Geophysical Journal International

Geophys. J. Int. (2013) **194,** 1640–1664
Advance Access publication 2013 May 30

doi: 10.1093/gji/ggt177

Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 1: imaging compressional wave speed, density and attenuation

Vincent Prieux,[1] Romain Brossier,[2] Stéphane Operto[1] and Jean Virieux[2]

[1] *Géoazur, Université Nice-Sophia Antipolis, CNRS, IRD, Observatoire de la Côte d'Azur, Valbonne, CGG Massy, France. E-mail: vincent.prieux@cgg.com*
[2] *ISTerre, Université Joseph Fourier Grenoble I, CNRS, France*

Accepted 2013 May 1. Received 2013 April 23; in original form 2012 July 30

SUMMARY

Multiparameter full waveform inversion (FWI) is a challenging quantitative seismic imaging method for lithological characterization and reservoir monitoring. The difficulties in multiparameter FWI arise from the variable influence of the different parameter classes on the phase and amplitude of the data, and the trade-off between these. In this framework, choosing a suitable parametrization of the subsurface and designing the suitable FWI workflow are two key methodological issues in non-linear waveform inversion. We assess frequency-domain visco-acoustic FWI to reconstruct the compressive velocity (V_P), the density (ρ) or the impedance (I_P) and the quality factor (Q_P), from the hydrophone component, using a synthetic data set that is representative of the Valhall oil field in the North Sea. We first assess which of the (V_P, ρ) and (V_P, I_P) parametrizations provides the most reliable FWI results when dealing with wide-aperture data. Contrary to widely accepted ideas, we show that the (V_P, ρ) parametrization allows a better reconstruction of both the V_P, ρ and I_P parameters, first because it favours the broad-band reconstruction of the dominant V_P parameter, and secondly because the trade-off effects between velocity and density at short-to-intermediate scattering angles can be removed by multiplication, to build an impedance model. This allows for the matching of the reflection amplitudes, while the broad-band velocity model accurately describes the kinematic attributes of both the diving waves and reflections. Then, we assess different inversion strategies to recover the quality factor Q_P, in addition to parameters V_P and ρ. A difficulty related to attenuation estimation arises because, on the one hand the values of Q_P are on average one order of magnitude smaller than those of V_P and ρ, and on the other hands model perturbations relative to the starting models can be much higher for Q_P than for V_P and ρ during FWI. In this framework, we show that an empirical tuning of the FWI regularization, which is adapted to each parameter class, is a key issue to correctly account for the attenuation in the inversion. We promote a hierarchical approach where the dominant parameter V_P is reconstructed first from the full data set (i.e. without any data preconditioning) to build a velocity model as kinematically accurate as possible before performing the joint update of the three parameter classes during a second step. This hierarchical imaging of compressive wave speed, density and attenuation is applied to a real wide-aperture ocean-bottom-cable data set from the Valhall oil field. Several geological features, such as accumulation of gas below barriers of claystone and soft quaternary sediment are interpreted in the FWI models of density and attenuation. The models of V_P, ρ and Q_P that have been developed by visco-acoustic FWI of the hydrophone data can be used as initial models to perform visco-elastic FWI of the geophone data for the joint update of the compressive and shear wave speeds.

Key words: Inverse theory; Controlled source seismology; Seismic attenuation; Computational seismology; Wave propagation.

INTRODUCTION

Full waveform inversion (FWI) is gradually being integrated into the seismic processing workflow for the building of high-resolution velocity models that can be used as improved background models for reverse time migration in complex environments (Plessix & Perkins 2010; Sirgue *et al.* 2010; Vigh *et al.* 2011). FWI is an optimization problem that seeks to minimize the misfit between the recorded and modelled seismic data (Tarantola 1984, 1987; Pratt *et al.* 1998; Virieux & Operto 2009). The synthetic data are computed with the two-way wave equation, to exploit the full information content of the data associated with diving waves, pre- and post-critical reflections, diffractions, and so on. However, several simplifications in the physical description of the subsurface are generally considered for seismic modelling and inversion. The most basic of these considers the subsurface to be an acoustic medium that can be parametrized by the compressional (P) wave speed only (e.g. Ravaut *et al.* 2004; Brenders & Pratt 2007). These simplifications reduce the computational burden of the seismic modelling as well as the non-linearity of the FWI. Although high-resolution *P*-wave velocity models can provide useful inferences for geological interpretation, a more realistic description of the physical properties of the subsurface, and in particular of the elastic properties, is required for realistic reservoir and fluid characterization (Shi *et al.* 2007).

One of these physical properties is the density, which is closely related to the porosity ϕ by the relation: $\rho = (1 - \phi)\rho_0 + \phi\rho_{fl}$, where ρ_0 and ρ_{fl} are the grain and pore fluid densities respectively. The velocity is not linearly related to the porosity, and thus not to the density. For example, the velocities can decrease because of cracks and faults, while the cracks have a null porosity (Mavko *et al.* 2009). This is why the empirical relations between the velocity and the density (Gardner *et al.* 1974; Castagna *et al.* 1993) are often more realistic within high-pressured and fluid-saturated rock. Another related physical property is the impedance, which controls the partitioning of the seismic energy at an interface. The impedance is traditionally reconstructed rather than the density, because of the potential trade-off between density and *P*-wave velocity to match amplitudes of reflected waves (Mora 1987; Tarantola 1987; Forgues & Lambaré 1997).

Generally speaking, the seismic reflection imaging workflow is subdivided into two distinct tasks, which respect the scale separation between the velocity macromodel and the subsurface reflectivity: a highly non-linear problem is first solved to build a large-scale velocity model from kinematic attributes, followed by a more linear problem that consists of the imaging of the reflectivity parametrized by the impedance from dynamic attributes (Tarantola 1986; Métivier *et al.* 2011). This scale separation occurs, because the intermediate wavelengths of the subsurface cannot be estimated from narrow-azimuth reflection data (Jannane *et al.* 1989). While the velocity-impedance parametrization makes sense in seismic reflection processing, is this parametrization the most suitable one for FWI of wide-aperture data? Indeed, FWI of wide-aperture data is a non-linear iterative process, that does not rely on scale separation, as a continuous spectrum of wavenumbers can theoretically be imaged from the broad illumination of scattering angles provided by wide-azimuth surveys (Pratt & Worthington 1990; Neves & Singh 1996; Sirgue & Pratt 2004). Therefore, the first issue that we address in this study aims to clarify whether velocity-impedance parametrization is more suitable than velocity-density parametrization for FWI of wide-aperture data. Two factors need to be taken into account for this assessment: the trade-off between the two parameter classes, and the resolution with which each parameter class can be reconstructed during FWI.

Attenuation is another physical property that, in addition to the *P*-wave velocity, provides useful inferences on the lithology, the physical state, and the degree of saturation, which are all very informative for reservoir characterization and monitoring (Toksöz & Johnston 1981; Best *et al.* 1994; Wang 2008). Intrinsic attenuation occurs when grains are not elastically bounded (Aki & Richards 1980), or due to the fluid 'squirt' within the pore space (Thomsen *et al.* 1997), or because of the intrapore clay content (Klimentos & McCann 1990). The footprint of the Q_P-factor translates in the data by amplitude attenuation and dispersion effects (Toverud & Ursin 2005). Numerous methods have been developed to extract the attenuation from seismic data: the spectral ratio method (Tonn 1991; Romero *et al.* 1997; Dasgupta & Clark 1998), the centroid frequency-shift method (Quan & Harris 1997; Plessix 2006a), least-squares migration (Ribodetti & Virieux 1998; Causse *et al.* 1999; Ribodetti *et al.* 2000) and FWI (Liao & McMechan 1995; Charara *et al.* 1996; Hicks & Pratt 2001; Askan *et al.* 2007; Malinowski *et al.* 2011). Ribodetti *et al.* (2000) demonstrated with the ray+Born migration/inversion approach that the reconstruction of the velocity and the attenuation is only possible when the reflector is illuminated from above and beneath, as the asymptotic Hessian is otherwise singular. Mulder & Hak (2009) and Hak & Mulder (2011) have shown that, when no dispersion is considered, for short-offset seismic reflection acquisitions, the velocity and the attenuation are related by a Hilbert transform, which prevents their independent reconstruction, with numerous combinations of the two parameters being able to match the data equally well. Only a few applications of FWI of real land and marine data have been presented for attenuation imaging. Most of these applications were performed for cross-well acquisitions (Song *et al.* 1995; Charara *et al.* 1996; Pratt *et al.* 2005; Kamei & Pratt 2008; Rao & Wang 2008), and to a lesser extent, for surface acquisitions (Hicks & Pratt 2001; Smithyman *et al.* 2009; Malinowski *et al.* 2011). Although Malinowski *et al.* (2011) showed that the joint reconstruction of *P*-wave velocity and attenuation is possible for very attenuating media, the attenuation reconstruction at low frequencies can be unstable, because of the small influence of the attenuation at these frequencies. Attenuation generally has a much smaller imprint on the data than wave speed, and in particular, in a marine environment. Therefore, the second issue that is addressed in this study is related to the design of a robust hierarchical optimization workflow, which can handle multiple classes of parameters with variable influence on the data, such as the *P*-wave velocity and attenuation.

FWI becomes highly non-linear when multiple classes of parameters with contrasted influences on the data are reconstructed. In this framework, a first key issue is the selection of a suitable parametrization of the subsurface for FWI, which allows the managing of this variable influence as well as the trade-off between parameters. Here, subsurface parametrization relates to set of independent parameters that fully describe the subsurface properties that govern wave propagation under some specific approximation (e.g. acoustic, elastic). In this study, the influence of one parameter class on the data for a given subsurface parametrization must be understood as the wavefield perturbation in terms of phase and/or amplitude that would be generated by a representative perturbation of this parameter class (i.e. a perturbation of size and amplitude that is expected to be reconstructed by FWI according to the intrinsic resolution power of FWI), while keeping fixed the other parameter classes of the parametrization. According to this definition, the influence of a given parameter class (e.g. the compressional wave

speed) will change depending on the other parameter classes involved in the parametrization (e.g. density or impedance) because of the potential trade-off between these parameters. Note also that the wavefield perturbation just described is not equivalent to the partial derivative of the wavefield with respect to the model parameter, the values of which are biased by the order of magnitude of the model parameter. Indeed, many subsurface parametrizations can be inferred from different non-linear combinations of parameters, and together with the scaling applied to the different parameter classes, this choice is not neutral in the framework of non-linear FWI.

A second issue is the design of a suitable hierarchical workflow to manage different data components (e.g. hydrophones, geophones) and multiple classes of parameters to reduce the non-linearity of FWI (Tarantola 1986; Sears *et al.* 2008; Jeong *et al.* 2012). As an illustrative example, the shear (S) wave velocity is generally considered as a parameter of secondary importance in terms of the influence on the data in marine environments where soft sediments can prevent significant P-to-S conversions on the sea bed (Tarantola 1986). This second-order parameter can be reconstructed from geophones, once the P-wave velocity (i.e. the dominant parameter) has been reconstructed from hydrophones.

This study is the first of a two-part series that investigates 2-D multiparameter FWI of multicomponent ocean-bottom-cable (OBC) data from the shallow-water Valhall field. These are designed to build visco-acoustic and visco-elastic subsurface models along two cables of the available 3-D acquisition. In this study, we discuss the feasibility of the reconstruction of the P-wave velocity, density or impedance, and the Q_P-factor from the hydrophone component in the visco-acoustic approximation. In the companion report, we will use the subsurface models that are built into this study as initial or background models to jointly reconstruct the P- and S-wave velocities from the hydrophone and geophone components using visco-elastic FWI.

In the first part of this study, we review the basics of the frequency-domain FWI method that we use. We stress two key algorithmic ingredients of multiparameter FWI: the first concerns the optimization algorithm, which should take into account the effects of the Hessian during inversion, to correctly scale the gradients of the misfit function associated with each parameter class, and the second key algorithmic ingredient is the regularization, which should help to reduce the non-linearity of the inversion without affecting the ability of the FWI to update parameters that have little influence on the data. In the second part of this study, we discuss the two methodological issues addressed here, with a realistic synthetic example that is representative of the Valhall field. We first show that the velocity–density parametrization is more suitable than the velocity–impedance counterpart for FWI of wide-aperture data. Secondly, we propose a hierarchical FWI workflow for the reconstruction of the P-wave velocity, density and attenuation from the hydrophone component, using visco-acoustic FWI. In the last part of this study, the FWI methodology that was developed from the synthetic case study is applied to a real data case study from Valhall, after which we draw some conclusions.

VISCO-ACOUSTIC FWI

Seismic modelling is performed in unstructured triangular meshes with a frequency-domain velocity–stress finite-element discontinuous Galerkin method (Brossier *et al.* 2010b). A forward-modelling operator based on the second-order wave equation for particle velocities and piecewise constant interpolation function (P0) in the finite-element method is used to build the diffraction kernel of the FWI from self-adjoint operators. Meanwhile, the incident and adjoint wavefields are accurately computed with a first-order velocity–stress formulation of the wave equation, with either P0, P1 (linear) or P2 (quadratic) interpolation functions (Brossier 2011). Of note, the velocity–stress wave equation allows us to introduce the pressure field in a natural way during the seismic modelling for both the incident and adjoint wavefields.

The misfit function $\mathcal{C}(\mathbf{m})$ is given by

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2}\Delta\mathbf{d}^{\dagger}\mathbf{W}_d\Delta\mathbf{d} + \frac{1}{2}\sum_{i=1}^{N_p}\lambda_i\left(\mathbf{m}_i - \mathbf{m}_{prior_i}\right)^{\dagger}\mathbf{W}_{m_i}\left(\mathbf{m}_i - \mathbf{m}_{prior_i}\right), \quad (1)$$

where the hydrophone data-residual vector is denoted by $\Delta\mathbf{d}$, and is weighted by the operator $\mathbf{W}_d$. The symbol $\dagger$ is the adjoint operator, the transpose conjugate. The matrix $\mathbf{W}_d$ can weight data residuals according to the standard error and/or according to the source–receiver offset (Ravaut *et al.* 2004; Operto *et al.* 2006). The multiparameter subsurface model is denoted by $\mathbf{m} = (\mathbf{m}_1, ..., \mathbf{m}_{N_p})$, where N_p denotes the number of parameter classes. Attenuation is implemented easily in frequency-domain modelling and FWI using a complex-valued velocity given by $\bar{c} = c\{1 - [i\ \mathrm{sgn}(\omega)]/(2Q_P)\}$ (Tarantola 1988). Contrary to the Kolski–Futterman model (Kolsky 1956), Q_P is independent of the frequency, because of the small range of inverted frequencies. In this study, we minimize the misfit function (eq. 1) with respect to normalized model parameters that are scaled by their mean value in the initial model, such that each class of parameter has the same order of magnitude. By doing so, the partial derivative of the wavefield with respect to the model parameters give some reasonable insights into the real influence of the parameters on the data in the sense that these normalized partial derivative wavefields represent wavefield perturbations generated by model perturbations defined as a fixed percentage of an average value of the parameters (Gholami *et al.* 2013b). This is a deliberate choice, and other parameter scaling can be viewed to perform FWI. Let us notice that we also apply this normalization to Q_P, although this parameter is dimensionless.

A Tikhonov regularization is applied to each class of parameter, i, through a roughness operator $\mathbf{W}_{m_i}$, which forces the difference between the model $\mathbf{m}_i$ and a prior model $\mathbf{m}_{prior_i}$ to be smoothed. The scalars λ_i control the influence of the data-space misfit function $\frac{1}{2}\Delta\mathbf{d}^{\dagger}\mathbf{W}_d\Delta\mathbf{d}$ relative to the model-space misfit functions $\frac{1}{2}(\mathbf{m}_i - \mathbf{m}_{prior_i})^{\dagger}\mathbf{W}_{m_i}(\mathbf{m}_i - \mathbf{m}_{prior_i})$. Of note, the scalar λ_i can be adapted to each parameter class i, which is helpful when the partial derivative wavefields with respect to different parameter classes have contrasted amplitudes, as we shall see. The prior model is the starting model at the first frequency group, and for each new frequency group, it is set to be the final model of the previous frequency group, in order to recover better-resolved models at the high frequencies.

FWI is performed in the frequency domain by local optimization where the gradient of the misfit function is computed with the adjoint-state method (Plessix 2006b; Chavent 2009). The expression of the perturbation model, which minimizes the misfit function at iteration k, is given by

$$\Delta\mathbf{m}^{(k)} = \Re\left[\widehat{\mathbf{W}}_m^{-1}\mathbf{J}^{(k)^{\dagger}}\mathbf{W}_d\mathbf{J}^{(k)} + \widehat{\mathbf{W}}_m^{-1}\left(\frac{\partial\mathbf{J}^{(k)^{\dagger}}}{\partial\mathbf{m}^T}\right)\left(\Delta\mathbf{d}^{(k)}\dots\Delta\mathbf{d}^{(k)}\right) + \Lambda\right]^{-1} \Re\left[\widehat{\mathbf{W}}_m^{-1}\mathbf{J}^{(k)^{\dagger}}\mathbf{W}_d\Delta\mathbf{d}^{(k)} + \Lambda\left(\mathbf{m}^{(k)} - \mathbf{m}_{prior}\right)\right] \quad (2)$$

where **J** denotes the sensitivity or Fréchet derivative matrix, and Λ is a block diagonal damping matrix

$$\Lambda = \begin{pmatrix} \lambda_1 \mathbf{I}_M & \dots & \mathbf{0} \\ \dots & \dots & \dots \\ \mathbf{0} & \dots & \lambda_{N_p} \mathbf{I}_M \end{pmatrix}, \tag{3}$$

and $\mathbf{I}_M$ is the identity matrix of dimension M, where M denotes the number of nodes in the computational mesh which is identical for each parameter class. In eq. (2), $\Re$ is the real part of a complex number. The matrix $\widehat{\mathbf{W}}_m$ is a $N_p \times N_p$ block diagonal matrix, where each block is formed by the $\mathbf{W}_{m_i}$ matrices. On the right-hand side of eq. (2), the term to be inverted is the full Hessian, which contains three terms. The first aims to correct the gradient for linear effects such as the limited bandwidth of the source, the limited spread of the acquisition geometry, and the geometrical spreading of the data. The second term corrects the Hessian for double-scattering effects in non-linear inverse problems (Pratt *et al.* 1998). The third regularization term dampens the deconvolution action of the first two terms of the Hessian to improve the conditioning of the Hessian matrix. The second term on the right-hand side of eq. (2) is the gradient of the misfit function, which is composed of two terms: the first represents the contribution of the data, and it is formed by the zero-lag correlations of the partial derivative wavefields at the receiver positions with the data residuals, while the second term is the model-space regularization term.

It is worth remembering that the partial derivative wavefields satisfy the differential equation,

$$\mathbf{B}\left[\omega, \mathbf{m}(\mathbf{x})\right] \frac{\partial \mathbf{v}}{\partial m_j} = -\frac{\partial \mathbf{B}\left[\omega, \mathbf{m}(\mathbf{x})\right]}{\partial m_j} \mathbf{v}, \tag{4}$$

where $\mathbf{B}[\omega, \mathbf{m}(\mathbf{x})]$ and $\mathbf{v}$ denote the forward-problem operator (in our case, the second-order wave-equation operator for particle velocities) and the incident particle velocity wavefield, respectively. The index j represents each parameter in the mesh, regardless of the class we consider. The right-hand side of eq. (4) is the secondary virtual source of the partial derivative wavefield, the spatial and temporal supports of which are centred on the position of the diffractor m_j and on the arrival time of the incident wavefield at the diffractor m_j, respectively (Pratt *et al.* 1998). The diffraction pattern of this virtual source is given by $\partial \mathbf{B}[\omega, \mathbf{m}(\mathbf{x})]/\partial m_j$, and it gives some insights into the sensitivity of the data to the parameter m_j for the parameter class i as a function of the scattering (or aperture) angle.

We use the quasi-Newton limited memory Broyden–Flechter–Goldfarb–Shanno (L-BFGS) optimization algorithm to solve eq. (2) (Nocedal 1980; Nocedal & Wright 1999). The L-BFGS algorithm computes recursively an approximation of the product of the inverse of the Hessian with the gradient, from a few gradients and a few solution vectors from previous iterations. As an initial guess of this iterative search, we use a diagonal approximation of the approximate Hessian (the linear term) damped by the Λ matrix,

$$H_0 = \left(\widehat{\mathbf{W}}_m^{-1} \mathrm{diag} \left\{ \mathbf{J}^{(k)^\dagger} \mathbf{W}_d \mathbf{J}^{(k)} \right\} + \Lambda \right)^{-1}. \tag{5}$$

Various expressions of eq. (2) exist (Greenhalgh *et al.* 2006), and these can be selected for convenience of implementation. Our implementation allows the initial estimation of the Hessian (eq. 5) to be diagonal, and hence easy to invert, because the matrix Λ is diagonal unlike $\widehat{\mathbf{W}}_m$.

The smoothing operators $\mathbf{W}_{m_i}^{-1}$ are Laplace functions that are given by

$$\mathbf{W}_{m_i}^{-1}(z, x, z', x') = \sigma_i^2(z, x) \exp\left(\frac{-|x - x'|}{\tau_x} \right) \exp\left(\frac{-|z - z'|}{\tau_z} \right), \tag{6}$$

where τ_x and τ_z denote the horizontal and vertical correlation lengths that are defined as a fraction of the local wavelength. The coefficients σ_i generally represent the standard error in the framework of Bayesian inversion (e.g. Gouveia & Scales 1998). In this study, we do not introduce standard deviations in the operators $\mathbf{W}_{m_i}$, $\sigma_i^2(z, x) = 1$, because we normalize the subsurface parameters by their mean values. A Laplace function is used for $\mathbf{W}_{m_i}^{-1}$, because its inverse in the expression of the misfit function can be computed analytically (Tarantola 1987, pp. 308–310).

The frequency-domain FWI is performed with a conventional multiscale approach, which proceeds from low to high frequencies (Pratt & Worthington 1990; Sirgue & Pratt 2004). A second level of data preconditioning with respect to time can be implemented through time damping of seismograms with a decaying exponential function (Brenders & Pratt 2007; Brossier *et al.* 2009; Shin & Cha 2009). This time damping is implemented in frequency-domain seismic modelling with complex-valued frequencies, the imaginary part of which controls the decay of the amplitude with time. This is shown by the expression of the Fourier transform of the damped function $p(t)\mathrm{e}^{-(t-t_0)/\tau}$ written as

$$p(\omega + \mathrm{i}/\tau)\mathrm{e}^{\frac{t_0}{\tau}} = \int_{-\infty}^{+\infty} p(t)\mathrm{e}^{\mathrm{i}\omega t}\mathrm{e}^{-(t-t_0)/\tau}\mathrm{d}t, \tag{7}$$

where τ will be referred to as the time-damping factor in the following.

The aim of this data pre-conditioning is the progressive injection of more complex or more resolving late-arriving phases into the inversion to reduce the risk of cycle skipping. An application of this heuristic approach to land data for the joint update of the *P*- and S-wave velocities by joint inversion of body waves and surface waves was presented by Brossier *et al.* (2009).

SYNTHETIC VALHALL CASE STUDY

Models and FWI set-up

We now present an application of visco-acoustic FWI to the synthetic Valhall model. The Valhall model is a shallow-water medium (water depth, 70 m), that is 16 km long and 5.2 km deep (Fig. 1a). We consider an OBC-like sea bottom acquisition with a maximum offset of 16 km, and a source and receiver spacing of 50 m. The depths of the explosive sources and the hydrophones are 6 and 71 m, respectively. Seismic modelling is performed on a hybrid P0–P1 triangular mesh where the P0 interpolation is set below 160 m in depth (Brossier 2011; Prieux *et al.* 2011). A discretization rule of 10 elements per minimum wavelength is used in the regular mesh for the P0 interpolation, leading to 20-m-long triangle edges. This mesh is kept the same over the iterations and the inverted frequencies. The *h*-adaptivity (the ability to adapt the size of the elements to the local wavelength through unstructured meshing) is not exploited below 160 m in depth for two reasons: first, the P0 interpolation order does not provide accurate solutions on unstructured meshes, and secondly P1 or P2 interpolation orders would not lead to computational savings because of the limited range of velocity in the Valhall subsurface ($Vp_{\max}/Vp_{\min} = 3500/1400 = 2.5$) (see Brossier *et al.*

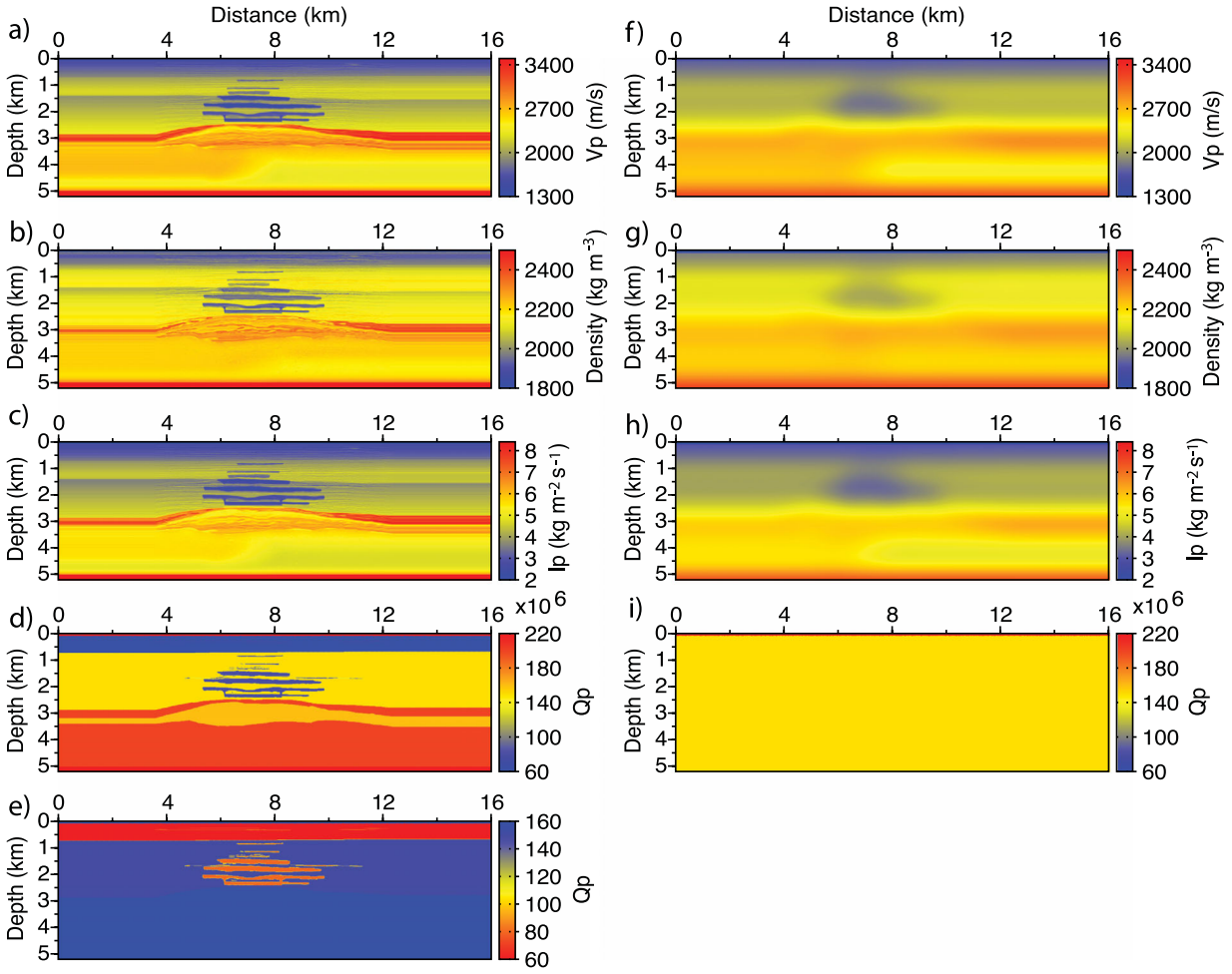

Figure 1. True (a–d) and corresponding starting models (f–i) for the *P*-wave velocity (a, f), density (b, g) and impedance (c, h), and true Q_P-factor model plotted with different colour scales (d, e). (i) The initial Q_P-factor model (i) is homogeneous, with $Q_P = 150$ and 1000 in the subsurface and in the water, respectively.

2010a; Brossier 2011, for a more extensive discussion on the use of the discontinuous Galerkin method in frequency-domain seismic modelling). However, a finer mesh is used in the first 160 m of the medium to allow for accurate positioning of the seismic devices.

The true Valhall models for *P*-wave velocity, density and impedance, were provided by the BP company, and they are shown in Figs 1(a)–(c). The Q_P-factor model is built by picking some geological interfaces on the velocity model (Fig. 1e). The values of Q_P are set to 80 where the velocity is smaller than 1600 m s^{-1} above 730 m in depth, and where the velocity is smaller than 1800 m s^{-1} below 730 m in depth. These two ranges of velocities are related to soft quaternary sediments and gas layers, respectively.

Five frequencies [3.5, 4, 5, 6.1, 7.1] Hz are successively inverted during FWI. The source excitation is estimated for each shot gather at each iteration of the FWI with an L2 norm, following the approach of Pratt (1999), while the the data functional for FWI is based on the L1 norm (Brossier *et al.* 2010c).

Smoothing of the gradient and of the diagonal Hessian is performed during the inversion to remove high-frequency artefacts through eq. (2). Before smoothing, gradient and diagonal Hessians are projected onto a Cartesian grid, for ease of implementation before projecting back the smoothed gradient and Hessian on the triangular mesh for step length estimation and model update. For both the synthetic and real case, the grid size of the Cartesian grid is 15 m, which is sufficiently small to avoid spatial aliasing (the best possible resolution would be $\lambda_{\min}/2 = 50$m). The vertical correlation length [τ_z, see eq. (6)] of the smoothing in each cell 'i' is defined as a fraction of the local wavelength, such that: $\tau_{z,i} = \frac{0.2\pi\ V_{P,i}}{\omega}$. The horizontal correlation length (τ_x) is set three times higher than the vertical correlation length as the vertical resolution power of surface acquisition is higher than the horizontal one in particular for horizontally stratified media.

For quality control of the FWI models, we compute the relative percentage of errors for the models [$\xi(\mathbf{m})$]

$$\xi(\mathbf{m}_{\text{cal}}) = \frac{100}{N_x.N_z} \sum_{i=1}^{N_x.N_z} \frac{|m_{\text{cal},i} - m_{\text{true},i}|}{m_{\text{true},i}}, \tag{8}$$

where N_x and N_z correspond to the number of nodes in the x and z direction after projection along a Cartesian grid. We compute the time-domain seismograms for three shots gather located at 2, 8 and 14 km in distance, and we take the mean data error to assess the data match. We define the data error as the expression

$$\xi(\mathbf{s}_{\text{cal}}) = \frac{100}{N_t.N_s} \sum_{i=1}^{N_s.N_t} |s_{\text{cal},i} - s_{\text{true},i}|, \tag{9}$$

where a time-domain shot gather is denoted by s_{cal}. The number of time samples and the number of seismograms per shot gather are denoted by N_t and N_s, respectively, while the horizontal and vertical numbers of nodes in the subsurface model are denoted by N_x and

N_z, respectively. We checked that the computed mean data misfit shows similar trend to those computed for other gathers located at different distances along the profile.

Velocity–density versus velocity–impedance imaging

We first discuss which of the two parametrizations, of (V_P, ρ) and (V_P, I_P), is the most relevant for acoustic FWI. Therefore, no attenuation is considered in the following discussion: the factor Q_P is homogeneous in both the true and the background models, and it is equal to 1000.

Radiation pattern analysis for the V_P, ρ and I_p parameters

Parametrization of the subsurface can be defined with different combinations of parameters for FWI (Forgues & Lambaré 1997; Plessix & Cao 2011). In non-linear waveform inversions such as FWI, the parameters of a given parametrization can be related non-linearly to the parameters of another parametrization. Therefore, the choice of the parametrization is not neutral (Tarantola 1986), and it will govern how the amplitude of a wave scattered by a point perturbation of one model parameter varies with the scattering angle. This scattered wave divided by the model perturbation corresponds to the partial derivative of the wavefield with respect to the parameter located at the diffractor point, as in eq. (4). In the following, with a slight abuse of language, the diffraction pattern of the secondary source of the partial derivative wavefield at the diffractor point, as the right-hand side of eq. (4), will be referred to as the diffraction pattern of the parameter. This diffraction pattern gives some insight into the sensitivity of the data to the parameter as a function of the scattering angle. As the scattering angle is closely related to the wavenumbers injected into the subsurface model during FWI (e.g. Sirgue & Pratt 2004), the diffraction pattern provides hints of the resolution with which a given parameter class is reconstructed. The strength of these secondary sources control the relative amplitudes of the partial derivatives of the wavefield with respect to the different parameter classes. Parameter scaling can be used to arbitrarily change the relative amplitude of these partial derivative wavefields and to deliberately steer the inversion towards some specific parameter classes.

The analytical expressions of the visco-acoustic/elastic diffraction patterns were developed asymptotically in the framework of the ray+Born approximation by Wu & Aki (1985), Tarantola (1986), Ribodetti & Virieux (1996) and Forgues & Lambaré (1997). Alternatively, these can be estimated numerically by computing the partial derivative wavefields in a finite-difference discrete sense, as shown in Malinowski *et al.* (2011) and Gholami *et al.* (2013b). In this study, we follow this numerical finite-difference approach, and derive diffraction patterns for three different parametrizations: (V_P, ρ), (V_P, I_P) and (I_P, ρ) (Fig. 2). We compute an incident monochromatic wavefield in a homogeneous model, which is subtracted from the wavefield computed in the same model in which we add a point diffractor associated with one parameter perturbation in the centre of the mesh. The resulting wavefield represents the wavefield scattered by the parameter perturbation in all directions for a given parametrization or, in other words, the partial derivative of the wavefield with respect to the model parameter located at the point diffractor. A key point is that during this simulation, the other parameters of the parametrization are left unchanged to the true values, and this procedure is repeated for each parameter class of the parametrization. The amplitude variations of the partial derivative wavefield around the point diffractor give some insight into the diffraction pattern of the model parameter. The diffraction pattern of a given parameter class (e.g. V_P) can change as a function of the parametrization (i.e. the other parameters involved in the subsurface description). This occurs because the partial derivative of the wavefield with respect to one model parameter is computed by keeping the other parameter of the parametrization constant by virtue of the chain rule of derivatives. This highlights why, for example, the V_P parameter does not have the same diffraction pattern (or Fréchet derivatives) when it is combined with density or impedance in the subsurface parametrization.

Table 1 outlines the values of the model parameters, that are used in the computation of the diffraction patterns. The source is a Dirac function and the frequency is 10 Hz. A first conclusion is that the diffraction pattern of the *P*-wave velocity parameter is isotropic in the (V_P, ρ) parametrization (Fig. 2a), while scattering of the *P*-wave velocity diffractor is only significant at large scattering angles in the (V_P, I_P) parametrization (Fig. 2c). A second conclusion is that the diffraction patterns of the *P*-wave velocity and the density overlap at small and intermediate scattering angles (Figs 2a and b). This suggests that many combinations of *P*-wave velocity and density can allow short-spread reflections to be matched equally well, hence highlighting the trade-off between these two parameters. In contrast, when the (V_P, I_P) parametrization is considered, the *P*-wave velocity and the impedance scatter waves at large and short scattering angles, respectively (Figs 2c and d). This highlights that the imaging of the long wavelengths is closely related to the reconstruction of the *P*-wave velocities from the diving waves and supercritical reflections, while the short wavelengths of the subsurface are closely related to impedance, which governs the amplitude of the short-spread reflections. This justifies the conventional workflow of seismic imaging based on the scale separation underlying the velocity model building task and migration. As the diffraction patterns of the *P*-wave velocity and impedance do not overlap significantly, no trade-off between these parameters is expected during FWI. In the (ρ, I_P) parametrization, the density and the impedance scatter waves for large and for all scattering angles, respectively (Figs 2e and f). Therefore, a trade-off between ρ and I_P is expected, similar to that between V_P and ρ.

Resolution analysis performed in the framework of diffraction tomography (Devaney 1982; Wu & Toksöz 1987; Sirgue & Pratt 2004) has shown that the large aperture angles between the source, the diffracting point, and the receiver allow for the recovery of the low spatial wavenumbers of the subsurface. Thus, the reconstruction of the *P*-wave velocity from the (V_P, I_P) parametrization is expected to be smoother than that inferred from the (V_P, ρ) parametrization as V_P scatters energy only for the large aperture angles in the first parametrization. This is a criterion that should be considered for choice of the best parametrization. The maximum amplitudes of the diffraction patterns shown in Fig. 2 are of the same order of magnitude. This implies that each parameter class is expected to have an influence on the data of similar strength over distinct or common range of scattering angles. Therefore, the main issue here is to manage the trade-off between parameters rather than to manage parameters with contrasting influence on the data (for an illustration of FWI involving parameters of contrasting influence on the data, see Gholami *et al.* (2013b) for imaging of VTI acoustic media). Note that the diffraction patterns of V_P and ρ are similar when they are computed in the acoustic approximation and in the elastic approximation (in the latter case for the *P–P* scattering mode) (Forgues & Lambaré 1997). Moreover, the diffraction pattern of the V_S parameter for the *P–P* scattering mode has a diffraction pattern with a smaller amplitude relative to the V_P and ρ amplitudes. Moreover,

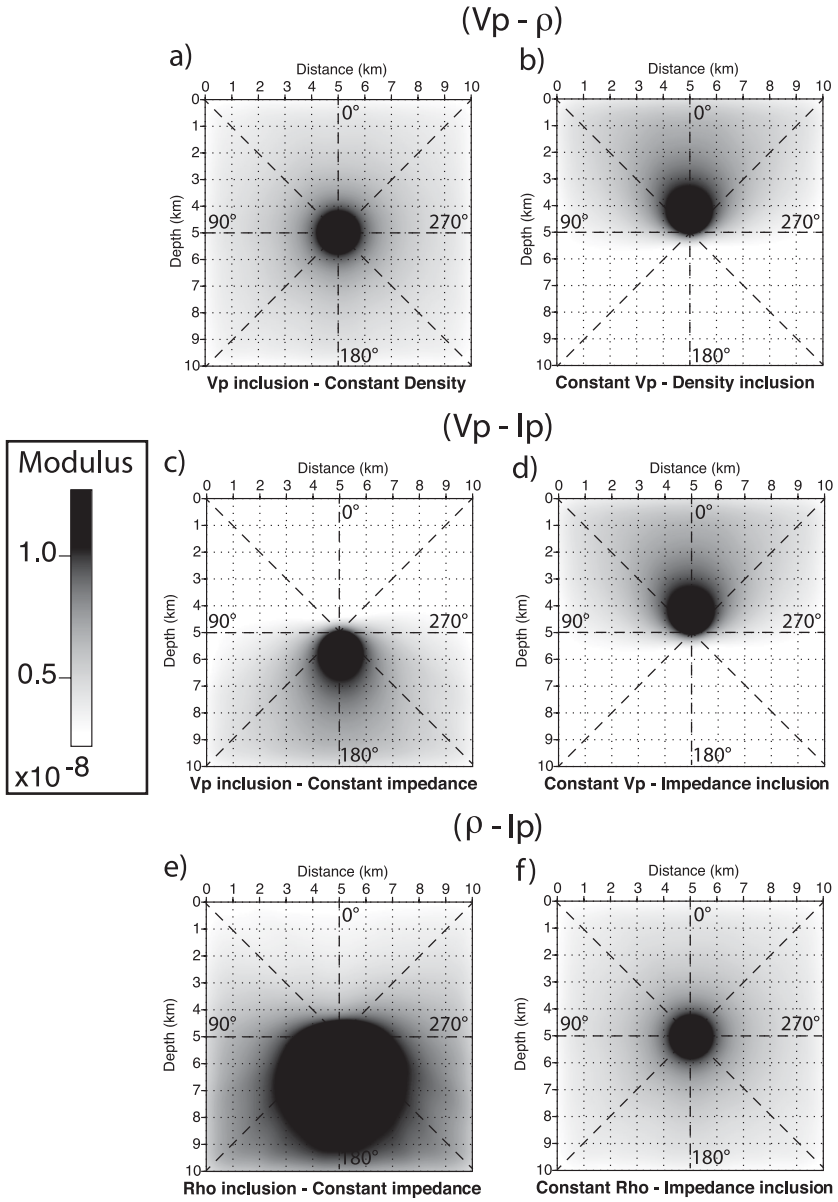

Figure 2. Wavefields scattered by a point diffractor located in the centre of the medium for three subsurface parametrizations: (V_P, ρ) (a, b), (V_P, I_P) (c, d) and (I_P, ρ) (e, f). The source is located at a distance of 5 km and a depth of 1 km.

Table 1. Values of the physical parameters in the homogeneous medium and the inclusion, as used to compute the diffraction patterns of Fig. 2 (see text for details).

Fig. 2 panels	Family	Inclusion of	Value in the	V_P (m s^{-1})	ρ (kg m^{-3})	I_P (kg m^{-2} s^{-1})
(a–f)	All	–	background	4000	2000	8×10^6
(a)	(V_P, ρ)	V_P	Inclusion	8400	2000	16.8×10^6
(b)		ρ		4000	4200	16.8×10^6
(c)	(V_P, I_P)	V_P		8400	952	8×10^6
(d)		I_P		4000	4200	16.8×10^6
(e)	(I_P, ρ)	I_P		8400	2000	16.8×10^6
(f)		ρ		1905	4200	8×10^6

the impact of V_S should be minor in soft seabed environments such as Valhall, as this should prevent significant *P*–*S* conversions. We therefore believe that the conclusions of our parametrization analysis drawn in the acoustic approximation will apply equally well to elastic media.

FWI results—joint reconstruction of two parameters

We first consider the joint reconstruction of the two parameter classes by FWI for the following two parametrizations: (V_P, ρ) and (V_P, I_P). For the joint reconstruction of the two parameter classes,

we use three time dampings, namely $\tau = 1, 3, 10$ s, and we allow for a maximum number of 30 iterations, to achieve time saving without preventing a significant decrease in the cost function (about 50 per cent). This leads to ten iterations per time damping. The stopping criterion of the iteration based on the amount of velocity perturbation at each iteration was never reached, and 30 iterations were effectively performed. We use the same optimization damping coefficients λ_i in the misfit function, eq. (1), for the *P*-wave velocity, the density and the impedance (4×10^{-18}), as we showed that the maximum amplitudes of the partial derivative wavefields associated with each parameter classes are similar (Fig. 2).

The initial velocity and density models for the FWI are built by Gaussian smoothing of the true models below the sea level, with horizontal and vertical correlation lengths of 500 m, as shown in Figs 1(f) and (g). The initial impedance model is built by multiplying the initial velocity and density models. This implies that the data residuals at the first iteration are the same when the (V_P, ρ) and the (V_P, I_P) parametrizations are used.

The velocity and density FWI models obtained with the (V_P, ρ) parametrization are shown in Figs 3(a), (b), 4(a) and (b), while the velocity and the impedance FWI models obtained with the (V_P, I_P) parametrization are shown in Figs 3(d), (e), 4(d) and (e).

The model error highlights not only the model inaccuracies associated with convergence towards a local minimum, but also the resolution with which a model is reconstructed according to the diffraction pattern of the parameter. Both sources of errors should be taken into account in the appraisal of the models. For example, we see in Table 2 that the percentage of errors for the density starting model (1.94 per cent) is smaller than the percentage of errors for the FWI density model (2.54 per cent), despite numerous fine structures created by the FWI. On the other hand, the percentage of errors on the seismograms computed in the starting models is almost three-times greater than that inferred from the FWI models.

The *P*-wave velocity reconstruction for the (V_P, ρ) parametrization is better resolved than that inferred from the (V_P, I_P) parametrization, in which some artefacts are shown down to 2.5 km in depth (Figs 3a, d, 4a and d). These inaccuracies might be related to the narrow diffraction pattern of the *P*-wave velocity when the (V_P, I_P) parametrization is used, which prevents broad-band reconstruction of the *P*-wave velocity model. The lack of high wavenumbers in the *P*-wave velocity model inferred from the (V_P, I_P) parametrization is more obvious in the upper structure, where the broader aperture illumination makes the filtering applied by the directivity of the virtual sources more visible.

The density model shows sharp reconstruction of the main reflector (Fig. 3b), with locally overestimated or underestimated perturbations (Fig. 4b). However, the impedance model inferred from the FWI (V_P, ρ) models (i.e. by multiplication of the V_P and ρ models) shows good agreement with the true impedance model, except at the reservoir level where a mispositioning of the reflector at 2.5 km in depth appears (Figs 3c and 4c). The good agreement between the true and the induced impedance models suggests that the amplitude errors of the density perturbations result from the trade-off between the velocity and density at short apertures. These trade-off artefacts are cancelled out when these two parameters are multiplied to build the impedance. This trade-off is highlighted by the opposite polarity of the velocity and density errors in Figs 4(a) and (b) (a positive velocity error is correlated with a negative density error and vice versa between 1.5 km and 2 km in depth).

The impedance model inferred from the (V_P, I_P) parametrization shows good agreement with the true impedance model (Figs 3e and 4e). However, the impedance model inferred from the velocity and the density FWI models obtained with the (V_P, ρ) parametrization (Fig. 3c) is of higher quality than the impedance model inferred from FWI with the (V_P, I_P) parametrization (Fig. 3e), with an error 0.32 per cent higher in the last case (Table 2). It is also illustrated by the comparison of the impedance logs at 1.3 km in depth in Figs 4(c) and (e). This might occur because the resolution of the impedance model might be hampered by its narrow radiation pattern when it is inferred from the (V_P, I_P) parametrization. In contrast, an improved impedance model can be inferred from the broad-band V_P model and the density model when the (V_P, ρ) parametrization is used. Furthermore, the density model inferred from the impedance and velocity FWI models obtained with the (V_P, I_P) parametrization is

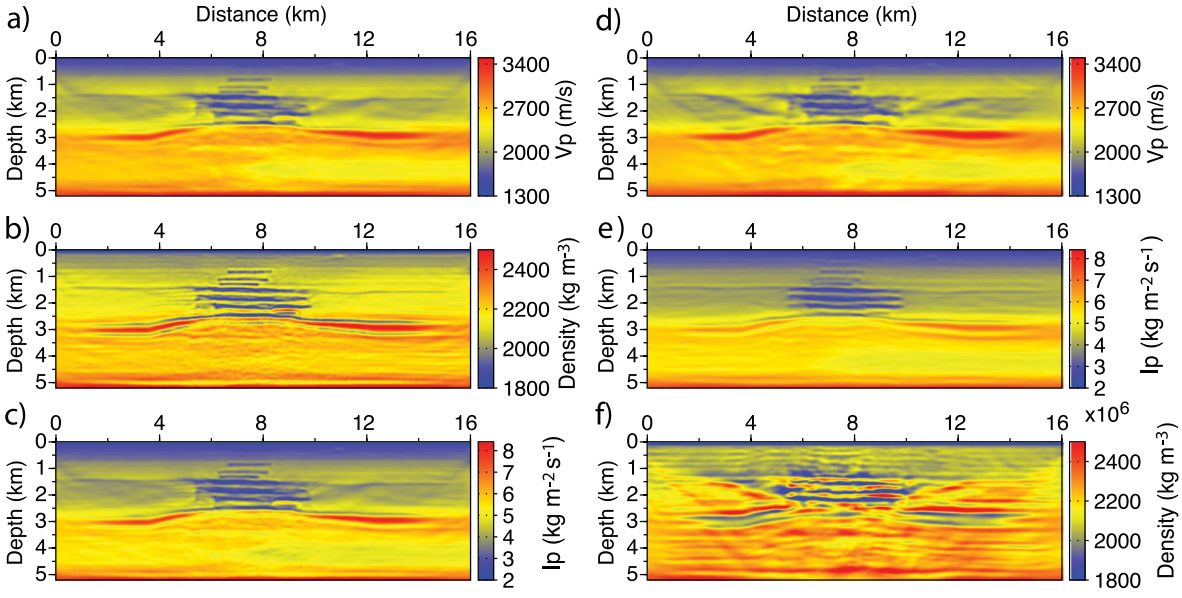

Figure 3. (V_P−ρ) versus (V_P−I_P) imaging. Joint update of two parameters (Table 2, tests 1 and 2). (a, b) Final V_P and ρ FWI models for the (V_P−ρ) parametrization (Table 2, test 1). (c) Impedance model inferred from the velocity and density models shown in (a, b). (d, e) Final V_P and I_P FWI models for the (V_P−I_P) parametrization (Table 2, test 2). (f) Density model inferred from the velocity and impedance models shown in (d, e).

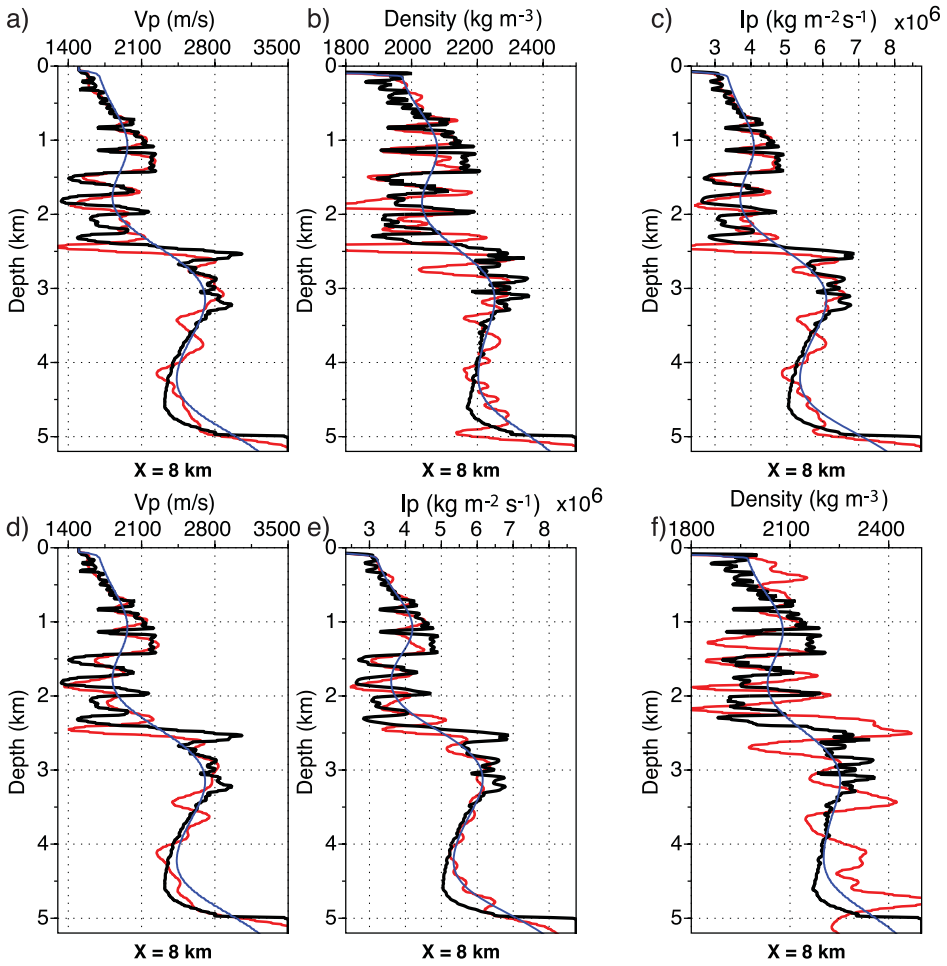

Figure 4. (V_P−ρ) versus (V_P−I_P) imaging. (a, b) Logs extracted from the final V_P and ρ FWI models for the (V_P, ρ) parametrization (Figs 3a, b). (c) Log extracted from the impedance model of Fig. 3(c). (d, e) Logs extracted from the final V_P and I_P FWI models for the (V_P, I_P) parametrization (Figs 3d, e). (f) Log extracted from the density model of Fig. 3(f). True models, starting models, and FWI models are plotted as black, blue and red curves, respectively.

Table 2. (V_P−ρ) versus (V_P−I_P) imaging. Model and data errors, [ξ(**m**)] and [ξ(**s**)] for the starting models and the FWI models obtained with the simultaneous or the hierarchical strategies, to recover the velocity, the density and the impedance. The smallest and the largest errors are written in red and blue, respectively. We added in subscript the parametrization of the tests, in the cells of the table where can occur ambiguities.

Test No.	Test step Step 1	Step 2	ξ(**m**) V_P	ρ	I_P	$\xi(s_{cal})$ $\times 10^{-5}$
	$V_{P,0}$, ρ_0, $I_{P,0}$		5.91	1.94	7.73	5.04
1	(Vp, ρ)		4.31	2.54	6.23	1.61
2	(V_P, I_P)		4.49	4.51	6.55	1.80
3	$V_{P_{(V_P,\rho)}}$		4.59			1.76
4	$V_{P_{(V_P,\rho)}}$	(V_P, ρ)	4.28	2.17	5.85	1.50
5	$V_{P_{(V_P,\rho)}}$	(V_P, I_P)	4.38	2.77	5.98	1.53

very unstable (Figs 3f and 4f), with the model error being almost 2 per cent higher than the error of the density model inferred from the (V_P, ρ) parametrization. This might occur because the density is deduced from two subsurface models that are parametrized by I_P and V_P which have two distinct wavenumber bandwidths with a limited overlap (Figs 2c and d). Moreover, a limited trade-off is expected between these two parameters. Therefore, no correlation is expected between the errors that impact of the two models. Another possible reason to explain the instability of the density model inferred from the (V_P, I_P) parametrization is related to the relative range of variations of each parameter, which is defined by (value max−value min)/value max. This is 60, 26 and 73 per cent for the velocity, density and impedance synthetic models, respectively. The level of these percentages shows the same trend as the magnitudes of errors computed in each of the starting models (Table 2). This percentage is higher for the impedance, because this is the product of the two other parameters, which vary within the same direction at all depths, as shown in the true logs of the velocity and density (Figs 4a and b, black lines). Small errors in the impedance or the velocity reconstruction will translate into a much greater relative error for the density inferred from those two models. In contrast, the impedance inferred from the velocity and the density is less sensitive to an error in the density. This statement will apply for all of the subsurface models where the velocity and density vary in the same direction, which is the most common scenario in geological targets (except for salt and coal).

An alternative that we did not implement for the (V_P, I_P) parametrization, would be to add some constraints in the misfit function to better stabilize the reconstruction of the density deduced from the two inverted parameters (V_P and I_P), by bounding the density according to empirical law, which relates velocity and density.

FWI results—hierarchical reconstruction of two parameters

We assess now whether the hierarchical reconstruction of the parameters is more robust than the simultaneous counterpart. In our hierarchical approach, during a first inversion, we update the *P*-wave velocity with a single time damping τ of 1 *s*, to cancel most of the residuals at large scattering angles. During a second inversion, we jointly update the *P*-wave velocity and the density or the impedance with three time dampings to reduce the remaining residuals at short scattering angles, for which the density and impedance have a significant influence on the data. During the second step, we jointly update the density or impedance and the velocity, as there can be a trade-off between those parameters.

During the first inversion, the *P*-wave velocity is updated with the (V_P, ρ) parametrization (Table 2, test 3), because the diffraction pattern of the *P*-wave velocity is broader when this parametrization is used. Some other tests where the (V_P, I_P) parametrization is used to recover V_P at the first step have shown stronger data errors than their counterparts for both the first and second steps. At the first step of the (V_P, ρ) parametrization (test 3), the trade-off between the *P*-wave velocity and density at short scattering angles should not impact on the imaging of the wave speed, because the inversion is limited to wide scattering angles through the use of aggressive time damping. The *P*-wave velocity model inferred from this first inversion step is shown in Figs 5(a), 6(a) and (d) (blue line). The resulting model is of satisfying quality, although its resolution is poorer than that of the model obtained during the joint reconstruction of velocity and density using three dampings (Figs 3a and 4a). The final FWI *P*-wave velocity model of the first inversion step is used as the initial model for the second inversion step. The initial density model is the smoothed version of the true model when the (V_P, ρ) parametrization is used during the second inversion step. The product of velocity inferred from the first inversion step with the smoothed density model provides the starting impedance model for the inversion performed with the (V_P, I_P) parametrization (Fig. 6e, blue line). Therefore, the data residuals are the same at the first iteration of the second FWI step for the two parametrizations.

The final models and logs of the second inversion step for the two parametrizations can be compared in Figs 5 and 6. The FWI density model obtained with the (V_P, ρ) parametrization (Fig. 5c) shows weaker perturbations than that obtained during the joint update of the *P*-wave velocity and density (Fig. 3b), which shows that many combinations of velocity and density models can match the data equally well. In addition, we show that the density is under-estimated down to 1.4 km in depth, and is polluted by ringing artefacts in the deep structure, which can result from the narrow aperture illumination at depth (Fig. 6b). The impedance model inferred from the *P*-wave velocity and density models shows good agreement with the exact model, except for the depth range of 2–2.5 km (Figs 5d and 6c).

Comparisons between seismograms computed in the true models and in the models obtained after the first inversion step (the first

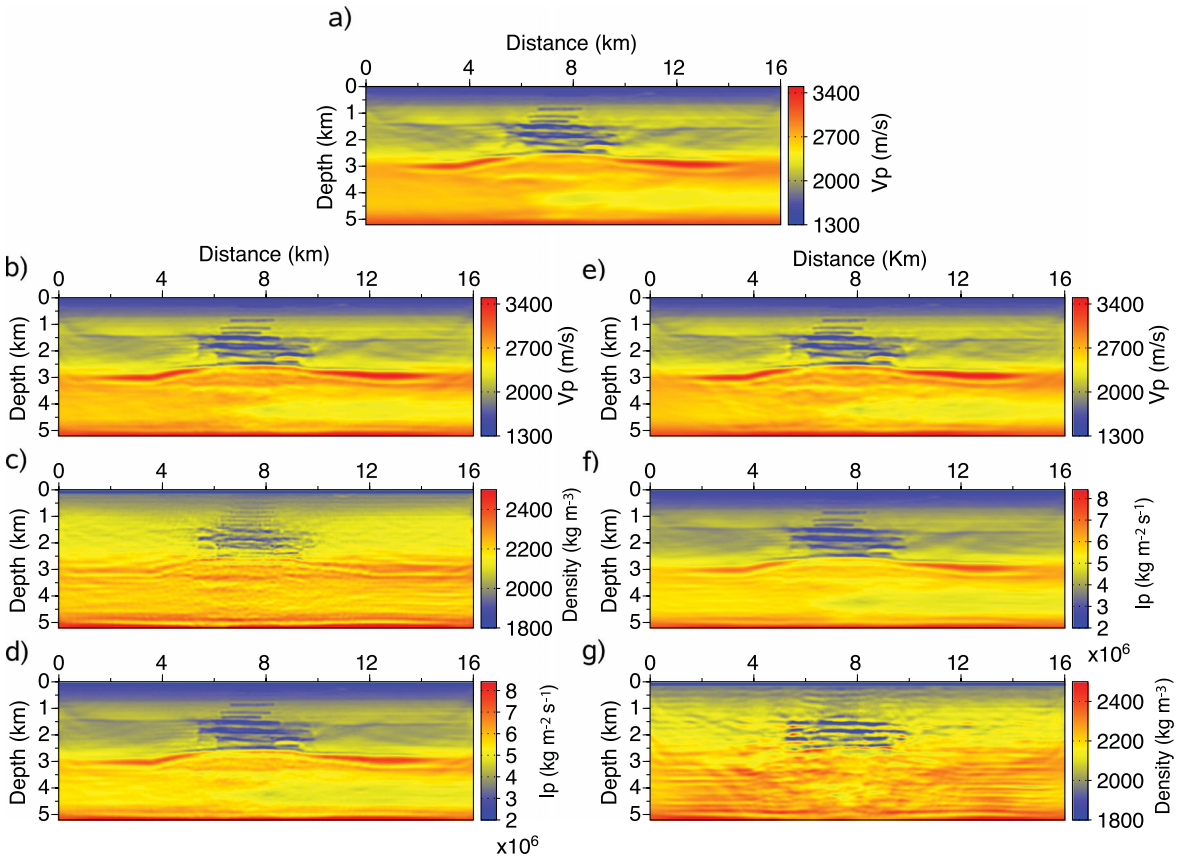

Figure 5. (V_P−ρ) versus (V_P−I_P) imaging. Hierarchical update of two parameters. (a) *P*-wave velocity model built by monoparameter inversion for the (V_P, ρ) parametrization (Table 2, test 3). A time damping of 1 s is applied to the data. (b, c) *P*-wave velocity (b) and density (c) FWI models for the (V_P, ρ) parametrization using the velocity model shown in (a) as initial model (Table 2, test 4). (d) Impedance model built from the velocity and density models shown in (b, c). (e, f) *P*-wave velocity and impedance FWI models for the (V_P, I_P) parametrization using the velocity model shown in (a) as initial model (Table 2, test 5). (g) Density model built from the velocity and impedance models shown in (e, f).

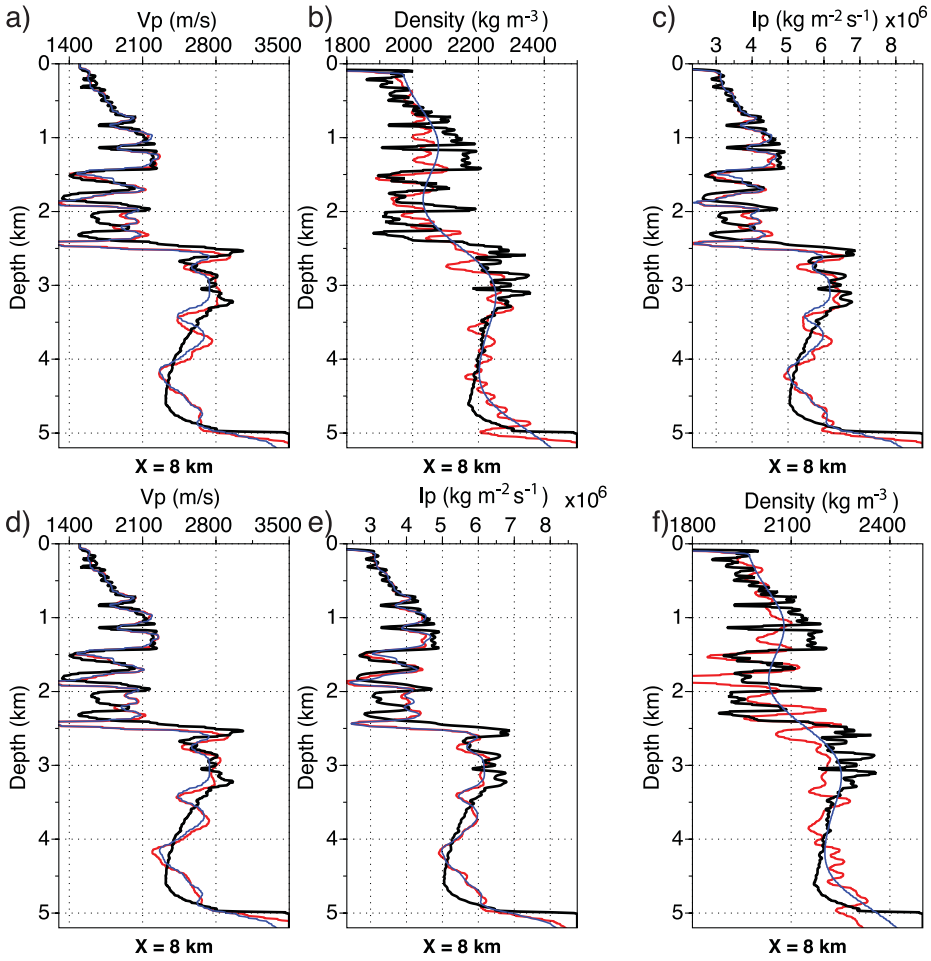

Figure 6. (V_P−ρ) versus (V_P−I_P) imaging. Hierarchical update of two parameters. (a, b) Logs extracted from the *P*-wave velocity (a) and density (b) FWI models of the second step of the hierarchical inversion for the (V_P, ρ) parametrization. (d, e) Logs extracted from the *P*-wave velocity (d) and impedance (e) FWI models of the second step of the hierarchical inversion for the (V_P, I_P) parametrization. (c, f) Impedance and density profiles inferred from the logs in (a, b) and (d, e) respectively. True models are plotted as black curves, starting models of the second-step inversion as blue curves, and FWI models of the second as red curves.

FWI V_P model, and the smooth initial density model) show that significant residuals associated with short-offset reflections were not cancelled out (Fig. 7e). In contrast, most of the residuals associated with diving waves and wide-aperture reflections were efficiently reduced (compared to the residuals in the starting model, Fig. 7d). This is consistent, as the first inversion was tuned to match the wide-aperture phases only. The seismograms computed in the final models of the second inversion step, for the (V_P, ρ) parametrization, and the corresponding residuals (Figs 7c and f) show that the match of the short-offset reflections was improved by incorporating the density effects into the inversion (Table 2).

During the second inversion step performed with the (V_P, I_P) parametrization, the velocity and impedance FWI models do not show significant perturbations, as their respective starting models are already quite close to the true models (Figs 5e, f, 6d and e). The density model inferred from the reconstructed V_P and I_P models is more accurate than the density model inferred from the simultaneous strategy with the (V_P, I_P) parametrization (Figs 5g and 6f). However, it is still less accurate than the density models inferred from (V_P, ρ) parametrization (in one or two steps, as confirmed by the density errors in Table 2).

The model errors for the (V_P, I_P) parametrization are higher than those for the (V_P, ρ) parametrization whatever the parameter class, although we note that the mean time residual is the same for the two parametrizations. This highlights the non-uniqueness of the multiparameter inversion, as several model combinations might provide an equivalent match of the data.

Considering this surface acquisition and this synthetic case study, all of the tests converge on the conclusion that the (V_P, ρ) parametrization is more suitable than the (V_P, I_P) counterpart for reliable reconstruction of velocity, density and impedance. This contradicts the widely accepted idea that promotes the reconstruction of velocity and impedance. One reason is the more limited bandwidth reconstruction of the velocity in the (V_P, I_P) parametrization, which might in turn impact on the impedance reconstruction. A second reason is that trade-off effects between velocity and density at short-to-intermediate scattering angles can be removed by multiplication to build an impedance model, which allows for the matching of the reflection amplitudes, while the broad-band velocity model accurately describes the kinematic attributes of both diving waves and reflections. Finally, this unexpected conclusion arises from the incorporation of wide-aperture data into the FWI analysis, which prompts us to favour a parametrization for which the dominant parameter V_P has a diffraction pattern as broad as possible. This is why we will only consider the (V_P, ρ) parametrization in the following of this study. We will also use the hierarchical strategy at the

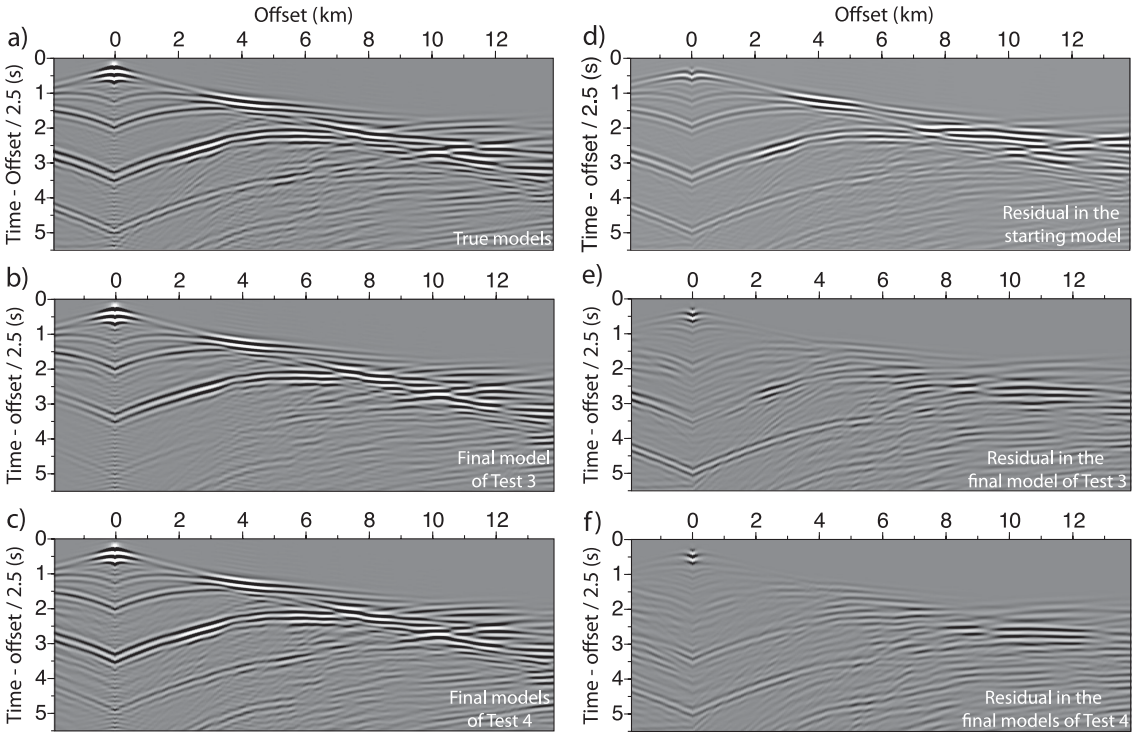

Figure 7. ($V_P-\rho$) versus (V_P-I_P) imaging. Synthetic seismograms computed in (a) the true V_P and ρ models; (b) the models obtained close of the first step of the hierarchical inversion when we only update the V_P parameter (Fig. 5a); (c) the models obtained at the end of the second step of the hierarchical strategy for the (V_P, ρ) parametrization (Figs 5b, c). (d) Residuals between the seismograms shown in (a) and the seismograms computed in the starting models of the first inversion step (smoothed version of V_P and ρ). (e) Residuals between the seismograms shown in (a) and (b). (f) Residuals between the seismograms shown in (a) and (c).

expense of the simultaneous counterpart, because the reconstruction of the density is more stable with the former approach (compare the density profiles of Figs 4b and 6b between 1.5 and 3 km depth); and because for a given parametrization, the data and model errors are smaller with the hierarchical inversion (compare test 1 with test 4).

Reconstruction of wave speed, density and Q_P-factor

In the following part of this study, we investigate the reconstruction of the Q_P-factor in addition to the *P*-wave velocity and density. This parameter can have a smaller imprint on the data than wave speed. However, values of Q_P as small as 10 are realistic in shallow soft sediments. If these low values of Q_P are not present in the initial model, the large error hampering the initial Q_P model can raise its influence in the inversion. Moreover, realistic values for the quality factor are around one order of magnitude lower than those of wave speeds or density. Therefore, fine tuning of the FWI regularization should be defined with the synthetic Valhall example before the real data case study is addressed. We set up this synthetic example such that it mimics as closely as possible the real data case study from Valhall presented later in this study. We consider a vertically transverse isotropic medium, that is parametrized by the vertical velocity (V_{P0}), the Thomsen parameters δ and ϵ, the density, and the quality factor. The Thomsen parameters provided by BP are kept fixed during the inversion, and are smoothed with a correlation length of 500 m to reproduce the inaccuracies of the Thomsen parameters observed in the real case (Prieux *et al.* 2011; Gholami *et al.* 2013a). The vertical velocity, the density, and the quality factor are updated during the FWI. The initial vertical velocity model is built by smoothing the true model with a Gaussian filter, where the correlation length is 250 m at the sea level, and 500 m at the bottom of the model. The initial density model is inferred from this velocity starting model using Gardner's law. The true model and the starting models of the Q_P-factor are presented in Figs 1(d) and (i). The smoothing of the model perturbations is the same for all of the parameters. Of note, the use of a stronger smoothing could have been applied to the attenuation, as shown by Malinowski *et al.* (2011), to stabilize the inversion when inverting the smallest frequencies for which the data are very weakly sensitive to the attenuation, although this was not judged necessary in our case as we did not observe instabilities.

On the sensitivity of the data to the density and Q_P-factor

Before showing the results of the multiparameter FWI, it is instructive to show how the inaccuracies of the vertical velocity, density, and Q_P-factor starting models translate into the data. For this purpose, we show in Figs 8(b)–(d) the differences between seismograms computed in the true (V_{P0}, ρ, Q_P) models (Fig. 8a) and those computed in the true models of two parameters and in the initial model of the remaining parameter. A smooth density model mainly impacts the match of the short-offset reflections, as a density diffractor leads to significant scattering at short apertures only (Fig. 8c). The footprint of the homogeneous Q_P-factor model is visible in the residual seismograms over the full offset range, although it is stronger at large offsets as more wavelengths are propagated (Fig. 8d). This is consistent with the isotropic diffraction pattern of the Q_P-factor parameter (Malinowski *et al.* 2011). The velocity smoothing leads to much stronger residuals relative to the density and Q_P-factor, hence this illustrates the dominant influence of this parameter on the data (Fig. 8b). All of these statements are also supported by the data

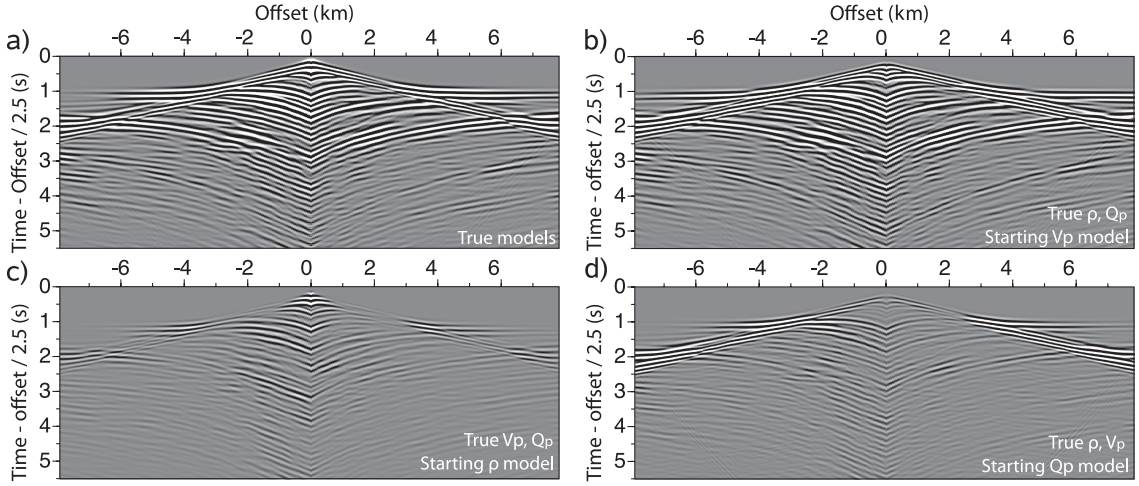

Figure 8. (V_P, ρ, Q_P) imaging. Sensitivity of the data to V_P, ρ and Q_P. (a) Synthetic seismograms computed in the true V_P, ρ and Q_P models (Figs 1a, b, d). (b) Residuals between seismograms computed in the true models and in a smooth velocity model, the true density model, and the true Q_P model. (c) Same as (b) for seismograms computed in a smooth density model, the true velocity model, and the true Q_P model. (d) Same as (b) for seismograms computed in the Q_P-factor model of 150 of Fig. 1(i), the true velocity model, and the true density model.

Table 3. (V_P, ρ, Q_P) imaging—data and model errors computed in the different FWI models presented in Fig. 11. Comparison of the three following strategies: (1) simultaneous reconstruction of V_P, ρ, and Q_P without time damping [$(V_P, \rho, Q_P)_{no-dp}$]; (2) reconstruction of V_P and Q_P with time damping of 1 s in a first step [$(V_P, Q_P)_{1dp}$], followed by simultaneous reconstruction of V_P, ρ, and Q_P in a second step [$(V_P, \rho, Q_P)_{no-dp}$]; (3) reconstruction of V_P without time damping in a first step [$(V_P)_{no-dp}$], followed by simultaneous reconstruction of V_P, ρ and Q_P without time damping.

Test No.	Test STEP 1	Test STEP 2	ξ(**m**) V_P	ρ	Q_P	I_P	$\xi_{mean,n}$ (**m**)	ξ_{mean} (**s**) $\times 10^{-5}$	Q_P	λ V_P	ρ
–	$V_{P,0}$		5.09	0	0			2.29			
–	ρ_0		0	2.26	0			0.43			
–	$Q_{P,0}$		0	0	32.66			0.95			
1	$(V_P, \rho, Q_P)_{no-dp}$		5.03	2.53	25.3	6.77	0.92	1.63	$\mathbf{8 \times 10^{-20}}$		
2	$(V_P, \rho, Q_P)_{no-dp}$		4.90	2.82	27.86	6.81	0.96	1.83	4×10^{-18}		
3	$(V_P, Q_P)_{1dp}$		5.20		28.89	6.82		1.94	$\mathbf{8 \times 10^{-20}}$		
4	$(V_P, Q_P)_{1dp}$	$(V_P, \rho, Q_P)_{no-dp}$	5.09	2.96	27.84	6.97	0.99	1.38	$\mathbf{8 \times 10^{-20}}$	4×10^{-18}	4×10^{-18}
5	$(V_P, Q_P)_{1dp}$	$(V_P, \rho, Q_P)_{no-dp}$	5.12	3.04	24.65	7.07	0.97	1.35	4×10^{-18}		
6	$(V_P)_{no-dp}$		4.94			6.45		1.70	$\mathbf{8 \times 10^{-20}}$		
7	$(V_P)_{no-dp}$	$(V_P, \rho, Q_P)_{no-dp}$	4.89	2.41	26.11	6.35	0.89	1.35	$\mathbf{8 \times 10^{-20}}$		
8	$(V_P)_{no-dp}$	$(V_P, \rho, Q_P)_{no-dp}$	5.00	2.39	23.54	6.43	0.88	1.16	4×10^{-18}		
–	Mean FWI models		4.72	2.41	25.08	6.31	0.88	1.25			

errors that are computed in the corresponding seismograms (Table 3), with an error for the vertical velocity that is 2.4-fold greater than for the attenuation, and 5.3-fold greater than for the density.

To assess the sensitivity of the monoparameter inversion for the vertical velocity to the inaccuracies of the density and the Q_P-factor background models, we compare the FWI velocity models, that are built without time damping, when the background density and Q_P-factor models are the true models and the smoothed models, respectively (Figs 9 and 10). The FWI models inferred from the smooth density and Q_P-factor background models are slightly noisier in the tertiary sediments, but remain close to the vertical velocity model inferred from the true density and Q_P-factor background models. This validates the hierarchical approach, where the vertical velocity is updated during a first inversion step.

Inversion strategies

We test three different inversion strategies to update the vertical velocity, the density and the Q_P-factor.

The 'first strategy', strategy 1, consists of the simultaneous update of the vertical velocity, density and Q_P-factor without any time-damping procedures. The simultaneous inversion should prevent artefacts that arise when hierarchical strategies are used: the reconstruction of the dominant parameter alone might contribute to reduce the residuals mainly associated with secondary parameters. The 'second strategy', strategy 2, is a multiscale procedure, that proceeds hierarchically over the data and the parameter classes. In a first step, we jointly update the large wavelengths of the *P*-wave velocity and the Q_P-factor using a strong time-damping factor τ of 1 *s*, as eq. (7), to minimize data residuals associated with the wide scattering angles before the simultaneous update of the shorter wavelengths of *P*-wave velocity, density and Q_P-factor without time damping to reduce the remaining residuals at short apertures. Following a multiscale approach, low wavenumbers are reconstructed first by updating the parameters V_P and Q_P, which are sensitive to the wide scattering angles, then by the reconstruction of the short wavelengths by updating the parameters that are sensitive to the short scattering angles (V_P, ρ and Q_P). In the 'third strategy',

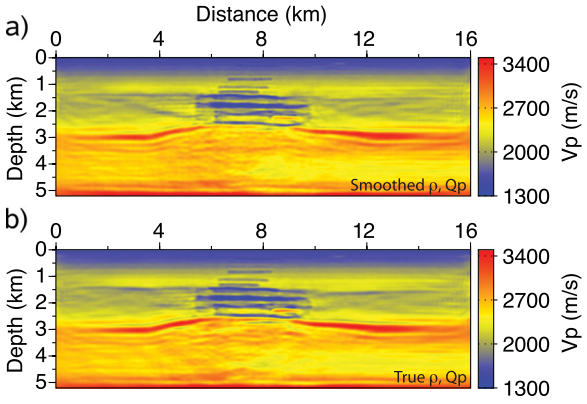

Figure 9. (V_P, ρ, Q_P) imaging. Sensitivity of monoparameter FWI for V_P to the accuracy of the density and Q_P-factor background models. (a, b) *P*-wave velocity model inferred from monoparameter FWI using three time dampings. (a) The background density and Q_P models are smoothed versions of the true models. (b) The background density and Q_P models are the true models.

strategy 3, we update the *P*-wave velocity without time dampings during a first step to cancel most of the residuals, and build a velocity model as kinematically accurate as possible before jointly updating the three parameter classes (without time damping). In this last strategy, we keep the same starting model of density for the first and the second inversion steps.

Compared to the update of the density or impedance presented in the previous section, we found that an additional difficulty with the reconstruction of the Q_P-factor arises because the partial derivative of the wavefield with respect to Q_P have much smaller amplitudes than the V_P and ρ counterparts, although we scaled the parameters by their mean value. The difficulty associated with this scaling comes from the fact the model perturbations relative to the starting models during FWI can be much higher for Q_P (100 per cent of its mean value) than for V_P and ρ (20 per cent of its mean value), in particular in shallow soft sediments where a strong decrease of Q occurs. This is highlighted by the diagonal terms of the Hessian computed for parameters scaled by their mean value. These diagonal terms are of the order of 10^{-17}, 5×10^{-18} and 10^{-21} for V_P, ρ and Q_P, respectively. Those values show in one hand the weaker influence of ρ and Q_P relative to V_P on the data as shown by Figs 9 and 10, and are in the other hand related to the dynamic of Q as mentioned just above. Applying the inverse of the Hessian to the gradient of the misfit function should theoretically correct for the variable influence of each parameter class on the data, if the inverse problem was not ill-posed, to reconstruct the true values of the model parameters. Due to the ill-conditioning of the Hessian, regularization of the inverse problem is needed. However, a difficulty arises when the damping regularization term λ_i, eq. (2), is added to the diagonal of the approximate Hessian for regularization. If the damping λ_i is too high, it can hamper the scaling effect of the Hessian during the updating of the parameter class i, and also impact on the reconstruction of the other parameters. In this framework, we found that a suitable choice of the damping λ_i is a key feature to properly scale the model perturbations associated with each parameter class. Alternatively, different parameter scaling can be viewed to balance the amplitudes of the partial derivative wavefields associated with the different parameter classes, and hence to design a better-conditioned Hessian. This should facilitate the setting of the hyperparameter λ_i, which should have a comparable order of magnitude. In this case, the risk is however to give a significant weight during the optimization process to a parameter that has a weak imprint on the data, this imprint being potentially dominated by data noise. In this case, the inversion can become unstable. The choice of the parameters and their scaling is thus critical as it should combine physical considerations (is the influence of the parameter greater than noise?) and numerical aspects (is the Hessian well-conditioned?). In this study, we focus on the strategy where the parameter classes are normalized by the mean value of the parameter. In this case, the inversion is steered towards the reconstruction of the parameters that have a dominant imprint on the data. We however seek to find a fine tuning of the λ_i such that the secondary parameters can be reconstructed in a stable way.

To illustrate the effects of the damping terms λ_i on the inversion, we perform a series of tests, an outline of which is presented below. First, we estimate a suitable value of the damping term for the velocity (λ_{V_P}) during a monoparameter inversion. As an illustrative example, a damping term of 10^{-6} leads to negligible velocity perturbations. An improved *P*-wave velocity model is obtained using a damping term of 4×10^{-8}. Secondly, we perform the

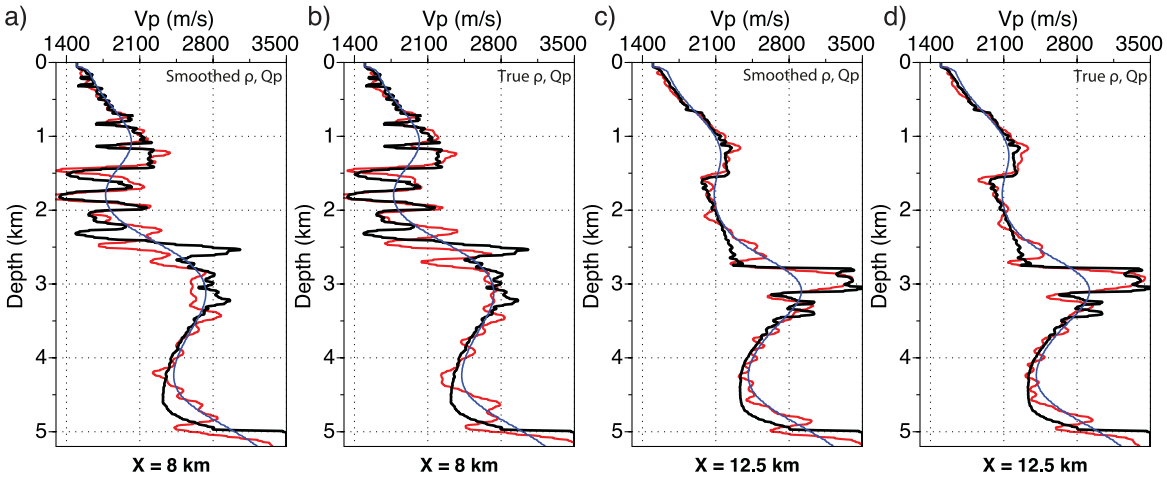

Figure 10. (V_P, ρ, Q_P) imaging. Sensitivity of monoparameter FWI for V_P to the accuracy of the density and Q_P-factor background models. Comparison between the true and FWI velocity profiles at 8 km (a, b) and 12.5 km (c, d) in distance. (a, c) The FWI logs are taken from the FWI model of Fig. 9(a). (b, d) The FWI logs are taken from the FWI model of Fig. 9(b). True models are plotted as black curves, starting models of the second-step inversion as blue curves, and FWI models of the second step as red curves.

simultaneous updates of the *P*-wave velocity, density and Q_P-factor using the same damping for each parameter class (4×10^{-8}). This inversion set-up leads to underestimation of the model perturbations, which prompted us to repeat the experiment with decreasing values of the damping until unstable reconstructions were found. The smallest value of the damping that leads to reasonable models is 4×10^{-18}, which is ten times smaller than the maximum coefficient value of the block in the approximate Hessian associated with the velocity parameter. Decreasing the damping to values as small as 8×10^{-20} for the three parameters leads to an unperturbed velocity model, a sharp model of density, and an unstable model of the Q_P-factor. We conclude from these experiments that a suitable approach to properly scale the model perturbations associated with each parameter class consists of adapting the value of the damping factor to each parameter class. We use a value of the damping term for Q_P, that is 50-fold smaller than the dampings used for the *P*-wave velocity and density. It is worth noting that, even in this setting, the damping term of the Q_P-factor is almost two orders of magnitude higher than the maximum coefficient of the diagonal block of the Hessian associated with the Q_P-factor, hence, making the contribution of this part of the approximate Hessian negligible. Note that after some iterations, estimation of the Hessian done by the L-BFGS is not damped anymore by the regularization term.

FWI results

We synthesize all of the FWI results (models and logs) in Figs 11 and 12, and to appraise these, the corresponding data and model errors [$\xi(\mathbf{s})$ and $\xi(\mathbf{m})$] are shown in Table 3. The FWI Q_P models can be compared with the true model represented with the appropriate colour scale in Fig. 1(e). We show the inversion results for the three strategies described in the previous section for two sets of dampings λ_i during the joint update of the *P*-wave velocity, density, and Q_P-factor: ($\lambda_{V_P} = 4 \times 10^{-18}$, $\lambda_\rho = 4 \times 10^{-18}$, $\lambda_{Q_P} = 8 \times 10^{-20}$), and ($\lambda_{V_P} = \lambda_\rho = \lambda_{Q_P} = 4 \times 10^{-18}$). A first overview of the whole scene shows large variations in the solutions, and in particular for the density and the Q_P-factor, which reminds us of the intrinsic ill-posedness of the multiparameter inversion. It is worth noting that the value of λ_{Q_P} not only impacts on the reconstruction of the Q_P-factor parameter, but also on the reconstruction of all of the other parameters because the sum in the model cost function is differently weighted.

Strategy 1: joint update of multiple parameters

For the *first strategy*, the time and model errors (except for the velocity) tend to promote the use of a small value of λ_{Q_P} (i.e. 8×10^{-20}) (Table 3). With this setting (Figs 11a–c), the density and the Q_P-factor perturbations are better balanced compared to

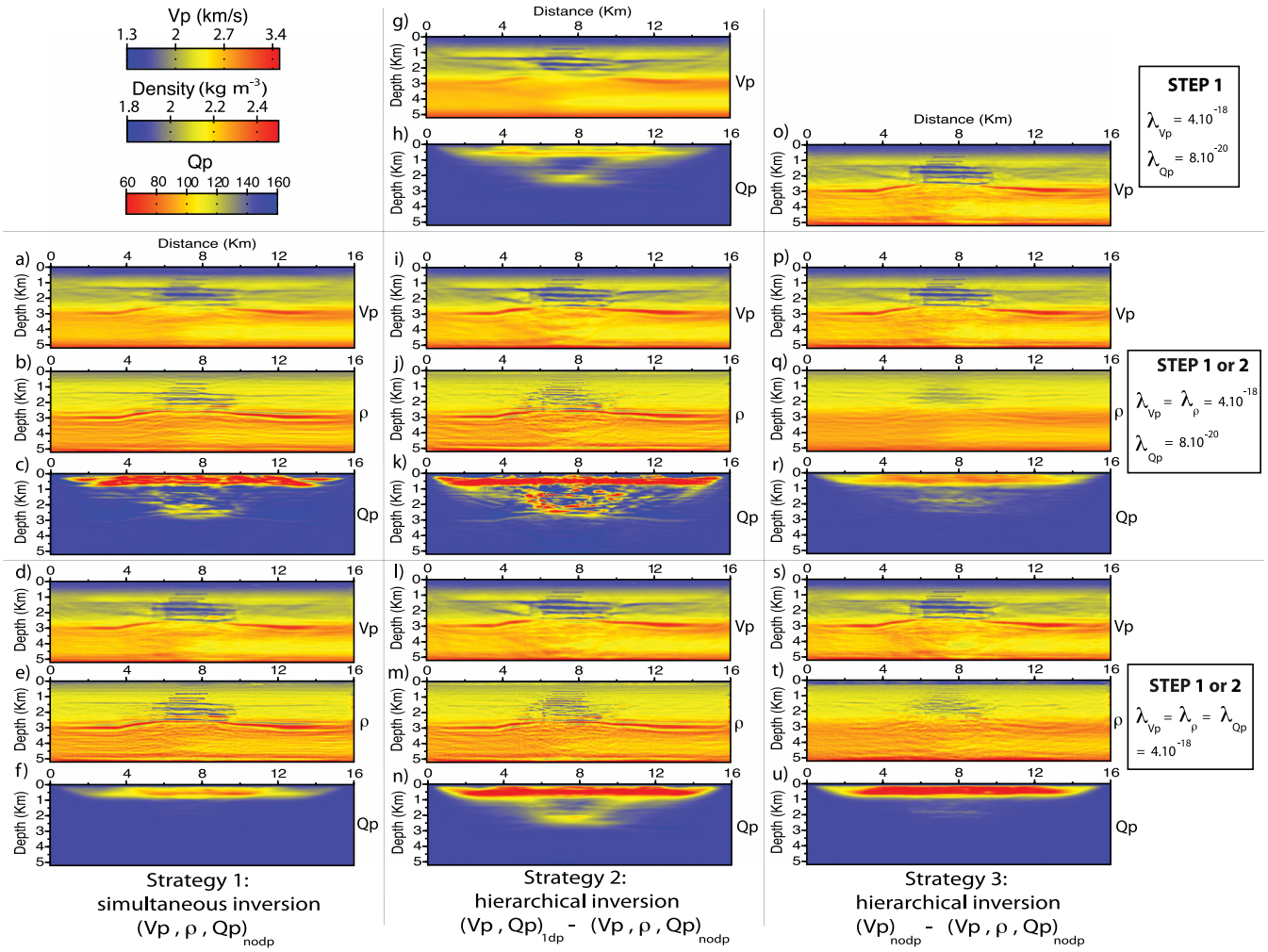

Figure 11. (V_P, ρ, Q_P) imaging. FWI results. (a–u) Velocity (a, d, g, i, l, o, p, s), density (b, e, j, m, q, t) and Q_P-factor (c, f, h, k, n, r, u) models inferred from three different inversion strategies (see text for details). For each strategy, the FWI was performed using the same damping ($= 4 \times 10^{-18}$) for all of the parameter classes (d–f, l–n, s–u), and using λ_{Q_P} 50-times smaller than λ_{V_P} and λ_ρ (a–c, g–k, o–r).

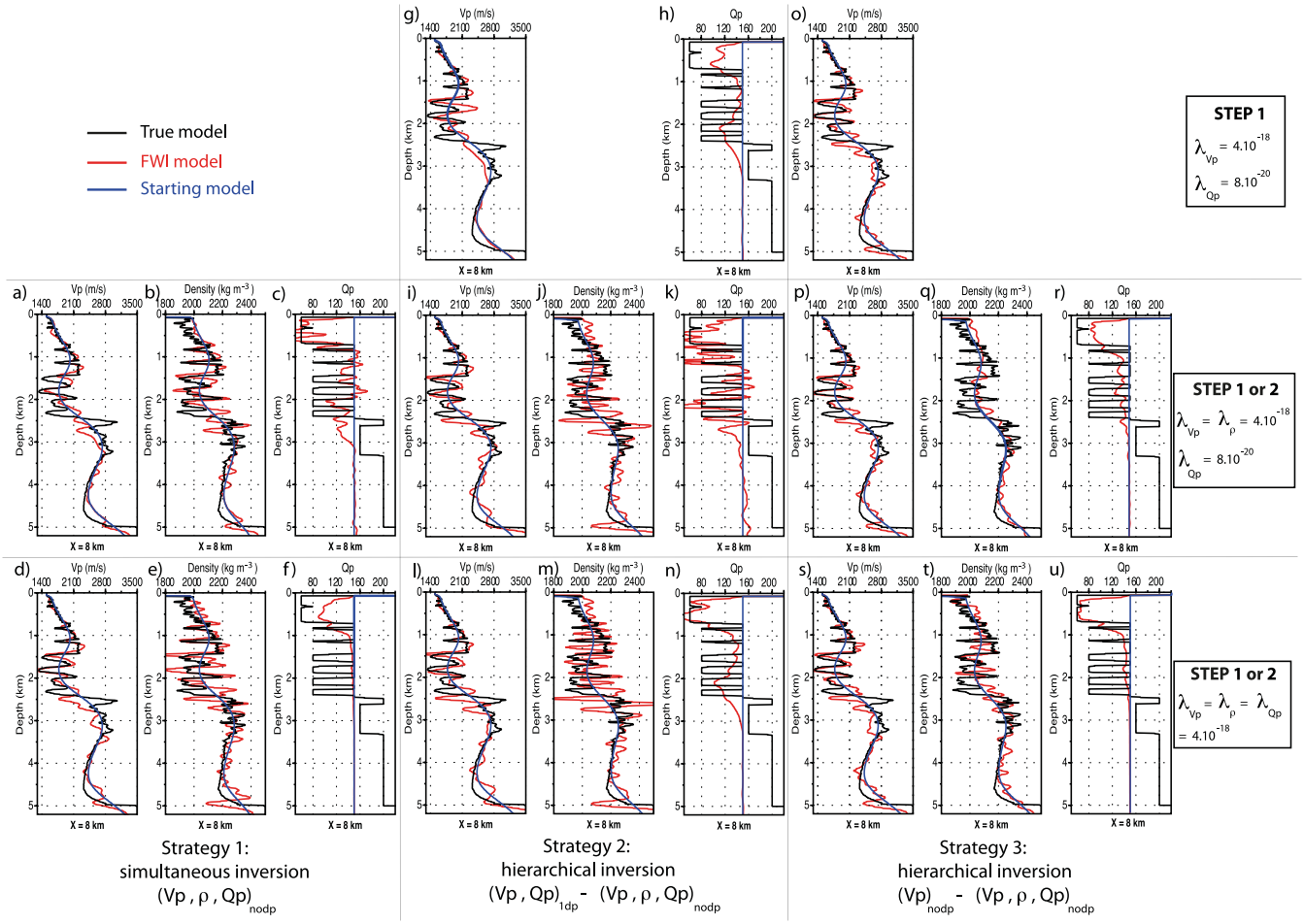

Figure 12. (V_P, ρ, Q_P) imaging. FWI results. (a–u) Velocity, density and Q_P-factor profiles extracted from the FWI models shown in Figs 11(a)–(u). True models are plotted as black curves, starting models of the second-step inversion as blue curves, and FWI models of the second step as red curves.

the case where the same damping is used for all of the parameters (Figs 11d–f). In this case, the Q_P-factor perturbations are clearly underestimated, while the density perturbations are overestimated (Figs 11a–f and 12a–f).

Strategies 2 and 3: hierarchical approaches

In strategies 2 and 3, we attempt to increase the quality of the starting models of the second-step inversion, the aim of which is to jointly update the three parameters. *Strategy 2* is intermediate between strategies 1 and 3, and while *strategy 1* does not include a preliminary inversion step, the first inversion step of *strategy 3* uses the full data space to update a reduced model space corresponding to the dominant parameter V_P, hence putting strong prior information into the initial V_P model of the second-step inversion. *Strategy 2* uses a reduced data space; namely, the wide-aperture components of the data, to update a reduced model space that involves two parameter classes during the first inversion step. Contrary to *strategy 1*, the data and model errors in strategies 2 and 3 tend to promote the use of the same damping term for all of the parameters. *Strategy 2* succeeds in the reconstruction of the smooth velocity and Q_P-factor models during the first step of the hierarchical approach (Figs 11 g and h). This should help to build a more accurate velocity model during the second step of *strategy 2* than during *strategy 1*. Although a mispositioning in depth of the top of the reservoir subsists, the velocities are better reconstructed in terms of amplitude (Figs 11i and l). However, the reconstruction of the density and the Q_P-factor is not successful during the second step, and shows noisy reconstructions in particular for the density (Figs 11j, k, m and n).

Strategy 3 (Figs 11o–u) leads to a final velocity model, that is close to that inferred from *strategy 2*. A more stable reconstruction of the Q_P-factor and density is, however, achieved with strategy 3 because most of the residuals were cancelled out during the update of the dominant parameter during the first inversion step (Fig. 11o), which was performed without any time damping. The underestimated perturbations in the density and Q_P-factor models highlight, however, the trade-off between velocity and these two secondary parameters. The trade-off between the two secondary parameters, ρ and Q_P, is highlighted by the results obtained with the two different values of λ_{Q_P}: as λ_{Q_P} is decreased, the Q_P-factor perturbations are strengthened at the expense of the density perturbations (Figs 11q, r, t and u).

Summary of the results of the synthetic example

This numerical investigation highlights the difficult problem of FWI tuning when multiple classes of parameters must be updated with potential trade-off between the parameters and the variable sensitivity of the data to the parameters. The choice of the best damping factor in the Hessian has been shown to be a key issue to guarantee reliable results. Our approach to estimate these dampings (one per parameter class) remains heuristic and relies on trial-and-error approaches. The L-curve method might be preferred to determine the best damping terms (Hansen 1992), although it would be very time consuming to do this for each of the three parameters. Despite our

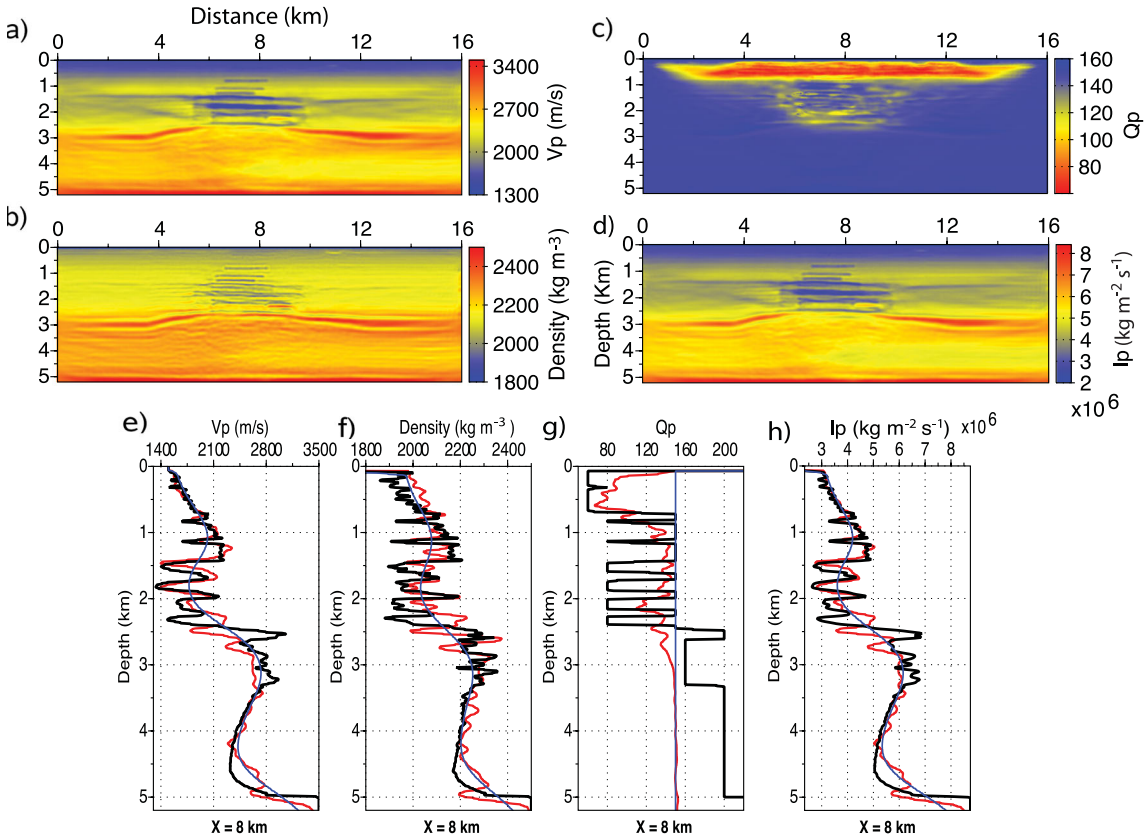

Figure 13. (V_P, ρ, Q_P) imaging. (a–c) Mean FWI results for the three parameters respectively obtained by averaging all of the models in Fig. 11. (d) Mean impedance model derived using (a) and (b). (e–h) Corresponding vertical profiles at a distance of 8 km.

indicators of quality, it is difficult to decide which model is the most satisfying. For real case studies, our prior geological knowledge and the prior information coming from well logs should help to choose the best parametrization. It might also be preferred to adopt the famous principle: '*When in doubt, smooth*' [Sir Harold Jeffreys, as quoted by Moritz & Sunkel (1978)], which is also known as the Occam principle. A Bayesian view to smooth the models is to take the mean over all of the derived FWI models. We applied this idea to the models presented in Fig. 11, which provided the mean models shown in Fig. 13. It is of interest to see how this process decreases the noise that was present in the original models. We see that the mean Q_P-factor allows us to accurately locate the depth of the gas layers, although the amplitude is still underestimated. With the models of test 8, the normalized mean error on these models is the smallest, and the data error is the second smallest value among all of the other tests. Of note, we have taken a simple mean of the models, but it is possible to compute a weighted mean, where the weights would be defined from the model and time errors to enhance the most realistic models (only possible for the synthetic case). For real cases, prior criteria like *in situ* petrophysical measurements, can be used to define the probability densities via Bayes Theorem (Gouveia & Scales 1997, 1998).

APPLICATION TO REAL OBC DATA FROM THE VALHALL FIELD

We now present the application of multiparameter anisotropic visco-acoustic FWI to wide-aperture data recorded by two ocean-bottom cables (21 and 29) of a 3-D survey on the Valhall field (Fig. 14). Valhall is an old Grabben that went into compression during the late Cretaceous, which led to the creation of an anticline that delineates the cap-rock of the reservoir at a depth of 2.6 km (Fig. 15a). The overburden is made of tertiary sediments, where gas is trapped in some layers. Line 21 is more centred on the axis of the anticline than line 29, and its data are thus more affected by the gas cloud attenuation effect. Line 21 is also closer to the platform, located as ($X = 6.2$ km, $Y = 11.2$ km), and consequently more sensitive to its noise. Acoustic isotropic FWI of the 3-D data set was presented by Sirgue *et al.* (2010). The resulting velocity model shows a complex network of channels 150 m below the surface as well as a gas cloud with gas-filled fractures at around 1 km in depth. The hydrophone data of cable 21 were processed by monoparameter anisotropic and isotropic acoustic FWI by Prieux *et al.* (2011) for an assessment of the footprint of anisotropy on the isotropic FWI. They showed that the horizontal velocities are reconstructed by isotropic FWI in the upper part of the target, where the FWI is mainly driven by the wide-aperture components of the data (diving waves and supercritical reflections). This can lead to underestimated velocities in the gas layers and/or to mispositioned reflectors at the reservoir level, which are required to match the short-spread reflections from the deep discontinuities. In the following, we will take anisotropy into account by using large-scale models of the Thomsen parameters δ and ϵ, which have already been built by anisotropic reflection traveltime tomography. Prieux *et al.* (2011) showed that, for the Valhall case study, the long wavelengths of the Thomsen parameters provide sufficiently accurate background models for the reliable reconstruction of the vertical velocity by FWI. The background model of horizontal velocity nevertheless presents a kinematic error, which reaches up to 0.1 s at 5.5 km offset (Prieux *et al.* 2011). The

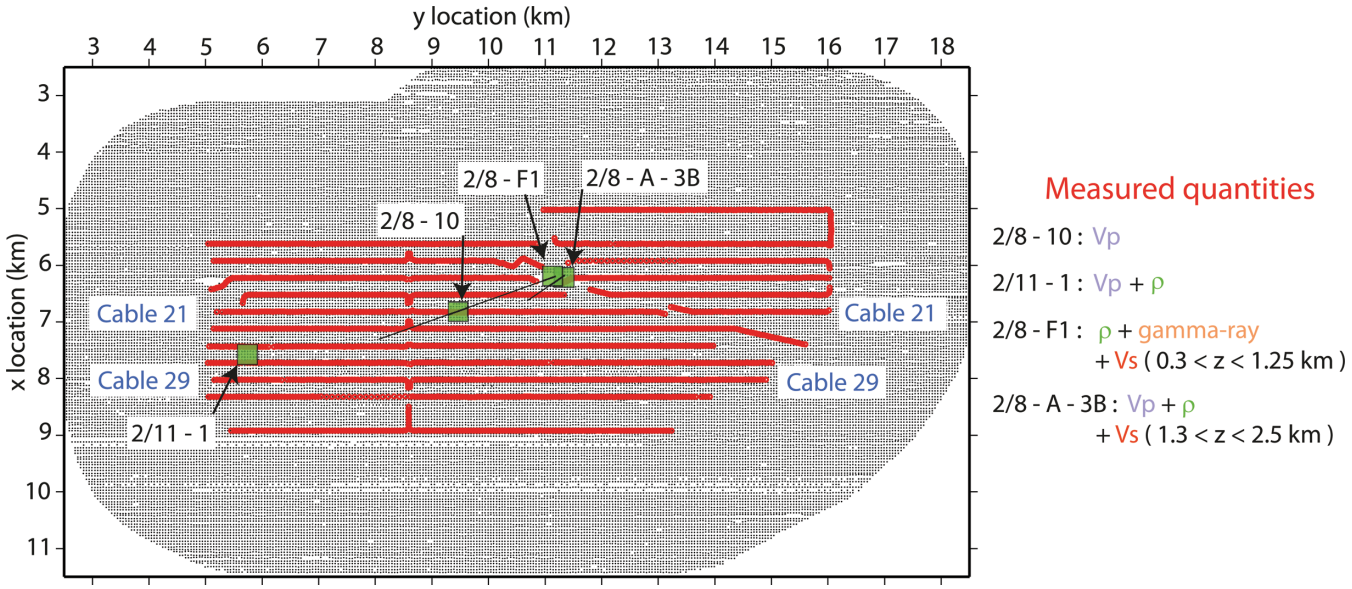

Figure 14. Real data case study. Valhall acquisition layout on which the locations of the well logs provided by BP are shown (green squares) with their different quantities measured (on right-hand side). Of note, wells '2/8-F1' and '2/8-A-3B' are close to the Valhall platform. Black line, route in depth of the well logs.

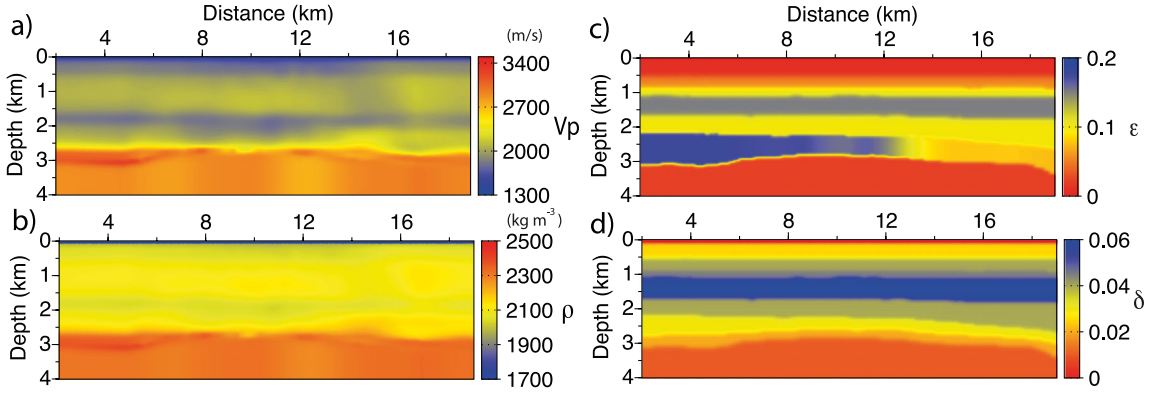

Figure 15. Real data case study. Starting models of line 29 for the vertical velocity (V_{P0}) (a), the density (b) derived from V_{P0} model using Gardner's law and for the Thomsen's parameters ϵ and δ (c, d).

shot and receiver spacings are 50 m and the maximum offset in the data is 13 km. A more detailed description of the anatomy of the Valhall OBC data is provided in Prieux *et al.* (2011).

Initial models and FWI set-up

Only the hydrophone component is considered for the FWI. We invert five overlapping groups of three frequencies: [3.5, 3.78, 4], [4, 4.3, 4.76], [4.76, 5, 5.25], [5.25, 5.6, 6] and [6, 6.35, 6.7] Hz. Contrary to the synthetic tests where only single frequencies were inverted, we adopt this strategy to be less sensitive to the noise. We apply *strategies 1* and *3*, which were assessed during the previous synthetic case study. *Strategy 2* provides negligible perturbations for the three parameters (not shown here) once the long-wavelength components of the V_p were reconstructed during the initial step: in this case the hierarchical inversion approach remained stuck in a local minimum because of the over-interpretation of the data during the early inversion step.

As for the previous synthetic example, the anisotropic visco-acoustic wave equation is parametrized by the vertical velocity, the density, the Q_P-factor, and the Thomsen parameters δ and ϵ, with these last two parameters kept fixed during the inversion (Prieux *et al.* 2011). The initial density model of one inversion step is inferred from the initial velocity model of the current inversion step, using Gardner's law. The starting models for the vertical velocity, the Thomsen parameters, and the densities of line 21 and line 29 in Prieux *et al.* (2011), were shown in their fig. 2, and are shown here in Fig. 15.

A two-layer Q_P-factor starting model was built by matching the amplitude versus offset trend of the first arrivals. Q_P is 150 below the sea bottom, while we use a value of 1000 in the water layer (Prieux *et al.* 2011).

FWI results and model appraisal

The final FWI models of lines 29 and 21 inferred from *strategies 1* (joint update) and *3* (hierarchical update) are shown in Figs 16 and 17, respectively. No time damping was applied in *strategies 1* and *3*, because reproducing the first step of *strategy 3* with three time dampings (equal to 1s, 3s, 5s) for line 21 have contributed to the amplification of artefacts in the top left of the model (Fig. 17d). One reason might be that the hierarchical inversions of limited (damped) subdata sets have more degrees of freedom to create artificial features in the subsurface model that aim to compensate

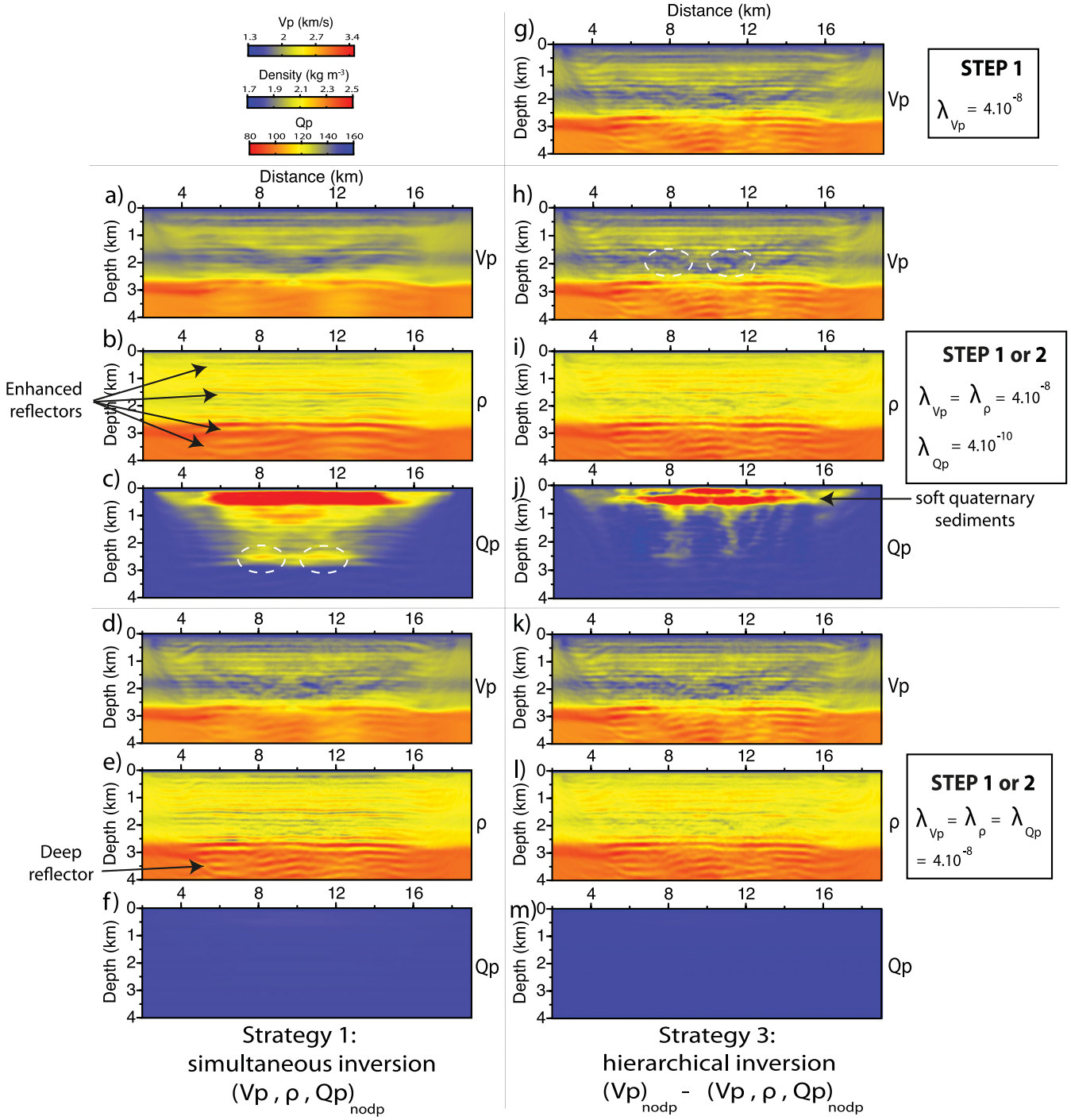

Figure 16. Real data case study. FWI models for vertical velocity (a, d, g, h, k), density (b, e, i, l), and Q_P-factor (c, f, j, m) for line 29. (a–f) Simultaneous update of multiple parameters. (g–m) Hierarchical update of multiple parameters. (d–f, k–m) The damping terms are the same for all of the the parameters. (a–c, h–j) The damping term associated with the Q_P-factor parameter is 100-times smaller than those used for the other parameters.

for the inaccuracies of the background models of the Thomsen parameters to fit the data. These artificial features are less visible when inverting the data in one go, namely, without time damping, because the optimization process seeks the best match for the whole data set.

Following the results of the synthetic experiment, two combinations of dampings λ_i are tested: the same damping is used for all of the parameters (Figs 16d–f and k–m) or the damping term of Q_P is 100-times smaller than the dampings for the velocity and density (Figs 16a–c and h–j).

We follow the same heuristic rule as for the synthetic experiment to define the suitable values of the dampings λ_i. Compared to the synthetic case, we note that the range of acceptable values of λ_{V_P} during the monoparameter inversion is much narrower than for the synthetic experiment. This different behaviour is probably due to the noise in the real data. The smallest acceptable values for the damping were of the order of 10^{-8}. It follows that the contribution of the diagonal blocks of the approximate Hessian used as an initial estimation, is negligible for each of the three parameters relative to the damping terms. We recall that this initial estimation of the

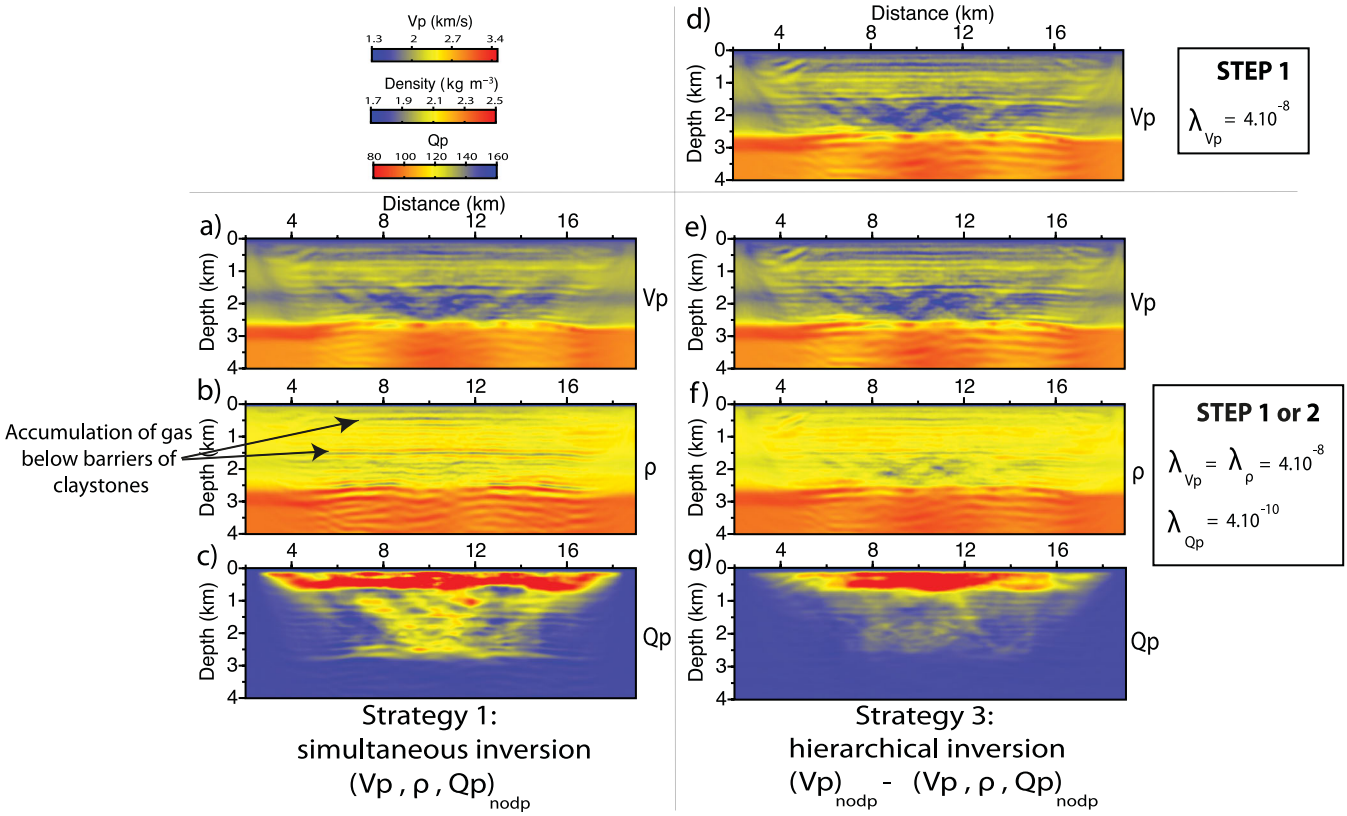

Figure 17. Real data case study. FWI models for vertical velocity (a, d, e), density (b, f), and Q_P-factor (c, g) for line 21, obtained by simultaneous (a–c) or hierarchical (d–g) update of the parameters. The damping term associated with the Q_P-factor parameter is 100-times smaller than those used for the other parameters.

Hessian is introduced by scaling the gradient provided to the L-BFGS by the sum of the approximate Hessian and of the damping parameter. At the first iteration, the approximate Hessian is thus not taken into account: the perturbation model is just a scaling of the gradient. At the next iterations however, the Hessian estimation realized by the L-BFGS is able to heterogeneously scale the gradient. To know the respective contribution of the diagonal blocks and of the off-diagonal blocks of the Hessian (estimated by the L-BFGS) to this scaling would deserve a more extensive study of the Hessian.

Of note, the data Hessian for the velocity, density, and Q_P-factor vary within the same proportion as for the synthetic case study. When the same damping is used for all of the parameters, the Q_P-factor model is left almost unchanged when *strategies 1* and *3* are used (Figs 16d–f and k–m). Moreover, the velocity model is noisier than that obtained with a small value of λ_{Q_P}. Therefore, we will focus on the interpretation of the results obtained with a small value of λ_{Q_P}.

FWI allows us to image several density contrasts along lines 21 and 29 (Fig. 16b, black arrows). In particular, we show a deep reflector at 3.5 km in depth, bending from the middle of the model to the bottom left, which has already been imaged by reverse time migration (Prieux *et al.* 2011, their fig. 8). The low-density contrasts at 0.6 and 1.5 km in depth are interpreted as an accumulation of gas below barriers of claystone. Note that these density contrasts are more pronounced for line 29 at around 7 and 11 km in distance, just above two low-velocity areas that are probably associated with gas-charged sediments (Fig. 16 h, white dashed circles). The FWI velocity model inferred from *strategy 1* is smoother than that inferred from *strategy 3*, which is balanced by more contrasted density and Q_P-factor models when *strategy 1* is used. This is consistent with the overall trend revealed by the synthetic experiment, and it highlights the trade-off between the parameters. The velocity models obtained after the first-step and second-step inversions are almost the same when *strategy 3* is used, which highlights the dominant weighting of the velocity in the inversion.

The Q_P-factor models show interesting features that are amenable to geological interpretations (Figs 16c and j), although values of the Q_P-factor should be interpreted with caution because of potential lack of resolution and trade-off with both velocity and density estimations. We show weak values of Q_P down to 600 m in depth in both models of lines 21 and 29, reaching values down to 50 (although the colour scale is limited to 80 in Figs 16 and 17). We interpret these attenuating zones as the signature of soft quaternary sediments. Small values of Q_P are also shown in the model of line 29 at distances of 8 and 11 km and at a depth of 2.5 km; these are the same locations where we observed smaller values of V_P (Fig. 16c, white circles). Line 21 is known to be located above tertiary sediments that are strongly charged with gas (Sirgue *et al.* 2010), as confirmed by the final velocity models of Figs 17(a) and (e). Consequently, the Q_P-factor has smaller values in the middle of the model of line 21, compared to the model of line 29.

The quality factor is much lower in the first 600 m layer than in the gas zone, even though we know that the intrinsic attenuation of the gas is very important. In addition to the lithological differences, this might arise because the inverse of the quality factor is strongly dependent on the effective pressure, which decreases by at least one order of magnitude between the ambient pressure and 40 MPa, which correspond to a burial depth of about 2000 m (Winkler & Nur 1982). However, according to the synthetic tests (Figs 11u and

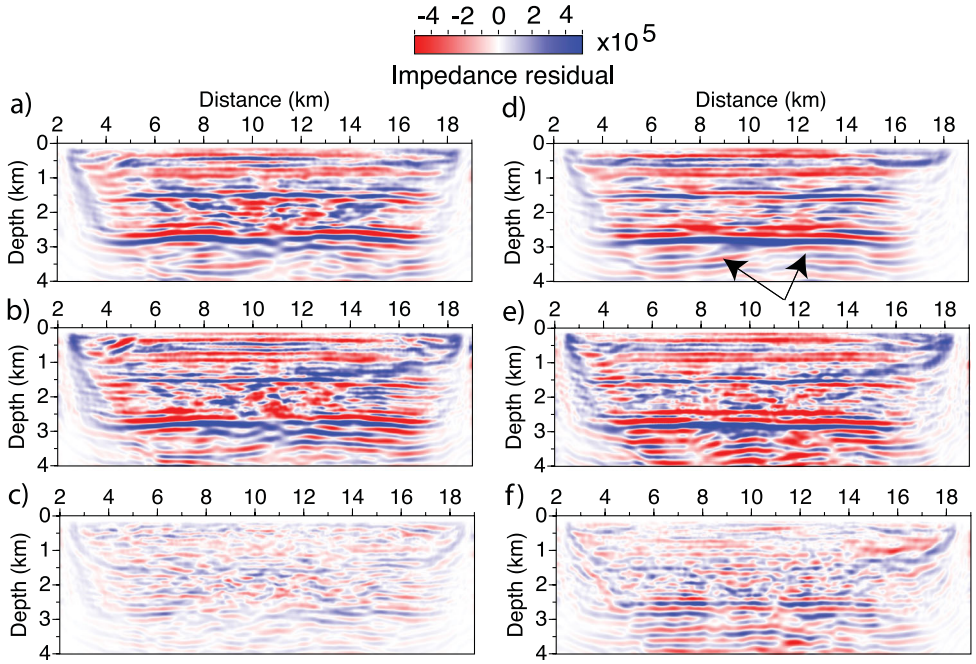

Figure 18. Real data case study. Impedance perturbation models for line 21 (a–c) and line 29 (d–f). (a, d) Difference obtained by subtracting the final model from the starting model of the simultaneous inversion. (b, c, e, f) For the hierarchical inversion, the perturbation models computed are distinguished by subtracting the final model of the second step from the starting model of the first step (b, e), and the perturbation models of the second inversion step (c, f). Black arrows in (d) point out the deep reflector at 3.5 km in depth.

12u for instance), the main reason is probably due to the intrinsic weak resolution of this parameter which dampens the Q values. Furthermore, the resolution strongly decreases with depth, because at depth the medium is only illuminated by reflections, which are not as sensitive as the diving waves to the attenuation as less wavelengths are propagated.

To enhance the information on the reflectivity provided by the reconstruction of the velocity and the density, we show the impedance perturbations for lines 21 and 29 and for the two strategies (Fig. 18). We note overall good lateral continuity of the reflectors. The simultaneous inversion mainly enhances two reflectors at 2.6 and 1.5 km in depth (Figs 18a and d). The amplitudes of these perturbations are greater than the amplitudes of the perturbations built by the second step of the hierarchical inversion alone (Figs 18c and f), especially for line 21, and they are of the same order of magnitude as the total perturbations of the hierarchical inversion (Fig. 18e). We conclude that the impedance starting model for the second inversion step of the hierarchical approach contains most of the high-wavenumber perturbations of the final impedance model. Although this strategy allows us to inject prior information, it might on the other hand drive the convergence towards a local minimum. We also note the impedance contrast associated with the deep reflector at 3.5 km in depth (Fig. 18b, black arrows), which was already shown in the FWI velocity model. The weaker impedance perturbations of line 21 compared to line 29 at the second inversion step can be related to the platform noise, which could hamper the convergence, or to the stronger presence of gas for line 21.

Comparison with sonic logs

To further appraise FWI results, we compare the FWI profiles with corresponding band-pass filtered well logs in the [0–14] Hz frequency range, the locations of which are indicated in Fig. 14. The match between the FWI and well logs needs to be assessed taking into account that most of these logs are not located on lines 21 and 29. Moreover, the logs are not rigorously vertical, as shown by the black line starting from the green square for the '2/8-F1' and '2/8-A-3B' wells in Fig. 14. At first glance, the distance between the density FWI models and the wells appears to be better for the first 1.5 km in depth, down to the top of the gas, than below this. This has to be related to the starting model, which roughly follow the same trend. On the other hand, the main discontinuities of the medium, which roughly follow the black dashed line in Fig. 19, and the few other layers seen in the "2/8-F1" well log down to 700 m in depth, are reconstructed by the FWI with an underestimated amplitude. These observations illustrate the notch in the lowest part of the wavenumber spectrum of the FWI density model, which correspond to a lack of sensitivity with respect to the density at wide apertures. The direct comparison is thus difficulty possible between the wells and FWI profiles. The lake of clear reflections in the gas is an additional reason why even the highest wavenumbers of the FWI profile are not compatible with those of the well log.

Data match

In Fig. 20, we show the direct comparisons between the recorded and the computed data within the final models obtained for the line 29 by simultaneous or hierarchical inversion. The source wavelet used for the seismic modelling is estimated in the normal moveout velocity model that was built by reflection traveltime tomography from offsets smaller than 2 km (Prieux *et al.* 2011). The long-offset refracted waves computed in the FWI model obtained after the second inversion step of the hierarchical approach arrive slightly earlier than the recorded arrivals, and amplitudes are weaker than those obtained after the first inversion step (Fig. 20a, ellipse). These two effects are probably related to the low values of Q_p created by the inversion down to 700 m in depth. The small changes in the velocities between the first and second steps are not significant enough to correct for this small phase shift (Fig. 20b, ellipse). In contrast, the amplitude and phase match of the diving waves in the models obtained by simultaneous inversion are slightly better for line 29 (Fig. 20c). This

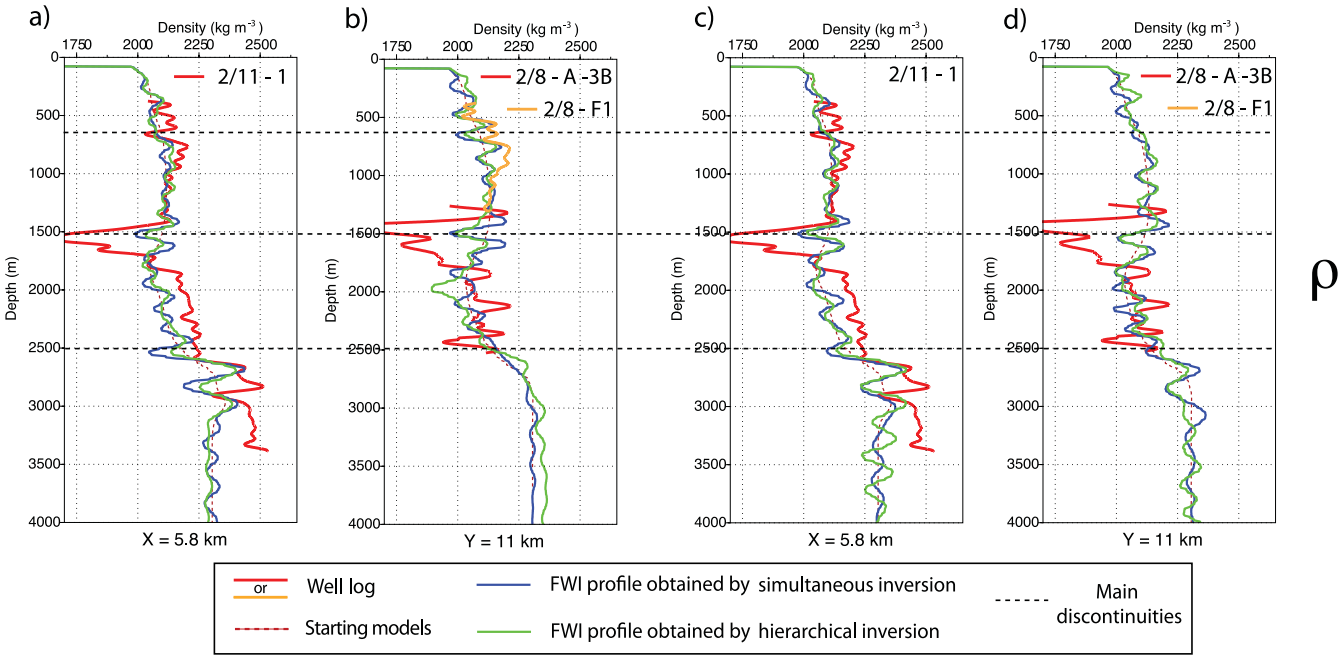

Figure 19. Real data case study. Comparison between the FWI profiles of the density for line 21 (a, b) and line 29 (c, d), and the corresponding well logs (red and orange) (see Fig. 14). The FWI profiles are extracted from the models obtained by simultaneous inversion (blue) or by hierarchical inversion (green).

improved match might be allowed by the simultaneous inversion update of the three parameter classes, which gives more freedom to the inversion to match the data. On the other hand, seismograms indicate that the second step inversion strengthens and improves the short-spread reflections, which also appear with stronger amplitude than in the seismograms from the simultaneous inversion (Figs 20a and d, ellipse).

The data match is also shown by the plot in Fig. 21 of the total misfit functions for line 29, for the tests where the damping term of Q_P is 100-times smaller than the dampings for velocity and density. We can compare the decrease of the misfit function for the simultaneous strategy (Fig. 21a) and the second step of the hierarchical strategy (Fig. 21b). We observe that the misfit decrease is relatively moderate (a mean of 15 per cent) and is equivalent for the two tests. It is interesting to note that the corresponding synthetic tests show an average misfit decrease of 30 per cent. We believe that this smaller decrease can be attributed to the presence of noise in the real data set, which gives more weight to the model space misfit function, to prevent the creation of sharp anomalies.

CONCLUSIONS

We have presented here an application of visco-acoustic FWI of hydrophone data for the imaging of the compressional wave speed, the density, and the Q_P-factor in marine environments. We first show that a subsurface parametrization that combines compressional wave speed and density leads to better FWI results than parametrization, which combines compressional wave speed and impedance. In the first case, the diffraction pattern of the compressional wave speed is isotropic, which leads to a high resolution reconstruction of the velocity, the parameter with the dominant imprint on the data. A second recommendation is that the velocity and density are updated in a hierarchical manner, rather than simultaneously. This hierarchical approach contributes to the update of the density during the second step of the inversion in a more stable manner than the simultaneous update of the velocity and the density in one go. We have found that judicious scaling of the model parameters and judicious tuning of the FWI regularization were key issues to retrieve reliable model perturbations of the dominant and secondary parameters. In the case of the imaging of the velocity, density, and Q_P-factor, we also recommend from this case study a hierarchical approach, where the dominant wave speed parameter is reconstructed first, before the simultaneous update of the three parameter classes. These conclusions are probably still valid for other marine case studies, but should differ for land environments, where the footprint of the attenuation and the shear wave velocity is much stronger. When numerous inversion models are recovered from different choices of damping, inversion strategies, and parametrizations, their mean, which can be weighted by prior information, can be of primary importance to better constrain the most probable solutions within the model space. The update of the *P*-wave velocity, density and Q_P-factor by visco-acoustic FWI from the hydrophone data recorded at the Valhall field provide reliable starting and background models for the subsequent elastic FWI for imaging compressional and shear wave speeds.

ACKNOWLEDGEMENTS

The authors are grateful to the two anonymous reviewers and to the associate editor for their comments that have helped us to improve the manuscript. This study was funded by the SEISCOPE consortium (http://seiscope.oca.eu), and sponsored by BP, CGG-VERITAS, ENI, EXXON-MOBIL, PETROBRAS, SAUDI ARAMCO, SHELL, STATOIL and TOTAL. The linear systems were solved with the MUMPS package, available on http://graal.ens-lyon.fr/MUMPS/index.html and http://mumps.enseeiht.fr. The mesh generation was performed with help of TRIANGLE, available on http://www.cs.cmu.edu/~quake/triangle.html. This study was granted access to the high-performance computing facilities of the SIGAMM (Observatoire de la Côte d'Azur) and of CINES/IDRIS under

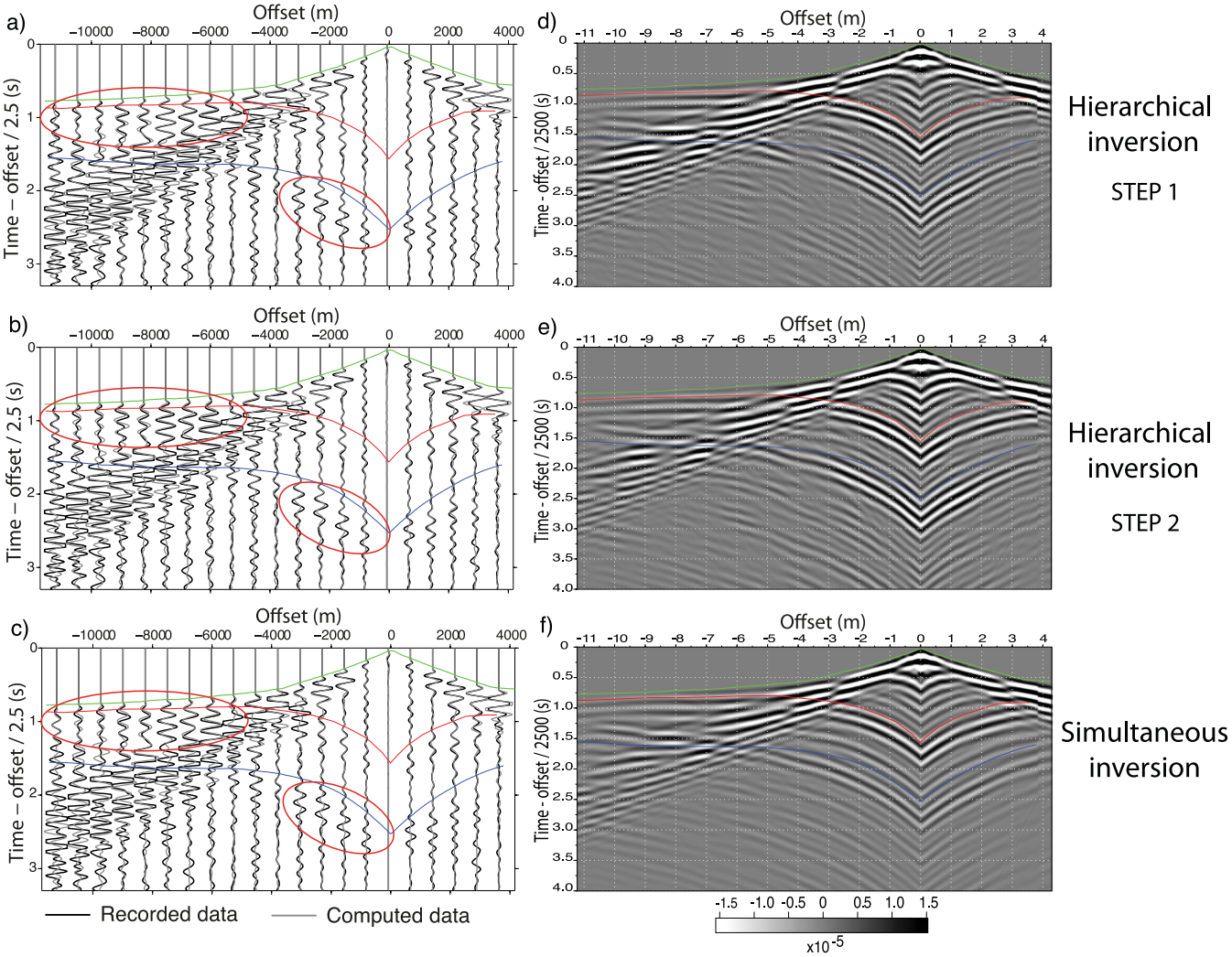

Figure 20. Real data case study. Seismograms represented with a reduction velocity of 2.5 km s^{-1}. (a–c) True amplitude comparison between recorded (black line) and modelled (gray line) seismograms. A gain with offset is used to represent these data. (a, b) Seismograms are computed in the final FWI models of the first (a) and second (b) steps of the hierarchical inversion (see Figs 16g–j). (c) Same as (a) for the seismograms modelled in the final FWI models of the simultaneous inversion (Figs 16a–c). (d–f) Modelled data shown in (a–c) but displayed with the black/white scale. Green curves, manually picked first arrival traveltimes; red curves, reflections from the top of the gas; blue curves, reflections from the bottom of the gas.

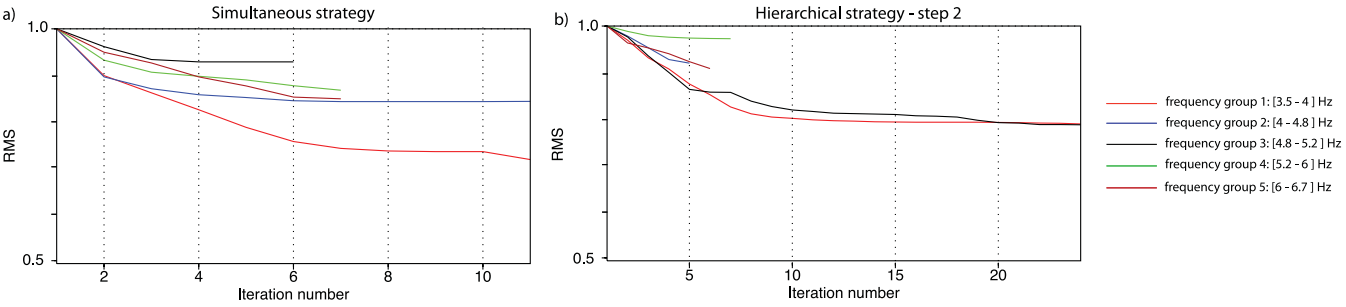

Figure 21. Normalized misfist functions with respect to the iterations for the five groups of frequencies (for identification, see right-hand side), from the tests realized on line 29 when the damping term associated with the Q_P factor was 100-times smaller than those used for the other parameters. (a) Simultaneous strategy (see Figs 16h–j) and (b) second step of the hierarchical strategy (see Figs 17e–g).

allocation 2010 [project gao2280] of GENCI (Grand Equipement National de Calcul Intensif), and we gratefully acknowledge both of these facilities and the support of their staff. We also thank BP Norge AS and their Valhall partner Hess Norge AS, for allowing access to the Valhall data set as well as the well-log velocities.

REFERENCES

Aki, K. & P., Richards, 1980. *Quantitative Seismology: Theory and Methods*, W. H. Freeman & Co., New York.

Askan, A., Akcelik, V., Bielak, J. & Ghattas, O., 2007. Full waveform inversion for seismic velocity and anelastic losses in heterogeneous structures, *Bull. seism. Soc. Am.*, **97,** 1990–2008.

Best, A.I., McCann, C. & Sothcott, J., 1994. The relationships between the velocities, attenuations and petrophysical properties of reservoir sedimentary rocks, *Geophys. Prospect.,* **42,** 151–178.

Brenders, A.J. & Pratt, R.G., 2007. Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model, *Geophys. J. Int.,* **168,** 133–151.

Brossier, R., 2011. Two-dimensional frequency-domain visco-elastic full waveform inversion: parallel algorithms, optimization and performance, *Comput. Geosci.,* **37,** 444–455.

Brossier, R., Etienne, V., Operto, S. & Virieux, J., 2010a. Frequency-domain numerical modelling of visco-acoustic waves based on finite-difference and finite-element discontinuous galerkin methods, in *Acoustic Waves,* pp. 125–158, ed. Dissanayake, D.W., SCIYO, Rijeka, Croatia.

Brossier, R., Gholami, Y., Virieux, J. & Operto, S., 2010b. 2D frequency-domain seismic wave modeling in VTI media based on a hp-adaptive discontinuous galerkin method, *in Proceedings of the 72nd Annual International Meeting,* EAGE, Barcelona, Expanded Abstracts, C046.

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

Brossier, R., Operto, S. & Virieux, J., 2010c. Which data residual norm for robust elastic frequency-domain full waveform inversion?, *Geophysics,* **75,** R37–R46.

Castagna, J.P., Batzle, M.L. & Kan, T.K., 1993. Rock physics—the link between rock properties and AVO response, in *Offset-Dependent Reflectivity—Theory and Practice of AVO Analysis,* pp. 135–171, eds Castagna, J.P. & Backus, M.M., SEG, Tulsa, OK.

Causse, E., Mittet, R. & Ursin, B., 1999. Preconditioning of full-waveform inversion in viscoacoustic media, *Geophysics,* **64,** 130–145.

Charara, M., Barnes, C. & Tarantola, A., 1996. The state of affairs in inversion of seismic data: an OVSP example, *in Proceedings of the SEG Meeting,* Expanded Abstracts, 1999–2002, SEG.

Chavent, G., 2009. *Nonlinear Least Squares for Inverse Problems,* Springer, Dordrecht/Heidelberg/London/New York.

Dasgupta, R. & Clark, R., 1998. Estimation of q from surface seismic reflection data, *Geophysics,* **63,** 2120–2128.

Devaney, A.J., 1982. A filtered backprojection algorithm for diffraction tomography, *Ultrasonic Imaging,* **4,** 336–350.

Forgues, E. & Lambaré, G., 1997. Parameterization study for acoustic and elastic ray+born inversion, *J. Seism. Explor.,* **6,** 253–278.

Gardner, G.H.F., Gardner, L.W. & Gregory, A.R., 1974. Formation velocity and density—the diagnostic basics for stratigraphic traps, *Geophysics,* **39,** 770–780.

Gholami, Y., Brossier, R., Operto, S., Prieux, V., Ribodetti, A. & Virieux, J., 2013a. Which parametrization is suitable for acoustic VTI full waveform inversion?—Part 2: application to Valhall, *Geophysics,* **2,** R107–R124.

Gholami, Y., Brossier, R., Operto, S., Ribodetti, A. & Virieux, J., 2013b, Which parametrization is suitable for acoustic VTI full waveform inversion?—Part 1: sensitivity and trade-off analysis, *Geophysics,* **2,** R81–R105.

Gouveia, W.P. & Scales, J.A., 1997. Resolution of seismic waveform inversion: Bayes versus OCCAM, *Inverse Problems,* **13,** 323–349.

Gouveia, W.P. & Scales, J.A., 1998. Bayesian seismic waveform inversion: parameter estimation and uncertainty analysis, *J. geophys. Res.,* **103,** 2759–2779.

Greenhalgh, S., Zhou, B. & Green, A., 2006. Solutions, algorithms and interrelations for local minimization search geophysical inversion, *J. geophys. Eng.,* **3,** 101–113.

Hak, B. & Mulder, W.A., 2011. Seismic attenuation imaging with causality, *Geophys. J. Int.,* **184,** 439–451.

Hansen, P.C., 1992. Analysis of discrete ill-posed problems by means of the L-curve, *SIAM Rev.,* **34,** 561–580.

Hicks, G.J. & Pratt, R.G., 2001. Reflection waveform inversion using local descent methods: estimating attenuation and velocity over a gas-sand deposit, *Geophysics,* **66,** 598–612.

Jannane, M. *et al.*, 1989. Wavelengths of Earth structures that can be resolved from seismic reflection data, *Geophysics,* **54,** 906–910.

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

Kamei, R. & Pratt, R.G., 2008. Waveform tomography strategies for imaging attenuation structure for cross-hole data, *in Proceedings of the 70th Annual International Meeting,* EAGE, Expanded Abstracts, F019.

Klimentos, T. & McCann, C., 1990. Relationships among compressional wave attenuation, porosity, clay content, and permeability in sandstones, *Geophysics,* **55,** 998–1014.

Kolsky, H., 1956. The propagation of stress pulses in viscoelastic solids, *Philos. Mag.,* **1,** 693–710.

Liao, O. & McMechan, G.A., 1995. 2.5D full-wavefield viscoacoustic inversion, *Geophys. Prospect.,* **43,** 1043–1059.

Malinowski, M., Operto, S. & Ribodetti, A., 2011. High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full waveform inversion, *Geophys. J. Int.,* **186,** 1179–1204.

Mavko, G., Mukerji, T. & Dvorkin, J., 2009. *The Rocks Physics Handbooks, Tools for Seismic Analysis in Porous Media,* 2nd ed, Cambridge University Press, Cambridge, UK.

Métivier, L., Lailly, P., Delprat-Jannaud, F. & Halpern, L., 2011. A 2D nonlinear inversion of well-seismic data, *Inverse Problems,* **27,** doi:10.1088/0266-5611/27/5/055005.

Mora, P.R., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data, *Geophysics,* **52,** 1211–1228.

Moritz, H. & Sunkel, H., 1978. *Approximation Methods in Geodesy,* H. Wichmann, Karlsruhe.

Mulder, W.A. & Hak, B., 2009. An ambiguity in attenuation scattering imaging, *Geophys. J. Int.,* **178,** 1614–1624.

Neves, F.A. & Singh, S.C., 1996. Sensitivity study of seismic reflection/refraction data, *Geophys. J. Int.,* **126,** 470–476.

Nocedal, J., 1980. Updating Quasi-Newton Matrices With Limited Storage, *Math.Comput.,* **35,** 773–782.

Nocedal, J. & Wright, S.J., 1999. *Numerical Optimization,* Springer, New York.

Operto, S., Virieux, J., Dessa, J.X. & Pascal, G., 2006. Crustal imaging from multifold ocean bottom seismometers data by frequency-domain full-waveform tomography: application to the eastern Nankai trough, *J. geophys. Res.,* **111,** doi:10.1029/2005JB003835.

Plessix, R.E., 2006a. Estimation of velocity and attenuation coefficient maps from crosswell seismic data, *Geophysics,* **71,** S235–S240.

Plessix, R.E., 2006b. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.,* **167,** 495–503.

Plessix, R.E. & Cao, Q., 2011. A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium, *Geophys. J. Int.,* **185,** 539–556.

Plessix, R.E. & Perkins, C., 2010. Full waveform inversion of a deep water ocean bottom seismometer dataset, *First Break,* **28,** 71–78.

Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain, part I: theory and verification in a physic scale model, *Geophysics,* **64,** 888–901.

Pratt, R.G., Hou, F., Bauer, K. & Weber, M., 2005. Waveform tomography images of velocity and inelastic attenuation from the Mallik 2002 crosshole seismic surveys, in *Scientific Results from the Mallik 2002 Gas Hydrate Production Research Well Program*, pp. 1–14, eds Dallimor, S.R. & Collett, T.S., Mackenzie Delta, Northwest Territories, Canada. Geological Survey of Canada.

Pratt, R.G., Shin, C. & Hicks, G.J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion, *Geophys. J. Int.,* **133,** 341–362.

Pratt, R.G. & Worthington, M.H., 1990. Inverse theory applied to multisource cross-hole tomography. Part I: acoustic wave-equation method: *Geophys. Prospect.,* **38,** 287–310.

Prieux, V., Brossier, R., Gholami, Y., Operto, S., Virieux, J., Barkved, O. & Kommedal, J., 2011. On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study, *Geophys. J. Int.,* **187,** 1495–1515.

Quan, Y. & Harris, J.M., 1997. Seismic attenuation tomography using the frequency shift method, *Geophysics,* **62,** 895–905.

Rao, Y. & Wang, Y., 2008. The strategies for attenuation inversion with waveform tomography, *Presented at the 70th Annual Conference,* EAGE, Roma.

Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A. & dell'Aversana, P., 2004. Multi-scale imaging of complex structures from multi-fold wide-aperture seismic data by frequency-domain full-wavefield inversions: application to a thrust belt: *Geophys. J. Int.,* **159,** 1032–1056.

Ribodetti, A., Operto, S., Virieux, J., Lambaré, G., Valéro, H.-P. & Gibert, D., 2000. Asymptotic viscoacoustic diffraction tomography of ultrasonic laboratory data: a tool for rock properties analysis, *Geophys. J. Int.,* **140,** 324–340.

Ribodetti, A. & Virieux, J., 1996. Asymptotic theory for imaging the attenuation factors Q_p and Q_s, *in Proceedings of the INRIA, Aix-les-Bains,* France, Expanded Abstracts, pp. 334–353, Springer-Verlag, Berlin, Heidelberg.

Ribodetti, A. & Virieux, J., 1998. Asymptotic theory for imaging the attenuation factor Q, *Geophysics,* **64,** 1–12.

Romero, A.E., McEvilly, T.V. & Majer, E.L., 1997. 3-D microearthquake attenuation tomography at the northwest geysers geothermal region, california, *Geophysics,* **62,** 149–167.

Sears, T., Singh, S. & Barton, P., 2008. Elastic full waveform inversion of multi-component OBC seismic data, *Geophys. Prospect.,* **56,** 843–862.

Shi, Y., Zhao, W. & Cao, H., 2007. Nonlinear process control of wave-equation inversion and its application in the detection of gas, *Geophysics,* **72**(1), R9–R18.

Shin, C. & Cha, Y.H., 2009. Waveform inversion in the Laplace-Fourier domain: *Geophys. J. Int.,* **177,** 1067–1079.

Sirgue, L., Barkved, O.I., Dellinger, J., Etgen, J., Albertin, U. & Kommedal, J.H., 2010. Full waveform inversion: the next leap forward in imaging at Valhall, *First Break,* **28,** 65–70.

Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics,* **69,** 231–248.

Smithyman, B., Pratt, R.G., Hayles, J. & Wittebolle, R., 2009. Detecting near-surface objects with seismic waveform tomography, *Geophysics,* **74,** WCC119–WCC127.

Song, Z., Williamson, P. & Pratt, G., 1995. Frequency-domain acoustic-wave modeling and inversion of crosshole data. Part 2: inversion method, synthetic experiments and real-data results, *Geophysics,* **60,** 786–809.

Tarantola, A., 1984. Linearized inversion of seismic reflection data, *Geophys. Prospect.,* **32,** 998–1015.

Tarantola, A., 1986. A strategy for non linear inversion of seismic reflection data, *Geophysics,* **51,** 1893–1903.

Tarantola, A., 1987. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation,* Elsevier, New York.

Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms including elasticity and attenuation, *Pure appl. Geophysics,* **128,** 365–399.

Thomsen, L.A., Barkved, O., Haggard, B., Kommedal, J.H. & Rosland, B., 1997. Converted wave imaging of Valhall reservoir, *Presented at the 59th EAGE Conference & Exhibition,* Geneva.

Toksöz, M.N. & Johnston, D.H., 1981. *Seismic Wave Attenuation,* Geophysics Reprint Series, No. 2, Society of Exploration Geophysicists, Tulsa, OK.

Tonn, R., 1991. The determination of seismic quality factor Q from VSP data: a comparison of different computational methods, *Geophys. Prospect.,* **39,** 1–27.

Toverud, T. & Ursin, B., 2005. Comparison of seismic attenuation models using zero-offset vertical seismic profiling (VSP) data, *Geophysics,* **70,** F17–F25.

Vigh, D., Kapoor, J., Moldoveanu, N. & Li, H., 2011. Breakthrough acquisition and technologies for subsalt imaging, *Geophysics,* **76,** WB41–WB51.

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

Wang, Y., 2008. *Seismic Inverse Q Filtering,* Wiley-Blackwell, Hoboken, New Jersey.

Winkler, K. & Nur, A., 1982. Seismic attenuation: effects of pore fluids and frictional sliding, *Geophysics,* **47,** 1–15.

Wu, R.S. & Aki, K., 1985. Scattering characteristics of elastic waves by an elastic heterogeneity, *Geophysics,* **50,** 582–595.

Wu, R.S. & Toksöz, M.N., 1987. Diffraction tomography and multisource holography applied to seismic imaging, *Geophysics,* **52,** 11–25.