building relying on tomography or analysis of surface wave dispersion is not needed. This observation is important for future near-surface FWI studies as it simplifies the FWI workflow greatly.

5.3 Elastic or viscoelastic FWI?

In Figs 18 and 19, we present the reconstructed V_S and V_P velocities using both elastic (a and c) and viscoelastic modelling (b and d) on two different frequency bands (65 Hz at the top and 45 Hz at the bottom), starting from the homogeneous initial model. The inversion on the highest frequency band (65 Hz) generates significantly different models close to the surface.

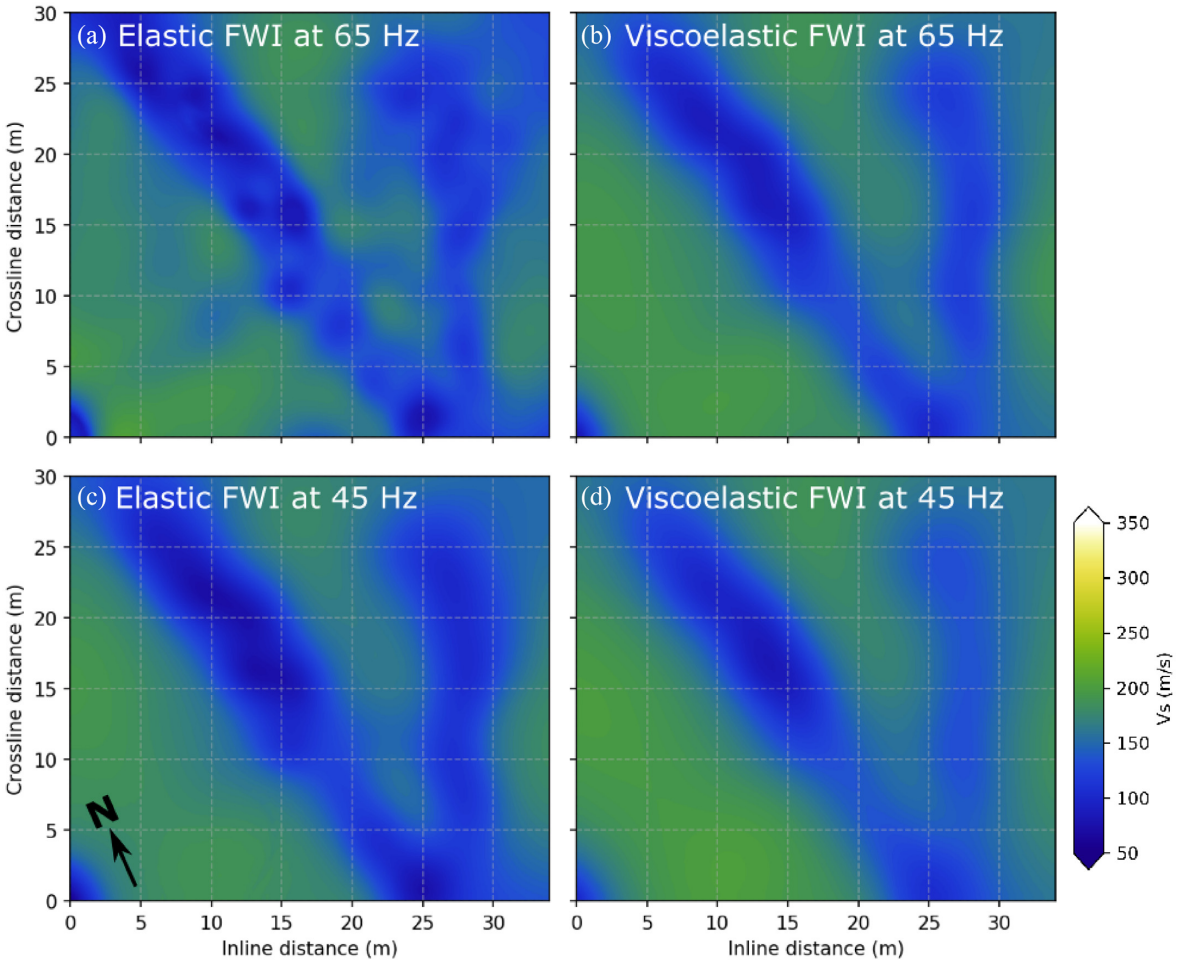

Figure 18. Horizontal slice of V_S from the FWI at 1 m depth for FWI with elastic medium (left-hand column) and viscoelastic medium (right-hand column). Top row images are the reconstructed V_S at 65 Hz, whereas 45 Hz for the bottom row images.

We present several seismograms comparing the synthetic data obtained using elastic and viscoelastic FWI with the observed data (Fig. 20). The amplitudes are normalized based on the maximum amplitude of each seismic section longest offset. We can see that the waveforms predicted from the two models are quite similar, showing that inversions converged towards equivalent model from the data fit side.

While sharing a similar background model, the elastic FWI result display some additional smaller-scale features. Contrary, the viscoelastic results exhibit a smoother and more spatially consistent aspect, with slightly higher velocity values. We also note that the divergence between elastic and viscoelastic FWI results appear only on the largest frequency band. Results obtained at 45 Hz are reasonably similar.

The similarity of the data fit, combined with small scale feature appearing in elastic inversion only at higher frequencies and well known intrinsic attenuation effect of shallow subsurface make us prefer parsimonious viscoelastic results and therefore suspect that small-scales feature of elastic inversion would be rather artifact, as already been shown in various FWI application at different scales (for instance in Kamath *et al.* 2020).

5.4 Other limitations and prospects

This study demonstrates the 3-D FWI application on the Ettlingen Line using 9C seismic data. However, there are several limitations related to the two-step strategy, initial model, source–time function, data coverage, and modelling approach. It is essential to discuss these limitations better to understand the advantages and disadvantages of our approach.

The two steps FWI (parameter binding and non-linear model constraint) are able to reconstruct V_P and V_S but leads to an increase in the relative computational cost related to the two FWI. The reconstruction of V_P itself is rather limited due to the stronger influence of V_S in the surface wave and the lack of *P*-wave information in the data.

The source–time function is one of the keys in obtaining a good model reconstruction during FWI. Our experiment assumes a known and fixed direction of the source–time function (Table 1). We have minimized the local dependency of the source–time function by performing normalization and the spatial weighting, but with better source estimation, we have potential to improve the result.

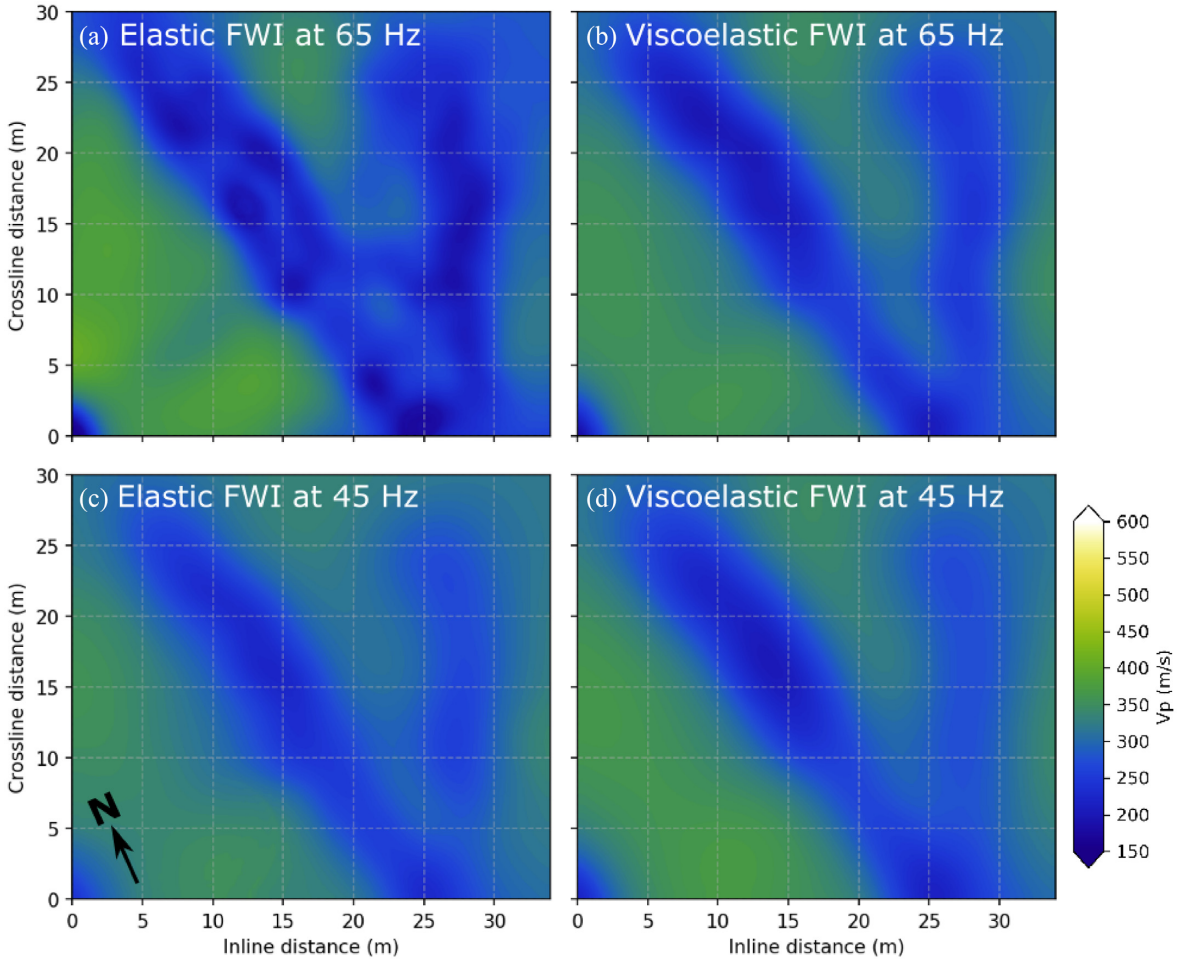

Figure 19. Horizontal slice of V_P from the FWI at 1 m depth for FWI with elastic medium (left-hand column) and viscoelastic medium (right-hand column). Top row images are the reconstructed V_P at 65 Hz, whereas 45 Hz for the bottom row images.

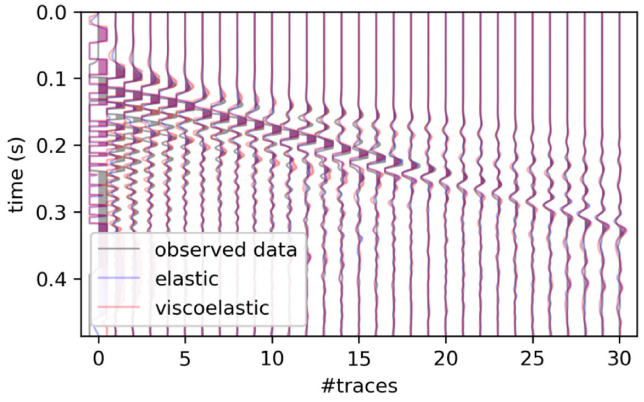

Figure 20. Synthetic data using elastic and viscoelastic medium compared with the observed data. The trace is normalized based on the maximum absolute amplitude of the longest offset data. The seismograms with the amplitude which overlapped with the neighboring seismogram are clipped.

Good coverage of sources and receivers also plays an essential factor in the accurate model reconstruction. The receiver spacing (1 m), which is way less than the target (∼5 m) width, ensures the information redundancy in terms of data, therefore enhancing the signal and helping the inversion. The availability of 9C data can also contribute towards the improvement of the model reconstruction. A more detailed study regarding multicomponent data and source and receiver setup is currently ongoing.

6 CONCLUSIONS

We demonstrate an application of 3-D FWI for shallow-seismic scale. The target is the Ettlingen Line, an ancient war-trench built during the War of The Spanish Succession in 1707.

With a limited number of equipment, we perform a 3-D dense acquisition patch by patch. Our objective is to interpret the data set as a whole, as if all the data subsets would have been recorded at the same time. This data correction can be performed using matching filter. This process decreases the potential computational cost (up to a factor 6 in our case) since the data can be treated as a single 3-D dense acquisition.

The FWI follows a multiscale approach with two specific strategies for each frequency band (the parameter binding and the non-linear model constraints). The inversion is performed up to 65 Hz with two initial models (homogeneous and MASW) and modelling approaches (elastic and viscoelastic). The result shows that we can start from a homogeneous initial model, thanks to sufficient low-frequency data. On the aspect of the modelling approach, both elastic and viscoelastic approaches are comparable up to 45 Hz. At higher frequency, we observe differences between both approaches, where the elastic modelling produces some additional small scale features. We hypothesize that these additional features in elastic FWI are artifacts due to the inappropriate physical model assumption, thus advocating for the use of viscoelastic FWI.

Utilizing FWI for a shallow seismic scale with a limited offset requires a proper strategy. In shallow seismic scale, the seismic data has generally sufficient low-frequency content with respect to the scale of investigation. This fact enables to use less accurate initial model, compared to what is required at the crustal scale for example. The dominant imprint of the surface waves can be mitigated by using a proper hierarchical strategy. A first step focusing on V_S reconstruction is performed, without involving V_P which would be underconstrained. Then, the two-parameters can be reconstructed simultaneously once the surface wave is well predicted by the updated V_S. This strategy requires however two inversions for each frequency scale, which lead to more computational cost. The reliability of the strategy that we proposed is assessed by considering two different physical model assumptions and two different initial models, and also discussed with respect to previous geophysical investigations in the same area.

From the archaeological point of view, our 3-D FWI results have reconstructed a high-resolution (up to 65 Hz) historical trench. The shape and geometry of the Ettlingen Line can be recovered. The width of the Ettlingen Line is around 5 m, and the depth is around 2 m. We are also able to reconstruct another trench-like structure, which was not detected previously. This structure width is around 3 m, and the depth is around 1.5 m.

With the successful application of FWI using 9C seismic data in the Ettlingen Line, we would like to expand our study to analyse the interest in using multicomponent data and each component contribution. Studying multicomponent data might allow us to better design strategies for acquisition, data processing and the FWI at such a scale.

ACKNOWLEDGEMENTS

The authors thank the collaboration between GPI KIT, UGA, GFZ Potsdam, ETH Zurich, which enable us to perform the acquisition, processing, and FWI. This study was partially funded by the SEISCOPE consortium (*http://seiscope2.osug.fr*), sponsored by AkeRBP, CGG, Chevron, Equinor, Exxon Mobil Corporation, JGI, Shell, SINOPEC, SISPROBE and Total; also partial financial support by the Deutsche Forschungsgemeinschaft (DFG) with project-ID 258734477-SFB 1173. This study was granted access to the HPC resources of the Froggy platform of the CIMENT infrastructure (*https://ciment.ujf-grenoble.fr*), which is supported by the Rhône-Alpes region (GRANT CPER07_13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56) and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the programme Investissements d'Avenir supervised by the Agence Nationale pour la Recherche, and the HPC resources of CINES/IDRIS/TGCC under the allocation 046091 made by GENCI.

DATA AVAILABILITY

The data underlying this paper are available in KITopen repository, at https://doi.org/10.5445/IR/1000125628

REFERENCES

Agudo, O.C., da Silva, N.V., Warner, M. & Morgan, J., 2018. Acoustic full-waveform inversion in an elastic world, *Geophysics,* **83**(3), 1–62, doi:10.1190/geo2017-0063.1.

Aki, K., Christoffersson, A., Husebye, E. & Powell, C., 1974. Three-dimensional seismic velocity anomalies in the crust and upper-mantle under the usgs, california seismic array, *EOS Trans. Am. Geophys. Un.,* **56,** 1145.

Allred, B., Daniels, J.J. & Ehsani, M.R., 2008. *Handbook of Agricultural Geophysics,* CRC Press.

Anderson, J.E., Tan, L. & Wang, D., 2012. Time-reversal checkpointing methods for RTM and FWI, *Geophysics,* **77**(4), S93–S103.

Beller, S., Monteiller, V., Operto, S., Nolet, G., Paul, A. & Zhao, L., 2018. Lithospheric architecture of the south-western alps revealed by multi-parameter teleseismic full-waveform inversion, *Geophys. J. Int.,* **212**(2), 1369–1388.

Boyle, J.P. & Dykstra, R.L., 1986. A method for finding projections onto the intersection of convex sets in Hilbert spaces, in *Advances in Order Restricted Statistical Inference: Proceedings of the Symposium on Order Restricted Statistical Inference,* held in Iowa City, IA, September 11–13, 1985, pp. 28–47, eds Dykstra, R., Robertson, T. & Wright, F.T., Springer New York.

Bretaudeau, F., Brossier, R., Leparoux, D., Abraham, O. & Virieux, J., 2013. 2D elastic full waveform imaging of the near surface: application to synthetic and a physical modelling data sets, *Near Surf. Geophys.,* **11,** 307–316.

Bull, S., 2014. *Trench: A History of Trench Warfare on the Western Front,* Bloomsbury Publishing.

Bunks, C., Salek, F.M., Zaleski, S. & Chavent, G., 1995. Multiscale seismic waveform inversion, *Geophysics,* **60**(5), 1457–1473.

Cardarelli, E., Cercato, M., Cerreto, A. & Di Filippo, G., 2010. Electrical resistivity and seismic refraction tomography to detect buried cavities, *Geophys. Prospect.,* **58**(4), 685–695.

Chapman, M., Martin, J., Olgun, C. & Beale, J., 2006. Site-response models for Charleston, South Carolina, and vicinity developed from shallow geotechnical investigations, *Bull. seism. Soc. Am.,* **96**(2), 467–489.

Deville, M., Fischer, P. & Mund, E., 2002. *High Order Methods for Incompressible Fluid Flow,* Cambridge Univ. Press.

Etienne, V., Hu, G., Operto, S., Virieux, J., Barkved, O. & Kommedal, J., 2012. Three-dimensional acoustic full waveform inversion: algorithm and application to Valhall, in *Proceedings of the 74th Annual EAGE Conference & Exhibition,* Expanded Abstracts, Copenhagen, EAGE.

Fathi, A., Poursartip, B., Stokoe, K.H.II & Kallivokas, L.F., 2016. Three-dimensional P-and S-wave velocity profiling of geotechnical sites using full-waveform inversion driven by field data, *Soil Dyn. Earthq. Eng.,* **87,** 63–81.

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.

Foti, S., Sambuelli, L., Socco, V.L. & Strobbia, C., 2003. Experiments of joint acquisition of seismic refraction and surface wave data, *Near Surf. Geophys.,* **1**(3), 119–129.

Gao, L., Pan, Y., Tian, G. & Xia, J., 2018. Estimating q factor from multimode shallow-seismic surface waves, *Pure appl. Geophys.,* **175**(8), 2609–2622.

Gao, L., Pan, Y. & Bohlen, T., 2020. 2-D multiparameter viscoelastic shallow-seismic full-waveform inversion: reconstruction tests and first field-data application, *Geophys. J. Int.,* **222**(1), 560–571.

Göktürkler, G., Balkaya, Ç. & Erhan, Z., 2008. Geophysical investigation of a landslide: the Altındağ Landslide Site, İZmir (Western Turkey), *J. appl. Geophys.,* **65**(2), 84–96.

Górszczyk, A., Operto, S. & Malinowski, M., 2017. Toward a robust workflow for deep crustal imaging by FWI of OBS data: the eastern Nankai trough revisited, *J. geophys. Res.,* **122**(6), 4601–4630.

Griewank, A. & Walther, A., 2000. Algorithm 799: revolve: an implementation of checkpointing for the reverse or adjoint mode of computational differentiation, *ACM Trans. Math. Software,* **26,** 19–45.

Hauksson, E. & Shearer, P.M., 2006. Attenuation models (Qp and Qs) in three dimensions of the southern California crust: inferred fluid saturation at seismogenic depths, *J. geophys. Res.,* **111**(B5), doi:10.1029/2005JB003947.

Häusler, M., Schmelzbach, C. & Sollberger, D., 2018. The Galperin source: a novel efficient multicomponent seismic source, *Geophysics,* **83**(6), P19–P27.

He, W., Brossier, R., Métivier, L. & Plessix, R.-É., 2019. Land seismic multi-parameter full waveform inversion in elastic VTI media by simultaneously interpreting body waves and surface waves with an optimal transport based objective function, *Geophys. J. Int.,* **219**(3), 1970–1988.

Irnaka, M., Brossier, R. & Métivier, L., 2018. 3x3C seismic's sensitivity analysis on near-surface towards full waveform inversion, in *Proceedings of the EAGE-HAGI 1st Asia Pacific Meeting on Near Surface Geoscience & Engineering,* Expanded Abstracts, Yogyakarta.

Kamath, N., Brossier, R., Metivier, L., Pladys, A. & Yang, P., 2020. Multiparameter full-waveform inversion of 3D OBC data from the Valhall field, *Geophysics,* **86**(1), doi:10.1190/geo2019-0705.1.

Kjartansson, E., 1979. Constant Q-wave propagation and attenuation, *J. geophys. Res.,* **84**(B9), 4737–4748.

Köhn, D., Meier, T., Fehr, M., De Nil, D. & Auras, M., 2016. Application of 2D elastic Rayleigh waveform inversion to ultrasonic laboratory and field data, *Near Surf. Geophys.,* **14**(5), 461–476.

Köhn, D., Wilken, D., De Nil, D., Wunderlich, T., Rabbel, W., Werther, L., Schmidt, J., Zielhofer, C. & Linzen, S., 2019. Comparison of time-domain sh waveform inversion strategies based on sequential low and bandpass filtered data for improved resolution in near-surface prospecting, *J. appl. Geophys.,* **160,** 69–83.

Komatitsch, D., Barnes, C. & Tromp, J., 2000. Simulation of anisotropic wave propagation based upon a spectral element method, *Geophysics,* **65**(4), 1251–1260.

Komatitsch, D., Xie, Z., Bozdağ, E., de Andrade, E.S., Peter, D., Liu, Q. & Tromp, J., 2016. Anelastic sensitivity kernels with parsimonious storage for adjoint tomography and full waveform inversion, *Geophys. J. Int.,* **206**(3), 1467–1478.

Kramer, S., 1996. *Geotechnical Earthquake Engineering,* Prentice-Hall, Inc, pp. 348–422.

Lamert, A. & Friederich, W., 2019. Full waveform inversion for advance exploration of ground properties in mechanized tunneling, *Int. J. Civil Eng.,* **17**(1), 19–32.

Lang, K. *et al.*, 1907. *Die Ettlinger Linien und ihre Geschichte,* Selbstverlag der Stadt Ettlingen.

Liu, Q. & Peter, D., 2019. Square-root variable metric based elastic full-waveform inversion-part 2: uncertainty estimation, *Geophys. J. Int.,* **218**(2), 1100–1120.

Lu, Y., Stehly, L., Paul, A. & Group, A.W., 2018. High-resolution surface wave tomography of the european crust and uppermost mantle from ambient seismic noise, *Geophys. J. Int.,* **214**(2), 1136–1150.

Lynn, J.A., 2013. *The Wars of Louis XIV 1667-1714,* Routledge.

Martínez, K. & Mendoza, J.A., 2011. Urban seismic site investigations for a new metro in central Copenhagen: near surface imaging using reflection, refraction and VSP methods, *Phys. Chem. Earth, Parts A/B/C,* **36**(16), 1228–1236.

Métivier, L. & Brossier, R., 2016. The seiscope optimization toolbox: a large-scale nonlinear optimization library based on reverse communication, *Geophysics,* **81**(2), F11–F25.

Modrak, R. & Tromp, J., 2016. Seismic waveform inversion best practices: regional, global and exploration test cases, *Geophys. J. Int.,* **206**(3), 1864–1889.

Modrak, R., Borisov, D., Lefebvre, M. & Tromp, J., 2018. The seisflows framework for waveform inversion: rapid prototyping and HPC applications, *Comput. Geosci.,* **115,** 88–95.

Nguyen, T.D. & Tran, K.T., 2018. Site characterization with 3D elastic full-waveform tomography, *Geophysics,* **83**(5), R389–R400.

Nocedal, J. & Wright, S.J., 2006. *Numerical Optimization,* 2nd edn, Springer,

Palmer, D., 1980. *The Generalized Reciprocal Method of Seismic Refraction Interpretation,* Society of Exploration Geophysicists.

Pan, Y., Schaneng, S., Steinweg, T. & Bohlen, T., 2018. Estimating S-wave velocities from 3D 9-component shallow seismic data using local Rayleigh-wave dispersion curves – a field study, *J. appl. Geophys.,* **159,** 532–539.

Pan, Y., 2021. 3D shallow-seismic data acquired at Rheinstetten, doi: 10.5445/IR/1000125628.

Park, C., Miller, R. & Xia, J., 1999. Multichannel analysis of surface waves, *Geophysics,* **64,** 800–808.

Patera, A.T., 1984. A spectral element method for fluid dynamics: laminar flow in a channel expansion, *J. Comput. Phys.,* **54,** 468–488.

Pegah, E. & Liu, H., 2016. Application of near-surface seismic refraction tomography and multichannel analysis of surface waves for geotechnical site characterizations: a case study, *Eng. Geol.,* **208,** 100–113.

Peter, D. *et al.*, 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.,* **186**(2), 721–739.

Peters, B. & Herrmann, F.J., 2017. Constraints versus penalties for edge-preserving full-waveform inversion, *Leading Edge,* **36**(1), 94–100.

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.

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., Operto, S. & Virieux, J., 2013. Multiparameter full waveform inversion of multicomponent OBC data from Valhall. Part 2: imaging compressional and shear-wave velocities, *Geophys. J. Int.,* **194**(3), 1665–1681.

Roberts, J.C. & Asten, M.W., 2004. Resolving a velocity inversion at the geotechnical scale using the microtremor (passive seismic) survey method, *Explor. Geophys.,* **35**(1), 14–18.

Samyn, K., Travelletti, J., Bitri, A., Grandjean, G. & Malet, J.-P., 2012. Characterization of a landslide geometry using 3D seismic refraction traveltime tomography: the La Valette landslide case history, *J. appl. Geophys.,* **86,** 120–132.

Signanini, P. & Torrese, P., 2004. Application of high resolution shear-wave seismic methods to a geotechnical problem, *Bull. Eng. Geol. Environ.,* **63**(4), 329–336.

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.

Smith, J.A., Borisov, D., Cudney, H., Miller, R.D., Modrak, R., Moran, M., Peterie, S.L., Sloan, S.D., Tromp, J. & Wang, Y., 2019. Tunnel detection at Yuma Proving Ground, Arizona, USA–Part 2: 3D full-waveform inversion experiments, *Geophysics,* **84**(1), B95–B108.

Sodani, A., Gramunt, R., Corbal, J., Kim, H.-S., Vinod, K., Chinthamani, S., Hutsell, S., Agarwal, R. & Liu, Y.-C., 2016. Knights landing: second-generation Intel Xeon Phi product, *Ieee micro,* **36**(2), 34–46.

Stewart, I., Williams, J. & Walker, C., 1997. Seismic refraction in relation to geotechnical information for (road) construction contracts, *Geol. Soc., Lond., Eng. Geol. Spec. Publ.,* **12**(1), 383–390.

Stokoe, K.H., Santamarina, J.C. *et al.*, 2000. Seismic-wave-based testing in geotechnical engineering, in *ISRM International Symposium,* International Society for Rock Mechanics and Rock Engineering.

Symes, W.W., 2007. Reverse time migration with optimal checkpointing, *Geophysics,* **72**(5), SM213–SM221.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.,* **180,** 433–462.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics,* **49**(8), 1259–1266.

Trinh, P., Brossier, R., Lemaistre, L., Métivier, L. & Virieux, J., 2019a. 3D elastic FWI with a non-linear model constraint: application to a real complex onshore dataset, in *Proceedings of the 81th Annual EAGE Conference & Exhibition,* London, Expanded Abstracts, EAGE.

Trinh, P.T., 2018. 3D multi-parameters Full Waveform Inversion for challenging land targets, *PhD thesis*, University Grenoble Alpes.

Trinh, P.T., Brossier, R., Métivier, L., Virieux, J. & Wellington, P., 2017. Bessel smoothing filter for spectral element mesh, *Geophys. J. Int.,* **209**(3), 1489–1512.

Trinh, P.T., Brossier, R., Métivier, L., Tavard, L. & Virieux, J., 2019b. Efficient 3D time-domain elastic and viscoelastic Full Waveform Inversion using a spectral-element method on flexible Cartesian-based mesh, *Geophysics,* **84**(1), R75–R97.

Vafidis, A., Manakou, M., Kritikakis, G., Voganatsis, D., Sarris, A. & Kalpaxis, T., 2003. Mapping the ancient port at the archaeological site of Itanos (Greece) using shallow seismic methods, *Archaeol. Prospect.,* **10**(3), 163–173.

Vigh, D., Jiao, K., Watts, D. & Sun, D., 2014. Elastic full-waveform inversion application using multicomponent measurements of seismic data collection, *Geophysics,* **79**(2), R63–R77.

Virieux, J. & Operto, S., 2009. An overview of full waveform inversion in exploration geophysics, *Geophysics,* **74**(6), WCC1–WCC26.

Wang, Y. *et al.*, 2019. Tunnel detection at Yuma Proving Ground, Arizona, USA–part 1: 2D full-waveform inversion experiment, *Geophysics,* **84**(1), B95–B105.

Wegscheider, S., 2017. Abbildung der Ettlinger Linie auf dem Segelflugplatz Rheinstetten mittels Georadar, *Master's thesis*, Karlsruhe Institute of Technology.

Weil, M.H., DeJong, J.T., Martinez, B.C. & Mortensen, B.M., 2012. Seismic and resistivity measurements for real-time monitoring of microbially induced calcite precipitation in sand, *Geotechn. Test. J.,* **35**(2), 330–341.

Wittkamp, F., Athanasopoulos, N. & Bohlen, T., 2018. Individual and joint 2-D elastic full-waveform inversion of Rayleigh and Love waves, *Geophys. J. Int.,* **216**(1), 350–364.

Wynn, J.C., 1986. A review of geophysical methods used in archaeology, *Geoarchaeology,* **1**(3), 245–257.

Yang, P., Brossier, R., Métivier, L. & Virieux, J., 2016a. Checkpointing-assisted reverse forward simulation: an optimal recomputation method for FWI and RTM, in *SEG Technical Program Expanded Abstracts 2016,* pp. 1089–1093.

Yang, P., Brossier, R., Métivier, L. & Virieux, J., 2016b. A review on the systematic formulation of 3D multiparameter full waveform inversion in viscoelastic medium, *Geophys. J. Int.,* **207**(1), 129–149.

APPENDIX A: DATA CORRECTION USING MATCHING FILTER

The 3-D seismic acquisition in the Ettlingen Line was performed using two different acquisition grids, the coarse and dense grid acquisition. The dense grid acquisition was performed part by part due to limited number of equipment. This type of acquisition raises some difficulties related to the potential differences in source-ground coupling, trigger time and surface condition. Our recorded data shows that this type of acquisition introduces discrepancies in terms of the amplitude and phase between receivers from different acquisition subsets. This discrepancy can be mitigated by taking advantage of the coarse grid acquisition that share acquisition locations with dense grid and the matching filter technique. In our experiment, some receivers locations are common between both acquisition grids. A matching filter $f_{i,\,sc,\,s}(t)$ can be computed such that the dense grid data $d_{i,\,sc,\,s,\,j,\,sr}(t)$ is matched to the coarse grid data reference $c_{i,\,sc,\,j,\,sr}(t)$

$$c_{i,sc,j,sr}(t) = d_{i,sc,s,j,sr}(t) * f_{i,sc,s}(t), \tag{A1}$$

where $*$ is the temporal convolution operator, $i =$ is source location index from 1 to N_s, $sc =$ is source component index with $sc \in \{U, V, W\}$, $j =$ is the receiver location index from 1 to N_r, $sr =$ is receiver component index with $sr \in \{X, Y, Z\}$ and s is the acquisition patch index from 1 to N_a. With such frame, one matching filter is expected per source location, source component and per acquisition patch. The computation of this filter relies on the solution of the following linear inverse deconvolution problem

$$\mathcal{C}(f_{i,sc,s}(t)) = \sum_{j,sr} \frac{1}{2} \left\| c_{i,sc,j,sr}(t) - f_{i,sc,s}(t) * d_{i,sc,s,j,sr}(t) \right\|^2, \tag{A2}$$

where $\mathcal{C}(f_{i,sc,s}(t))$ is the least-squares misfit function to be minimized. Writing the equation in the frequency domain gives

$$\mathcal{C}(f_{i,sc,s}(\omega)) = \sum_{j,sr} \frac{1}{2} \left\| c_{i,sc,j,sr}(\omega) - \widehat{f}_{i,sc,s}(\omega)\hat{d}_{i,sc,s,j,sr}(\omega) \right\|^2. \tag{A3}$$

The optimal solution of the matching filter minimizing eq. (A3) can be determined by zeroing its gradient, giving

$$\begin{aligned} \frac{\partial \mathcal{C}(f_{i,sc,s}(\omega))}{\partial \widehat{f}_{i,sc,s}(\omega)} = \sum_{j,sr} & -\overline{\hat{d}_{i,sc,s,j,sr}(\omega)}(c_{i,sc,j,sr}(\omega) \\ & - \widehat{f}_{i,sc,s}(\omega)\hat{d}_{i,sc,s,j,sr}(\omega)) = 0. \end{aligned} \tag{A4}$$

The matching filter can be calculated by rearranging eq. (A4) into

$$\widehat{f}_{i,sc,s}(\omega) = \frac{\sum_{j,sr} \overline{\hat{d}_{i,sc,s,j,sr}(\omega)}\hat{c}_{i,sc,j,sr}(\omega)}{\sum_{j,sr} \overline{\hat{d}_{i,sc,s,j,sr}(\omega)}\hat{d}_{i,sc,s,j,sr}(\omega) + \varepsilon}, \tag{A5}$$

where ε is the stabilization factor for the matching filter. Finally, the corrected dense acquisition data d_{ij_p} can be calculated through the convolution of the rough dense grid data d_{ij_t} and the filter

$$d_{ij_p} = d_{ij_t} * f_{i,sc,s}(t). \tag{A6}$$

APPENDIX B: NON-LINEAR MODEL CONSTRAINTS FWI

The objective of our experiment is to reconstruct simultaneously V_P and V_S. Parameter binding strategy which is proposed as the first step will only reconstruct V_S. While V_S is updated, it depends on the hard-constrained V_P/V_S value. On the second step, we have a multiparameter inversion (V_P and V_S). In addition to the hard-bound constraints of V_P and V_S, we also impose another non-linear constraint on the relationship between V_P and V_S. This constraint limit the V_P/V_S value over a specific interval. In the case of non-linear constraint, we would like to solve the following minimization problem

$$\begin{aligned} &\min_{V_P,V_S} f(V_P,V_S) = \frac{(\overline{V_P} - V_P)^2}{V_P^2} + \frac{(\overline{V_S} - V_S)^2}{V_S^2}, \\ &\text{subject to} \qquad r_1 \le V_P/V_S \le r_2, \end{aligned} \tag{B1}$$

where $\overline{V_P}$ and $\overline{V_S}$ are the updated V_P and V_S value given by the optimization for each iteration, V_P and V_S represent a projected value with respect to the r_1 and r_2 boundary. eq. (B1) demonstrates that the projected value of V_P and V_S must be within the given interval, and the value of V_P and V_S must be as close as possible with respect to the $\overline{V_P}$ and $\overline{V_S}$. In order to determine the V_P and V_S, there exists an analytical solution when $V_P/V_S > r_2$

$$V_P = r_2 V_S \qquad \text{and} \qquad V_S = \frac{\overline{V_S V_P}(r_2\overline{V_S} + \overline{V_P})}{(r_2\overline{V_S})^2 + \overline{V_P}^2}, \tag{B2}$$

and when $V_P/V_S < r_1$

$$V_P = r_1 V_S \qquad \text{and} \qquad V_S = \frac{\overline{V_S V_P}(r_1\overline{V_S} + \overline{V_P})}{(r_1\overline{V_S})^2 + \overline{V_P}^2}. \tag{B3}$$

Following Trinh (2018), both constraints $\mathcal{C}_1$ and $\mathcal{C}_2$ (bound constraints, ratio constraints) can be solved using Dykstra's algorithm (Boyle & Dykstra 1986).

APPENDIX C: SOURCE ESTIMATION

In shallow-seismic data where the source is manually performed by the sledgehammer, there are several potential problems affecting the source–time function. The problems include a potential unbalance in the relative source energies and the risk to incorporate subsurface-related contribution into the source wavelet. Those problems could lead to both unbalance contribution of each source and insensitive inversion.

From this status, our source–time function estimation procedure is the following, for each source location and orientation:

(i) Raw source estimation $s_{i,sc}(t)$ following Pratt (1999) approach. We assume that the data $d_{i,sc,j,sr}$ for source location index i, source component sc, receiver location index j and receiver component sr, can be written as the convolution between $g_{i,sc,s,j,sr}(t)$ the Green function for the same set of source and receiver for an impulse Dirac delta function and $s_{i,sc}(t)$ the source–time function $d_{i,sc,s,j,sr}(t) = g_{i,sc,s,j,sr}(t) * s_{i,sc}(t)$.

The unknown source–time function can be estimated in the frequency-domain by solving the linear inverse problem, whose misfit function is given by

$$\mathcal{C}(s_{i,sc}(\omega)) = \sum_{j,sr} \frac{1}{2} \left\| d_{obs_{i,sc,j,sr}}(\omega) - g_{cal_{i,sc,j,sr}}(\omega) \times s_{i,sc}(\omega) \right\|^2. \tag{C1}$$

Zeroing the derivative of $\mathcal{C}(s_{i,sc}(\omega))$ with respect to $s_{i,sc}(\omega)$ gives

$$s_{i,sc}(\omega) = \frac{\sum_{j,sr} \overline{g_{cal_{i,sc,j,sr}}}(\omega) \times d_{obs_{i,sc,j,sr}}(\omega)}{\sum_{j,sr} \overline{g_{cal_{i,sc,j,sr}}}(\omega) \times g_{cal_{i,sc,j,sr}}(\omega) + \epsilon}, \tag{C2}$$

where $\overline{a}$ is the complex conjugate of a, and ϵ is a small value for stabilization of the deconvolution.

Table A1. Computational cost for 3-D FWI using elastic medium approximation.

No.	FWI	*Niter*	Time/∇	Seq. time	Memory	Storage	NCPU	Elap. time
1	15 Hz V_S	1	3.67	7.33	7559	7197	108	0.07
2	15 Hz V_S and V_P	5	3.67	32.99	7559	7191	108	0.31
3	25 Hz V_S	5	23.99	143.96	18 530	45 996	432	0.33
4	25 Hz V_S and V_P	17	23.47	610.23	18 530	45 996	432	1.41
5	35 Hz V_S	9	58.98	766.70	30 472	118 541	648	1.18
6	35 Hz V_S and V_P	9	59.11	591.08	30 472	118 541	648	0.91
7	45 Hz V_S	15	144.45	2311.14	47193	283 256	972	2.38
8	45 Hz V_S and V_P	15	139.28	4874.95	47 193	283 256	972	5.02
9	55 Hz V_S	33	342.10	11973.38	75 037	702 491	1296	9.24
10	55 Hz V_S and V_P	25	337.99	10139.69	75 037	702 491	1296	7.82
11	65 Hz V_S	28	473.52	13731.96	88 232	851 449	3456	3.97
12	65 Hz V_S and V_P	12	460.32	23936.43	88 232	851 449	3456	6.93

Note. The unit time is in hours. The total sequential CPU time (seq. time) is 69 119 hr for all frequency bands on Irene KNL HPC (TGCC) to calculate from the homogeneous initial model. The memory and storage unit is in megabytes (MB), and the indicated frequency is the highest frequency band starting from 3 Hz. The elapsed time (elap. time) is the time needed to perform the FWI in parallel. Time/∇ indicated the time (in hours) needed to calculate a single gradient of misfit function.

After estimating each frequency-domain component of $s_{i,sc}(\omega)$, the time domain $s_{i,sc}(t)$ version is obtained by inverse Fourier transform.

(ii) Amplitude normalization, which aims to equalize the energy contribution on each source. We calculate the scaling factor $c_{i,sc}$ for each source–time function $s_{i,sc}(t)$ as

$$c_{i,sc} = \frac{n}{\sum_t s_{i,sc}^2(t)} \tag{C3}$$

where n is an arbitrary constant, and the denominator represent the energy of the source–time function. Each source–time function is then corrected by $\tilde{s}_{i,sc}(t) = s_{i,sc}(t) \times c_{i,sc}$ and the observed data $d_{obs_{i,sc}}$ is also corrected by the same factor consistently.

(iii) Spatial weighting average, which aims to reduce dissimilarities between neighbouring sources. For each source $\tilde{s}_{i,sc}(t)$, a Gaussian-weighted average from the neighbourood is computed to preserve some characteristics of each source while maintaining some similarities properties in space. In our application, a standard deviation of 6 m is chosen for the Gaussian weights.

APPENDIX D: COMPUTATIONAL COST

The calculation of this FWI experiment is performed using Irene KNL. It is a high performance computer (HPC) managed by TGCC in France. Irene KNL is based on Intel KNL processor with Knightlanding architecture which use Many Integrated Core (MIC) architecture (Sodani *et al.* 2016). For each node, it has 68 cores, with 1.4 GHz CPU clock and 1.4GB RAM for each core. In practice, due to the limited memory of Irene KNL, we store the wavefield on a very high bandwidth disk (60 GB s^{-1}). The bandwidth is shared amongst the users therefore the Input/Output performance differ from time to time and affect the computational time.

The computational time needed to perform the whole set of FWI is 69 119 and 41 415 sequential hours for elastic and viscoelastic FWI, respectively (Table A1 and Table A2). Both tables represent the FWI which are started from the homogeneous initial model. The sequential hour represents the computational time needed to calculate a task only using a single processor, disregarding its memory requirement. V_S FWI represents FWI with parameter binding, whereas V_S and V_P represents FWI with non-linear model constraints. The CPU time which is required to estimate a single gradient increase with the increase of the degree of freedom, time step, and frequency band. The number of element and the time step can be seen in Table A3 for elastic case and Table A4 for viscoelastic case. The global CPU time, on the other hand, is highly influenced by the total number of the gradient as well as the time needed to perform the communication between the processor and the disk. Since we stored the resampled wavefield directly on the disk, the performance of the I/O also depends on the global I/O load on the whole HPC cluster.

The total CPU time is more expensive for shallow-seismic applications compared with other methodologies (FATT, MASW). Nevertheless, when we look at the table, for the elastic FWI case almost 87 per cent of the total computational hours were spent on the last two frequency bands. It means, with 9338 sequential hours, 47 GB of RAM, and 283 GB of fast storage for homogeneous initial model, we can already compute FWI up to 45 Hz and obtain resolution beyond FATT and MASW. It is currently not cheap for a personal workstation, but it is still feasible to be done and with the added benefit of the reconstruction of V_P.

Table A2. Computational cost for 3-D FWI using viscoelastic medium approximation.

No.	FWI	N_{iter}	Time/∇	Seq. time	Memory	Storage	NCPU	Elap. time
1	15 Hz V_S	1	7.81	15.63	9273	50 382	108	0.14
2	15 Hz V_S and V_P	4	7.83	62.62	9273	50 382	108	0.58
3	25 Hz V_S	23	25.83	697.51	15 254	137 909	864	0.81
4	25 Hz V_S and V_P	6	26.33	158.00	15 254	137 909	864	0.18
5	35 Hz V_S	23	109.27	2731.66	27 672	494 284	1296	2.11
6	35 Hz V_S and V_P	10	92.30	2769.14	27 672	494 284	1296	2.14
7	45 Hz V_S	10	285.66	2856.57	42 878	1140 998	1296	2.20
8	45 Hz V_S and V_P	11	327.20	3926.42	42 878	1140 998	1296	3.03
9	55 Hz V_S	14	424.94	7224.00	64 184	2355 901	2592	2.79
10	55 Hz V_S and V_P	1	537.10	537.10	64 184	2355 901	2592	0.21
11	65 Hz V_S	2	812.20	19492.88	81 566	3497 627	3240	6.02
12	65 Hz V_S and V_P	1	944.43	944.43	81 566	3497 627	3240	0.29

Note The unit time is in hours. The total sequential CPU time (seq. time) is 41 415 hr for all frequency bands on Irene KNL HPC (TGCC) to calculate from the homogeneous initial model. The memory and storage unit is in megabytes (MB), and the indicated frequency is the highest frequency band starting from 3 Hz. The elapsed time (elap. time) is the time needed to perform the FWI in parallel. Time/∇ indicated the time (in hours) needed to calculate a single gradient of misfit function.

Table A3. Seismic wavefield simulation information for 3-D FWI using elastic medium approximation.

No.	FWI	N_Z	N_X	N_Y	Time step	N_T
1	15 Hz	5	11	10	4.39039e-04	1120
3	25 Hz	7	18	15	1.49172e-04	3320
5	35 Hz	10	22	20	1.01776e-04	4880
7	45 Hz	11	29	26	6.94629e-05	7160
9	55 Hz	12	38	34	4.59136e-05	10 880
11	65 Hz	13	43	38	4.99168e-05	10 000

Note. N_Z, N_X and N_Y is the number of element for each Z, X and Y direction, respectively. The sponge boundary elements are not included. Time steps are indicated in second. N_T is the number of time step.

Table A4. Seismic wavefield simulation information for 3-D FWI using viscoelastic medium approximation.

No	FWI	N_Z	N_X	N_Y	Time step	N_T
1	15 Hz	5	11	10	4.39039e-04	1120
3	25 Hz	6	15	13	2.62543e-04	1880
5	35 Hz	8	19	17	1.22664e-04	4040
7	45 Hz	9	25	23	8.64095e-05	5760
9	55 Hz	12	38	34	4.59136e-05	10 880
11	65 Hz	11	31	28	6.71441e-05	7440

Note. N_Z, N_X and N_Y is the number of element for each Z, X and Y direction, respectively. The sponge boundary elements are not included. Time steps are indicated in second. N_T is the number of time step.