Subsoil density field reconstruction through 3-D FWI
a systematic comparison between vertical- and horizontal-f or ce seismic sources

Kawasaki, Yusuke; Minato, Shohei; Ghose, Ranajit

DOI
10.1093/gji/ggad445

Publication date
2023

Document Version
Final published version

Published in
Geophysical Journal International

Citation (APA)
Kawasaki, Y., Minato, S., & Ghose, R. (2023). Subsoil density field reconstruction through 3-D FWI: a systematic comparison between vertical- and horizontal-f or ce seismic sources. *Geophysical Journal International*, *236*(2), 727-747. https://doi.org/10.1093/gji/ggad445

Geophysical Journal International

Geophys. J. Int. (2024) **236,** 727–747 https://doi.org/10.1093/gji/ggad445
Advance Access publication 2023 November 15
GJI Applied and Marine Geophysics

Subsoil density field reconstruction through 3-D FWI: a systematic comparison between vertical- and horizontal-force seismic sources

Yusuke Kawasaki,[1,2] Shohei Minato[1,3] and Ranajit Ghose[1]

[1]*Department of Geoscience and Engineering, Delft University of Technology, Stevinweg* 1, 2628 *CN Delft, The Netherlands. E-mail: Y.Kawasaki@tudelft.nl*
[2]*OYO Corporation, Engineering Headquarters,* 1-66-2 *Miyahara-cho, Kita-ku, Saitama-shi, Saitama* 331-0812, *Japan*
[3]*National Institute of Advanced Industrial Science and Technology (AIST), Geological Survey of Japan, Tsukuba Central* 7, 1-1-1, *Higashi, Tsukuba, Ibaraki* 305-8567, *Japan*

Accepted 2023 November 8. Received 2023 September 30; in original form 2023 May 12

SUMMARY

Bulk-density (ρ) of soil is an important indicator of soil compaction and type. A knowledge of the spatial variability of *in situ* soil density is important in geotechnical engineering, hydrology and agriculture. Surface geophysical methods have so far shown limited success in providing an accurate and high-resolution image of 3-D soil-density distribution. In this pursuit, 3-D seismic full-waveform inversion (FWI) is promising, provided the robustness and accuracy of density inversion via this approach can be established in the near-surface scale. However, simultaneous reconstruction of ρ and seismic wave velocities through multiparameter FWI remains a challenging task. Near-surface seismic data are commonly dominated by dispersive surface waves whose velocities are controlled by the value and distribution of shear-wave velocity (V_S). One major difficulty in estimating reliably ρ from near-surface seismic data is due to the relatively low sensitivity of the seismic wavefield to ρ compared to seismic velocities. Additionally, the accuracy of the estimated ρ decreases due to error in V_S—an issue known as parameter coupling. Parameter coupling makes it difficult to estimate accurately ρ within the framework of conventional gradient-based FWI. More sophisticated optimization approaches (e.g. truncated Newton) can reduce the effect of parameter coupling, but these approaches are commonly not affordable in near-surface applications due to heavy computational burden. In this research, we have investigated how choosing correctly the force direction of the seismic source can contribute to a higher accuracy of ρ estimates through 3-D FWI. Using scattered wavefields, the Hessian, and inversion tests, an in-depth and systematic investigation of data sets corresponding to different force directions has been carried out. A comparison of the scattered wavefields due to a point-localized ρ perturbation for different force directions shows the robustness of the horizontal-force data set to noise compared to the vertical-force data set. Furthermore, for a point-scatterer model, an analysis of the gradients of the misfit function using the Hessian shows that utilizing a horizontal-force source enables one to reconstruct the high-resolution gradient with relatively small parameter coupling. Finally, inversion tests for two different subsoil models demonstrate that 3-D FWI on a horizontal-force-source seismic data set is capable of providing a more accurate 3-D ρ distribution in soil compared to a vertical-force-source data set. Our results show that the use of a horizontal-force source might allow avoiding computationally demanding, costly optimization approaches in 3-D FWI.

Key words: Numerical modelling; Surface waves and free oscillations; Waveform inversion.

1 INTRODUCTION

The distribution of density in the subsoil is generally quite heterogeneous. A knowledge of the 3-D distribution of soil density is beneficial in numerous engineering, and environmental applications, in urban planning and constructions, and in natural hazard assessments. Bulk density (ρ) controls the soil's ability to provide structural support and determine water/solute movement and soil aeration. Bulk density is regarded as a key factor controlling soil compaction. Subsoil bulk density is closely linked to physical, chemical and biological properties of the soil layers. Different soil types and soil textures/structures correspond to different soil density. Density is also related to key geotechnical or hydrological parameters like porosity or void ratio, hydraulic conductivity and small-strain shear modulus.

Soil density is conventionally measured by direct methods, for example excavation sampling, core and clod methods (Vanremortel & Shields 1993) or by indirect methods, for example radiation and regression approaches (Lobsey & Viscarra Rossel 2016). Although the direct methods are affected by sampling disturbances, both methods are generally expensive and time-consuming. These conventional approaches are limited to a given point from where the soil sample is collected or to a 1-D profile along a borehole. It is challenging to infer the 3-D spatial variability of subsoil density.

Surface geophysical methods have occasionally been used to map the distribution of soil density. These methods are advantageous because they are mostly non-invasive, are generally useful for mapping large areas and allow temporal monitoring. Microgravity method has been used for mapping variations in subsoil density related to cavities and voids and geological structures (e.g. Tuckwell *et al.* 2008; Arisona *et al.* 2018). Electrical and electromagnetic properties have been correlated to density and compaction distribution in the ground, mostly in the very shallow ($<2-3$ m) soil layers which are relevant for agriculture and hydrogeology (e.g. Weihnacht & Börner 2007; Allred *et al.* 2008; Franko & Grote 2013). The bulk density inferred from these approaches is an empirical estimate, with limited spatial resolution and accuracy. Seismic wave velocities, both for P and S waves, have also been correlated to compaction and bulk density distribution in soil (Donohue *et al.* 2012; Anbazhagan *et al.* 2016; Romero-Ruiz *et al.* 2021).

More recently, seismic full-waveform inversion (FWI) has been used to obtain the 2-D spatial variability of density in the near-surface (e.g. Dokter *et al.* 2017; Gao *et al.* 2020; Chen *et al.* 2021; Mecking *et al.* 2021). FWI has proven to be a powerful tool to reconstruct in high-resolution the subsurface properties by fitting the observed seismic data with the synthetic data. With the increase of computation power, 3-D FWI in the near-surface scale has lately been plausible (e.g. Tran *et al.* 2019, 2020; Irnaka *et al.* 2022; Irnaka 2022). 3-D FWI can capture subsurface heterogeneities more accurately than 2-D FWI because of incorporation of actual 3-D wave propagation in the subsurface (Butzer *et al.* 2013; Irnaka *et al.* 2022). However, most 3-D FWI studies so far concentrated on estimation of the high-resolution, near-surface seismic velocity field, and not density (e.g. Tran *et al.* 2019; Smith *et al.* 2019; Tran *et al.* 2020; Teodor *et al.* 2021; Irnaka *et al.* 2022; Irnaka 2022).

FWI is capable of evaluating simultaneously multiple parameters, namely seismic velocities (V_P, V_S), density (ρ) and attenuation (Q_P, Q_S), across a variety of spatial scales. Herein we will focus on 3-D ρ estimation in the near-surface scale using active seismic sources. The difficulties in resolving the ρ distribution with reasonable accuracy arise mainly from the low sensitivity of the seismic wavefield to ρ (compared to V_P and V_S) and strong coupling among multiple parameters, especially between ρ and V_S (Pan *et al.* 2018a). Near-surface seismic data are typically dominated by surface waves which are most sensitive to the V_S distribution and much less sensitive to ρ. As a result, any differences in waveform due to ρ perturbations can easily be hidden by noise in real-world data.

Several approaches have so far been proposed to mitigate the parameter coupling effect in FWI (Köhn *et al.* 2012; Métivier *et al.* 2015; Wang *et al.* 2016; Yang *et al.* 2016; Pan *et al.* 2018a; Gao *et al.* 2021). However, these approaches are mostly computationally prohibitive for near-surface applications. For instance, taking the shape of the misfit function between observed and synthetic seismic waveforms into consideration, Métivier *et al.* (2015) and Gao *et al.* (2021) have applied truncated Newton method with an accurate Hessian inverse. They have shown, using 2-D synthetic seismic data, the superiority of this approach over gradient-based optimization approaches such as non-linear conjugate gradient (NCG) and Broyden-Fletcher–Goldfarb–Shanno (BFGS) for multiparameter FWI. However, solving the Newton's equation iteratively using the second-order adjoint-state method requires at each non-linear iteration up to several tens of forward simulation more than the gradient-based inversion (Métivier *et al.* 2013). Because a typical unconsolidated and fully saturated soil column including very soft peat and clay layers can have V_S as low as 50 m s^{-1} and V_P about 1500 m s^{-1}, the forward calculation of the seismic wavefield requires very fine spatial and temporal sampling. This results in a very high computational cost. This computational burden, in addition to the low sensitivity of seismic data to ρ, have been the limiting factors for application of 3-D FWI for estimating high-resolution ρ distribution in the subsoil.

The parameter coupling effect has been conventionally evaluated by comparing the theoretical radiation pattern of the scattered wavefields between different parameter classes based on the Born approximation (e.g. Virieux & Operto 2009; Operto *et al.* 2013). This theoretical approach does not take the various factors included in seismic data into account, such as the complexity of a subsurface model, the effect of finite frequency and the acquisition geometry. This limitation can bring misunderstanding and/or misinterpretation of the parameter coupling effect. In order to quantitatively evaluate the parameter coupling on the gradient of the misfit function in a realistic situation, Pan *et al.* (2018a) has proposed a method that uses numerically calculated Hessian-vector products. However, this technique has not been applied to surface-wave dominated, 3-D near-surface seismic data.

To mitigate the low sensitivity and parameter coupling issues in FWI for ρ estimation, another idea is to make a judicious choice of the force direction while generating the seismic wavefield. Previous studies have shown that, compared to a vertical-force seismic source, using a horizontal-force source results in less parameter coupling in 2-D FWI and higher resolution of the obtained V_S and ρ models (Dokter *et al.* 2017; Wittkamp *et al.* 2019). If the parameter coupling issue is less severe in case a horizontal-force source is used, then that will lead to computational efficiency. Smith *et al.* (2019) and Irnaka (2022) have also shown that using a horizontal-force source has helped increasing the accuracy in the deeper part of the near-surface model obtained via 3-D FWI. However, the reasons why a horizontal-force source should offer better results in FWI (greater accuracy and less parameter coupling) than a vertical-force source are not yet sufficiently understood.

In this research, we systematically investigate the effect of using different force directions on the accuracy of the 3-D FWI. We first investigate the waveform differences due to a ρ perturbation, and the impact of different source-force directions on the noise robustness. To evaluate the parameter coupling effect, we then investigate the differences in the shape of the misfit function (i.e. the gradient and Hessian). We use the numerically calculated Hessian for this purpose. The Hessian is also used to compare the resolution of the estimated density. Finally, we perform 3-D FWI using synthetic data considering realistic near-surface subsoil models and different force directions.

This paper is organized as follows. First, we briefly explain the theory of FWI that is relevant for comparing systematically the three factors mentioned above (i.e. noise robustness, parameter coupling, and resolution) for the purpose of 3-D ρ estimation. We then explore in detail the effect of force directions on these three factors for a point-scatterer model. Lastly, we perform 3-D FWI for a spatially uncorrelated model and a realistic subsoil model derived from field measurements, and compare the accuracy of the reconstructed ρ for different force directions.

2 METHODOLOGY

2.1 3-D FWI

FWI solves a non-linear optimization problem to reconstruct the subsurface model by minimizing the misfit function E defined by the synthetic and the observed seismic data (e.g. Pratt *et al.* 1998). In this study, we use the square of l_2 norm of the waveform differences as the misfit function E:

$$E(m) = \frac{1}{2}\|u_{\text{syn}}(m) - u_{\text{obs}}\|^2, \quad (1)$$

where $u_{\text{syn}}(m)$ is the synthetic seismic (displacement) data calculated assuming a subsurface model m, and u_{obs} is the observed seismic data. The model parameters m are the elastic properties (V_P, V_S and ρ) that are functions of space in 3-D. We use the finite-difference time-domain (FDTD) method to calculate the synthetic seismic data (Virieux 1986; Levander 1988). We generate body force in different directions (f_x,f_y,f_z) in the Cartesian coordinate, at the same position and with the same amplitude (unit: N m^{-3}).

The minimization of the misfit function E in eq. (1) and finding an optimal solution for m over the large parameter space is a non-linear, large-scale problem. Therefore, a local optimization approach is commonly used to solve this problem (e.g. Mora 1987; Pratt *et al.* 1998; Brossier *et al.* 2009). The local optimization approach iteratively updates the model parameters m using the following formula:

$$m_{n+1} = m_n + \alpha_n \Delta m_n, \quad (2)$$

where m_n is the model parameter at nth iteration, Δm_n is the descent direction of the misfit function E and α_n is the step length that is estimated by a line search method (Nocedal & Wright 2006). Among various non-linear optimization approaches to calculate Δm_n, the gradient-based approaches, such as steepest descent (SD), non-linear conjugate gradient (NCG) and limited-memory Broyden–Fletcher–Goldfarb–Shanno (l-BFGS), are considered in this study, given the respective computation costs. In the case of SD, the following simple formula is used for the calculation of Δm_n:

$$\Delta m_n = -P_m \nabla_m E_n, \quad (3)$$

where $\nabla_m E_n$ is the gradient of the misfit function calculated by the adjoint-state method (Plessix 2006), and P_m is the preconditioning filter. The other approaches (NCG and l-BFGS) use the same formula (eq. 3) at the first iteration, and after that they use Δm_n which is calculated using $\nabla_m E_n$ at the current iteration n and the previous history of $\nabla_m E_n$. Therefore, the characteristics of $P_m \nabla_m E_n$ are the key elements for effectively solving the optimization problem using these approaches (SD, NCG and l-BFGS).

To reconstruct ρ models, we apply 3-D FWI to the synthetic data sets generated using different force directions for the seismic source. First, to reduce the memory requirements, we adopt the time-frequency approach (Sirgue *et al.* 2008, 2010): the seismic wavefield u_{syn} is simulated in the time domain, but the descent direction of the misfit function Δm_n is calculated in the frequency domain. During the non-linear inversion, the descent direction Δm_n is calculated using the NCG method (Nocedal & Wright 2006). Δm_n is then normalized by the maximum value of each parameter class, followed by scaling with the representative value of each model parameter. The same step length α_n is assumed for all parameter classes; α_n is estimated by a line search method in order to invert all model parameters simultaneously.

The preconditioning filter P_m in eq. (3) is also essential in controlling accuracy and efficiency of FWI. For P_m, we take the diagonal elements of the approximate Hessian for each parameter class (Appendix A). In FWI, this filter is conventionally used to compensate for the effect of the limited illumination on $\nabla_m E_n$, such as geometrical spreading (e.g. Ravaut *et al.* 2004; Operto *et al.* 2004, 2006). This filter also calculates suitable scaling for each parameter class, as the diagonal elements of the approximate Hessian contain the information of the scattering radiation pattern (for details see Appendix A). Finally, in order to reduce the computational cost to calculate the Hessian, we estimate P_m only at the first iteration, and use the same filter for the rest of the inversion. Note that this strategy can decrease the performance of FWI, for example convergence speed and accuracy of an estimated model, when the initial model is very far from an optimum model.

2.2 Approach to evaluate different force directions in 3-D FWI for density

We investigate the capability of 3-D FWI to estimate ρ by concentrating on three factors, namely noise robustness, parameter coupling, and resolution (Section 3). We then evaluate the reconstructed ρ models using different force directions (Section 4). In this subsection, we discuss first our approach of assessing the efficacy of these different force directions, considering the above-mentioned three factors and the theory introduced in Section 2.1.

2.2.1 Noise robustness

For real-world applications of FWI on noisy field data, the robustness is a crucial factor. As shown in eq. (1), FWI finds the model parameter m that minimizes the square of the l_2 norm of the waveform residuals. Suppose the observed data with noise can be divided into data calculated using the true model and noise, that is $u_{\text{obs}} = u_{\text{syn}}(m_{\text{true}}) + u_{\text{noise}}$. In this case, when we consider a specific model parameter $m = m_0$ ($\neq m_{\text{true}}$), eq. (1) can be written as,

$$E(m_0) = \frac{1}{2}\|u_{\text{sct}}(m_0) + u_{\text{noise}}\|^2, \quad (4)$$

where the scattered wavefield $u_{\text{sct}}(m_0)$ due to the model perturbation $(m_{\text{true}} - m_0)$ is defined as $u_{\text{sct}}(m_0) = u_{\text{syn}}(m_{\text{true}}) - u_{\text{syn}}(m_0)$. Eq. 4 shows that the value of E is determined by the amplitude of noise when the amplitude of u_{noise} is larger than that of $u_{\text{sct}}(m_0)$. This makes it difficult to effectively distinguish the difference between m_{true} and m_0.

We consider a specific pair of the model (m_0, m_{true}) where the difference is only in ρ, and other properties (V_P and V_S) remain the same. Then, we calculate numerically the scattered wavefield $u_{\text{sct}}(m_0)$ and the squared norm of the amplitudes. Here, seismic sources and receivers are distributed on the free surface. We compare the noise robustness for different force directions using the squared norm of the amplitude in the scattered wavefields.

2.2.2 Parameter coupling

As shown in eqs (2) and (3), the preconditioned gradient $P_m\nabla_m E_n$ contains key information to obtain the accurate descent direction Δm_n close to the true model-update direction at the first iteration when using SD, NCG and *l*-BFGS. The gradient $P_m\nabla_m E_n$ partly represents the shape of the misfit function E in the model domain. Generally, evaluating the shape of E is crucial in order to solve the inverse problem advantageously in terms of both accuracy and efficiency. For example, $-\nabla_m E_n$ represents the direction in which E decreases, but it is not always identical to the direction toward the true model: assuming that the misfit function E is a quadratic function, there is a possibility that the descent direction to an optimum solution of a certain parameter class (e.g. V_S) contaminates the descent direction of another parameter class (e.g. ρ). In this research, we delve into this parameter coupling contained in the gradient. As illustrated in eqs (2) and (3), the inversion at the first iteration seeks an optimum solution along the direction of $-P_m\nabla_m E_n$ using a line search method. This procedure implies that a significant parameter coupling effect in $-P_m\nabla_m E_n$ can make the model parameter fall into the local minima at early iterations, and the inversion fails to converge to an optimum solution. For FWI, using data sets corresponding to different force directions implies optimizing E having different shapes in the model domain. We, therefore, investigate the shape of E and explore the parameter coupling effect in $-P_m\nabla_m E_n$ due to the force direction.

First, we define the parameter coupling effect in $\nabla_m E$. When the misfit function E is quasi-linear, that is m_n is close to the optimum solution, E can be approximated as a quadratic function of m. In this case, $\nabla_m E$ follows the Newton's equation:

$$-\nabla_m E(m_n) = H(m_n)\Delta m^N, \tag{5}$$

where $H(m_n)$ is the Hessian of E at $m = m_n$, and Δm^N is the model-update direction toward the optimum solution. Eq. (5) indicates that the update direction based on the gradient (i.e. $-\nabla_m E(m_n)/|\nabla_m E(m_n)|$) is not identical to the Newton-step direction (i.e. $\Delta m^N/|\Delta m^N|$). This difference is the parameter coupling considered in this study, which is characterized by the Hessian $H(m_n)$ in eq. (5). Here, the Hessian has a multiparameter form $H_{m_i m_j}$ and is written using Jacobian ($\partial u_{\mathrm{syn}}/\partial m$) as follows:

$$\begin{aligned} H_{m_i m_j} &= \Re\left\{\left(\frac{\partial u_{\mathrm{syn}}}{\partial m_i}\right)^*\left(\frac{\partial u_{\mathrm{syn}}}{\partial m_j}\right) + \left(\frac{\partial^2 u_{\mathrm{syn}}}{\partial m_i \partial m_j}\right)^*(u_{\mathrm{syn}} - u_{\mathrm{obs}})\right\} \\ &= H^a_{m_i m_j} + R_{m_i m_j}, \end{aligned} \tag{6}$$

where m_i and m_j are model parameters (V_P, V_S, ρ), and the symbol * denotes the complex conjugate in the frequency domain. The first term $H^a_{m_i m_j}$ on the right-hand side of eq. (6) is the approximate Hessian, which is the cross-correlation between Jacobians with respect to m_i and m_j. The second term $R_{m_i m_j}$ on the right-hand side of eq. (6) is the cross-correlation between the second-order partial derivative wavefield and the data residual, which represents the second-order scattering (Pratt *et al.* 1998). In this study, assuming that the residuals are small due to the small model perturbations, the second term $R_{m_i m_j}$ is neglected and only the approximate Hessian $H^a_{m_i m_j}$ is considered. Thus, eq. (5) can be rewritten as follows:

$$-\begin{bmatrix} \nabla_{V_P} E \\ \nabla_{V_S} E \\ \nabla_\rho E \end{bmatrix} \approx \begin{bmatrix} H^a_{V_P V_P} & H^a_{V_P V_S} & H^a_{V_P \rho} \\ H^a_{V_S V_P} & H^a_{V_S V_S} & H^a_{V_S \rho} \\ H^a_{\rho V_P} & H^a_{\rho V_S} & H^a_{\rho\rho} \end{bmatrix} \begin{bmatrix} \Delta V_P{}^N \\ \Delta V_S{}^N \\ \Delta \rho^N \end{bmatrix}. \tag{7}$$

The diagonal elements of the matrix on the right-hand side of eq. (7) (e.g. $H^a_{\rho\rho}$) represent the coefficients for the Newton step (e.g. $\Delta\rho^N$) in the gradient of the same parameter class (e.g. $\nabla_\rho E$). On the other hand, the off-diagonal elements represent how the gradient for a certain parameter class (e.g. $\nabla_\rho E$) is influenced by the Newton step of the other parameter classes (e.g. $\Delta V_S{}^N$), that is the parameter coupling effect. Therefore, the effect in the gradient can be evaluated by investigating the relative strength of the off-diagonal elements with respect to the diagonal elements. Note that the diagonal elements are the auto-correlation of Jacobians for a certain model parameter class (i.e. $m_i = m_j$), and the off-diagonal elements are the cross-correlation between Jacobians for two different model parameter classes (i.e. $m_i \neq m_j$).

We evaluate eq. (7) using numerically calculated Hessian for a specific model in order to compare the parameter coupling effect for different force directions (Section 3.2). Furthermore, to take into account the preconditioned gradient in the actual implementation of FWI (eq. 3), we consider the preconditioning filter P_m in eq. (7). By multiplying both sides of eq. (7) with P_m, we have the following relations:

$$\begin{aligned} -P_{V_P}\nabla_{V_P} E &\approx P_{V_P} H^a_{V_P V_P}\Delta V_P{}^N + P_{V_P} H^a_{V_P V_S}\Delta V_S{}^N + P_{V_P} H^a_{V_P \rho}\Delta\rho^N \\ &= K_{V_P \leftrightarrow V_P} + K_{V_S \to V_P} + K_{\rho \to V_P}, \end{aligned} \tag{8}$$

$$\begin{aligned} -P_{V_S}\nabla_{V_S} E &\approx P_{V_S} H^a_{V_S V_P}\Delta V_P{}^N + P_{V_S} H^a_{V_S V_S}\Delta V_S{}^N + P_{V_S} H^a_{V_S \rho}\Delta\rho^N \\ &= K_{V_P \to V_S} + K_{V_S \leftrightarrow V_S} + K_{\rho \to V_S}, \end{aligned} \tag{9}$$

$$\begin{aligned} -P_\rho\nabla_\rho E &\approx P_\rho H^a_{\rho V_P}\Delta V_P{}^N + P_\rho H^a_{\rho V_S}\Delta V_S{}^N + P_\rho H^a_{\rho\rho}\Delta\rho^N \\ &= K_{V_P \to \rho} + K_{V_S \to \rho} + K_{\rho \leftrightarrow \rho}, \end{aligned} \tag{10}$$

where we call $K_{m_i \leftrightarrow m_i}$ and $K_{m_j \to m_i (m_i \neq m_j)}$ the update kernel and the contamination kernel, respectively (Pan *et al.* 2018a). Note that Pan *et al.* (2018a) do not include the preconditioning filter in the definition of the kernels. Furthermore, when Δm^N is localized at a single point, the kernels are identical to the point spread functions (PSFs, Fichtner & Trampert 2011; Fichtner & Leeuwen 2015; Pan *et al.* 2018a, 2019), which is used for evaluating the resolution in this study (see Section 2.2.3). In this research, we numerically calculate kernels in eqs (8)–(10) using a point scatterer model considering typical near-surface seismic acquisition parameters. We then investigate how the relative strength between the update and the contamination kernels differs due to different force directions.

2.2.3 Resolution

Resolution is another important indicator to evaluate the performance of FWI. For instance, in near-surface engineering problems, it is sometimes important to detect the existence of a thin layer in the subsurface or to image the accurate shape of an anomaly like a void. We investigate here how the resolution of FWI in such cases might differ for the different force directions.

First, we define the resolution using kernels introduced in Section 2.2.2. Given a point-localized perturbation as Δm^N in eqs (8)–(10), the spatial distribution of the value of the update kernel in the model domain represents how Δm^N is smeared around a point scatterer in the preconditioned gradient. On the other hand, the distribution of the values of the contamination kernel in the model domain represents the spatial extent where the parameter coupling prevails around a point scatterer in the preconditioned gradient. We define the spreading width of the kernels as the distance between the point scatterer and the location where the value of the kernels is half of the maximum value of the update kernel. A large spreading width of the update kernel will indicate low resolution of the reconstructed model, while a large spreading width of the contamination kernel

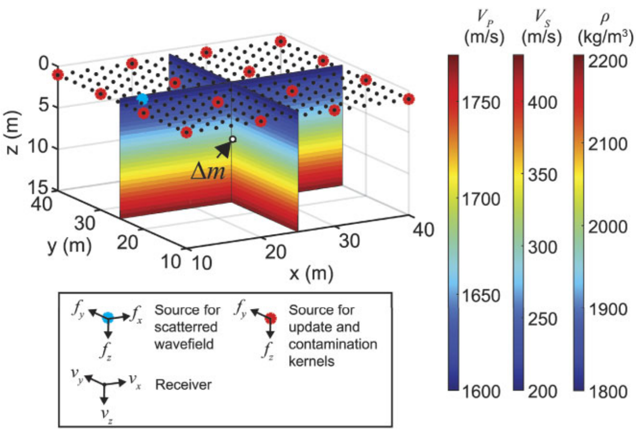

Figure 1. A synthetic model for investigating noise robustness, parameter coupling, and resolution.

will imply occurrence of spatial parameter-coupling in a wide area around the point scatterer.

Evaluating the resolution using the kernels defined above is similar to the resolution analysis in Pan *et al.* (2018a), except that we use the inverse of the diagonal elements of the approximate Hessian as the generalized inverse of the Hessian in the resolution matrix [see eq. 54 in Pan *et al.* (2018a)]. As before, we numerically calculate the kernels using a point scatterer model considering typical near-surface seismic acquisition parameters and then investigate how the spreading widths of the kernels differ for different force directions (Section 3.3).

3 RESULTS OF 3-D FWI FOR DIFFERENT FORCE DIRECTIONS: NOISE, PARAMETER COUPLING AND RESOLUTION

For these numerical tests, we consider a depth-dependent 3-D model with a point scatterer at a location where V_S and ρ perturbations are allocated at a single grid ($[x, y, z] = [25$ m, 25 m, 7.5 m$]$ in Fig. 1). The model is discretized with $100 \times 100 \times 50$ gridpoints in x-, y- and z-direction with a grid spacing of 0.5 m. The background values of V_P, V_S and ρ are kept within the typical ranges found for soils, assuming fully water saturated condition (V_P: 1600–1800 m s^{-1}, V_S: 200–450 m s^{-1}, ρ: 1800–2200 kg m^{-3}). Sources and receivers are distributed on the free surface as shown in Fig. 1. Such a setup for the numerical study allows investigating the scattered wavefields due to a point scatterer. Given a good initial model, FWI aims to reconstruct the true model using the scattered wavefield dominated by first-order scattering due to small perturbations in medium properties. The following discussion using a point scatterer can be applied to a complex wavefield in the real world, as such a wavefield can be approximated as superposition of scattered wavefields due to point scatterers distributed over the entire model space.

3.1 Noise robustness

We investigate for different source directions the noise robustness of ρ estimation using wavefields scattered due to a perturbation in ρ only (Section 2.2.1). For calculating the scattered wavefields, the background models without $\Delta\rho$ and with $\Delta\rho$ located at $[x, y, z] = [25$ m, 25 m, 7.5 m$]$ are denoted as m_0 and m_{true}, respectively (Fig. 1). As model perturbation, a $\Delta\rho$ which is -10 per cent of the background value is considered. We calculate 3-component particle-velocity wavefields (i.e. v_x, v_y, v_z) due to a seismic source, denoted by the cyan symbol in Fig. 1. For a comparison between different force directions, three sources with three different force directions (i.e. f_x, f_y, f_z) are considered at the same location. Here, the directions for the two horizontal-force sources are either in the 2-D plane containing both of the source and a point scatterer (x-direction) or perpendicular to it (y-direction), which enables one to take into account scattered wavefields generated by significantly different incident wavefields (i.e. SV/Rayleigh waves for an f_x source and SH/Love waves for an f_y source). We use the following Fuchs–Müller wavelet (Fuchs & Müller 1971) with a central frequency $f_c = 20$ Hz:

$$f(t) = \sin(2\pi t f_c) - 0.5\sin(4\pi t f_c). \tag{11}$$

To keep the computation time for the 3-D FWI manageable, this frequency is not chosen to a higher value. Sources and receivers are located on the free surface.

In Fig. 2, we show the snapshots at 0.14 s, illustrating different characteristics of the wavefield due to different force directions. With horizontal-force sources, the energy of the scattered wavefields at the location of $\Delta\rho$ dominates in the backward direction (see Figs 2a, c and e). These wavefields are generated by the incident SV/Rayleigh and SH/Love waves for the f_x and f_y sources, respectively; the surface waves (Rayleigh/Love waves) propagate parallel to the free surface toward the point scatterer, while the body waves (SV/SH waves) have incident angles of $\sim 60^\circ$. When using a vertical-force source (f_z), the incoming Rayleigh wave is back-scattered with a large amplitude (see Fig. 2g), while some energy is present also in the forward direction (see Figs 2g and i).

To address the noise robustness, we calculate the squared norm of the amplitude in the scattered wavefields over the depth slice shown in Fig. 2. A comparison of the sum of the squared norm (Table 1) shows that the scattered wavefields due to horizontal-force sources (f_x or f_y) have more energy than those due to a vertical-force source (f_z). Additionally, Table 1 shows that the scattered wavefield at the same receiver component as the source (i.e. f_x–v_x, f_y–v_y and f_z–v_z) has, as expected, the largest energy. Furthermore, the scattered wavefields for the horizontal sources (f_x–v_x or f_y–v_y) have much larger energy than the vertical force (f_z–v_z). This large energy in f_x–v_x or f_y–v_y wavefields are due to the presence of backward scattering with large amplitudes (see Figs 2a and e). Thus, in the context of ρ estimation, 3-D FWI applied to a horizontal-force data set is more robust to noise than that applied to a vertical-force data set (see Section 2.2.1), assuming same amplitude and source–time function for all sources. As the scattered wavefields generated by different incident wavefields from the two orthogonal horizontal-force sources (i.e. SV/Rayleigh waves for an f_x source and SH/Love waves for an f_y source) have larger energy than those due to a vertical-force source, the conclusion on noise robustness remains valid for the case when the source is located at any arbitrary location on the free surface.

3.2 Parameter coupling

Here we investigate the parameter coupling effect in the preconditioned gradient for ρ using the update and the contamination kernels calculated for the same point-scatterer model as in Section 3.1 (see Section 2.2.2). We compare the relative strength between these two kernels for different force directions.

For calculation of the kernels, 16 seismic sources located at 10 m spacing and 256 3-component receivers at 2 m spacing are used;

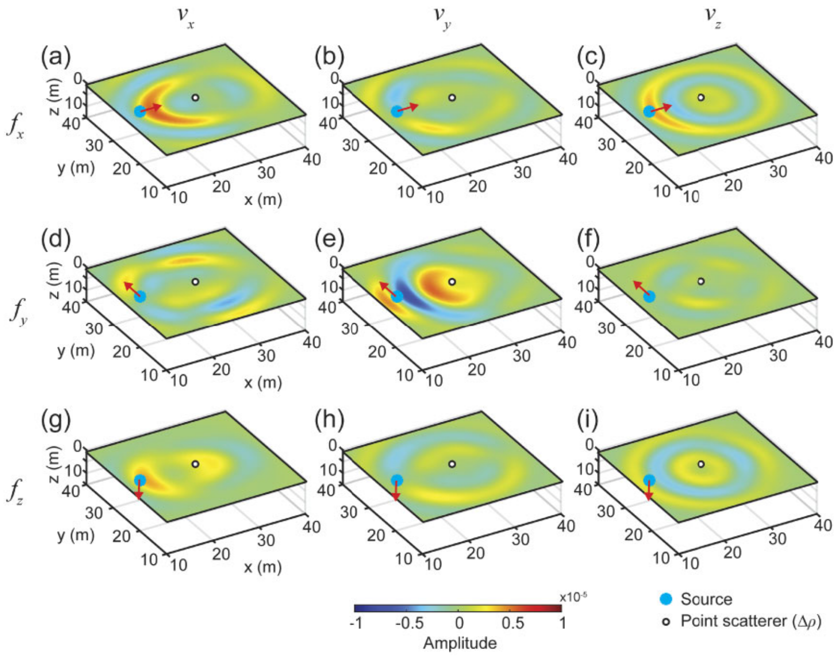

Figure 2. Depth slice of 9-component scattered wavefields (3-component source: f_x, f_y, f_z, 3-component particle velocity: v_x, v_y, v_z) due to a point scatterer $\Delta\rho$.

Table 1. Squared norm of the amplitude in the 9-component scattered wavefields shown in Fig. 2.

	v_x	v_y	v_z	Total
f_x	10.6e-09	4.1e-09	7.9e-09	22.6e-09
f_y	7.4e-09	16.7e-09	1.2e-09	25.2e-09
f_z	5.7e-09	4.3e-09	6.5e-09	16.4e-09

they are respectively denoted by red and black symbols in Fig. 1. The vertical-force sources (i.e. f_z) and horizontal-force sources (i.e. f_y) have the same source–time function (eq. 11), as in Section 3.1. We consider only one horizontal-force direction (i.e. f_y) because the model and the acquisition geometry are symmetric. The kernels (eqs 8–10) are calculated in the frequency domain using 8 monochromatic frequencies (i.e. 10, 14, 18, 22, 26, 30, 34 and 38 Hz). This setup enables us to evaluate the impact of a typical near-surface seismic acquisition geometry on the kernels. The kernels are calculated assuming a perturbation (−10 per cent of the background value) as the Newton step (Δm^N) at a point scatterer.

Fig. 3 presents a comparison of the preconditioned gradients for ρ calculated using the adjoint-state method (i.e. $-P_\rho \nabla_\rho E$ in eq. 10) and the update and contamination kernels for ρ (i.e. $K_{\rho\leftrightarrow\rho}$ and $K_{V_S\to\rho}$ in eq. 10) between data sets corresponding to different source-force directions. The relative strength between the update and the contamination kernels is compared based on the maximum magnitude of the kernels shown in Fig. 3. When using the f_z data set, $K_{V_S\to\rho}$ (Fig. 3c) has a larger relative strength than $K_{\rho\leftrightarrow\rho}$ (Fig. 3b). In other words, the effect of the perturbation of V_S (i.e. ΔV_S^N) dominates in $-P_\rho \nabla_\rho E$ (Fig. 3a); the strong parameter coupling occurs when updating the ρ model in the direction along $-P_\rho \nabla_\rho E$. On the contrary, $K_{V_S\to\rho}$ for the f_y data set (Fig. 3f) has a smaller relative strength than $K_{\rho\leftrightarrow\rho}$ (Fig. 3e). This implies that $-P_\rho \nabla_\rho E$ (Fig. 3d) is not significantly influenced by ΔV_S^N, and therefore the ρ model can be updated with a relatively weak parameter-coupling problem. From the above findings, the shape of E for the f_y data set has more favourable characteristics for ρ estimation using 3-D FWI (i.e. weak parameter coupling) than the shape of E for the f_z data set. Choosing f_y as a force direction would result in reconstructing a more accurate ρ model within the framework of gradient-based FWI, without taking the Hessian inverse into account.

The update and the contamination kernels for V_S are also calculated (Fig. 4). As opposed to ρ, the relative strengths of the contamination kernels ($K_{\rho\to V_S}$) shown in Figs 4(c) and (f) are smaller than those of the update kernels ($K_{V_S\leftrightarrow V_S}$) shown in Figs 4(b) and (e), regardless of the force directions. Thus, both data sets equally allow reconstruction of an accurate V_S model without suffering from a strong parameter coupling due to $\Delta\rho^N$.

3.3 Resolution

In this subsection, we investigate the resolution of ρ estimates using the update and contamination kernels calculated in Section 3.2. We compare the spreading widths (Section 2.2.3) between these two kernels for different force directions.

The spreading widths of the kernels for ρ are shown by yellow lines and black arrows in Fig. 3. The spreading widths of $K_{\rho\leftrightarrow\rho}$ do not significantly differ between f_z and f_y data sets (Figs 3b and e): the maximum widths for f_z and f_y data sets are 2.2 and 2.1 m, respectively. This indicates that the resolution in the reconstructed ρ model would be quite similar for different force directions if the contamination kernel $K_{V_S\to\rho}$ is small. On the contrary, the spreading width of $K_{V_S\to\rho}$ for the f_z data set (Fig. 3c) is much larger (3.8 m) than that for the f_y data set (1.9 m, Fig. 3f). Therefore, the use of the f_z data set can cause artefacts due to parameter coupling (ΔV_S^N) in a wide area in the reconstructed ρ model.

Next, we examine the spreading widths of the update and the contamination kernels for V_S (yellow lines and black arrows in Fig. 4) to evaluate how their characteristics differ from those for ρ.

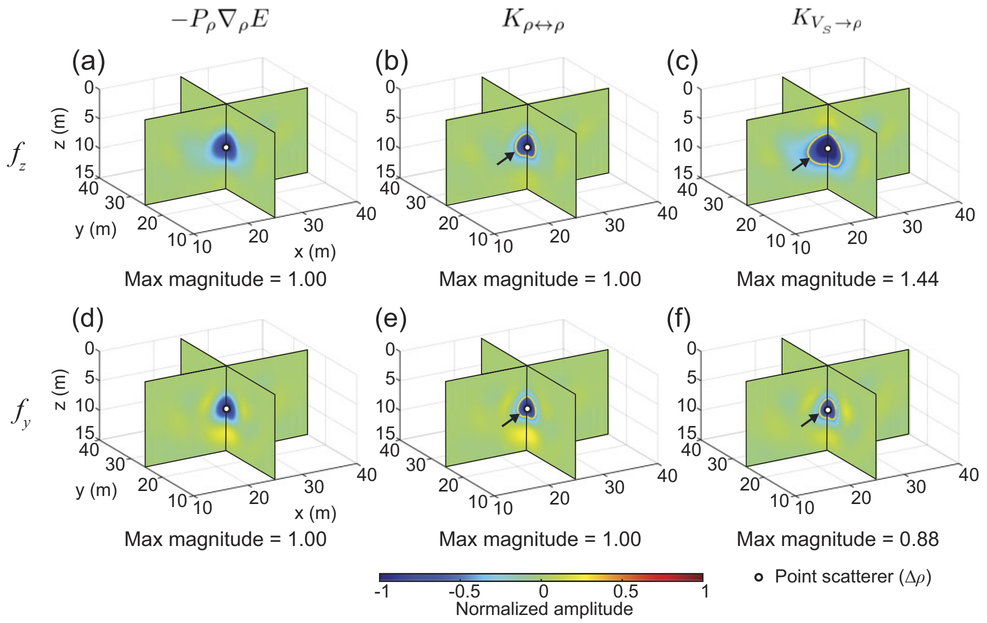

Figure 3. Preconditioned gradients, update kernels and contamination kernels for ρ. Preconditioned gradients in (a) and (d) are normalized by their maximum magnitude, while the kernels in (b), (c), (e) and (f) are normalized by the maximum magnitude of update kernels. Yellow lines and black arrows in update and contamination kernels denote the spreading widths of the kernels.

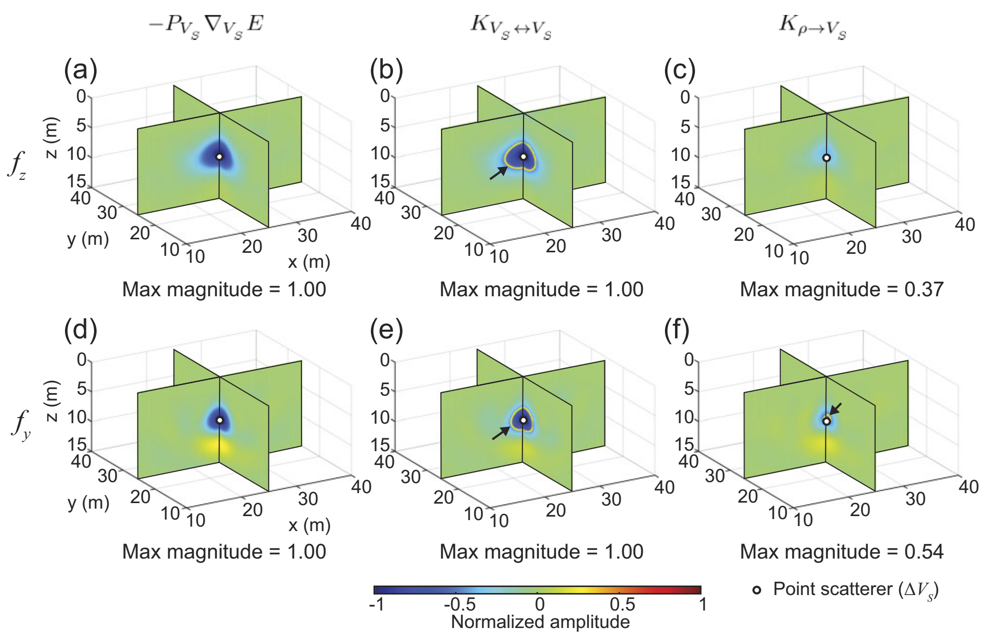

Figure 4. Preconditioned gradients, update kernels and contamination kernels for V_S. Preconditioned gradients in (a) and (d) are normalized by their maximum magnitude, while the kernels in (b), (c), (e) and (f) are normalized by the maximum magnitude of update kernels. Yellow lines and black arrows in update and contamination kernels denote the spreading widths of the kernels.

The spreading width of $K_{V_S\leftrightarrow V_S}$ for the f_z data set is larger than that for the f_y data set (Figs 4b and e): the maximum width for the f_z data set is 3.4 m, while that for the f_y data set is 2.5 m. Therefore, the use of the f_y data set will enable constructing a preconditioned gradient which is more focused at ΔV_S^N than using the f_z data set. Contrary to ρ, the spreading widths of $K_{\rho\rightarrow V_S}$ for both data sets are significantly small (<0.5 and 0.7 m, see Figs 4c and f), which can result in few contaminations due to $\Delta\rho^N$ in the V_S estimates for both force-direction data sets.

4 NEAR-SURFACE MODELS: RESULTS OF 3-D FWI FOR DIFFERENT FORCE DIRECTIONS

The results in the previous section have demonstrated that FWI applied to horizontal-force data sets can reconstruct more accurate ρ and V_S models than FWI applied to vertical-force data sets, in terms of the noise robustness, weak parameter coupling, and high-resolution gradient for ρ. For ρ estimation using gradient-based FWI, the relatively weak parameter coupling in the preconditioned gradient is the key factor which contributes to better results. In this section, we perform 3-D FWI using specific near-surface models to examine this difference.

Two near-surface models are used for different purposes. First, to visualize the impact of parameter coupling on the reconstructed result, we consider a spatially uncorrelated model where the anomalous zone for each parameter class is located at a separate position (Section 4.1). Next, to approximate the field condition, we build a realistic near-surface model containing 3-D heterogeneities (Section 4.2). The model is derived from actual downhole data acquired in the field. We incorporate anelastic attenuation in the synthetic data through viscoelastic forward modelling. Random noise is also added to the data.

4.1 FWI results for a spatially uncorrelated model

4.1.1 Model building and inversion setup

Our spatially uncorrelated model represents a special case where Δm^N in eqs (8)–(10) is spatially decomposed into parameter classes. Data using this particular model produce the update and the contamination kernels at different locations in the preconditioned gradient. The kernels, spatially separated in the preconditioned gradient for a certain parameter class, produce the artefacts due to different parameter classes at locations that differ from the location of the correct anomaly. This helps distinguishing the parameter coupling effect in the results.

We consider the 10 m × 5 m box-shaped anomalies for V_P, V_S and ρ located at different positions but at the same depth (Fig. 5). The background model varies only in depth (Fig. 1). Note that Figs 5(b) and (c) are created by subtracting the initial background model (Fig. 1) from the true model to clearly visualize the box-shaped anomalies. Although the anomalies – ΔV_S^N and $\Delta\rho^N$ – have values which are −20 per cent of the background, the ΔV_P^N is −10 per cent of the background value.

We perform gradient-based 3-D FWI on vertical-force (f_z) and horizontal-force (f_y) data sets. The acquisition geometry and the source wavelet are the same as those in Sections 3.2 and 3.3. We use the background model (Fig. 1) as the initial model in 3-D FWI, and assume the source wavelet as known. During the inversion, the model parameters (V_P, V_S, ρ) are simultaneously updated. The same monochromatic frequencies as in Sections 3.2 and 3.3 are used to calculate the preconditioned gradients. To compare the different data sets using vertical- and horizontal-force sources, we use the same stop-criterion for inversion: the iteration is stopped when the relative change between the misfit of the current iteration step and that of the second to the last iteration step is less than 1 per cent. Since the data is not sensitive to ΔV_P because of the very long wavelength of *P* wave (∼40 m) relative to the size of the anomaly, we discuss only about the reconstructed V_S and ρ models.

4.1.2 Parameter coupling: after the first iteration

We first investigate the preconditioned gradients for ρ (i.e. $-P_\rho\nabla_\rho E$) at the first iteration using the vertical-force source (Fig. 6a) and the horizontal-force source (Fig. 6b). Because of the spatially uncorrelated anomalies, the update kernels ($K_{\rho\leftrightarrow\rho}$) are imaged around the location of the anomaly of ρ (blue boxes in Figs 6a and b), while the contamination kernels ($K_{V_S\to\rho}$) are distributed around the location of the anomaly of V_S (green boxes in Figs 6a and b). When using the f_z data set, $K_{V_S\to\rho}$ dominates significantly in $-P_\rho\nabla_\rho E$ compared to $K_{\rho\leftrightarrow\rho}$ (see Fig. 6a). On the contrary, the f_y data set produces $K_{\rho\leftrightarrow\rho}$ with larger magnitude than $K_{V_S\to\rho}$ in $-P_\rho\nabla_\rho E$ (see Fig. 6b). As a consequence, 3-D FWI using the f_y data set can estimate $\Delta\rho^N$ more accurately than FWI using the f_z data set, with a weak parameter coupling due to ΔV_S^N at least at the first iteration.

As opposed to ρ, the preconditioned gradients for V_S (i.e. $-P_{V_S}\nabla_{V_S}E$) at the first iteration (Figs 6c and d) do not contain significant artefacts due to $K_{\rho\to V_S}$ (blue boxes in Figs 6c and d). This result is consistent with the results for a point scatterer model (see Section 3.2). Besides, one can see that the amplitude distribution of $-P_\rho\nabla_\rho E$ for the f_z data set is similar to that of $-P_{V_S}\nabla_{V_S}E$ (see Figs 6a and c). This suggests that $-P_\rho\nabla_\rho E$ for the f_z data set is contaminated by ΔV_S^N.

4.1.3 Results after all iterations

During the inversion, the ρ models are iteratively updated (Fig. 7). In the first iteration, both data sets produce artefacts around the location of the anomaly of V_S associated with $K_{V_S\to\rho}$ (green boxes in Figs 7a and d). The artefacts are generated because the ρ models are updated in the directions along $-P_\rho\nabla_\rho E$ in the first iteration (Figs 6a and b) based on eqs (2) and (3). With iterations, however, the nonlinear optimization process (i.e. NCG) reduces these artefacts for both data sets, resulting in gradual reconstruction of the anomaly of ρ at the correct location [see the results after iteration 10 in Figs 7(b) and (e) and after iteration 40 in Figs 7(c) and (f)]. Finally, at the convergence (i.e. after 69 iterations for the f_z data set and after 71 iterations for the f_y data set), the reconstructed ρ model using the f_y data set no longer exhibits any significant artefacts that were produced at the earlier iterations (green box in Fig. 8b). On the other hand, the finally obtained result of 3-D FWI using the f_z data set shows some remaining artefacts due to parameter coupling (see the green box in Fig. 8a). Furthermore, a comparison of the 1D profiles (Fig. 8c) shows that utilizing the f_y data set provides better ρ estimates (values closer to the true values) than the f_z data set. The likely reason for this difference is the strong parameter coupling effect in the f_z data set in the preconditioned gradient (Fig. 6a). In this case, after the first iteration, the FWI estimates a ρ model which is very far from the true model (Fig. 7a), and as a consequence, at the end the final model also falls into a local minimum located in the vicinity of the solution of the first iteration.

We also examine the reconstructed V_S models after the inversion for the vertical-force data set (Fig. 8d) and the horizontal-force data set (Fig. 8e). Contrary to ρ, the anomaly of V_S is successfully reconstructed without serious artefacts regardless of the force directions (Figs 8d and e). Furthermore, a comparison of 1-D profiles (Fig. 8f) shows that both data sets estimate almost identical V_S models.

4.2 FWI results for a realistic subsoil model

In Section 4.1 the initial model is different from the true model at locations that differ among the parameter classes. That helped distinguishing the parameter coupling effect in the results. However, this assumption is obviously unrealistic. For FWI using field data where the accuracy of the initial model can only be controlled to a limited extent, one has to choose a suitable approach for building the initial model, considering the data quality and/or any available prior information. Besides, the inversion tests in Section 4.1 have been carried out, assuming purely elastic wavefields and noise-free data.

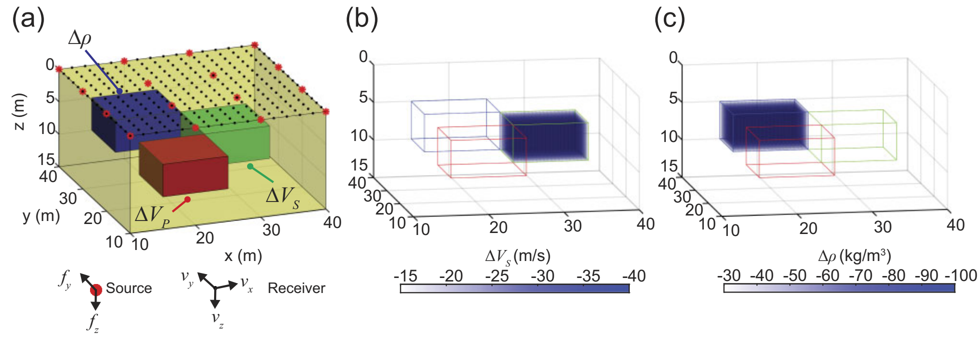

Figure 5. The spatially uncorrelated model for inversion tests. The red, green and blue boxes represent the locations of the anomalies for V_P, V_S and ρ, respectively. The models in (b) and (c) are visualized by subtracting the initial model from the true model.

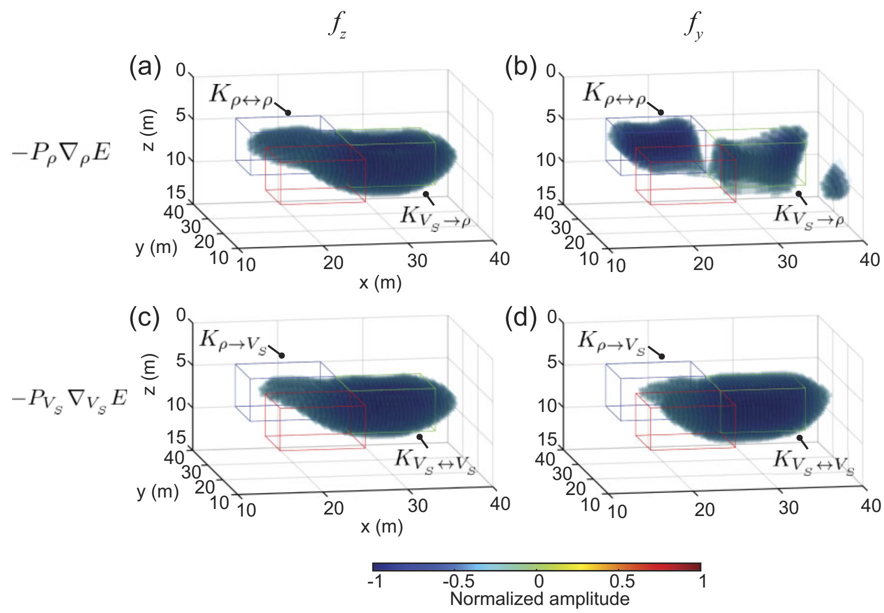

Figure 6. The preconditioned gradients for (a), (b) ρ and (c), (d) V_S.

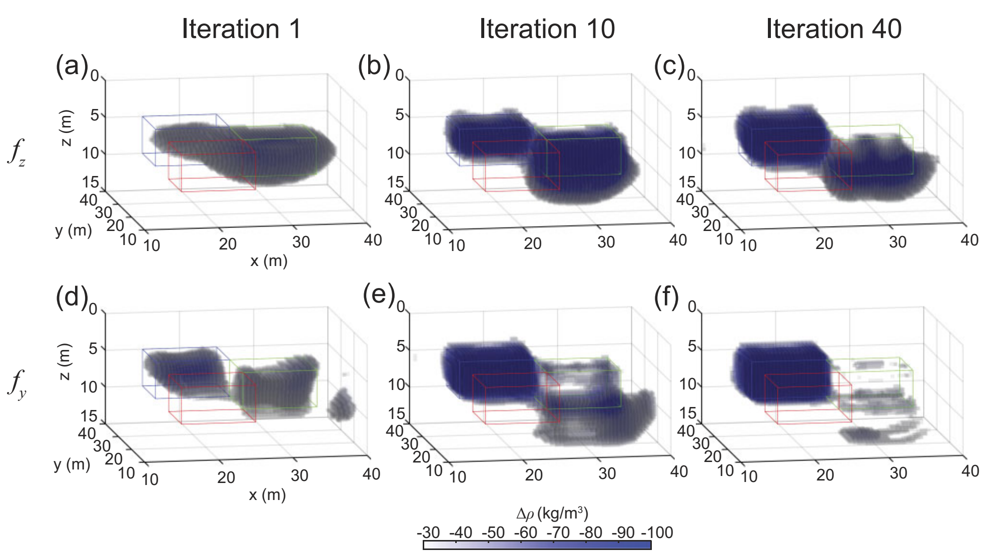

Figure 7. The comparison of the iteratively updated ρ models.

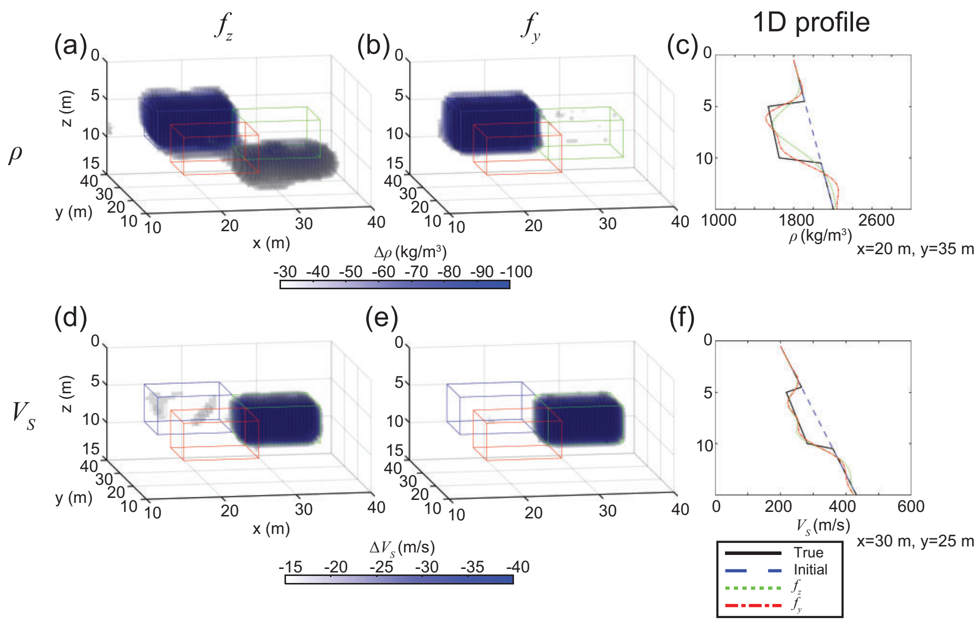

Figure 8. The comparison of the finally obtained (a)–(c) ρ and (d)–(f) V_S models.

In this subsection, we take into consideration a more realistic situation. The subsoil model is built on downhole information obtained in the field. Unlike in Section 4.1, we work here with data having random noise and viscoelastic wavefield instead of elastic wavefield.

4.2.1 Model building and pseudo-seismic data

To build the realistic 3-D subsoil model, we use V_S and porosity (ϕ) data measured at a soft-soil site in the western part of the Netherlands (Fig. 9a). In this earlier study, V_S was measured at 25 cm interval in depth via SCPT (Seismic Cone Penetration Test) and the porosity from laboratory tests on soil samples collected in boreholes (Ghose 2007; Zhubayev & Ghose 2012). From continuous soil sampling performed at this site, the soil-layer composition is known (Fig. 9a). The soil column is composed of alternating layers of clay and sand. Also, a peat layer located at ~5 m depth shows a very high porosity (~0.83), indicative of low density.

Using this data set, we have constructed a five-layered model of the near-surface (till 15 m depth; Fig. 9b) for testing FWI. The ρ values in Fig. 9(b) are calculated from ϕ as follows:

$$\rho = \rho_s(1 - \phi) + \phi S_f \rho_f, \tag{12}$$

where ρ_s is solid grain density, ρ_f is fluid density and S_f represents the degree of water saturation. For ρ_s and ρ_f, we use the representative values for such soil types in this area (Table 2). Since the water table at this site is located at 1.4 m depth, we consider an unsaturated condition until 1.4 m, and a fully water-saturated condition below this depth. For the unsaturated and fully saturated conditions, we assume $S_f = 50$ per cent and $S_f = 100$ per cent, respectively. Since we do not have V_P data, the V_P values at the fully water-saturated condition are calculated based on Gassmann's equation (e.g. Mavko *et al.* 2009), while the unsaturated V_P values are assumed to be 800 m s^{-1}. For the soil properties required in the Gassmann's equation, we consider the representative values (Table 2) based on past research (Inci *et al.* 2003; Emerson & Foray 2006; Chesworth 2008; Kumar & Madhusudhan 2012). Finally, extending this model (Fig. 9b), 3-D heterogeneity is introduced as shown in Figs 10(a)–(c). The complexity of the 3-D model is characterized by a peat layer (second layer in Fig. 9b) that gradually thins out in x- and y-directions and a sand layer (fourth layer in Fig. 9b) that gently dips in 3-D. Note that the 1-D layered model in Fig. 9(b) is located at the centre of the 3-D model (at $[x, y] =$ [13.5 m, 13.5 m]). In order to simulate the viscoelastic wavefields, we assume a constant value of 20 for both Q_P and Q_S, for the entire model.

For this realistic 3-D subsoil model, the viscoelastic wavefields are simulated using vertical-force (f_z) and horizontal-force (f_y) sources. Since the 3-D FWI applied to an f_x data set gives almost the same results as those for 3-D FWI applied to an f_y data set for this model (see Section S1 in the Supporting Information), we here discuss only the results for the f_y data set. The model is discretized using 180 × 180 × 120 gridpoints in x-, y- and z-direction, with a grid spacing of 0.15 m. We consider a symmetric acquisition geometry with 9 seismic sources at 10 m spacing and a fixed array of 121 3-component receivers planted at 2 m spacing, as shown in Fig. 10. The sources generated a Fuchs–Müller wavelet (eq. 11) with a central frequency of 20 Hz. As mentioned earlier, this frequency is not chosen to a higher value only to keep the computation time for the 3-D FWI manageable, while still achieving sufficiently the goals of this research. It is also the same reason why we restricted the depth to only 15 m. We finally add identical random noise to both force data sets. The S/N ratio in the y-component data (v_y) at the farthest source–receiver offset (28.3 m) due to the f_y source is ~8, while that in the z-component data (v_z) at the same offset due to the f_z source is ~60. The big difference in the S/N ratio reflects the difference in the amplitudes of the dominated wavefields in each data set.

The example of the simulated data (particle velocity), representing field seismic data using different force directions for the seismic source, is shown in Fig. 11. As expected, for the f_z source, the energy is predominantly present in the vertical-component

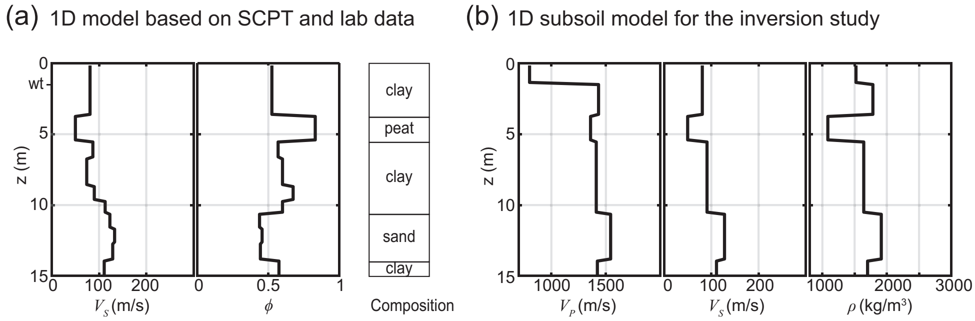

Figure 9. (a) The 1D layered V_S and ϕ models and soil composition (Ghose 2007; Zhubayev & Ghose 2012). (b) The 1D layered V_P, V_S, and ρ models for the inversion test in this research.

Table 2. Values of soil properties (Inci *et al.* 2003; Emerson & Foray 2006; Chesworth 2008; Kumar & Madhusudhan 2012) for calculating ρ and V_P for each soil layer (ρ_s: solid density, ρ_f: fluid density, K_s: solid bulk modulus, K_f: fluid bulk modulus, ν_{sk}: Poisson's ratio of the soil skeleton).

Soil type	$\rho_s\,(\mathrm{kg\,m^{-3}})$	$\rho_f\,(\mathrm{kg\,m^{-3}})$	K_s (GPa)	K_f (GPa)	ν_{sk}
Clay	2650	1000	14	2.18	0.10
Sand	2650	1000	30	2.18	0.23
Peat	1500	1000	1.4	2.18	0.15

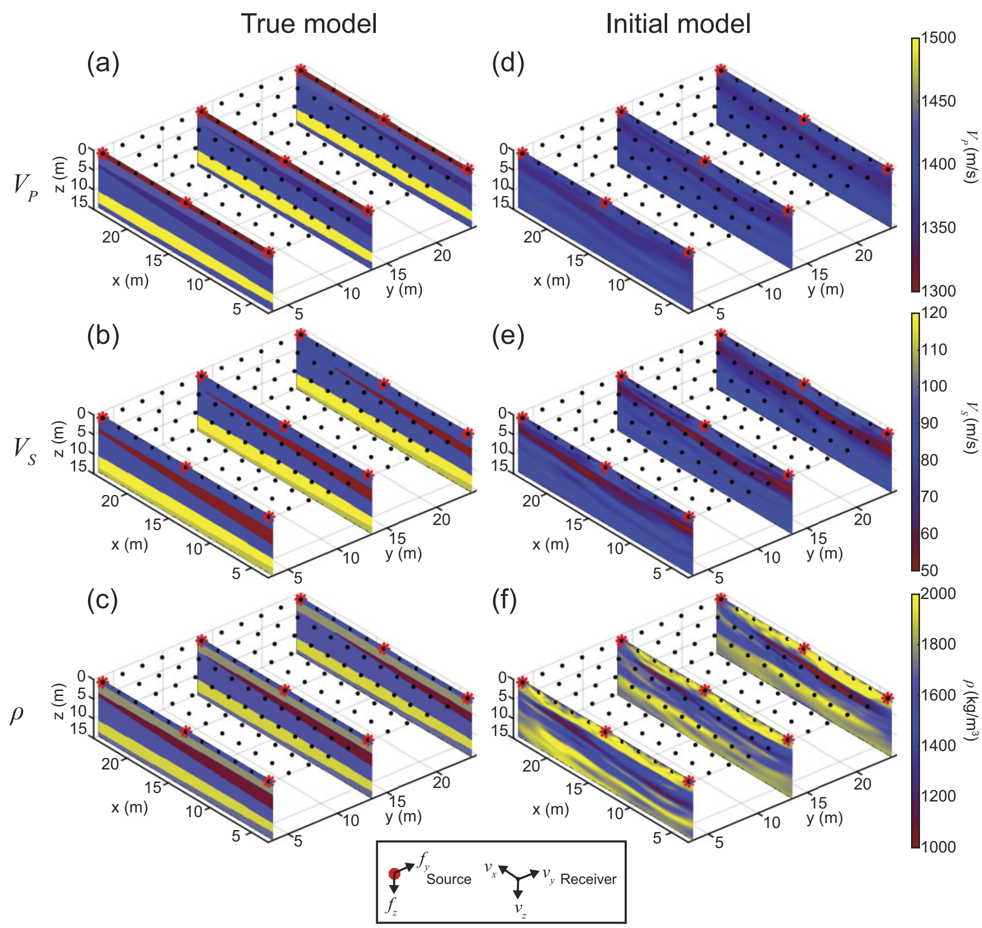

Figure 10. The 3-D complex subsoil model for the inversion tests. The true (a) V_P, (b) V_S and (c) ρ models. The initial (d) V_P, (e) V_S and (f) ρ models.

receiver (Fig. 11c), while for the same source limited amount of seismic energy present in the horizontal-component receiver (Figs 11a and b) is due to the Rayleigh waves. On the contrary, for the f_y source there is more energy in the horizontal-component receivers (Figs 11d and e) than in the vertical-component (Fig. 11f). Additionally, the f_y data show reflected arrival from the lower boundary of the peat layer at ∼0.3 s, which is less visible in the f_z data.

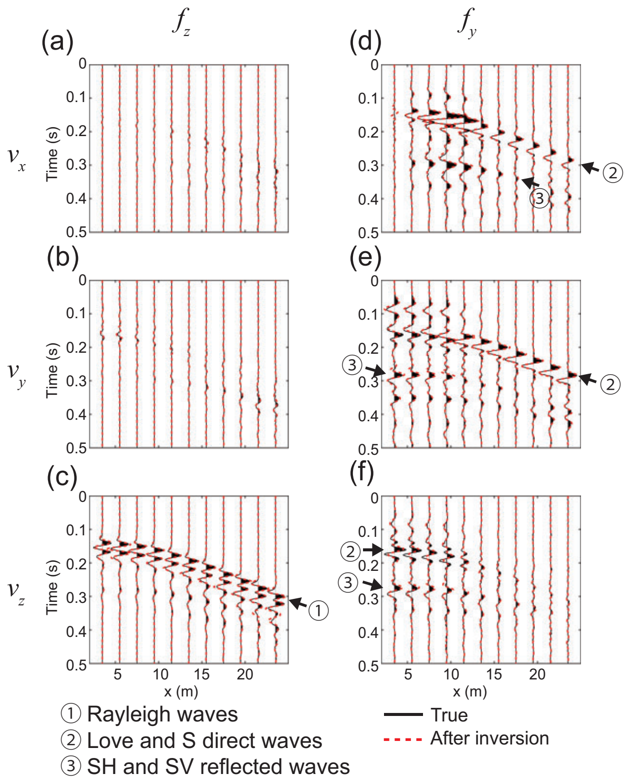

Figure 11. The example of the pseudo-seismic data calculated for the model shown in Figs 10(a)–(c). The 3-component records at $y = 3.5$ m (a)–(c) using an f_z source and (d)–(f) using an f_y source. The source is located at $[x, y] = [3.5$ m, 13.5 m$]$. For visualization, each trace is normalized by the maximum amplitude at a specific receiver component corresponding to a force component.

4.2.2 *Initial model*

We perform at first 2-D SH FWI using the 2-D (v_y) data set. This data set is derived from the f_y data set along the three 2-D lines at $y = 3.5$, 13.5 and 23.5 m (Fig. 10). For an initial V_S model for the 2-D SH FWI, we build a homogeneous model using the first-arrival travel times. An initial ρ model is constructed by first using an empirical relationship between V_S and V_P (Kitsunezaki *et al.* 1990) and then using a relationship between V_P and ρ (Ludwig 1970).

For 2-D SH FWI, we use a global-correlation-based misfit function (Choi & Alkhalifah 2012): the model is updated to maximize the cross-correlation between the normalized observed and calculated seismograms. This misfit function does not focus on the amplitudes but on the similarities (phase matching) between the seismograms, which enables reconstructing a rough model using 2-D FWI without compensating for the difference between the 3-D and 2-D geometrical spreading. Therefore, at this point, we apply 3-D to 2-D point-to-line source phase correction by convolving each trace of the 3-D seismic data set with $\sqrt{t^{-1}}$, where t is the recording time (Forbriger *et al.* 2014; Schäfer *et al.* 2014; Liu *et al.* 2022). We assume the source–time function to be known. The V_S and ρ models are updated simultaneously during the inversion. The attenuation Q_S is assumed to be known. The estimation of Q_S is possible, for example, by using the passive-viscoelastic FWI approach (Groos *et al.* 2014): the optimum Q value with the minimum misfit function for an initial model is chosen through the grid search.

The V_S and ρ sections estimated by 2-D SH FWI along the three 2-D lines are used, followed by spline interpolation, to build the 3-D initial model (Figs 10e and f). The V_P initial model is built using the empirical relationship between V_P and V_S (Kitsunezaki *et al.* 1990; Fig. 10d). Note that the initial model in Figs 10(d)–(f) is used for 3-D FWI using both f_z and f_y data sets, although building an initial model using SH FWI is possible only for the f_y data set. This allows us to investigate the sole impact of the force direction of the seismic source on the reconstructed ρ model using 3-D FWI.

4.2.3 *Parameter coupling: after the first iteration*

The preconditioned gradients for ρ $(-P_\rho \nabla_\rho E)$ calculated at the first iteration for the f_z data set (Fig. 12a) and the f_y data set (Fig. 12b) exhibit significant differences. Contrary to Section 4.1, the update and contamination kernels are not spatially separated in the preconditioned gradient. This is because the true model perturbations (Δm^{true}) for each parameter class, that is the differences between the true and initial models, are distributed over the whole model, and are superimposed on each other (Figs 12g and h). Also, a large Δm^{true} (Figs 12g and h) implies that it would not be possible to assume the Newton step (Δm^N) in eqs (8)–(10) as we do in Section 4.1.2. The Newton step Δm^N could be estimated by investigating the detailed shape of the misfit function E; however, the computational cost for this analysis is prohibitively expensive. Due to these constraints, it is challenging to visualize the kernels and evaluate the parameter coupling using eqs (8)–(10).

In order to overcome this difficulty, we directly calculate the approximate Hessian ($H^a_{m_i m_j}$), which is a component of the kernels in eqs (8)–(10). The effect of the off-diagonal elements of this Hessian (i.e. the spatial cross-correlation of the Jacobians between different grids) on the preconditioned gradient varies with the spatial distribution of Δm^N. Since Δm^N is unknown, we take the relative magnitude of the coefficients of Δm^N in eqs (8)–(10). For this purpose, we assume a simple case where $H^a_{m_i m_j}$ is diagonally dominant. In this case, eq. (10) can be written as:

$$
\begin{aligned}
-\frac{1}{diag\left(H^a_{\rho\rho}\right)+\epsilon_\rho}\nabla_\rho E \approx\ & \frac{diag\left(H^a_{\rho V_P}\right)}{diag\left(H^a_{\rho\rho}\right)+\epsilon_\rho}\Delta V_P^N \\
& + \frac{diag\left(H^a_{\rho V_S}\right)}{diag\left(H^a_{\rho\rho}\right)+\epsilon_\rho}\Delta V_S^N \\
& + \frac{diag\left(H^a_{\rho\rho}\right)}{diag\left(H^a_{\rho\rho}\right)+\epsilon_\rho}\Delta \rho^N,
\end{aligned}
\qquad (13)
$$

where $diag\left(H^a_{m_i m_j}\right)$ is the diagonal element of $H^a_{m_i m_j}$, and $diag\left(H^a_{\rho\rho}\right)+\epsilon_\rho$ represents the preconditioning filter for ρ (i.e. P_ρ) (see Appendix A for details). Evaluation of eq. (13) enables one to investigate the parameter coupling grid by grid, while the spatial parameter coupling between different grids is ignored.

The second term on the right-hand side of eq. (13) represents the parameter coupling due to V_S, and the third term represents the appropriate update direction for ρ. Because the coefficient of $\Delta\rho^N$ (the third term) is almost 1, investigating the coefficient of ΔV_S^N (the second term) allows appreciating how ΔV_S^N is amplified and how it contaminates the appropriate update direction for ρ in the preconditioned gradient. For brevity, hereafter, we call the coefficient for the second term, $diag\left(H^a_{\rho V_S}\right) / \left(diag\left(H^a_{\rho\rho}\right)+\epsilon_\rho\right)$, as the coupling coefficient of ρ.

The coupling coefficients of ρ show that their values are smaller for the f_y data set (Fig. 12e) than for the f_z data set (Fig. 12d). Around

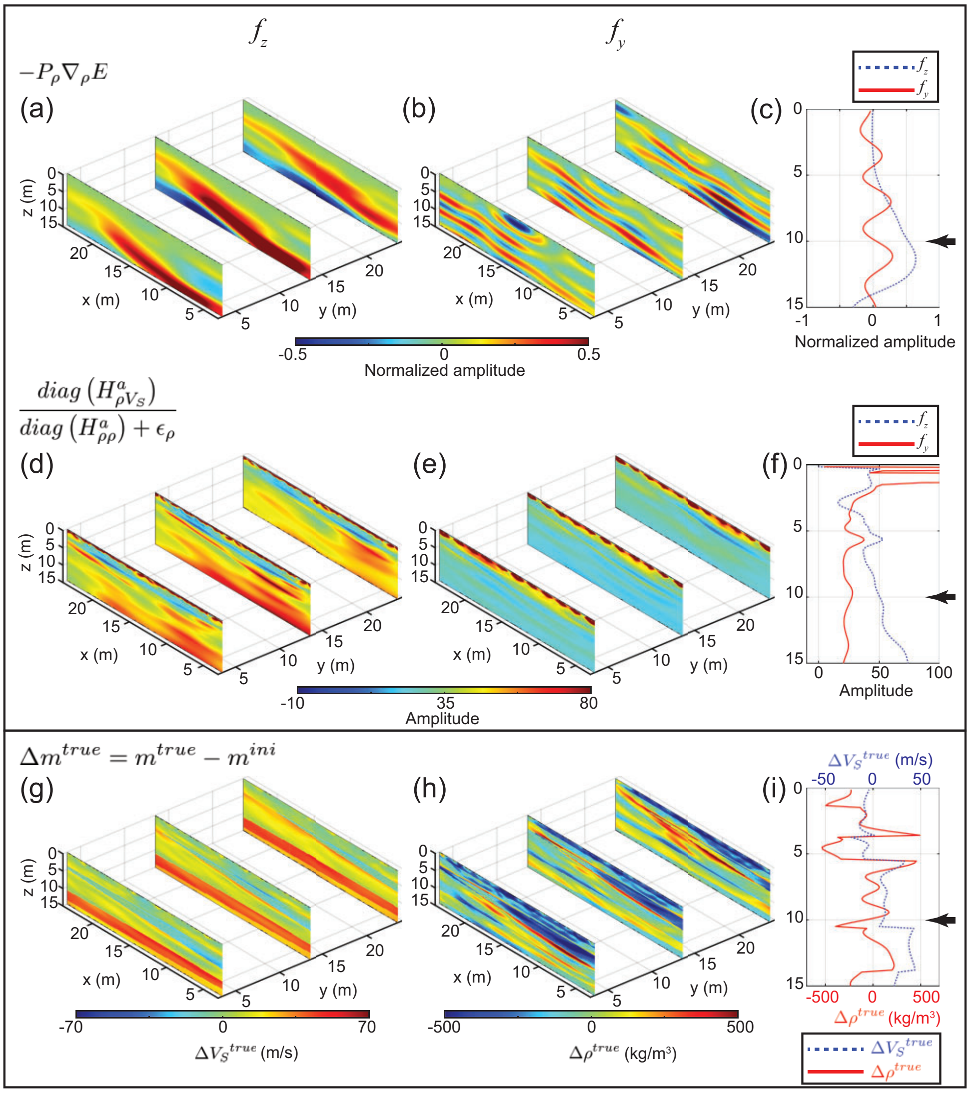

Figure 12. (a)–(c) The preconditioned gradients for ρ and (d)–(f) the coupling coefficients of ρ for different force data sets at the first iteration of the stage 1 shown in Table 3. (g)–(i) The differences between true and initial models for V_S and ρ.

z = 10 m at the centre of the model, for instance, the coupling coefficient of ρ for the f_y data set is approximately 25, while that for the f_z data set is approximately 50 (see the black arrow in Fig. 12f). Assuming that ΔV_S^N and $\Delta\rho^N$ do not significantly differ between the force directions, this difference in the amplitude of the coupling coefficients suggests that using the f_z data set results in twice larger the parameter coupling than using the f_y data set (see eq. 13). We can draw the same conclusion if the ratio of ΔV_S^N to $\Delta\rho^N$ is almost the same for the different force data sets, because the preconditioned gradient is normalized by its maximum value in the implementation of FWI here (see Section 2.1).

Next, we look more carefully into the preconditioned gradients of ρ at the centre of the model. The 1-D profiles illustrate different characteristics for different force directions (e.g. Fig. 12c). Around z = 10 m, the values for the f_y data set oscillate around zero, while almost all values for the f_z data set are positive (see the black arrow in Fig. 12c). To investigate the possible reason, we calculate the true V_S and ρ perturbations ($\Delta V_S{}^{\mathrm{true}}$ and $\Delta\rho^{\mathrm{true}}$) in Fig. 12(i). At the same location (at the centre of the model around z = 10 m), the true model perturbations show characteristics that are similar to the preconditioned gradient: $\Delta\rho^{\mathrm{true}}$ oscillates around zero, while $\Delta V_S{}^{\mathrm{true}}$ is mostly on the positive side (see the black arrow in Fig. 12i). Although the true model perturbations would not accurately represent the Newton step, the similarity between $\Delta\rho^{\mathrm{true}}$ (red line in Fig. 12i) and the preconditioned gradient for ρ for the f_y data set (red line in Fig. 12c) permits one to anticipate that the gradient for the f_y data set would more closely reflect $\Delta\rho^N$, compared to the gradient for the f_z data set. On the other hand, the similarity between $\Delta V_S{}^{\mathrm{true}}$ (blue dashed line in Fig. 12i) and the preconditioned gradient for ρ for the f_z data set (blue dashed line in Fig. 12c) suggests that the gradient is probably heavily contaminated by ΔV_S^N. The use of the f_y data set offers the possibility to obtain the preconditioned gradient for ρ with few artefacts due to ΔV_S^N, and hence leads to a more accurate ρ model from 3-D FWI.

We also investigate the parameter coupling in the preconditioned gradient for V_S (Appendix B). Unlike the preconditioned gradient for ρ, the different force data sets do not show significant differences in the parameter coupling. But the f_y data set can offer higher-resolution V_S model than the f_z data set (see Appendix B for more details).

4.2.4 Results after all iterations

During the iterations in FWI, V_S and ρ are updated simultaneously, while V_P remains fixed to the initial model (Fig. 10d); the data is not sensitive to V_P because of the very long wavelength of P wave ($\sim$40 m) relative to the heterogeneous structure. Similar to 2-D SH wave FWI (Section 4.2.2), Q_P, Q_S and the source–time function are assumed to be known. To avoid cycle skipping, we perform multiscale inversion (Bunks *et al.* 1995), where we gradually increase the maximum frequency for inversion (Table 3). We use four inversion stages; the frequencies at each stage are selected such that they continuously cover the vertical wavenumbers (Sirgue & Pratt 2004).

When the convergence is achieved, the synthetic waveforms fit well with the observed waveforms (red dashed lines in Fig. 11). A comparison of the ρ models (Figs 13a–c) shows that the use of the f_y data set gives a better resolution of the ρ values in the clay layer than when the f_z data set is used (see green arrows in Figs 13a–c). We indeed see a sharp transition from the clay layer (third layer) to the sand layer (fourth layer) in case of the f_y data set. The 1-D profiles at the centre of the model also illustrate that the ρ estimates for the f_y data set are more accurate than that for the f_z data set, especially at $\sim$10 m depth (see the black dashed line in Fig. 13d). This is probably because for the f_z data set there is a relatively strong parameter coupling at around 10 m depth, while for the f_y data set, the parameter coupling is rather weak, as demonstrated also in Section 4.2.3. Furthermore, the use of the f_y data set in 3-D FWI allows estimating the low ρ in the peat layer more accurately than the use of the f_z data set (see the black solid arrow in Fig. 13d).

Next, in order to examine the overall accuracy of the inverted ρ models for different force data sets, we define the rate of the model-error change as follows:

$$D_m = \frac{m_{\mathrm{est}}^{\mathrm{error}} - m_{\mathrm{ini}}^{\mathrm{error}}}{m_{\mathrm{ini}}^{\mathrm{error}}} \times 100, \tag{14}$$

where $m_{\mathrm{ini}}^{\mathrm{error}} = |m_{\mathrm{true}} - m_{\mathrm{ini}}|$ and $m_{\mathrm{est}}^{\mathrm{error}} = |m_{\mathrm{true}} - m_{\mathrm{est}}|$ represent the model errors for the initial model (m_{ini}) and for the estimated model (m_{est}), respectively. Eq. (14) indicates that D_m is 0 at the beginning of the inversion (i.e. $m_{\mathrm{est}} = m_{\mathrm{ini}}$), becomes a negative value if m_{est} approaches the true value (i.e. m_{true}), and becomes -100 if the model is perfectly reconstructed (i.e. $m_{\mathrm{est}} = m_{\mathrm{true}}$). On the other hand, D_m becomes a positive number if the model is updated in the opposite direction to the true value from the initial model.

To compare the overall trends of ρ estimates for f_z and f_y data sets, we calculate grid by grid D_ρ using eq. (14). The result is shown in histograms in Fig. 14. For both f_z and f_y data sets, we note incorrect ρ updates in the opposite direction to the true model (i.e. $D_\rho > 0$). This could be due to the presence of side-lobe in the vertical section of the ρ model at the layer boundaries (note the surrounding of the peat layer in Fig. 13d), due to the limited frequency bandwidth in our data, limited acquisition geometry (e.g. Virieux & Operto 2009; Li & Demanet 2016), and/or the convergence to the local minima. Such side-lobe increases the model error over the whole model during the inversion as shown in Section S2 in Supporting Information. Nonetheless, the use of the horizontal-force sources does lead to a ρ structure which correlates well with the true ρ when the convergence is achieved (see Section S2 in the Supporting Information). The comparison of the histograms of D_ρ also illustrates that 3-D FWI applied to the f_y data set is beneficial for estimating the accurate ρ: there are correct ρ updates for the f_y data set at many gridpoints (see the grids showing $D_\rho < 0$ in Fig. 14), while the use of the f_z data set produces incorrect ρ updates at many points (see the grids where $0 < D_\rho < 40$ in Fig. 14).

For the noise-contaminated data, we note that the horizontal-force data set has a lower S/N ratio than the vertical-force data set (see Section 4.2.1). Nevertheless, using the f_y data gives much better ρ estimates compared to using the f_z data. This suggests that horizontal-force data might be more robust to noise in the context of ρ estimation compared to vertical-force data, which is also found in Section 3.1.

Lastly, we compare the reconstructed V_S models using f_z and f_y data sets (Appendix C). As opposed to ρ, their accuracy is not significantly different for the two force directions, except for a slight difference in resolution (see Appendix C for more details). The choice of the force directions of the seismic source does not significantly affect the overall accuracy of the V_S estimates; however, small changes in accuracy can still be caused by different resolution of the preconditioned gradient for V_S. The small improvement in accuracy leads to a slightly higher correlation between the true and the estimated models (see Section S2 in the Supporting Information).

5 DISCUSSION

5.1 Reason behind the superiority of a horizontal-force source for density reconstruction

In Section 4, the inversion study using two different near-surface synthetic models has demonstrated that 3-D FWI applied to a horizontal-force data set gives more accurate ρ distribution compared to that applied to a vertical-force data set. There are two possible reasons behind the superiority of a horizontal-force source over a vertical-force source: parameter coupling between V_S and ρ is relatively weak and/or the observed data are relatively more sensitive to ρ. Based on the results shown in this paper, we conclude that the different parameter coupling effects contained in the preconditioned gradient can cause the difference in the accuracy of the finally estimated ρ for each force direction. Given a simple point scatterer model, the comparison of the relative strengths between the update and the contamination kernels has shown that the contamination in the preconditioned gradient for ρ due to a V_S perturbation is less severe for a horizontal-force data set than for a vertical force data set (see Section 3.2). Such a benefit of using a horizontal-force data set has been demonstrated also for two realistic near-surface models, at least at the first iteration (see Figs 6 and 12). In the context of the gradient-based inversion, the less contaminated, preconditioned gradient at the first iteration can lead to a ρ value which is close to the global minimum. This allows one to avoid falling into a local minimum during the non-linear inversion. This can offer finally more accurate ρ estimates after the convergence. The other possible reason, that is the difference in sensitivity to ρ of data sets of different-force directions, may have a rather small contribution to our results, because the spatial variability of the sensitivity is scaled for each force data set using the diagonal elements of the approximate Hessian (see Appendix A).

5.2 Limitation of the analyses

Our findings on parameter coupling are based on numerical investigations. Results of numerical investigations are influenced by the choice of the subsurface model, acquisition geometry, inversion method, and parametrization (e.g. Köhn *et al.* 2012; Métivier *et al.* 2015; Yang *et al.* 2016; Pan *et al.* 2018b, 2019;

Table 3. Monochromatic frequencies for multiscale inversion.

Stage	Monochromatic frequencies (Hz)
1	2.5, 3.8, 5.0, 6.3, 7.5, 8.8, 10.0
2	2.5, 3.8, 5.0, 6.3, 7.5, 8.8, 10.0, 11.3, 12.5, 13.8, 15.0, 16.3, 17.5, 18.8, 20.0
3	3.1, 4.3, 5.6, 6.8, 8.1, 9.3, 10.6, 11.8, 13.1, 14.3, 15.6, 16.8, 18.1, 19.3, 20.6 21.8, 23.1, 24.3, 25.6, 27.0, 28.5, 30.0
4	2.4, 3.6, 4.9, 6.1, 7.4, 8.6, 9.9, 11.1, 12.4, 13.6, 14.9, 16.1, 17.4, 18.6, 19.9, 21.1 22.4, 23.6, 24.9, 26.2, 27.7, 29.2, 30.7, 32.4, 34.2, 36.0, 37.9, 40.0

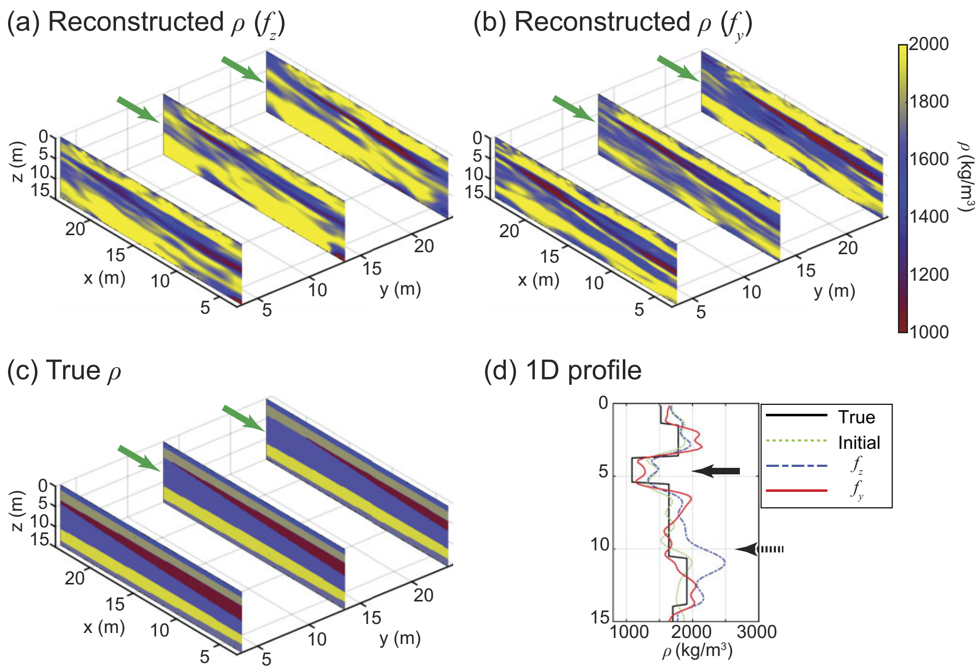

Figure 13. The reconstructed ρ models (a) when using an f_z data set and (b) when using an f_y data set. (c) The true ρ model. The green arrow shows the clay layer. (d) The comparison of 1-D profiles for the two different force-direction data sets at $[x, y] = [13.5\text{ m}, 13.5\text{ m}]$.

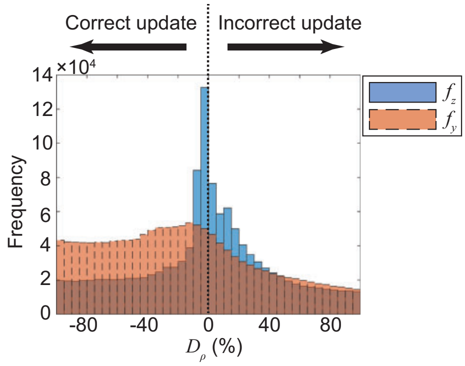

Figure 14. The comparison of the histograms of D_ρ for two different force-direction data sets.

Gao *et al.* 2021). In this regard, we have considered seismic data in the scale of near-surface exploration (i.e. several tens of meters in the horizontal direction and ∼15 m in the vertical direction). Such data sets are dominated by surface waves. The impact of parametrization (e.g. seismic velocities versus impedances) on the parameter coupling effect and the variation of this effect depending on the force direction of the source remain still an open question. This needs to be investigated in the future.

For inversion, we have used the non-linear conjugate gradient method (NCG) with a preconditioning filter of low computational cost (Appendix A). As mentioned in Section 1, other sophisticated optimization approaches, which take the accurate Hessian into consideration (e.g. truncated Newton method), can help reduce the parameter coupling (Métivier *et al.* 2015; Yang *et al.* 2016; Gao *et al.* 2021), but they generally increase the computational cost dramatically. Another approach, which may accurately estimate ρ, would exploit hierarchical inversion (Jeong *et al.* 2012; Prieux *et al.* 2013; Ren & Liu 2016): seismic velocities (i.e. V_P and V_S) are updated with the fixed ρ at the first step, and then all parameters are simultaneously inverted at the second step. This hierarchical approach was earlier tested using 2-D FWI in an exploration scale, and our results (Section 3.2) imply efficacy of this approach also for 3-D FWI at the near-surface scale. Since the V_S estimate has a small parameter coupling effect due to error in ρ (see, for instance, Fig. 4), V_S can be obtained in the first step with reasonable accuracy, which can then be used to help reducing the parameter coupling for ρ in the second step. Note, however, that the computational cost for this hierarchical approach can be much greater than the simultaneous inversion, especially for 3-D FWI. Our results have demonstrated that the use of the horizontal-force data sets would enable one to obtain accurate ρ estimates at a low cost using simultaneous inversion for both ρ and V_S, without suffering from the strong parameter-coupling problem.

Also, the difference in the parameter coupling due to different wave types is not yet well understood. Typical near-surface seismic data contain several prominent wave types—surface waves and various body waves. Each wave type is scattered due to a perturbation with different amplitude distribution, resulting in different parameter coupling effect (e.g. Pan *et al.* 2018a). In Section 3.1, for example, the scattering pattern differs due to the different incident wave types (SV/Rayleigh waves or SH/Love waves) produced by different force directions and recorded at different receiver components (see Fig. 2). Besides, Rayleigh wave propagation involves elliptical motion, either in retrograde or prograde sense, depending on its mode and depth. Gao *et al.* (2021) has shown that such different motions associated with Rayleigh waves cause different scattering patterns. Thus, the discussions in Section 3 are limited to the model considered in our numerical study. In this regard, a horizontal-force source generates both Rayleigh and SV waves propagating with a specific energy distribution over the whole model, which makes it difficult to evaluate the effect of the depth of a point scatterer on parameter coupling. This means that the wave type is crucial in order to estimate the extent of parameter coupling for a given subsurface model. Our numerical investigations, however, have considered the net influence on parameter coupling of all these wave types generated by a body force, and it is not possible to evaluate the contribution of one particular wave type here. Addressing this problem might help optimizing near-surface seismic acquisition. For instance, one may choose a combination of source–receiver components or an arbitrary force direction in 3-D for a source such that the wave type with the smallest parameter coupling dominates the observed data. In order to evaluate the contribution from each wave type to parameter coupling effect, the theoretical scattering pattern based on Born approximation for the surface waves as well as for the body waves would be needed (e.g. Snieder 1986).

Finally, the FWI result can be improved by simultaneously and/or sequentially using different force data sets during the inversion. An example can be that of a horizontal-force data set, which is not much influenced by parameter coupling; such a data set can be used to estimate reasonably accurate V_S and ρ at the first step, and then joint inversion of horizontal-force and vertical-force data sets may enhance the result at the second step. Furthermore, especially in the presence of significant lateral heterogeneity in the subsurface (e.g. Section 4.2), using two orthogonal horizontal-force sources (e.g. f_x and f_y) for the inversion can help improving the result: the use of the full wavefields produced by the horizontal-force sources may reduce parameter coupling further. Such simultaneous or sequential inversion, however, increases the computation time greatly as the data volume increases. To face such challenges, more cost-effective optimization approaches, like a mini-batch method (e.g. van Leeuwen & Herrmann 2013), might be of advantage.

5.3 Practical considerations for horizontal-force sources

We have shown the benefit of using a seismic data set with a horizontal-force source in 3-D FWI for density. However, in the current practice of near-surface seismic investigations, use of a horizontal-force source is less common than a vertical-force source. One reason is that generating efficiently horizontally polarized seismic wave is more difficult from a practical point of view. The slipping of the friction plate is a well-known problem for traditional horizontal-force (shear-wave) sources. This makes it difficult to generate a comparably large force as a vertical-force source, which results in reduction of the investigation depth. Attenuation of shorter-wavelength *S* waves in the low-velocity formations is further responsible for lower signal-to-noise ratio at a given depth, compared to *P* waves that enrich preferentially the vertical-force data. In the recent decades, however, shear-wave vibrators with small footprints have become more accessible for subsurface investigation in the near-surface exploration scale (e.g. Ghose *et al.* 1996; Ghose & Goudswaard 2004; Drijkoningen *et al.* 2006; Krawczyk *et al.* 2013; Burschil *et al.* 2022). Such vibrators can help generate a stable source signal with a relatively large frequency bandwidth, which should alleviate the cycle-skipping issue and increase source repeatability. Another possible effective source is an inclined impact source such as the Galperin source (Häusler *et al.* 2018) which has been developed for multicomponent seismic data acquisition in the near-surface scale. This source can help improve the source coupling with the ground by mitigating the slipping of the friction plate. Although horizontal-component records with a high signal-to-noise ratio (produced by a strong seismic source) are crucial for the success of 3-D FWI applied to a horizontal-force data set (see e.g. Fig. 11), such data sets might be more noisy at high frequencies in the horizontal components than in the vertical-component records. This is due to the difficulty in achieving adequate receiver coupling with the ground over a large frequency bandwidth (Krohn 1984). Further verification of the conclusions drawn in this article is now under progress through field studies involving 3-component sources and receivers.

6 CONCLUSIONS

In this paper, we have looked into the performance of gradient-based 3-D FWI applied to vertical force and horizontal force seismic data sets, in order to estimate the 3-D subsoil density distribution. We have investigated three important factors, namely noise robustness, parameter coupling and resolution, through numerical tests. We have simulated realistic near-surface seismic data dominated by surface waves.

Our investigations using the scattered wavefield due to a point-localized density perturbation have shown that the wavefield energy for a horizontal-force source is larger than that for a vertical-force source. This contributes to robustness with respect to noise of 3-D FWI using a horizontal-force data set, for estimating the subsurface density distribution. Furthermore, our investigation on the update and the contamination kernels for a point-scatterer model has indicated that the use of a horizontal-force data set allows reconstructing the high-resolution, preconditioned gradient for density with a relatively small parameter coupling. This is beneficial for estimating density at a low computational cost using a gradient-based FWI. Finally, inversion studies for two different near-surface models have demonstrated that 3-D FWI using a horizontal-force data set can reconstruct the density distribution more accurately than that using a vertical-force data set. A horizontal-force source is suitable for obtaining 3-D subsoil density distribution through FWI, without a heavy computational burden. The subsoil density variability obtained from 3-D FWI will be useful in many important applications in the near-surface scale.

ACKNOWLEDGMENTS

Y. Kawasaki is grateful to OYO Corporation for sponsoring his research at Delft University of Technology (TU Delft). The computations were carried out using the facilities at TU Delft (Delft High

Performance Computing Centre (DHPC) 2022). We thank Karlsruhe Institute of Technology (KIT), Germany for their open source FWI code IFOS2D for 2-D SH FWI (Köhn 2011) and IFOS3D for 3-D FWI (Butzer *et al.* 2013; Butzer 2015).

DATA AVAILABILITY

The synthetic models and data presented in this study are available upon request to the first author.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Supplementary_material.pdf

Section S1. FWI results using an f_x data set for a realistic subsoil model
Section S2. In-depth evaluation of the estimated density

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

REFERENCES

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

Anbazhagan, P., Uday, A., Moustafa, S.S. & Al-Arifi, N.S., 2016. Correlation of densities with shear wave velocities and SPT N values, *J. geophys. Eng.,* **13**(3), 320–341.

Arisona, A., Nawawi, M., Khalil, A.E. & Abdulrahman, A., 2018. Assessment of microgravity anomalies of soil structure for geotechnical 2-D models, *J. Geosci., Eng., Environ., Technol.,* **3**(3), 151–154.

Brossier, R., Operto, S. & Virieux, J., 2009. Seismic imaging of complex onshore structures by 2-D elastic frequency-domain full-waveform inversion, *Geophysics,* **74**(6), WCC105–WCC118.

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

Burschil, T., Buness, H. & Schmelzbach, C., 2022. Near-surface three-dimensional multicomponent source and receiver S-wave survey in the Tannwald Basin, Germany: acquisition and data processing, *Near Surf. Geophys.,* **20**(4), 331–348.

Butzer, S., 2015. 3-D elastic time-frequency full-waveform inversion, *PhD thesis*, Karlsruhe Institute of Technology, Germany.

Butzer, S., Kurzmann, A. & Bohlen, T., 2013. 3-D elastic full-waveform inversion of small-scale heterogeneities in transmission geometry, *Geophys. Prospect.,* **61**(6), 1238–1251.

Chen, R., Tran, K.T. & Wang, Y., 2021. 2-D time-domain full-waveform inversion of SH-and Love-waves for geotechnical site characterization, *Near Surf. Geophys.,* **19**(3), 283–295.

Chesworth, W., 2008. *Encyclopedia of Soil Science,* Springer Dordrecht.

Choi, Y. & Alkhalifah, T., 2012. Application of multi-source waveform inversion to marine streamer data using the global correlation norm, *Geophys. Prospect.,* **60**(4), 748–758.

Delft High Performance Computing Centre (DHPC), 2022. DelftBlue Supercomputer (Phase 1), https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase1.

Dokter, E., Köhn, D., Wilken, D., De Nil, D. & Rabbel, W., 2017. Full waveform inversion of SH-and Love-wave data in near-surface prospecting, *Geophys. Prospect.,* **65**(S1), 216–236.

Donohue, S., Long, M., Peter, O., Helle, T.E., Pfaffhuber, A.A. & Rømoen, M., 2012. Multi-method geophysical mapping of quick clay, *Near Surf. Geophys.,* **10**(3), 207–219.

Drijkoningen, G., Veltman, A., Hendrix, W., Faber, K., Brouwer, J. & Hemstede, G., 2006. A linear motor as seismic horizontal vibrator, in *Proceedings of the 68th EAGE Conference and Exhibition incorporating SPE EUROPEC 2006,* European Association of Geoscientists & Engineers.

Emerson, M. & Foray, P., 2006. Laboratory P-wave measurements in dry and saturated sand, *Acta Geotech.,* **1**(3), 167–177.

Fichtner, A., &, 2015. Resolution analysis by random probing, *J. geophys. Res.,* **120**(8), 5549–5573.

Fichtner, A. & Trampert, J., 2011. Resolution analysis in full waveform inversion, *Geophys. J. Int.,* **187**(3), 1604–1624.

Forbriger, T., Groos, L. & Schäfer, M., 2014. Line-source simulation for shallow-seismic data. Part 1: theoretical background, *Geophys. J. Int.,* **198**(3), 1387–1404.

Franko, K.M. & Grote, K.R., 2013. Estimation of near-surface soil density using electrical and electromagnetic geophysical techniques, in *Proceedings of the 125th Anniversary Annual Meeting & Expo,* Denver, CO, USA, 27-30 October 2013, The Geological Society of America.

Fuchs, K. & Müller, G., 1971. Computation of synthetic seismograms with the reflectivity method and comparison with observations, *Geophys. J. Int.,* **23**(4), 417–433.

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.

Gao, L., Pan, Y., Rieder, A. & Bohlen, T., 2021. Multiparameter viscoelastic full-waveform inversion of shallow-seismic surface waves with a pre-conditioned truncated Newton method, *Geophys. J. Int.,* **227**(3), 2044–2057.

Ghose, R., 2007. Integrating shear-wave seismic and cone penetration testing to derive laterally varying geotechnical information, in *Project report DAR.5761,* p. 85, Dutch Technology Foundation (STW).

Ghose, R. & Goudswaard, J., 2004. Integrating S-wave seismic-reflection data and cone-penetration-test data using a multiangle multiscale approach, *Geophysics,* **69**(2), 440–459.

Ghose, R., Brouwer, J. & Nijhof, V., 1996. A portable S-wave vibrator for high-resolution imaging of the shallow subsurface, in *Proceedings of the 58th EAGE Conference and Exhibition, cp-48,* EAGE Publications BV.

Groos, L., Schäfer, M., Forbriger, T. & Bohlen, T., 2014. The role of attenuation in 2-D full-waveform inversion of shallow-seismic body and Rayleigh waves, *Geophysics,* **79**(6), R247–R261.

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

Inci, G., Yesiller, N. & Kagawa, T., 2003. Experimental investigation of dynamic response of compacted clayey soils, *Geotech. Test. J.,* **26**(2), 125.

Irnaka, T.M., 2022. *3-D elastic full waveform inversion for subsurface characterization: study of a shallow seismic multicomponent field data, PhD thesis*, Université Grenoble Alpes, France.

Irnaka, T.M., Brossier, R., Métivier, L., Bohlen, T. & Pan, Y., 2022. 3-D multicomponent full waveform inversion for shallow-seismic target: Ettlingen Line case study, *Geophys. J. Int.,* **229**(2), 1017–1040.

Jeong, W., Lee, H.-Y. & Min, D.-J., 2012. Full waveform inversion strategy for density in the frequency domain, *Geophys. J. Int.,* **188**(3), 1221–1242.

Kitsunezaki, C. *et al.*, 1990. Estimation of P- and S-wave velocities in deep soil deposits for evaluating ground vibrations in earthquakes, *Sizen-Saigai-Kagaku (in Japanese),* **9**(3), 1–17.

Köhn, D., 2011. *Time domain 2-D elastic full waveform tomography, PhD thesis*, Kiel University, Germany.

Köhn, D., De Nil, D., Kurzmann, A., Przebindowska, A. & Bohlen, T., 2012. On the influence of model parametrization in elastic full waveform tomography, *Geophys. J. Int.,* **191**(1), 325–345.

Krawczyk, C.M., Polom, U. & Beilecke, T., 2013. Shear-wave reflection seismics as a valuable tool for near-surface urban applications, *Leading Edge,* **32**(3), 256–263.

Krohn, C.E., 1984. Geophone ground coupling, *Geophysics,* **49**(6), 722–731.

Kumar, J. & Madhusudhan, B., 2012. Dynamic properties of sand from dry to fully saturated states, *Géotechnique,* **62**(1), 45–54.

Levander, A.R., 1988. Fourth-order finite-difference P-SV seismograms, *Geophysics,* **53**(11), 1425–1436.
Li, Y.E. & Demanet, L., 2016. Full-waveform inversion with extrapolated low-frequency data, *Geophysics,* **81**(6), R339–R348.
Liu, J., Ghose, R. & Draganov, D., 2022. Characterizing near-surface structures at the Ostia archaeological site based on instantaneous-phase coherency inversion, *Geophysics,* **87**(4), R337–R348.
Lobsey, C. & Viscarra Rossel, R., 2016. Sensing of soil bulk density for more accurate carbon accounting, *Eur. J. Soil Sci.,* **67**(4), 504–513.
Ludwig, W.J., 1970. Seismic refraction, *The Sea,* **4,** 53–84.
Mavko, G., Mukerji, T. & Dvorkin, J., 2009. *The Rock Physics Handbook: Tools for Seismic Analysis of Porous Media,* 2nd edn, Cambridge Univ. Press.
Mecking, R., Köhn, D., Meinecke, M. & Rabbel, W., 2021. Cavity detection by SH-wave full-waveform inversion – a reflection-focused approach, *Geophysics,* **86**(3), WA123–WA137.
Métivier, L., Brossier, R., Virieux, J. & Operto, S., 2013. Full waveform inversion and the truncated Newton method, *SIAM J. Sci. Comput.,* **35**(2), B401–B437.
Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data, *Geophysics,* **52**(9), 1211–1228.
Métivier, L., Brossier, R., Operto, S. & Virieux, J., 2015. Acoustic multi-parameter FWI for the reconstruction of P-wave velocity, density and attenuation: preconditioned truncated Newton approach, in *Proceedings of the SEG Technical Program Expanded Abstracts 2015,* pp. 1198–1203.
Nocedal, J. & Wright, S.J., 2006. *Numerical Optimization,* 2nd edn, Springer.
Nuber, A., Manukyan, E. & Maurer, H., 2015. Enhancement of near-surface elastic full waveform inversion results in regions of low sensitivities, *J. appl. Geophys.,* **122,** 192–201.
Operto, S., Ravaut, C., Improta, L., Virieux, J., Herrero, A. & Dell'Aversana, P., 2004. Quantitative imaging of complex structures from dense wide-aperture seismic data by multiscale traveltime and waveform inversions: a case study, *Geophys. Prospect.,* **52**(6), 625–651.
Operto, S., Virieux, J., Dessa, J.-X. & Pascal, G., 2006. Crustal seismic imaging from multifold ocean bottom seismometer data by frequency domain full waveform tomography: application to the eastern Nankai trough, *Geophys. J. Res.,* **111**(B9), doi:10.1029/2005JB003835.
Operto, S., Gholami, Y., Prieux, V., Ribodetti, A., Brossier, R., Metivier, L. & Virieux, J., 2013. A guided tour of multiparameter full-waveform inversion with multicomponent data: from theory to practice, *Leading Edge,* **32**(9), 1040–1054.
Pan, W., Geng, Y. & Innanen, K.A., 2018a. Interparameter trade-off quantification and reduction in isotropic-elastic full-waveform inversion: synthetic experiments and Hussar land data set application, *Geophys. J. Int.,* **213**(2), 1305–1333.
Pan, W., Innanen, K.A. & Geng, Y., 2018b. Elastic full-waveform inversion and parametrization analysis applied to walk-away vertical seismic profile data for unconventional (heavy oil) reservoir characterization, *Geophys. J. Int.,* **213**(3), 1934–1968.
Pan, W., Innanen, K.A., Geng, Y. & Li, J., 2019. Interparameter trade-off quantification for isotropic-elastic full-waveform inversion with various model parameterizations, *Geophysics,* **84**(2), R185–R206.
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., Shin, C. & Hick, G., 1998. Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion, *Geophys. J. Int.,* **133**(2), 341–362.
Prieux, V., Brossier, R., Operto, S. & Virieux, J., 2013. Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 1: imaging compressional wave speed, density and attenuation, *Geophys. J. Int.,* **194**(3), 1640–1664.
Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A. & Dell'Aversana, P., 2004. Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt, *Geophys. J. Int.,* **159**(3), 1032–1056.
Ren, Z. & Liu, Y., 2016. A hierarchical elastic full-waveform inversion scheme based on wavefield separation and the multistep-length approach, *Geophysics,* **81**(3), R99–R123.
Romero-Ruiz, A., Linde, N., Baron, L., Solazzi, S.G., Keller, T. & Or, D., 2021. Seismic signatures reveal persistence of soil compaction, *Vadose Zone J.,* **20**(4), e20140.
Schäfer, M., Groos, L., Forbriger, T. & Bohlen, T., 2014. Line-source simulation for shallow-seismic data. Part 2: full-waveform inversion—a synthetic 2-D case study, *Geophys. J. Int.,* **198**(3), 1405–1418.
Shin, C., Jang, S. & Min, D.-J., 2001. Improved amplitude preservation for prestack depth migration by inverse scattering theory, *Geophys. Prospect.,* **49**(5), 592–606.
Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics,* **69**(1), 231–248.
Sirgue, L., Etgen, J. & Albertin, U., 2008. 3-D frequency domain waveform inversion using time domain finite difference methods, in *Proceedings of the 70th EAGE Conference and Exhibition incorporating SPE EUROPEC 2008,* cp-40, European Association of Geoscientists & Engineers.
Sirgue, L., Etgen, J.T., Albertin, U. & Brandsberg-Dahl, S., 2010. System and method for 3-D frequency domain waveform inversion based on 3-D time-domain forward modeling, United States Patent. US 7,725,266 B2.
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: 3-D full-waveform inversion experiments, *Geophysics,* **84**(1), B95–B108.
Snieder, R., 1986. 3-D linearized scattering of surface waves and a formalism for surface wave holography, *Geophys. J. Int.,* **84**(3), 581–605.
Teodor, D., Comina, C., Khosro Anjom, F., Brossier, R., Valentina Socco, L. & Virieux, J., 2021. Challenges in shallow target reconstruction by 3-D elastic full-waveform inversion–which initial model?, *Geophysics,* **86**(4), R433–R446.
Tran, K.T., Mirzanejad, M., McVay, M. & Horhota, D., 2019. 3-D time-domain Gauss–Newton full waveform inversion for near-surface site characterization, *Geophys. J. Int.,* **217**(1), 206–218.
Tran, K.T., Nguyen, T.D., Hiltunen, D.R., Stokoe, K. & Menq, F., 2020. 3-D full-waveform inversion in time-frequency domain: field data application, *J. appl. Geophys.,* **178,** doi:10.1016/j.jappgeo.2020.104078.
Tuckwell, G., Grossey, T., Owen, S. & Stearns, P., 2008. The use of micro-gravity to detect small distributed voids and low-density ground, *Quart. J. Eng. Geol. Hydrogeol.,* **41**(3), 371–380.
van Leeuwen, T. & Herrmann, F.J., 2013. Fast waveform inversion without source-encoding, *Geophys. Prospect.,* **61,** 10–19.
Vanremortel, R. & Shields, D., 1993. Comparison of clod and core methods for determination of soil bulk density, *Commun. Soil Sci. Plant Anal.,* **24**(17–18), 2517–2528.
Virieux, J., 1986. P-SV wave propagation in heterogeneous media: velocity-stress finite-difference method, *Geophysics,* **51**(4), 889–901.
Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics,* **74**(6), WCC1–WCC26.
Wang, Y., Dong, L., Liu, Y. & Yang, J., 2016. 2-D frequency-domain elastic full-waveform inversion using the block-diagonal pseudo-Hessian approximation, *Geophysics,* **81**(5), R247–R259.
Weihnacht, B. & Börner, F., 2007. Multi-method geophysical measurements for soil science investigations in the vadose zone, *Hydrol. Earth Syst. Sci. Discuss.,* **4**(4), 2659–2681.
Wittkamp, F., Athanasopoulos, N. & Bohlen, T., 2019. Individual and joint 2-D elastic full-waveform inversion of Rayleigh and Love waves, *Geophys. J. Int.,* **216**(1), 350–364.
Yan, Y., Wang, Z., Li, J., Huai, N., Liang, Y., Song, S., Zhang, J. & Zhang, L., 2020. Elastic SH-and Love-wave full-waveform inversion for shallow shear wave velocity with a preconditioned technique, *J. appl. Geophys.,* **173,** doi:10.1016/j.jappgeo.2020.103947.
Yang, J., Liu, Y. & Dong, L., 2016. Simultaneous estimation of velocity and density in acoustic multiparameter full-waveform inversion using an improved scattering-integral approach, *Geophysics,* **81**(6), R399–R415.
Zhubayev, A. & Ghose, R., 2012. Physics of shear-wave intrinsic dispersion and estimation of in situ soil properties: a synthetic VSP appraisal, *Near Surf. Geophys.,* **10**(6), 613–629.

APPENDIX A: PRECONDITIONING FILTER BASED ON DIAGONAL HESSIAN

We use the diagonal elements of the approximate Hessian (H^a_{mm} in eq. 6) for the preconditioning filter P_m in eq. (3). The diagonal Hessian is calculated by cross-correlation of the Jacobian with respect to a certain parameter class m (i.e. $\partial u_{\mathrm{syn}}/\partial m$). Since a whole model is divided into gridpoints, H^a_{mm} can be written in the $N \times N$ matrix form, where N is the number of grids. Here, the diagonal elements of H^a_{mm} [i.e. $diag(H^a_{mm})$] represent the cross-correlation between the Jacobians at the same grid, while the off-diagonal elements represent the spatial cross-correlation between the Jacobians at different grids. Since H^a_{mm} is typically assumed to be diagonally dominant and banded due to the finite-frequency effect (Pratt *et al.* 1998), $diag(H^a_{mm})$ or its approximated form (e.g. pseudo-Hessian) has been commonly used for the preconditioning filter (e.g. Shin *et al.* 2001; Ravaut *et al.* 2004; Operto *et al.* 2004, 2006). In this research, we follow the approach of Butzer (2015), where the inverse of $diag(H^a_{mm})$ is used for P_m as follows:

$$P_m = (diag(H^a_{mm}) + \varepsilon_m)^{-1}, \tag{A1}$$

where the stabilization factor (i.e. ε_m) is introduced to avoid the division by the very small values of $diag(H^a_{mm})$.

Using the preconditioning filter (eq. A1), one can compensate for the effect of geometrical spreading. Since the gradient ($\nabla_m E$) is calculated based on the adjoint-state method (Plessix 2006) using the wavefield affected by geometrical spreading, $\nabla_m E$ has limited spatial distribution of the amplitudes. In the near-surface exploration scale, this indicates that the amplitudes of $\nabla_m E$ are concentrated around the free surface due to the dominance of the surface waves in the observed seismic data. Thus, without a preconditioning filter, the model around the free surface is preferentially updated during the inversion, which leads to a slow convergence (e.g. Nuber *et al.* 2015; Yan *et al.* 2020). Since $diag(H^a_{mm})$ is calculated using the Jacobian which includes the effect of geometrical spreading, its spatial amplitude distribution (energy) is concentrated near the free surface. Therefore, by using the inverse of $diag(H^a_{mm})$ as a preconditioning filter, one can compensate for the limited spatial amplitude distribution of $\nabla_m E$. This makes it possible to uniformly update the whole model from the shallow to the deep parts, resulting in a fast convergence.

There is another advantage of using the diagonal Hessian as a preconditioning filter. In multiparameter FWI, it is necessary to estimate a suitable scaling factor for each parameter class in compensating for the amplitude distribution of $\nabla_m E$. This is because the Jacobian or the radiation pattern of the wavefield scattered due to a model perturbation differs for the different parameter classes (e.g. Virieux & Operto 2009; Operto *et al.* 2013; Gao *et al.* 2021). In this research, we calculate the diagonal Hessian for each parameter class. The preconditioning filter, therefore, automatically balances the spatial energy of the gradient for each parameter class.

Finally, in order to calculate the exact $diag(H^a_{mm})$, the Green's functions associated with each receiver position and each receiver component are required: we need to carry out $N_{\mathrm{r}} \times N_{\mathrm{rc}}$ simulations for each iteration, where N_{r} is the number of receivers and N_{rc} is the number of receiver components. Unfortunately, the computational cost for this simulation is prohibitively expensive. Therefore, to reduce the computational burden, we calculate the receiver-side Green's function only for a specific receiver component corresponding to a force component. Moreover, P_m is calculated at the first iteration and kept constant for the rest of iterations (e.g. Operto *et al.* 2006; Butzer 2015), which results in the additional reduction of the computational cost.

APPENDIX B: PARAMETER COUPLING ANALYSIS FOR V_S USING A REALISTIC SUBSOIL MODEL

As in Section 4.2.3, we investigate here the parameter coupling in the preconditioned gradient for V_S. We consider the following equation:

$$-\frac{1}{diag\left(H^a_{V_SV_S}\right)+\epsilon_{V_S}}\nabla_{V_S}E \approx \frac{diag\left(H^a_{V_SV_P}\right)}{diag\left(H^a_{V_SV_S}\right)+\epsilon_{V_S}}\Delta V_P^N + \frac{diag\left(H^a_{V_SV_S}\right)}{diag\left(H^a_{V_SV_S}\right)+\epsilon_{V_S}}\Delta V_S^N + \frac{diag\left(H^a_{V_S\rho}\right)}{diag\left(H^a_{V_SV_S}\right)+\epsilon_{V_S}}\Delta\rho^N, \tag{B1}$$

where $diag\left(H^a_{V_SV_S}\right)+\epsilon_{V_S}$ represents the preconditioning filter for V_S (P_{V_S} in eq. 9). We compare the coupling coefficients of V_S (i.e. $diag\left(H^a_{V_S\rho}\right)/\left(diag\left(H^a_{V_SV_S}\right)+\epsilon_{V_S}\right)$) for the different force data sets (Figs B1d–f).

The coupling coefficients of V_S show extremely small values ($\sim 10^{-2}$) below 2 m depth (Figs B1d–f). Note that the value is much smaller than the coefficient for ΔV_S^N (≈ 1). This indicates that the effect of $\Delta\rho^N$ on the preconditioned gradients for V_S is 10^{-2} smaller than that of ΔV_S^N: the preconditioned gradients for V_S are not significantly contaminated by the artefacts due to $\Delta\rho^N$, regardless of the force directions.

Below $z = 5$ m at the centre of the model, the gradient values for V_S for the f_z data set are mostly positive (blue dashed line in Fig. B1c), which is similar to the gradient values for ρ (see the blue dashed line in Fig. 12c). Although the overall trends for the gradient for the f_y data set is characterized by an oscillatory nature, one can recognize the large positive value around 11 m depth (see the red line and the black arrow in Fig. B1c). The true model differences ($\Delta V_S^{\mathrm{true}}$ and $\Delta\rho^{\mathrm{true}}$) show the oscillation of $\Delta\rho^{\mathrm{true}}$ around zero and large positive values for $\Delta V_S^{\mathrm{true}}$ below $z = 5$ m (see the black arrow in Fig. B1i). Thus, the large positive values of the gradient around 11 m depth for both force-direction data sets (Fig. B1c) indicate that the preconditioned gradients for V_S reflect the direction along ΔV_S^N without significant artefacts due to $\Delta\rho^N$. Besides, the oscillatory nature of the preconditioned gradient for the f_y data set suggests the possibility of achieving higher-resolution images than with f_z data set (Figs B1a–c), which is discussed in Section 3.3.

APPENDIX C: INVERSION RESULTS FOR V_S USING A REALISTIC SUBSOIL MODEL

Similar to ρ (Section 4.2.4), we compare the reconstructed V_S models using f_z and f_y data sets (Figs C1a and b) with the true V_S model (Fig. C1c). Both inversions successfully reconstruct the overall V_S structure. However, the results for the f_z data set show a few artefacts in the clay layer (see green dashed circles in Fig. C1a), while using the f_y data set does not produce such artefacts in the same layer (see green dashed circles in Fig. C1b).

The comparison of 1-D profiles illustrates better V_S estimates for the f_y data set in the peat layer than for the f_z data set (see the solid black arrow in Fig. C1d). Additionally, there is a larger oscillation in the clay layer when using the f_z data set than using the f_y data set

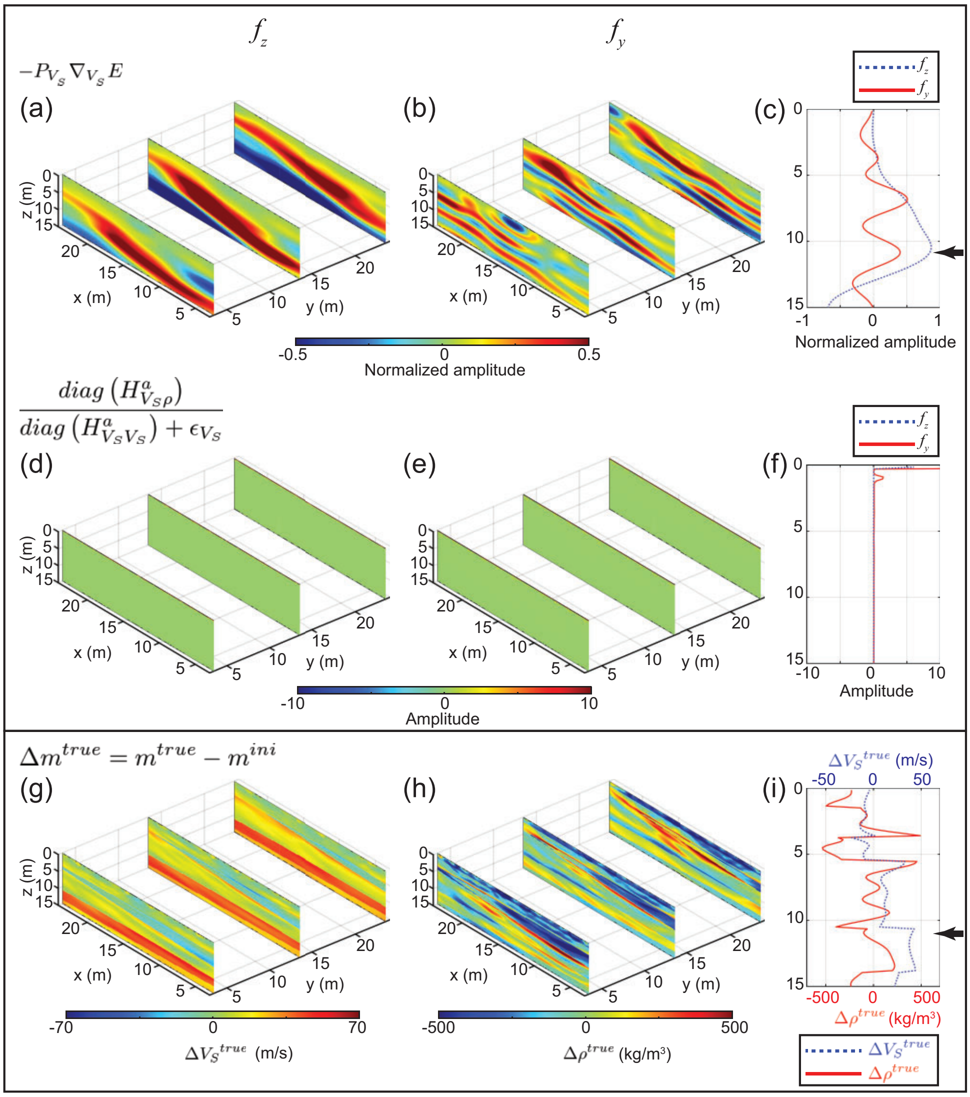

Figure B1. (a)–(c) The preconditioned gradients for V_S and (d)–(f) the coupling coefficients of V_S for different force data sets at the first iteration of the stage 1 shown in Table 3. (g)–(i) The differences between true and initial models for V_S and ρ.

(see the black dashed arrow in Fig. C1d). These differences reflect difference in resolution capabilities of the preconditioned gradient for the two different source directions, as discussed in Appendix B.

We calculate the histograms of D_{V_S} (eq. 14) for f_z and f_y data sets (Fig. C2). Unlike for ρ (Fig. 14), one cannot recognize any significant difference among the overall trends of D_{V_S} for the different force-direction data sets. In other words, an almost identical V_S model is estimated for all different force directions of the seismic source.

These results demonstrate that the choice of the force direction does not significantly affect the overall accuracy of the V_S estimates; such a choice might cause only a small difference in the accuracy due to different resolution of the preconditioned gradient for V_S.

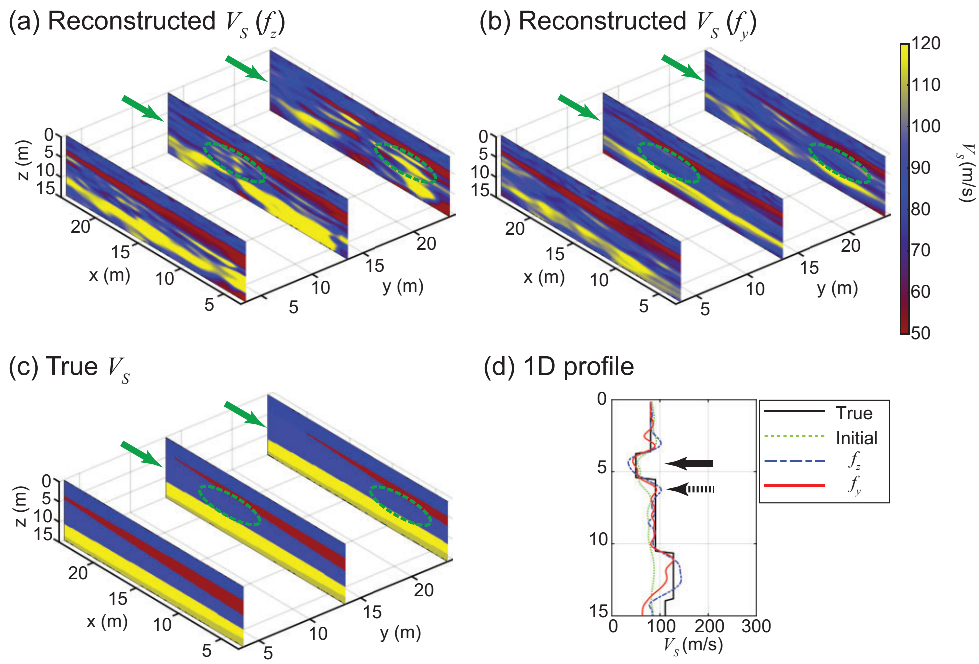

Figure C1. The reconstructed V_S models (a) when using an f_z data set, and (b) when using an f_y data set. (c) The true V_S model. The green arrow shows the clay layer. (d) The comparison of 1-D profiles for the different force data sets at $[x, y] = [13.5 \text{ m}, 13.5 \text{ m}]$.

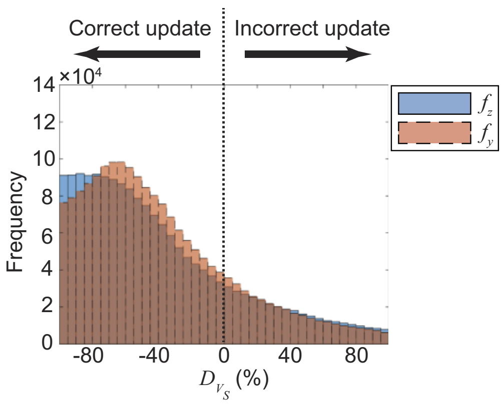

Figure C2. The comparison of the histograms of D_{V_S} for two different force-direction data sets.