High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full-waveform inversion

M. Malinowski,[1] S. Operto[2] and A. Ribodetti[2]

[1]*Institute of Geophysics, Polish Academy of Sciences, Ks. Janusza 64, 01-452 Warsaw, Poland. E-mail: michalm@igf.edu.pl*
[2]*UMR GéoAzur, CNRS-UNSA, 06235 Villefranche-sur-mer, France*

SUMMARY

Here we assess the potential of the visco-acoustic frequency domain full-waveform inversion (FWI) to reconstruct P-wave velocity (V_P) and P-wave attenuation factor (Q) from surface onshore seismic data. First, we perform a sensitivity analysis of the FWI based upon a grid search analysis of the misfit function and several synthetic FWI examples using velocity and Q models of increasing complexity. Subsequently, we applied both the acoustic and visco-acoustic FWI to real surface wide-aperture onshore seismic data from the Polish Basin, where a strong attenuation of the seismic data is observed. The sensitivity analysis of the visco-acoustic FWI suggests that the FWI can jointly reconstruct the velocity and the attenuation factor if the signature of the attenuation is sufficiently strong in the data. A synthetic example corresponding to a homogeneous background model with an inclusion shows a reliable reconstruction of V_P and Q in the inclusion, when Q is as small as 90 and 50 in the background model and in the inclusion, respectively. A first application of acoustic FWI to real data shows that a heuristic normalization of the data with offset allows us to compensate for the effect of the attenuation in the data and reconstruct a reliable velocity model. Alternatively, we show that visco-acoustic FWI allows us to reconstruct jointly both a reliable velocity model and a Q model from the true-amplitude data. We propose a pragmatical approach based upon seismic modelling and source wavelet estimation to infer the best starting homogeneous Q model for visco-acoustic FWI. We find the source wavelet estimation quite sensitive to the quality of the velocity and attenuation models used for the estimation. For example, source-to-source wavelets are significantly more consistent when computed in the final FWI model than in the initial one. A good kinematic and amplitude match between the early-arriving phases of the real and time-domain synthetic seismograms computed in the final FWI model provides an additional evidence of the reliability of the final FWI model. We find the recovered velocity and attenuation models consistent with the expected lithology and stratigraphy in the study area. We link high-attenuation zones with the increased clay content and the presence of the mineralized fluids.

Key words: Numerical solutions; Inverse theory; Controlled source seismology; Seismic attenuation; Seismic tomography.

GJI Seismology

1 INTRODUCTION

In recent years there has been an increasing interest in the use of full waveform inversion (FWI) for imaging nuclear waste deposit sites, monitoring CO_2 sequestration, assessing natural hazards and characterising geothermal and oil and gas reservoirs. Since the pioneering work of Tarantola (1984, 1987), FWI has been intensively developed both in the time domain and in the frequency domain (Virieux & Operto 2009, for a recent review). FWI attempts to exploit the full information content of the data by minimising the misfit between the recorded and the modelled wavefields. To achieve this goal, the full solution of the (two-way) wave equation is computed during the forward problem with numerical methods such as finite-difference (FD) or finite-element methods (Marfurt 1984). In 2-D and 3-D, the inverse problem is solved with iterative local optimization methods such as gradient-like methods to manage the computational burden of the multisource forward problem and the high-dimensionality of the model space. Iterations are performed in a nonlinear way, where the updated model at each iteration is used as the initial model for the next iteration. If suitable

acquisition geometry allows for recording of both short- and wide-aperture arrivals, FWI combines transmission-like imaging with migration-like imaging in a single algorithm to build quantitative broadband model of one or several physical parameters of the sub-surface (*P* and *S* wave speeds, density, attenuation and anisotropy Pratt *et al.* 1996).

FWI has been developed both in the time domain and in the frequency domain. Frequency-domain FWI (Pratt *et al.* 1998) has shown distinct advantages over time-domain FWI (Tarantola 1984), as it introduces a natural multiscale approach by proceeding hierarchically from the low frequencies to the higher ones. Another advantage of frequency-domain FWI is the ability to manage and process compact volumes of data by limiting the inversion to a limited number of frequency components of wide-aperture/wide-azimuth data (Sirgue & Pratt 2004). A third key advantage of frequency-domain inversion in the framework of this study is to allow for a straightforward implementation of attenuation in the forward problem by means of complex-valued velocities (Toksöz & Johnston 1981) where the real part of the velocity mainly describes the propagative properties of the subsurface, and the imaginary part of the velocity describes its diffusive properties. This straightforward implementation of attenuation in the forward problem can be exploited in FWI to perform the reconstruction of both the real and imaginary part of the complex velocity without extra computational cost (Liao & McMechan 1995; Song & Williamson 1995; Song *et al.* 1995; Liao & McMechan 1996).

Several 2-D and 3-D applications of FWI to real surface data have been reported recently in both offshore (e.g. Shipp & Singh 2002; Operto *et al.* 2006; Shi *et al.* 2007; Plessix 2009; Sirgue *et al.* 2010) and onshore (e.g. Ravaut *et al.* 2004; Bleibinhaus *et al.* 2007; Malinowski & Operto 2008; Jaiswal *et al.* 2009) environments. However, most of them have been performed under the acoustic approximation of the wave equation for the reconstruction of the *P*-wave velocity. Beyond computational aspects, inaccuracy of the amplitude modelling performed with the acoustic approximation can be one reason why FWI is generally limited to quite low frequencies, typically, smaller than 10 Hz. Hence, one crucial point for the success of this methodology, is to evolve beyond the acoustic approximation to take into account more realistic physical models of the subsurface. Such models have to include absorption and dispersion effects. During the last decade, the interest in the attenuation of seismic waves has increased, and was motivated by the improvements in the characterization and monitoring of hydrocarbon reservoirs and, more generally, of the earth models which can be obtained when accounting for the anelasticity of rocks (e.g. Wang 2008). These improvements concern the lithologic description of the earth, the analysis of its physical state and the degree of fluid saturation in reservoir rocks.

The objective of this study is to provide new insights on the ability of frequency-domain FWI to jointly reconstruct the *P*-wave velocity and the attenuation factor *Q* in the framework of the visco-acoustic approximation. One key issue associated with the reconstruction of multiple classes of parameters by FWI is related to the increased ill-posedness in term of non-uniqueness of the solution introduced by the multiple parameter classes. It is therefore crucial to assess the sensitivity of the data to each classes of parameters and the trade-off that may exist between them. Mulder & Hak (2009) have concluded that the joint reconstruction of velocity and attenuation was almost impossible from short-aperture reflection data by linear migration/inversion because velocity and attenuation perturbations are related by a Hilbert transform. Therefore, many combinations of velocity and attenuation models will lead to the same data. Ribodetti

et al. (2000) develop visco-acoustic ray+Born migration/inversion and concluded that velocity and attenuation can be reliably reconstructed only when a reflector is illuminated by reflected waves from above and beneath. If this condition is not satisfied, the asymptotic Hessian is singular. Hak & Mulder (2011) showed however that using an attenuation model with a causality correction term allows to successfully reconstruct the attenuation by non-linear FWI, provided a sufficiently high number of iterations is performed.

Very few attempts have been performed so far to reconstruct attenuation from synthetic and real data by FWI. An application of visco-acoustic time-domain FWI to a realistic synthetic data set at a crustal scale has been presented by Askan *et al.* (2007), while Liao & McMechan (1995, 1996) have presented the first attempts to reconstruct velocity and attenuation by frequency-domain FWI of synthetic data. Most of the applications to real data have been presented for a cross-hole acquisition (Song *et al.* 1995; Pratt *et al.* 2005; Kamei & Pratt 2008; Rao & Wang 2008). Fewer applications to real surface data exist, such as Hicks & Pratt (2001) for a streamer data set and Smithyman *et al.* (2009) for near surface onshore data.

In this study, we present one of the first application of visco-acoustic frequency-domain FWI to real onshore surface seismic data that were recorded in the Polish Basin. This case study follows the application presented by Malinowski & Operto (2008), that was limited to the acoustic FWI for velocity reconstruction. After a short review of our implementation of visco-acoustic frequency-domain FWI, we first present a sensitivity and trade-off analysis of visco-acoustic FWI based upon the analysis of the partial derivative wavefields with respect to velocity and attenuation and a grid-search analysis of the misfit function performed for a simple visco-acoustic inclusion model. Application of visco-acoustic FWI to synthetic models of increasing complexities supports a view that the velocity and attenuation can be reliably reconstructed from surface wide-aperture seismic data by non-linear inversion. For the application to real data, we first show how attenuation effects can be heuristically removed from the data before performing acoustic FWI for *P*-wave velocity. Alternatively, amplitude variations can be preserved in the data to perform visco-acoustic FWI for *P*-wave velocity and attenuation. In the final section, the relevance of the velocity and attenuation models are discussed based upon checkerboard tests, source wavelet estimation, seismic modelling and geological and petrophysical interpretations.

2 FREQUENCY-DOMAIN VISCO-ACOUSTIC FULL-WAVEFORM INVERSION

2.1 Forward modelling

For seismic modelling, we consider the 2-D visco-acoustic second-order wave equation in the frequency-domain.

$$\frac{\omega^2}{\kappa(\mathbf{x})}p(\mathbf{x},\omega) + \frac{\partial}{\partial x}b(\mathbf{x})\frac{\partial p(\mathbf{x},\omega)}{\partial x} + \frac{\partial}{\partial z}b(\mathbf{x})\frac{\partial p(\mathbf{x},\omega)}{\partial z} = s(\mathbf{x},\omega), \quad (1)$$

where *p* is the pressure wavefield and *s* is the source. The angular frequency is denoted by ω, the buoyancy by *b* (the inverse of the density ρ) and the bulk modulus by κ with $\kappa = \rho c^2$, where *c* (or V_P) denotes the wave speed of compressional waves. Attenuation is classically implemented with complex-valued wave speeds (Toksöz & Johnston 1981). In this study, we use for the complex-valued

wave speed $\tilde{c}$

$$\tilde{c} = c\left[1 - i\,\text{sign}(\omega)\frac{1}{2Q}\right], \tag{2}$$

where Q is the attenuation factor and $i = \sqrt{-1}$.

In order to solve numerically eq. (1), we used a FD frequency-domain method based upon the parsimonious staggered-grid mixed-grid stencil described in Hustedt *et al.* (2004) with perfectly matched layer (PML) absorbing boundary conditions (Berenger 1994). To match land data recorded on the surface, we need to infer from the pressure wavefield computed with eq. (1) the vertical particle velocities recorded by vertical geophones on the surface. To perform the pressure to particle velocity conversion at the receiver positions, we use the *sinc* interpolation method proposed by Hicks (2002) to compute the vertical derivative of the pressure wavefield on the surface: instead of using an interpolating *sinc* function to approximate a spatial Dirac delta function and compute the pressure at the receiver position, we use the vertical derivative of the *sinc* function to compute the vertical derivative of the pressure at the receiver position. For the FWI application presented in this study, we do not make any attempt to model accurately the free-surface effects by cancelling the pressure on the surface. Instead, we set a layer above the topography with a constant velocity of 1600 m s^{-1} in the manner of Bleibinhaus & Rondenay (2009).

Eq. (1) can be recast in matrix form as

$$\mathbf{A}(\mathbf{m})\mathbf{p} = \mathbf{s}, \tag{3}$$

where the coefficients of the sparse impedance matrix $\mathbf{A}$ depend on the frequency ω and on the medium properties b and κ.

2.2 Full-waveform inversion algorithm

We review the main features of the frequency-domain FWI algorithm that will be applied in this study. A complementary description of the algorithm can be found in Sourbier *et al.* (2009a,b). The reader is also referred to Pratt *et al.* (1998), Pratt (1999), Sirgue & Pratt (2004) and Virieux & Operto (2009) for a more exhaustive overview of theoretical and practical aspects of efficient frequency-domain FWI.

The weighted least-squares misfit function that is minimized is given by

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2}\Delta\mathbf{d}^{\dagger}\mathbf{W}_d\Delta\mathbf{d}, \tag{4}$$

where $\Delta\mathbf{d}$ is the data residual vector (the difference between the recorded and the modelled data) associated with each source–receiver couple and each simultaneously inverted frequency. In this study, $\mathbf{W}_d$ is a diagonal weighting operator, which applies a gain with the source–receiver offset to the data misfit vector. The symbol $\dagger$ denotes the complex conjugate operator. We seek to minimize the misfit function in the vicinity of a starting model $\mathbf{m}^{(k)}$ by iterative nonlinear local optimization, where k is the iteration number. The model vector $\mathbf{m}$ is parametrized by the wave speed c and the attenuation factor Q at each node of the FD grid, $\mathbf{m} = (c_1, Q_1, c_2, Q_2, \ldots, c_N, Q_N)$, where N is the number of nodes in the FD grid. The updated model $\mathbf{m}^{(k+1)}$ is related to the starting model $\mathbf{m}^{(k)}$ and the perturbation model $\Delta\mathbf{m}^{(k)}$ by

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \Delta\mathbf{m}^{(k)}. \tag{5}$$

In this study, we use either a preconditioned steepest-descent algorithm or a quasi-Newton L-BFGS algorithm for optimization

(Tarantola 1987; Nocedal & Wright 1999). In both cases, the model perturbation can be written as

$$\Delta\mathbf{m}^{(k)} = -\alpha^{(k)}\mathcal{H}^{(k)}\gamma^{(k)}, \tag{6}$$

where α is the step length, $\mathcal{H}^{(k)}$ is an approximation of the curvature operator, and $\gamma^{(k)}$ is the ascent direction (Tarantola 1987). In this study, γ is related to the gradient of the misfit function by

$$\gamma = \mathbf{C}_m\nabla\mathcal{C}, \tag{7}$$

where $\mathbf{C}_m$ is a smoothing operator (the equivalent to the covariance operator in the model space) which guarantees a desired level of smoothness in the reconstructed model (Tarantola 1987, p. 219). It is implemented in the form of a 2-D Gaussian function parametrized by vertical and horizontal correlation lengths for velocity and attenuation parameters. In our algorithm, the correlation lengths are defined for each modelled frequency as a fraction of the mean propagated wavelength (Ravaut *et al.* 2004).

The gradient of the misfit function for the model parameter m_j is computed efficiently with the adjoint-state method (Chavent 1974; Tarantola 1984; Plessix 2006; Chavent 2009), which gives

$$\begin{aligned}
\nabla\mathcal{C}_{m_j} &= \sum_{k=1}^{N_\omega}\sum_{l=1}^{N_s}\mathbf{p}_{l,k}^T\left(\frac{\partial\mathbf{A}_k}{\partial m_j}\right)^T\mathbf{A}_k^{-1}\mathbf{W}_d\Delta\mathbf{d}_{l,k}^* \\
&= \sum_{k=1}^{N_\omega}\sum_{l=1}^{N_s}\mathbf{p}_{l,k}^T\left(\frac{\partial\mathbf{A}_k}{\partial m_j}\right)^T\mathbf{r}_{l,k}^*,
\end{aligned} \tag{8}$$

where * denotes the complex conjugate, ^{T} the transpose operator, N_s the number of sources, and N_ω the number of frequencies simultaneously inverted. The monochromatic incident wavefield associated with frequency k and source l is denoted by $\mathbf{p}_{l,k}$. The so-called adjoint wavefield denoted by $\mathbf{r}_{l,k}$ corresponds to the conjugate of the wavefield backpropagated from the receiver positions using the weighted data residuals as the source term. Note that, in the framework of land data recorded by vertical geophones, the sources of the adjoint wavefield are not explosions as for hydrophones but vertical forces. To model vertical forces on the surface from the second-order acoustic wave equation in pressure, eq. (1), we use the *sinc* interpolation method of Hicks (2002) in the same manner as for the computation of the vertical particle velocity at the receiver positions from the pressure wavefield.

In the framework of generalized diffraction tomography, the sparse matrix $\frac{\partial\mathbf{A}_k}{\partial m_j}$ can be interpreted as the radiation pattern of the virtual secondary source located at position m_j which generates the partial derivative wavefield $\frac{\partial\mathbf{p}}{\partial m_j}$ (Pratt *et al.* 1998, their eqs 15 and 16).

In the preconditioned steepest-descent algorithm, we use the inverse of the diagonal terms of the approximate Hessian for the preconditioning operator $\mathcal{H}$ (e.g. Ravaut *et al.* 2004; Operto *et al.* 2006).

$$\mathcal{H} \approx \left[\text{Diag}(\mathbf{J}^T W_d\mathbf{J}) + \lambda\left\{\text{Diag}(\mathbf{J}^T W_d\mathbf{J})\right\}_{\text{max}}\right]^{-1}, \tag{9}$$

where $\mathbf{J}$ is the Fréchet derivative or the sensitivity matrix, which must be built explicitly if the diagonal terms of the approximate Hessian are needed. The scalar λ is a damping or prewhitening factor which prevent instabilities, and which is estimated by trial-and-error (Ravaut *et al.* 2004).

The element of the Fréchet derivative matrix associated with the frequency k, the source–receiver pair (l–r) and the parameter m_j is given by (see also Virieux & Operto 2009, their eq. 26)

$$\mathbf{J}_{k,l,r;j} = \mathbf{p}_{l,k}^T\left[\frac{\partial\mathbf{A}_k^T}{\partial m_j}\right]\mathbf{A}_k^{-1}\frac{\partial\delta_r}{\partial z}, \tag{10}$$

where δ denotes the Dirac delta function. As shown by eq. (10), the explicit computation of **J** requires to perform one seismic modelling for each non-redundant position of the source l and receiver r, while the gradient computed with the adjoint state method (eq. 8) requires to perform two seismic modellings per source (one to compute the incident wavefield $\mathbf{p}_{l,k}$ and one to compute the adjoint wavefield $\mathbf{r}_{l,k}$). To mitigate the computational burden associated with the explicit estimation of **J**, we compute the diagonal terms of the approximate Hessian only for the first iteration of the inverted frequency and keep it constant over the subsequent iterations (Ravaut *et al.* 2004; Operto *et al.* 2006). Alternatively, we use the diagonal terms of the approximate Hessian (eq. 9) as a preconditioner of the L-BFGS algorithm, which provides an estimate of the product of the inverse of the curvature operator with the steepest-ascent direction vector using M vectors $\gamma^{(m)}$ and M model vectors $\mathbf{m}^{(m)}$ from the previous iterations ($m = k - M, k - 1$) (Nocedal & Wright 1999). In this study, we use $M = 5$. The L-BFGS algorithm is provided in Nocedal & Wright (1999) and it was adapted for FWI applications by Brossier *et al.* (2009) and Brossier (2010). In the steepest-descent algorithm, the descent direction is given by the opposite of the steepest-ascent direction (since $\mathcal{H}$ is diagonal). In the L-BFGS algorithm, the descent direction is controlled both by the steepest-ascent direction γ^k and the curvature information provided by the approximate Hessian, which should result in faster and improved convergence than the steepest-descent algorithm (Brossier *et al.* (2009) for an illustration on an onshore synthetic case study).

The step length α is estimated by a parabolic interpolation using three values of the misfit function, one of which is computed in the starting model of iteration k. Any approximation of the inverse of the Hessian provides a useful scaling of the gradients associated with the wave speed and attenuation parameter classes, before estimation of the step length.

3 SENSITIVITY AND TRADE-OFF ANALYSIS OF JOINT VELOCITY AND ATTENUATION IMAGING

Before showing FWI examples, we present a sensitivity and trade-off analysis of the joint recovery of c and Q in the framework of FWI.

3.1 Radiation pattern analysis

In order to gain some insights on the sensitivity of the seismic data to c and Q model parameters, we compute by finite differences the partial derivative wavefield $\frac{\partial \mathbf{p}}{\partial m}$ with respect to c and Q in a homogeneous background model. The partial derivative wavefield associated with source l, frequency k and model parameter m_j satisfies (Pratt *et al.* 1998, their eq. 15).

$$\mathbf{A}_k \frac{\partial \mathbf{p}_{l,k}}{\partial m_j} = -\frac{\partial \mathbf{A}_k}{\partial m_j} \mathbf{p}_{l,k} = \mathbf{f}_{j,l,k} \tag{11}$$

By analogy with the forward problem (eq. 3), the partial derivative wavefield is the solution of the wave equation for the virtual source $\mathbf{f}_{j,l,k}$. This source is the product of the incident wavefield $\mathbf{p}_{l,k}$ with a sparse radiation pattern operator $\frac{\partial \mathbf{A}_k}{\partial m_j}$. Since the wave speed occurs only in the mass term of the wave equation (i.e. the wave speed is located only on the diagonal of the impedance matrix $\mathbf{A}_k$, eq. 1), the radiation pattern matrix associated with the c_j or Q_j parameter reduces to a complex-valued scalar d_j, which corresponds to the

partial derivative of the mass term $\frac{\omega^2}{\kappa(\mathbf{x})}$ with respect to c_j and Q_j, respectively.

This implies that the partial derivative wavefield with respect to c_j or Q_j is generated by an explosive (isotropic) secondary source, the signature d_j of which satisfies

$$\mathbf{A}_k \frac{\partial \mathbf{p}_{l,k}}{\partial m_j} = -d_{j,k} p_{l,k} \delta(\mathbf{x} - \mathbf{x}_j), \tag{12}$$

where δ denotes the Dirac delta function and $\mathbf{x}_j$ denotes the coordinates of node j in the FD grid. The fact that the radiation pattern of the c and Q model parameter is isotropic implies that the amplitude of the secondary sources associated with c and Q does not depend on the aperture angle. Therefore, we should expect the same resolution for the reconstructions of c and Q, if both parameters have a sufficiently strong signature in the data. The strength of this signature can be assessed for each parameter by the amplitude of $d_{j,k}$.

For the complex-valued wave speed given by eq. (2), the partial derivative of the mass term with respect to c and Q gives, respectively, for d

$$d_c = \frac{-2\omega^2}{\rho c^3 (1 - i/2Q)^2} \approx \frac{-2\omega^2 (1 + i/Q)}{\rho c^3} \tag{13}$$

and

$$d_Q = \frac{-i\omega^2}{\rho \, c^2 \, Q^2 \, (1 - i/2Q)^3} \approx \frac{-i\omega^2 (1 + i3/2Q)}{\rho \, c^2 \, Q^2}. \tag{14}$$

The approximate expression of d shows that the virtual source associated with c and Q are approximately related by a Hilbert transform through the imaginary term i. This prompt Mulder & Hak (2009), Hak & Mulder (2010) and Hak & Mulder (2011) to conclude that many combinations of c and Q can produce nearly identical data, and, therefore, cannot be retrieved by linear waveform inversion, in particular from short-aperture seismic reflection data. However, when Q decreases, the relationships between the phase of the radiation pattern of c and Q might become more linearly independent (through the complex coefficients $(1 + i/Q)$ and $(1 + i3/2Q)$ in eqs 13 and 14), that can help to unambiguously reconstruct c and Q in particular when wide-aperture data are considered and nonlinear inversion is performed. This statement is supported with the numerical experiments and the real data case study presented in this study.

The two basic conclusions that can be inferred from the expression of the virtual source terms (eqs 13 and 14) are as follows. As Q decreases to very small values, the sensitivity of the data to Q increases relatively to the sensitivity of the data to c as suggested by the value of c^3 and $c^2 Q^2$ in the denominator of eqs (13) and (14). Second, as the frequency decreases, the sensitivity of the data to Q decreases. Therefore, the reconstruction of Q for weakly attenuating media at low frequencies can be unstable. The modulus and the phase of monochromatic partial derivative wavefields computed by finite differences for dimensionless parameter ($\bar{c} = c/c_0$ and $\bar{Q} = Q/Q_0$) in a 10 km $\times$ 10 km homogeneous background model are shown in Fig. 1. The wave speed c_0 and the attenuation factor Q_0 in the background model are 4000 m s^{-1} and 200, respectively. The source coordinates are $x = 5$ km; $z = 1$ km. The frequency is 7 Hz. To compute the partial derivative wavefield with respect to $\bar{c} = c/c_0$, we first apply a perturbation Δc in the value of c_0 ($\Delta c = 400$ m s^{-1} leading to $c_0 + \Delta c = 4400$ m s^{-1}) at the grid point located in the middle of the model, compute the wavefield in the perturbed model and take the difference with the wavefield computed in the homogeneous background model. The resulting differential wavefield is divided by Δc and is multiplied by c_0 to obtain the
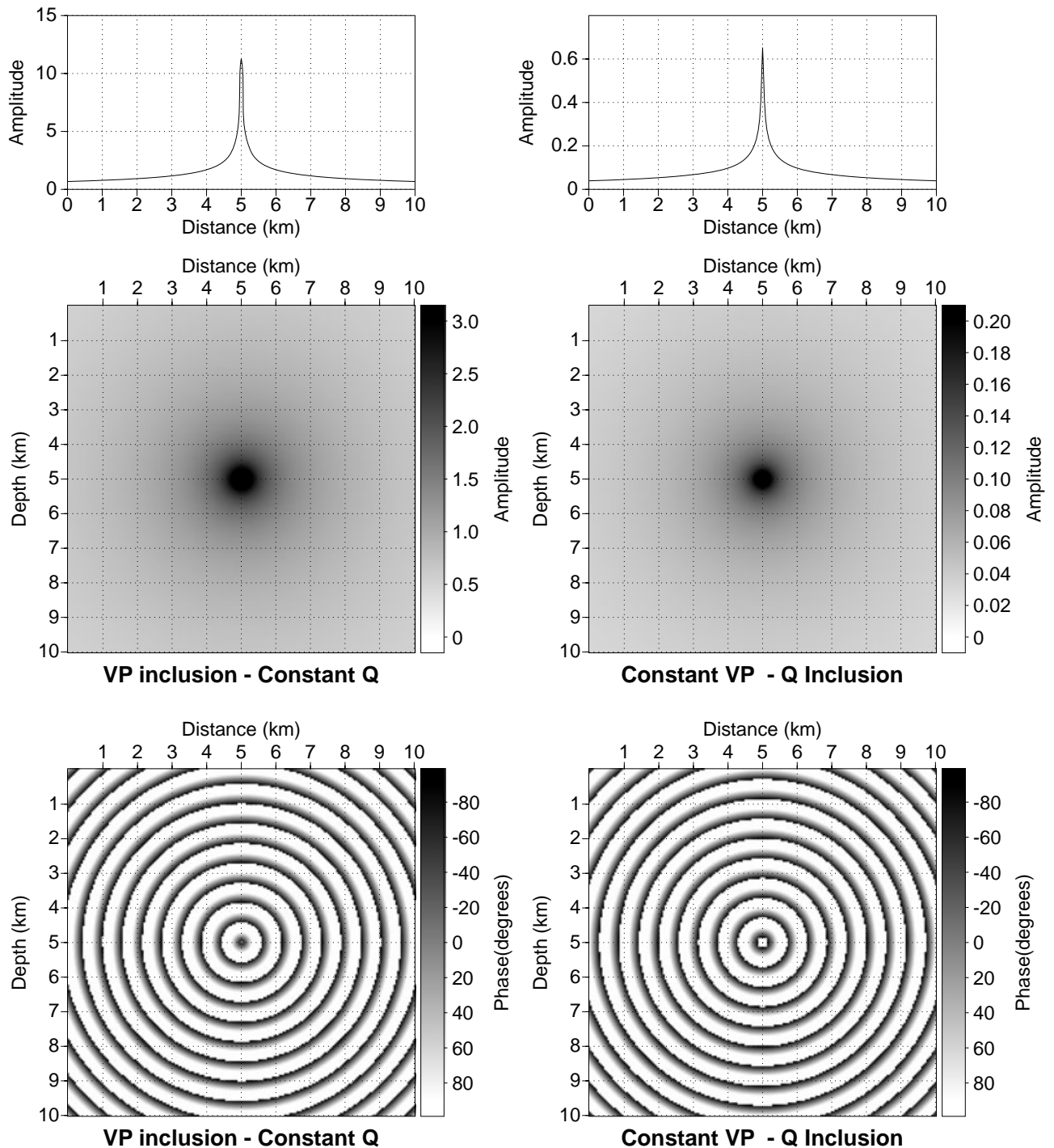
Figure 1. Partial derivative of the pressure wavefield with respect to c and Q parameter. (a–b) Modulus (a) and phase (b) of the partial derivative wavefield with respect to $\bar{c} = c/c_0$. (c-d) Modulus (c) and phase (d) of the partial derivative wavefield with respect to $\bar{Q} = Q/Q_0$. In the background model, $c_0 = 4000 \text{ m s}^{-1}$ and $Q_0 = 200$. The incident wavefield is computed for a source located at x = 5 km and z = 1 km. The partial derivative wavefields are computed for punctual c and Q perturbations in the middle of the model with $c = 4400 \text{ m s}^{-1}$ and $Q = 10$. The amplitude of the partial derivative wavefield with respect to $\bar{c}$ at the position of the diffractor point is higher that of the partial derivative wavefield with respect to $\bar{Q}$ by one order of magnitude. There radiation pattern of both parameters is isotropic (i.e. it does not vary with the aperture angle). Note the phase shift between the two partial derivative wavefields.

partial derivative of the wavefield with respect to the dimensionless wave speed $\bar{c}$. The same exercise is repeated for a perturbation ΔQ in the value of $Q_0(\Delta Q = -190$ leading to $Q_0 + \Delta Q = 10)$.

For these values of Q_0 and ΔQ, the amplitudes of the partial derivative wavefields with respect to $\bar{c}$ have an amplitude which is higher that of the partial derivative wavefields with respect to $\bar{Q}$ by one order of magnitude. This highlights that the seismic data are significantly more sensitive to velocity perturbations than to attenuation ones. For higher values of Q (i.e. weaker attenuation) and smaller frequencies (not shown here), the relative sensitivity of the data to Q will decrease and the reconstruction of Q_P will become strongly undetermined. The modulus of the partial derivative

wavefields shows the same isotropic radiation pattern as expected (Figs 1a–c). Only the phase shift between the two partial derivative wavefields can help to unambiguously reconstruct the two parameter classes (Figs 1b–d).

3.2 Grid-search analysis

We present now a grid search analysis of the misfit function for a simple model parametrized by only two parameters (the velocity and attenuation factor in an inclusion embedded in an homogeneous background) and an ideal acquisition with sources and receivers all around the target. Analysis of the variations of the misfit function

as a function of the velocity and attenuation errors in the inclusion gives some insights on the convexity of the misfit function in the neighbourhood of the global minimum, on the sensitivity of the misfit function to the two parameter classes and on the potential trade-off between the two sets of parameter. The velocity and the attenuation factors are $c_0 = 3500 \text{ m s}^{-1}$ and $Q_0 = 90$, respectively. The circular inclusion has a radius of 300 m, and is located in the middle of the model. The velocity and attenuation in the inclusion are $c = 3700 \text{ m s}^{-1}$ and $Q = 50$, respectively. The acquisition involves sources and receivers all along the edges of the square model. This corresponds to an ideal acquisition, where both transmitted and reflected wavefields sample the heterogeneity with all the possible
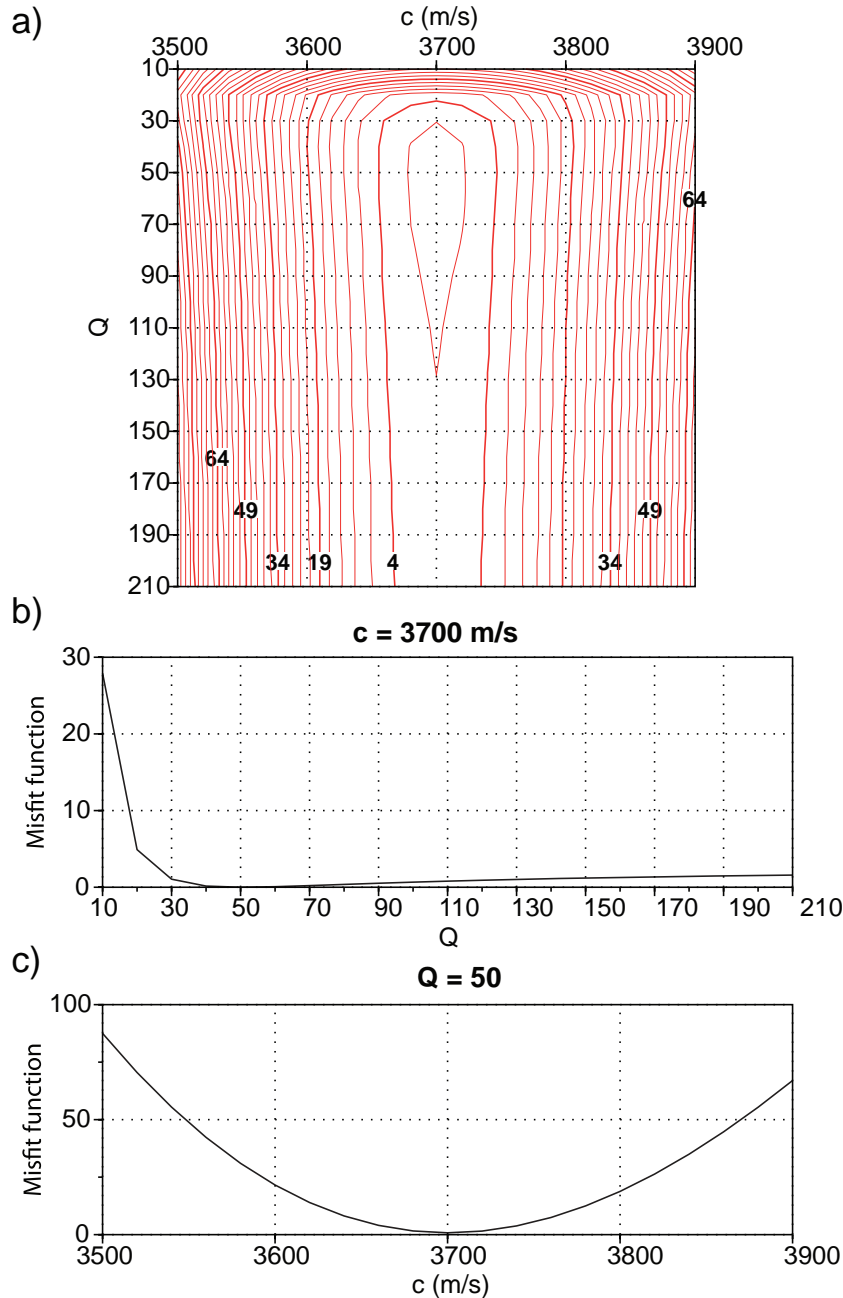

Figure 2. Sensitivity and trade-off analysis of c and Q parameters by a grid search. (a) Misfit function as a function of c and Q in the inclusion model (see text for details). The correct model corresponds to $c = 3700 \text{ m s}^{-1}$ and $Q = 50$ in the inclusion. (b) Profile of the misfit function surface for $c = 3700 \text{ m s}^{-1}$. (c) Profile of the misfit function surface for $Q = 50$. Note, how the sensitivity of the misfit function decreases when Q increases.

scattering angles. Data computed in the inclusion model for seven frequencies between 4 and 20 Hz are used as the recorded data set. The misfit function is then computed for models associated with different values of c and Q in the inclusion. The velocity in the inclusion ranges between 3500 and 3900 m s^{-1} while Q ranges between 10 and 210. The iso-contours of the misfit function in the $c - Q$ plane are shown in Fig. 2(a). The misfit function shows a well localized minimum. For this range of c and Q values where propagative effects dominate diffusive ones, the wave speed has a high sensitivity in the data, shown by the well-convex shape of the cross-section of the misfit function across the correct value of Q (Fig. 2c). For the Q parameter, the misfit function shows a significant sensitivity for Q smaller than the true value, while the misfit function is slowly varying for Q higher than the true value (Fig. 2b). This illustrates the fact that for high values of Q, the modelled data are mainly controlled by amplitude attenuation, while for low values of Q, both the amplitude and phase of the wavefield scattered by the inclusion change significantly with Q because the term $1/Q$ becomes non-negligible with respect to 1 in eq. (14).

3.3 Numerical verification in a canonical model

In order to verify the conclusions inferred from the grid-search analysis, we perform visco-acoustic FWI using the same acquisition and inclusion geometry as for the grid-search analysis. We invert successively seven frequencies between 4 and 20 Hz with the L-BFGS optimization and perform 50 iterations per frequency. The wave speed and the attenuation factor are jointly reconstructed. We perform two FWI experiments corresponding to two different attenuation in the background, $Q_0 = 300$ and 90. In both cases, the starting model for FWI is a homogeneous background model without the inclusion. During FWI, both velocity and Q gradients are smoothed with horizontal and vertical correlation lengths corresponding to 20 per cent of the propagated wavelength. The final FWI models for c and Q are shown in Figs 3(a) and (b). The velocity model is equally well recovered for $Q_0 = 300$ and 90. The attenuation model is well reconstructed in particular in terms of Q amplitude for $Q_0 = 300$ and 90. However, some low wavenumber artefacts in the Q model are shown when $Q_0 = 300$. These artefacts are created during the inversion of the low frequencies, that is illustrated by the fact that the inversion failed to perform 50 iterations at these frequencies (Fig. 3b). To remove these artefacts, we smooth the gradient of Q with horizontal and vertical correlation lengths corresponding to 40 per cent of the propagated wavelength instead of 20 per cent (Fig. 3c). This more aggressive regularization of the Q imaging filters out the low wavenumber artefacts, and allows the misfit function to decrease over the 50 iterations. The low wavenumber artefacts which appear only for $Q_0 = 300$ illustrates again the decrease of the sensitivity of the FWI at low frequencies to Q for weakly attenuating media.
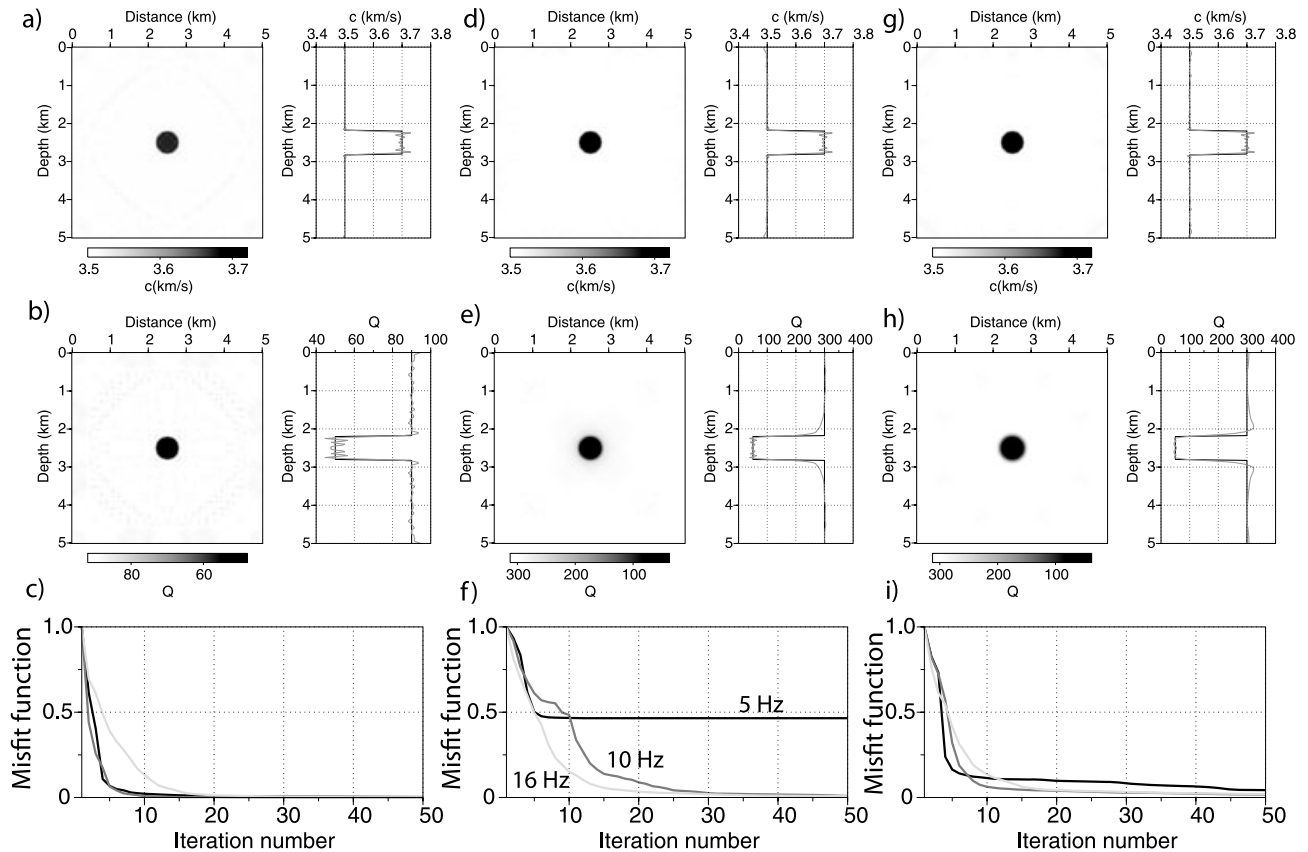

Figure 3. FWI synthetic example—inclusion model. (a–b) FWI velocity (a) and Q (b) model for a $Q = 90$ attenuating background model. The right-hand panels are vertical profiles across the inclusion. (c) Misfit functions versus iteration number for three inverted frequencies (5 (black), 10 (dark gray) and 16 (light gray) Hz). (d–e) FWI velocity (d) and Q (e) model for a $Q = 300$ attenuating background model. Note the low wavenumber artefacts in the Q profiles. (f) As in (c) for the $Q = 300$ attenuating background model. Note how the inversion fails to perform 50 iterations at 5 Hz. (g–h) Same as (d–e) but smoothing regularization of the Q gradient was strengthened. The low wavenumbers were filtered out by the smoothing regularization. (i) Misfit functions versus iteration number for the strongly regularized inversion.

4 SYNTHETIC EXAMPLES WITH REALISTIC SURFACE ACQUISITION

4.1 Simple layered model

We now assess the trade-off between model parameters c and Q for a more realistic setting similar to that of the GRUNDY 2003 experiment. The surface wide-aperture acquisition consists of 32 shot-points with a 1.5 km spacing and a constant recording spread of 501 receivers with a 100 m interval. The velocity model, 50 km long and 10 km deep, includes a 5 per cent c anomaly in a 500-m-thick layer at 2 km depth and a 10 per cent c anomaly at 3 km

depth (Fig. 4a). These two layers are embedded in a velocity gradient background. The background attenuation model is homogeneous with $Q = 100$. The top layer has an attenuation factor of 20, while the bottom layer does not have attenuation perturbation ($Q = 100$). Synthetic data were computed using a 10-Hz Ricker wavelet. Fourteen frequency components ranging between 3 and 16 Hz were successively inverted. The starting velocity and Q models for visco-acoustic FWI are the background models described above. The density model was obtained from either the true velocity model (for computing recorded data) or smooth background velocity model (for modelling during inversion) using Nafe-Drake curve.
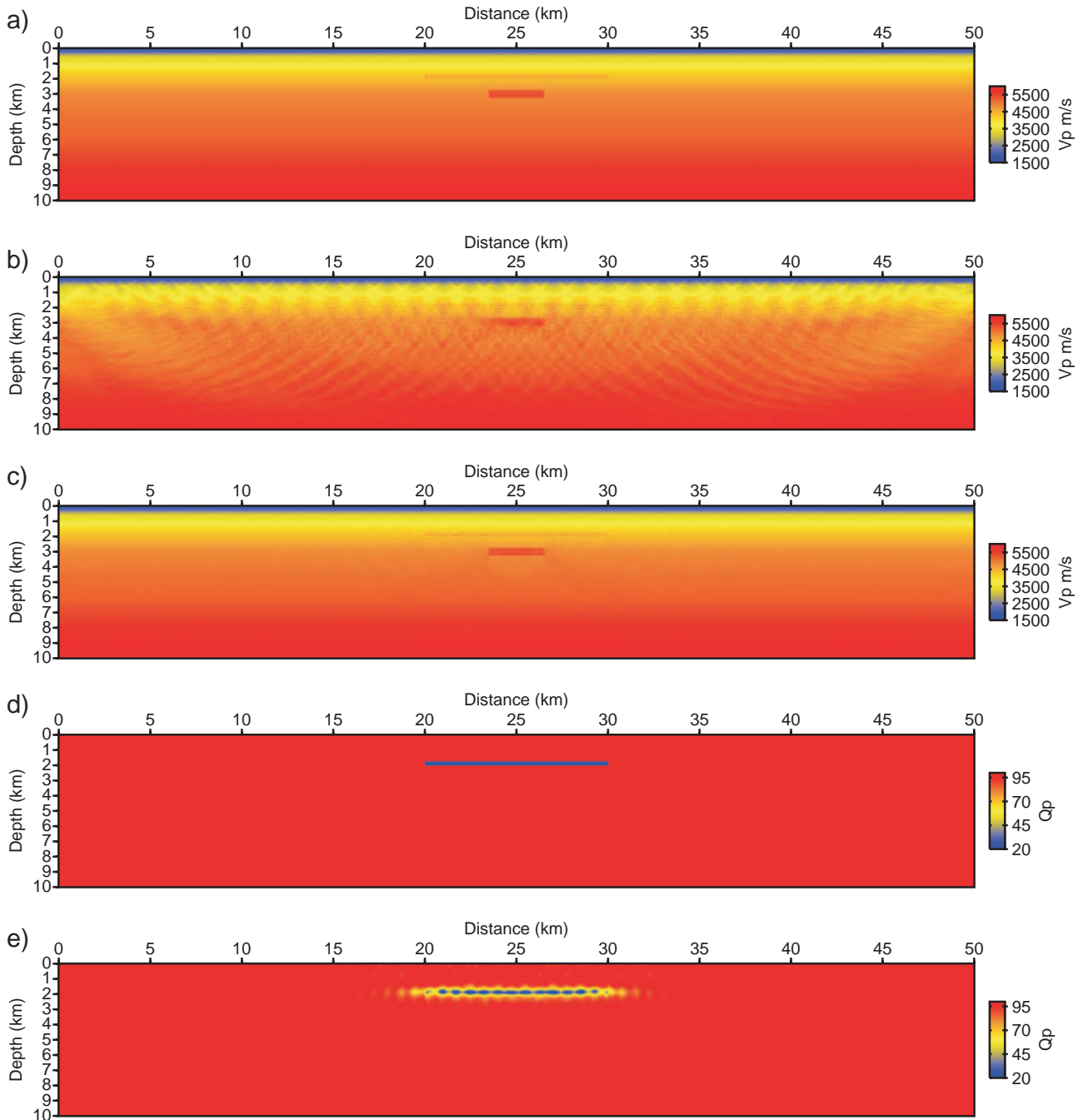
Figure 4. FWI synthetic example—coarse surface acquisition. (a) True model. (b) Velocity model from the acoustic FWI of visco-acoustic data. Note the footprint of the coarse acquisition. (c) Velocity model from the visco-acoustic FWI of visco-acoustic data. (d) True Q model. (e) Recovered Q model.
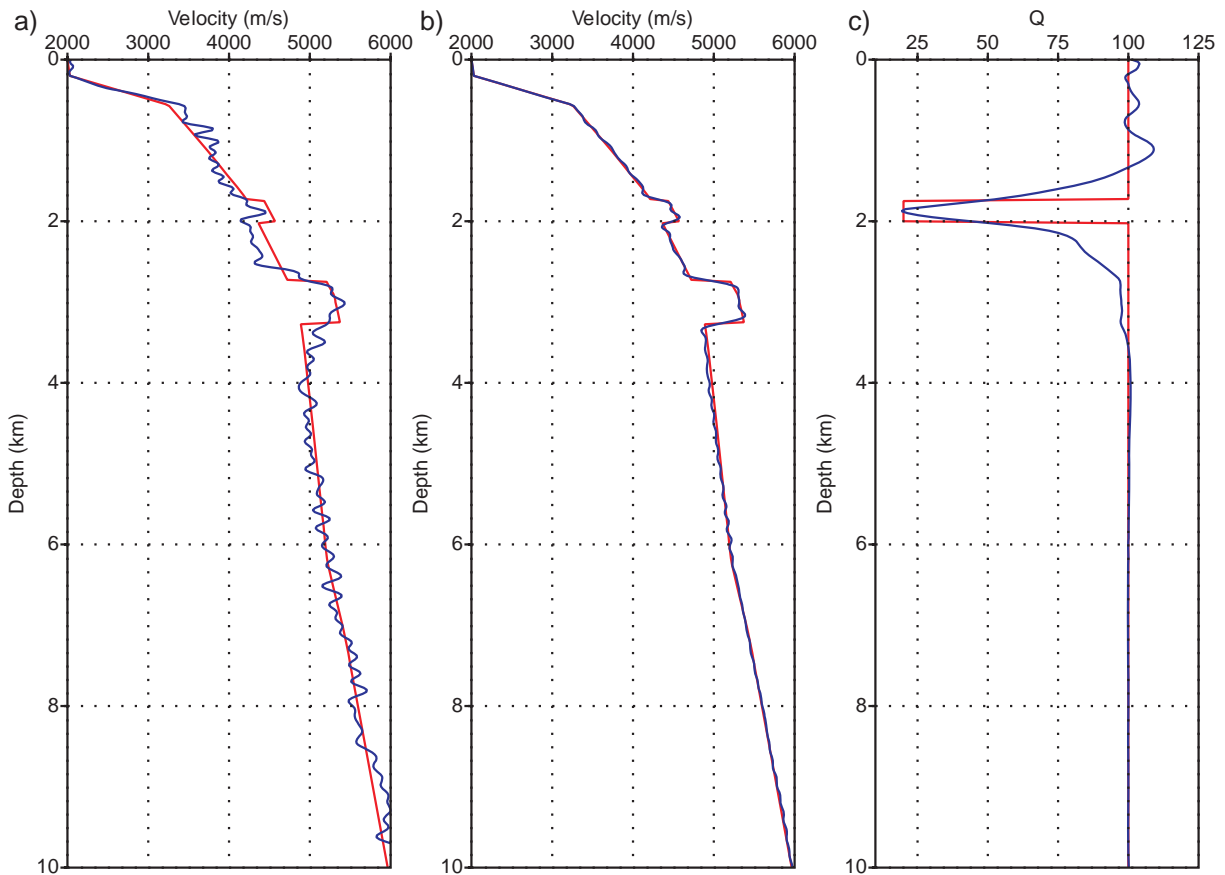
Figure 5. FWI synthetic example—coarse surface acquisition. (a) Comparison between vertical velocity profiles extracted at $X = 25$ km from the true model (red line) and the acoustic FWI model (blue line). (b) Comparison between vertical velocity profiles extracted from the true model (red line) and the visco-acoustic FWI model (blue line). (c) Comparison between vertical Q profiles extracted from the true model (red line) and the visco-acoustic FWI model (blue line). Amplitudes of Q are quite well reconstructed, but low-wavenumber artefacts are visible.

First, a purely acoustic FWI for c only was applied (setting $Q = 10\,000$). Although the FWI was able to detect the deeper velocity anomaly, the shallow layer remained almost unresolved (Figs 4b and 5a). Acoustic FWI did not produce sharp velocity contrasts and the overall picture is quite noisy with the clear footprint of a relatively sparse acquisition layout (Figs 4b and 5a). In contrast, the visco-acoustic FWI for the joint reconstruction of c and Q resolves quite well the velocity anomalies (Figs 4c and 5b).

The Q anomaly was properly localized with the absolute amplitude well estimated after 40 FWI iterations per frequency (Fig. 5c). However, the images are not so sharp in comparison to velocity recovery, probably because of the slower convergence of the Q reconstruction. We notice also some low wavenumber artefacts in the vertical profile of Fig. 5(c), which have probably a similar origin that the ones shown in Fig. 3(c). When only 20 iterations of FWI per frequency are performed, we observed also some leakage of Q into the c estimation, which disappeared when the number of iterations was doubled. Results of this test point out that the c and Q are potentially decoupled when a sufficient number of non-linear iterations is performed as suggested by Hak & Mulder (2011).

4.2 Realistic geological model

We repeat similar synthetic modelling exercise for a more realistic model representative for the Polish Basin area (Figs 6a and d). In

order to simulate weathered layer and hence make the test more realistic, we put a 100-m thick low-velocity ($c = 1800$ m s^{-1}) and low-Q ($Q = 20$) layer on the top of the model. The same frequency range was used as in the previous test. The starting velocity model is a smoothed version of the true one (smoothed from below the low-velocity layer, Fig. 7a) and the starting Q model is homogeneous with $Q = 75$.

The density model used to compute the synthetic data was obtained from the true velocity model using Nafe-Drake relation. In Fig. 6, we use the same density model as background model for FWI to eliminate the impact of density on Q and Vp recovery. In order to check that the layered background density model does not pose significant structural *a priori* constraints in the FWI, we test also two other inversion setups related to density. First, a homogeneous density equal to 1000 kg m^{-3} is used to compute both the synthetic data in the true model (inverted data) and in the FWI models. The results (not shown here) are slightly worse that obtained with the layered density model. We interpret this result by the fact that homogeneous density decreases the reflection coefficient at interfaces, and, hence, the amplitudes of the reflections. These lower amplitudes make in turn the inversion of the reflections less well-conditioned. Second, we consider the layered density model inferred from the Nafe-Drake relation as the true density model, while the background density model for FWI is a smoothed version of the true one. We note that the smooth background density model does not impact significantly the reconstructed velocity model (not shown
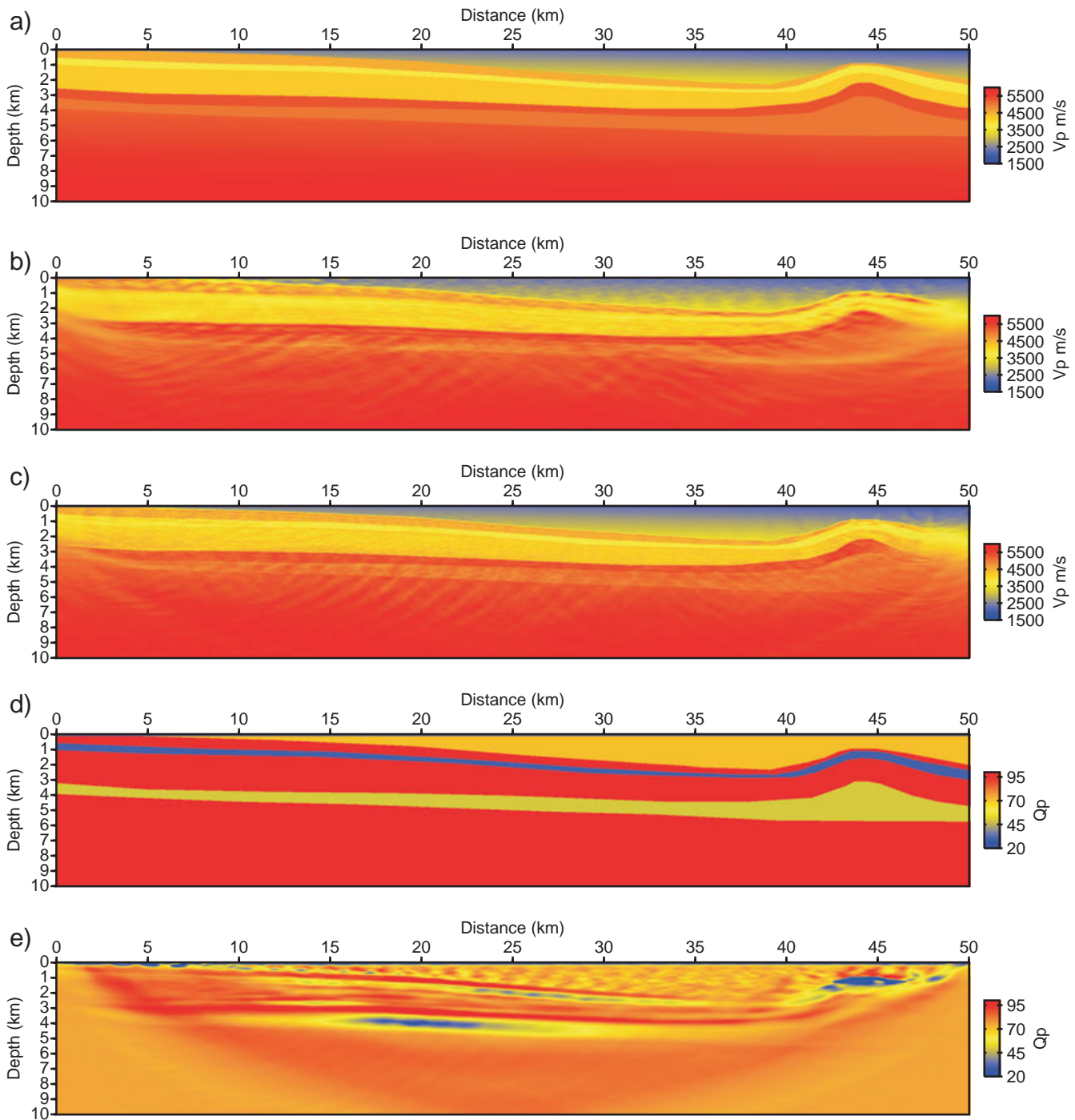
a)

b)

c)

d)

e)

Figure 6. FWI synthetic example—realistic geological model. (a) True model. (b) Velocity model from the acoustic FWI of visco-acoustic data. (c) Velocity model from the visco-acoustic FWI of visco-acoustic data; (d) True Q model; (e) Recovered Q model.

here), but that the attenuation model exhibits overestimated Q in the upper structure (Fig. 7d). We interpret this result as the footprint of the trade-off between the density and Q at short apertures: the inversion balances the underestimated reflection amplitudes resulting from the smoothing of the density by an underestimation of the attenuation effects. Of note, this trade-off should not impact the inversion of the real data shown in the following of this study, because the inversion will be limited to the wide aperture components of the data (i.e. the early arriving phases). The wide aperture components of the data are not sensitive to the density variations, when the

density is combined with the velocity in the model parametrization (Virieux & Operto 2009, their fig. 13).

As expected, the acoustic FWI failed to image the velocity model properly (Fig. 6b). The coarse acquisition scheme was accommodated by the model and there are clear artefacts related to the large shot spacing. The vertical profile from the velocity model (Fig. 7a) shows that the acoustic FWI creates an artificial velocity jump on top of the model to mimic the effects of the attenuation in the weathered layer. The erroneous reconstruction of the near surface leads to a noisy reconstruction of the deep part of the model. In contrast,
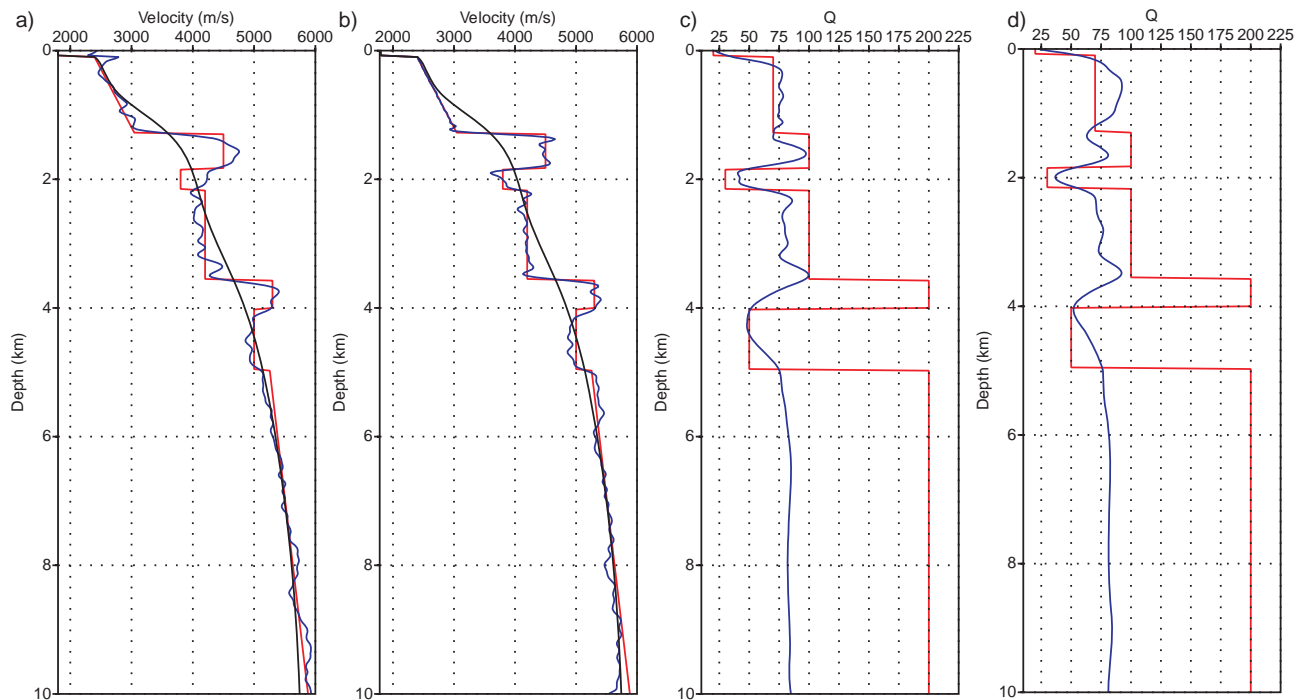
Figure 7. FWI synthetic example—realistic geological model. (a) Comparison between vertical velocity profiles extracted at $X = 25$ km from the true model (red line) and the acoustic FWI model (blue line). Starting velocity model is shown in black. (b) Comparison between vertical velocity profiles extracted from the true model (red line) and the visco-acoustic FWI model (blue line). Starting velocity model is shown in black. (c) Comparison between vertical Q profiles extracted from the true model (red line) and the visco-acoustic FWI model (blue line). The starting Q model is homogeneous ($Q_0 = 75$) and does not contain the weathered layer. (d) Same as (c) except that the background density model used for FWI is a smoothed version of the true density model (see text for details). Note the overestimated Q in the shallow part of the log.

the visco-acoustic FWI resolved perfectly the strong attenuation in the near-surface weathered layer, which was the key to properly resolving the velocity model in depth (Figs 6c and 7b). The recovery of the Q model is also reasonably good down to 2 km depth considering that the starting Q model is homogeneous (Figs 6e and 7c). Similarly to the previous results, the Q recovery was significantly improved when the number of iterations per frequency was increased from 20 to 40. It is interesting to note, that in the case of the visco-acoustic FWI, the footprint of the sparse acquisition is visible in the Q model not in the velocity model.

5 APPLICATION TO THE GRUNDY 2003 DATA

5.1 Data acquisition

The area of investigation is located in the SW part of the Polish Basin (PB), an easternmost part of the Central European Permian-Mesozoic Basin system (Fig. 8). The thickness of the Mesozoic and Permian (Zechstein) sedimentary cover in the survey area increases towards the NE from 4.5 to 6 km. The GRUNDY 2003 experiment (Malinowski *et al.* 2007) was aimed at the recognition of the pre-Permian strata, which are impenetrable for typical reflection seismics due to the shielding effect of Zechstein salts and anhydrites. Thus, for a successful investigation relatively low frequencies (of the order of 4 Hz) and long offsets (up to 50 km) were recorded.

In the 50 by 10 km rectangular area a total number of 786 RefTek-125 'Texan' stations with 4.5 Hz geophones were deployed (Fig. 8), forming a high-density central line (receiver spacing 100 m, 50 km

total length, referred as G01 line) and 4 additional parallel profiles with mean receiver spacing of 600 m. Thirty shots were fired along the G01 profile and 7 shots were fired in the side-profiles. The mean charge size of the shot points was 40 kg of TNT explosives. Shots were deployed in two 30-m deep boreholes providing consistent source signature. The data were recorded both inline and crossline allowing for 3-D tomographic imaging of the whole target area and a CDP processing along G01 line (Malinowski *et al.* 2007). Here we focus on the 2-D data recorded along the central profile.

5.2 FWI data pre-processing

Data pre-processing is one of the key issues in FWI, especially in the framework of the acoustic and visco-acoustic approximations, where a proper data windowing is necessary to eliminate elastic effects like ground roll and converted waves. Here we use the same pre-processing flow as previously applied to GRUNDY 2003 data set (Malinowski & Operto 2008), which incorporates the following steps:

(i) 3-D to 2-D correction by multiplying the amplitudes by $\sqrt{t}$;

(ii) Predictive spectral whitening (frequency domain deconvolution);

(iii) Butterworth bandpass filtering (2–25 Hz);

(iv) QC and bad trace removal;

(v) Muting centred on the calculated first-arrival traveltimes.

There is a one notable exception in comparison to the pre-processing used by Malinowski & Operto (2008). Namely, the spectral whitening was performed using a normalized deconvolution operator, which allowed for preservation of the amplitude-versus-offset
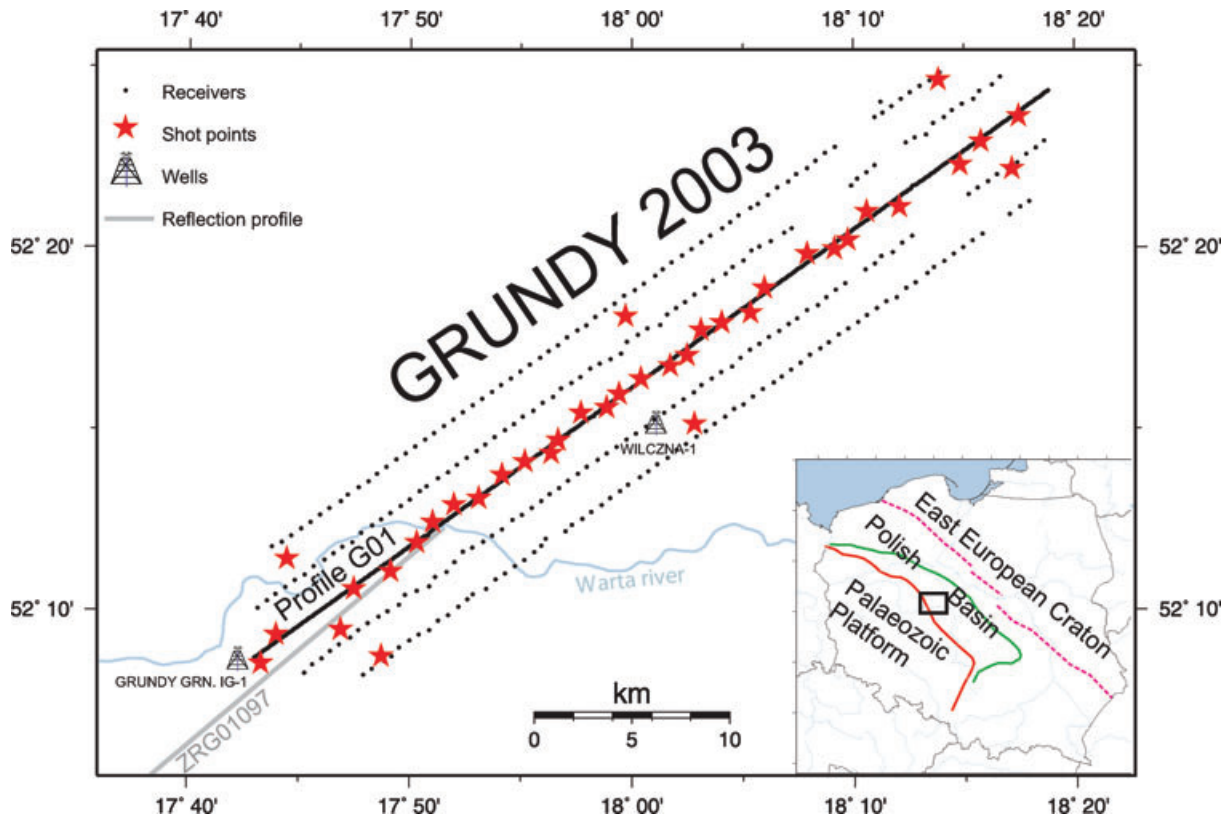
Figure 8. Location of the GRUNDY 2003 seismic experiment.

(AVO) response. The deconvolution operator applied to the trace j that we use is given by

$$s(\omega) = \left\{ \bar{A}_j(\omega) + \epsilon \left[\bar{A}_j(\omega) \right]_{\max} \right\} e^{i\phi_j(\omega)}, \tag{15}$$

where $\bar{A}_j(\omega)$ denotes a smoothed version of the amplitude spectrum of the recorded trace j, averaged over 10 neighbouring traces, ϵ is a prewhitening factor, and $\phi_j(\omega)$ is the minimum-phase spectrum of the trace j. To preserve the AVO response during the deconvolution process, we simply normalize the amplitudes of the deconvolution operator $\bar{A}_j(\omega)$ by its maximum amplitude for each trace.

Throughout the paper, we will use the term non-normalized data for the data with the preserved AVO response and the term normalized data for the data with a flat AVO response (i.e. when the deconvolution operator is not normalized). The effect of the two strategies for amplitude processing is shown in Fig. 9, where a shot gather is shown before and after whitening performed with and without normalization of the deconvolution operator. We first notice the sharp wide-angle reflection from the Triassic-Zechstein level at depths between 3 and 4 km (Malinowski & Operto 2008). A sharp attenuation of waves from below this level at offsets greater than 10–15 km is clearly visible in the raw and in the non-normalized shot gathers, and highlights the dramatic impact of attenuation on this land data set (Figs 9a and b). The normalized shot gather allows one to follow the arrivals up to the maximum offset of 40 km. This amplitude normalization shows how the deconvolution process helps to heuristically remove attenuation effects from the data and therefore allows for acoustic FWI of the normalized data (Malinowski & Operto 2008). Note also that the average amplitude spectrum of the non-normalized shot gather shows a higher

high-frequency content than the normalized data (Figs 9b and c). This results from the fact that the amplitudes of the low-frequency long-offset normalized data were accentuated by the deconvolution process.

The chosen mute window starts 0.1 s before the predicted first arrival traveltimes and ends from 0.1 to 0.5 s after calculated traveltimes depending on the offset range (Malinowski & Operto 2008, dashed line in Fig. 9). Typically, this time windowing preserved in the data the first arrivals and the critical and super-critical reflected wavefields whose arrival times are close to that of the first arrivals. Supercritical reflections preserved in this time window originate at the high velocity layers within the Jurassic and Lower Triassic as well as the Zechstein and sub-Zechstein (Upper Palaeozoic) strata. Time windowing can not be applied to the synthetic data during frequency-domain FWI. Therefore, the same pre-processing cannot be applied to the recorded and modelled data during FWI, if the real data are muted. Alternatively, we could have used complex-valued frequencies $(\omega + i\tau)$ in the frequency domain to apply some time damping from the first arrival (Brenders & Pratt 2007). It has been demonstrated on synthetic data examples (Brenders & Pratt 2007; Brossier *et al.* 2009) that using progressively increasing τ values (i.e. less damping) provides an additional level of data preconditioning in multiscale frequency-domain FWI by injecting progressively decreasing aperture angles during the inversion. However, we found the tuning of the τ damping difficult in case of the real data inversion. This might result from the fact that time damping requires to accurately mute the data before the first arrival to avoid the time damping to increase the noise level. So far, we failed to show that the time damping strategy was providing superior results than the more basic time windowing (muting).
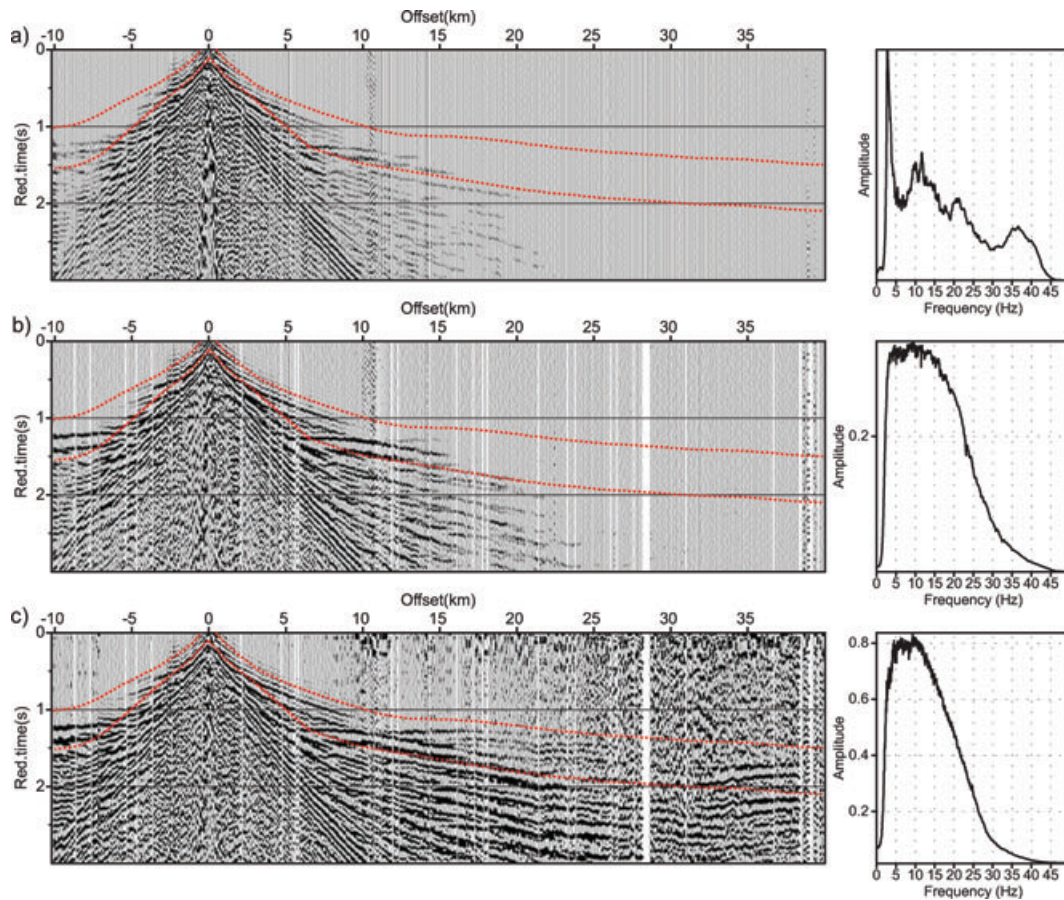
Figure 9. A representative shot-gather from the GRUNDY 2003 experiment with the associated amplitude spectra. (a) Raw data. (b) Data after FWI pre-processing including spectral whitening preserving AVO response. (c) Data after FWI pre-processing including spectral whitening with no-AVO preservation. Seismic sections are plotted with a 2–16 Hz bandpass filter and a reduction velocity of $5500\,\mathrm{m\,s^{-1}}$ applied. Note the strong attenuation of the arrival amplitudes for offsets greater than 10 km in (a-b). The dashed lines delineate the upper and lower limits of the mute used for FWI. The right-hand panels show the amplitude spectrum of the shot gather averaged over all the traces.

5.3 Starting models

We use a homogeneous starting Q model for visco-acoustic FWI. In order to estimate by trial-and-error the best homogeneous Q starting model for FWI, we compared the AVO responses of the real data with the response of the synthetic data calculated in the starting velocity model using different constant Q models (Fig. 10; Pratt 1999; Gao *et al.* 2006). The starting velocity model was derived by first-arrival traveltime tomography (FAT; Malinowski *et al.* 2007; Malinowski & Operto 2008, Fig. 11a). The density model used in the inversion was converted from the FAT model by a Nafe-Drake relation. We compute the visco-acoustic synthetic seismograms in the time domain using a wavelet inferred from the final model of the acoustic FWI. The source wavelet in the time domain was estimated by solving a linear inverse problem for each frequency of the source bandwidth followed by an inverse Fourier transform (Pratt 1999, his eq. 17). The modelled visco-acoustic AVO response calculated with $Q = 50$ matches well with the observed AVO curve (Fig. 10). Note that we do not apply any amplitude preconditioning like the artificially matching AVO behaviour of the data with an empirical function as proposed by Brenders & Pratt (2007).

5.4 FWI results

Similarly to Malinowski & Operto (2008), we invert successively 13 equally spaced frequencies ranging from 4 to 16 Hz. The starting
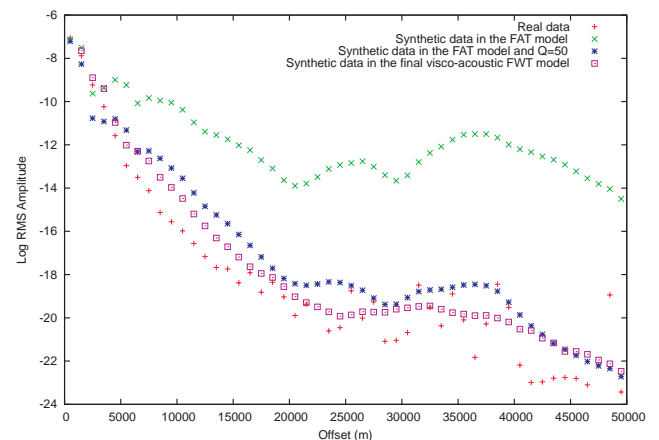

Figure 10. Comparison between RMS amplitude curves computed for real data and synthetic data. RMS amplitudes are averaged over all shot gathers and binned into 1-km offset bins. The synthetic data were computed in the starting FAT model in the acoustic ($Q = 10\,000$) and in the visco-acoustic ($Q = 50$) approximations, and in the final visco-acoustic FWI model.

frequency should be kept as low as possible to mitigate the cycle-skipping artefacts, but it should be adapted to the frequency content of the data. Due to the recording instrumentation used (4.5 Hz geophones) we chose 4 Hz as the starting frequency. Although
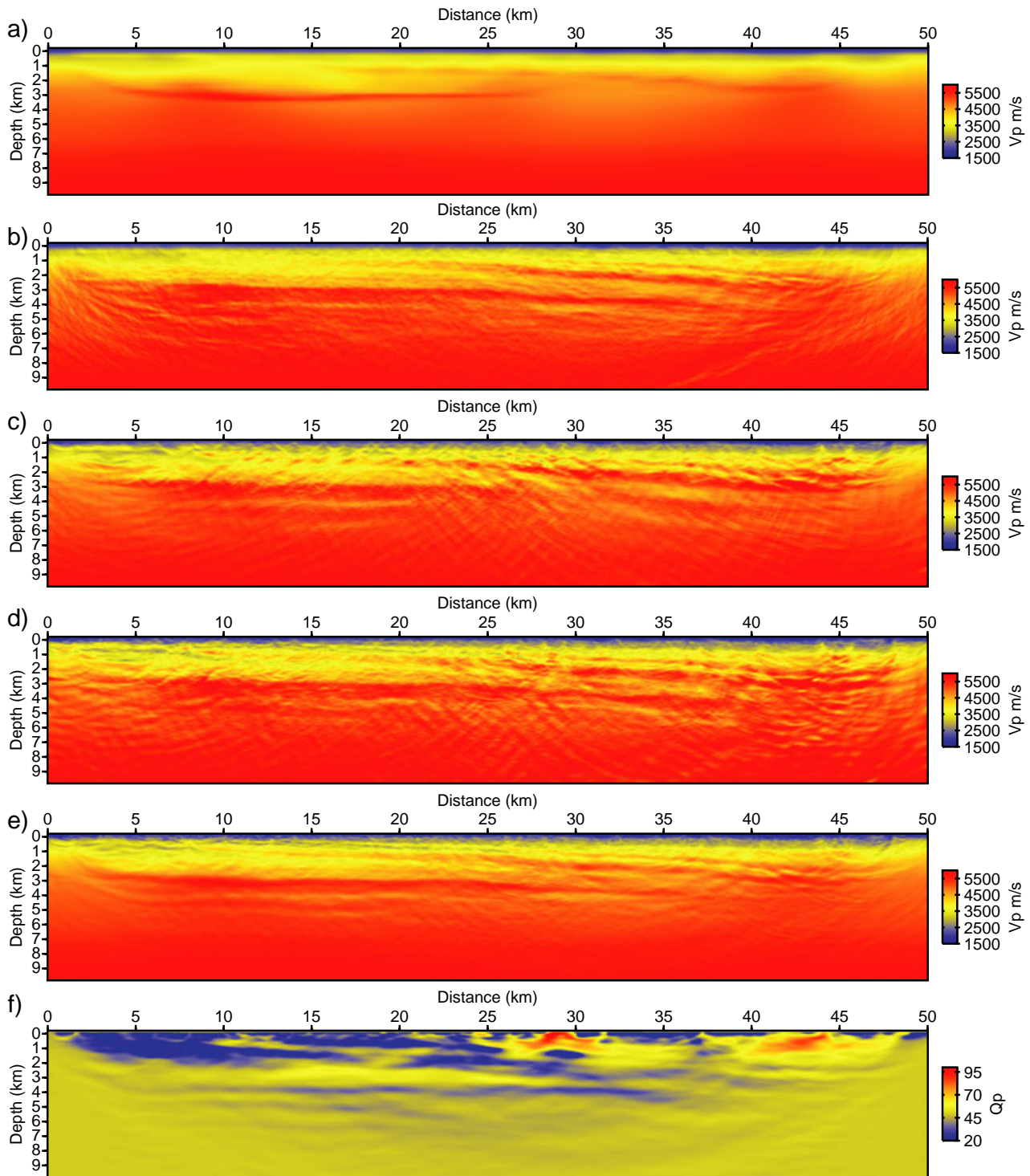
Figure 11. Starting velocity model from FAT (a) and final models from the FWI of the GRUNDY 2003 data (at final frequency of 16 Hz). (b) Velocity model from the acoustic FWI of normalized data. (c) Velocity model from the acoustic FWI of non-normalized data (constant $Q = 10\,000$). (d) Velocity model from the acoustic FWI of non-normalized data (constant $Q = 50$). (e) Velocity model from the final visco-acoustic FWI. (f) Attenuation (Q) model from the final visco-acoustic FWI.

the data contains free surface multiples we do not handle them: a layer with a constant wave speed of 1600 m s^{-1} is set above the topography. We assume that the free-surface multiples are confined to the top layers only and they do not affect significantly the wide-angle wavefield contained in our narrow mute window. The ghost effects at the free surface are partially taken into account by the

source wavelet estimation performed during each iteration of FWI using the approach of Pratt (1999). During the FWI, one single source wavelet is estimated from the full data volume.

In order to smooth the gradients of the misfit function, the correlation lengths for Q were twice as large as for c parameter. A more aggressive smoothing of the attenuation gradient was used

Table 1. Input parameters for the FWT.

Number of shot gathers (traces)	29 (12 754)
Number of traces for Hessian estimation	every 4th trace
Number of inverted frequencies	13 (4–16 Hz)
Model dimensions (km)	50 × 10
Forward gridstep (km)	0.025
Gain with offset (g)	1
ϵ	0.001
Iterations per frequency	20

to account for the lower sensitivity of the data to the attenuation parameter. For the data preconditioning performed by the weighting operator $\mathbf{W}_d$ (eq. 4), we apply a linear gain with offset to the data residuals. We found heuristically that this data weighting was providing the best scaling in depth of the gradients of the misfit function.

We use the L2 norm in the misfit function, although Brossier *et al.* (2010) and Brossier *et al.* (2009) concluded that the L1 norm provide a more robust alternative (at least for noisy synthetic data). The minimization of the objective function was done using the steepest-descent method. We also tested the L-BFGS optimization method, which provides much faster convergence in case of synthetic data inversion as described by Brossier *et al.* (2009). However, we observe that in case of our real data, the L-BFGS method indeed speeds up the convergence but it also provides much noisier results as compared to the classic steepest-descent approach. Other parameters of the modelling are summarized in Table 1.

First, we apply a purely acoustic FWI, setting Q to a high value (10 000) and inverting only for *P*-wave velocity using normalized data. Resulting velocity model is well-focused (Fig. 11b) and will be used as a reference in comparing other FWI results. In the next step we applied acoustic FWI to the non-normalized data. As
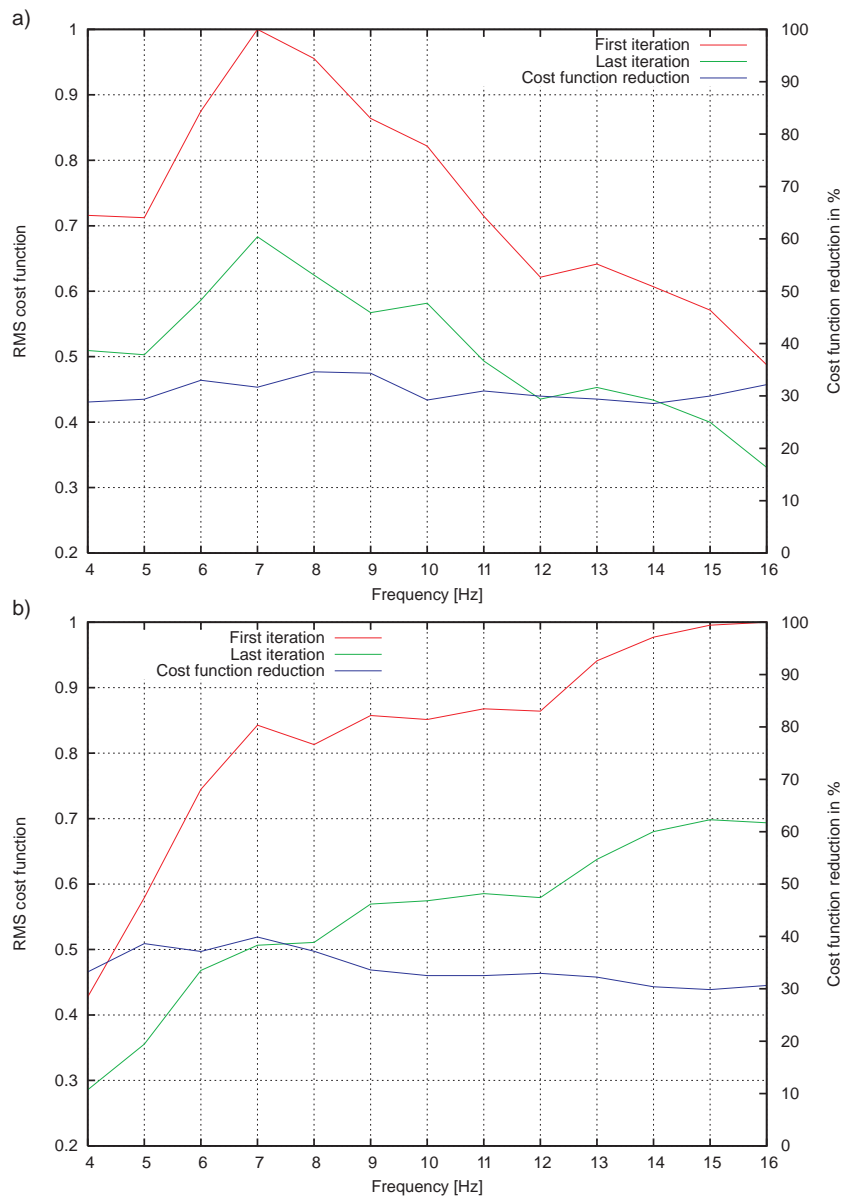

Figure 12. (a) Misfit function at first and last iterations versus inverted frequency for the acoustic FWI of normalized data. Right-hand axis shows misfit function reduction in percentage. (b) same as (a) for the visco-acoustic FWI of non-normalized data. Not the different trend of the curves as a function of iteration numbers in (a) and (b) (see text for details).

expected, the strong amplitude variations with offset were not properly handled by the simple acoustic FWI (Fig. 11c). The image is also polluted by the artefacts related to the sparse acquisition geometry. As already explained, matching of the AVO decay necessitates the introduction of attenuation. Subsequently, we applied the visco-acoustic FWI for velocity only using non-normalized data and a constant $Q = 50$ attenuation model. The results (Fig. 11d) are much better than the previous ones (Fig. 11c), but still remain noisy, which means that the amplitudes were not fully fitted during the FWI.

The velocity model obtained from the final visco-acoustic FWI of non-normalized data (Fig. 11e) is of a comparable quality to that of the normalized-data acoustic FWI (Fig. 11b). Simultaneously, we obtained also an attenuation model (Fig. 11f), which in some parts differs significantly from the starting $Q = 50$ model. The AVO-decay curve obtained from the synthetic data simulated for the final visco-acoustic FWI results is indeed closer to the AVO response of the real data (Fig. 10).

The effectiveness of the inversion procedure can be illustrated by looking at the misfit function reduction at each frequency (Fig. 12). For both acoustic FWI of normalized data and visco-acoustic FWI of non-normalized data, we obtained similar percentage of the misfit function reduction (30–40 per cent) with a slightly better performance of the latter. However, we show a very different behaviour of the misfit function trends for the two FWI applications. For nor-

malized data, the maximum value of the misfit function is reached for the frequency of 7 Hz, and then the misfit function decreases. In contrast, the misfit function is increasing with frequency for non-normalized data. This trend can be attributed to the fact that the AVO normalization of the data has increased the weight of the low-frequency long-offset data in the misfit function. In contrast, the misfit function is dominated by the high-frequency short-offset arrivals in the non-normalized data. It is worth noting that the visco-acoustic FWI succeeded in performing a quite significant misfit function reduction for the higher inverted frequencies (30–35 per cent), while it is generally acknowledged that it is quite challenging to perform reliable acoustic FWI at high frequencies due to the inaccuracy of the amplitude modelling and the increasing risk of cycle skipping artefacts. The significant reduction of the misfit function at the higher frequencies can be explained both by the more accurate modelling of amplitudes resulting from the accounting of the attenuation and by the relatively low starting frequency (4 Hz) that we used for inversion. This increasing weight of the long-offset data in the misfit function associated with the normalized data also probably explains why the FWI model inferred from acoustic FWI of normalized data is slightly more contrasted in the deep part than the FWI model inferred from visco-acoustic FWI of non-normalized data (compare Figs 11b and e). This different frequency weighting was already noticed in the spectra of the data shown in Figs 9b–c).
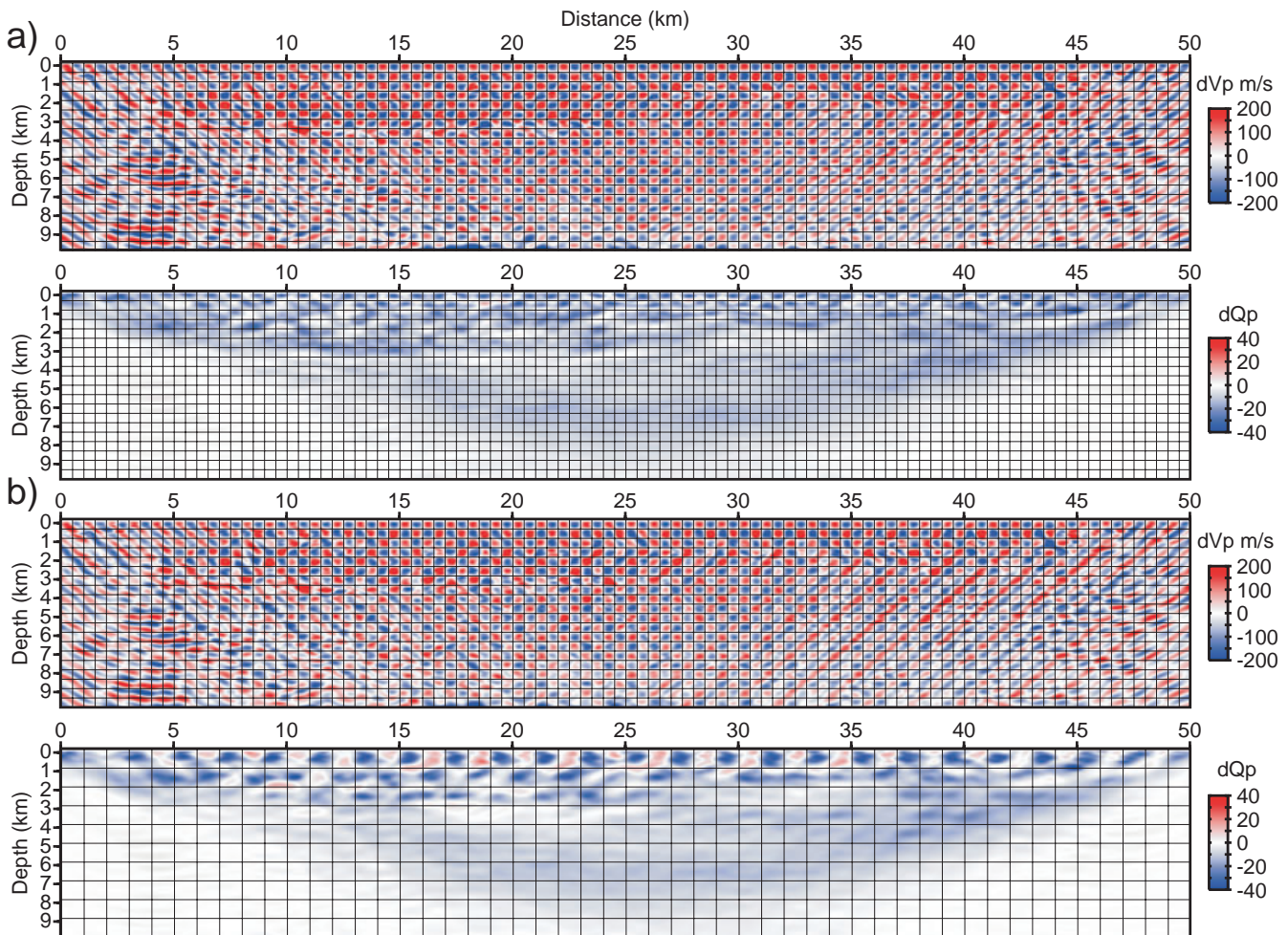

Figure 13. Checkerboard tests for velocity and attenuation: (a) Reconstructed checkerboard perturbations for *c* and *Q*. The size of the true perturbations is 500 m. (b) Same as (a) except that the size of the true perturbations for *Q* is 1 km.

5.5 Model appraisal

In this section, we discuss the relevance of the FWI model based upon checkerboard tests, source wavelet estimation, synthetic seismogram modelling and geological interpretation.

5.5.1 Resolution analysis

We assess the resolution of the final c and Q FWI models with a checkerboard test (Jaiswal *et al.* 2009). We add checkerboard perturbations to the initial FWI c and Q models, then we compute data in the perturbed models for the real acquisition geometry and we perform the multiscale FWI of these data starting from the initial c and Q models. The ability of the FWI to locally reconstruct perturbations of a given size provides some insights on the local spatial resolution of the FWI. The inversion setup (frequency bandwidth and sampling, regularization) for the checkerboard test is the same as the one used for the inversion of the real data. The amplitude of the perturbations are $\Delta c = \pm 200$ m s^{-1} and $\Delta Q = \pm 40$ for c and Q, respectively, and are consistent with that reconstructed by the FWI of the real data. Note that the amplitudes and the sizes of the c and Q perturbations in the checkerboard should also allow one to satisfy the assumptions underlying the local optimization based upon the Born approximation (i.e. the size and the amplitude of the perturbations are sufficiently small to avoid cycle skipping artefacts). We first test perturbations of dimension 500 m for both c and Q (Fig. 13a). The final FWI models show that velocity perturbations of this size can be reconstructed down to 10 km depth in the central part of the model. Q perturbations can be imaged down to 1 km depth, although only 50 per cent of the amplitude of the Q perturbations are reconstructed. We perform a second test where the size
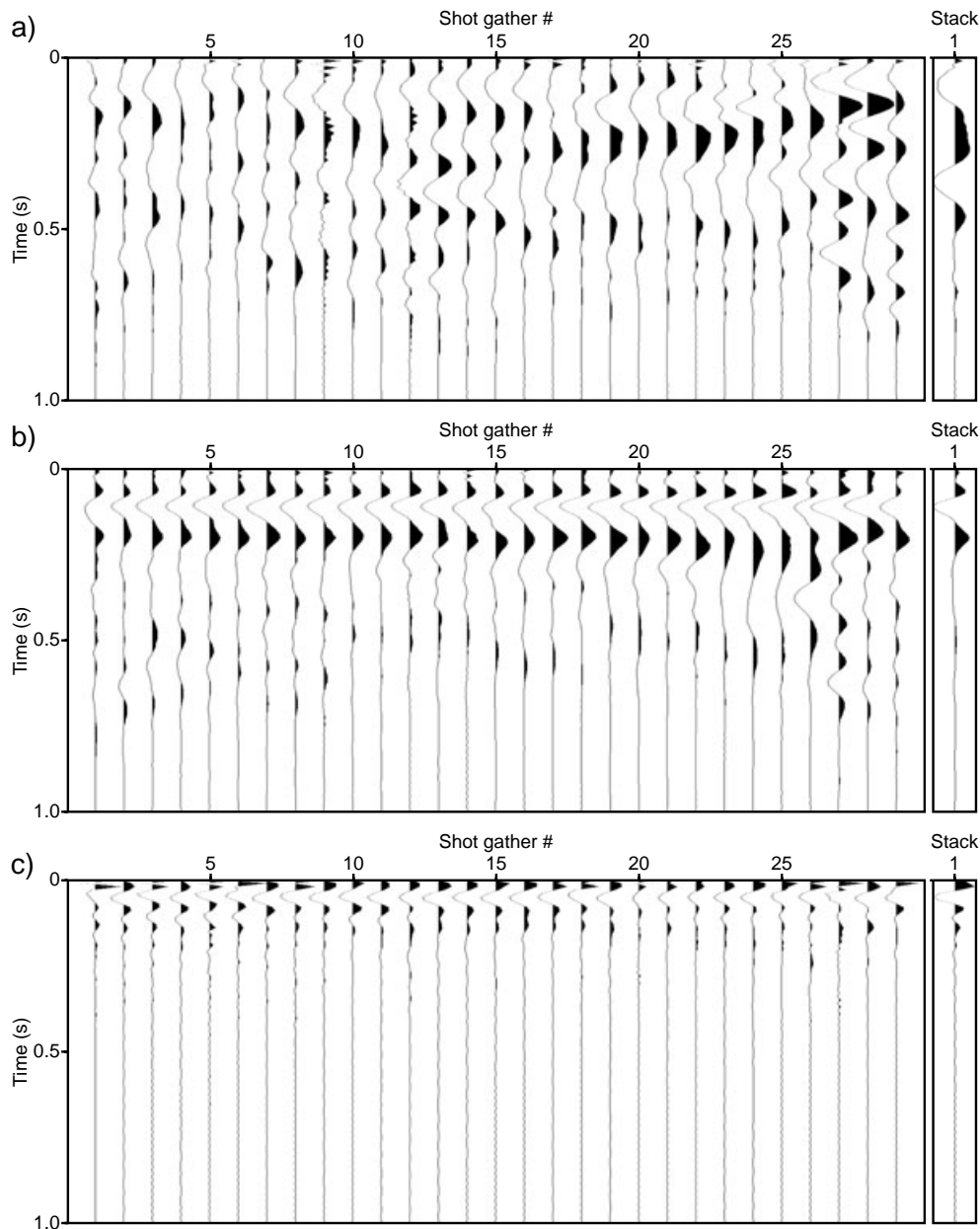

Figure 14. Time-domain source signature for each shot-gather. Source wavelets are estimated from (a) the starting FAT model, (b) from the final acoustic FWI model, and (c) from the final visco-acoustic FWI model (see text for more details). The right-hand panels represent the stacked wavelet.

of the Q perturbations was increased to 1 km, while the size of the c perturbations remains equal to 500 m (Fig. 13b). The resolution of the velocity model does not change significantly compared to the previous experiment. The amplitude of the Q perturbations are full recovered down to 1 km depth, while Q perturbations can be imaged down to 2 km depth with partial amplitudes. Note that, near the ends of the model, the perturbations are smeared along lines dipping towards the centre of the model. These artefacts are related to the limited data coverage near the end of the model, and look like those shown in (Jaiswal *et al.* 2009, their fig. 6).

5.5.2 Source wavelet estimation

We compute a source wavelet per shot gather in the time domain for the starting and the final FWI models using the approach of Pratt (1999, his eq. 17). The source term in eq. (3) is estimated for each discrete frequency of the source bandwidth before inverse Fourier transform which provides the wavelet in the time domain. The focusing and the consistency of the source wavelet from one shot to another provide some insights on the relevance of the FWI models.
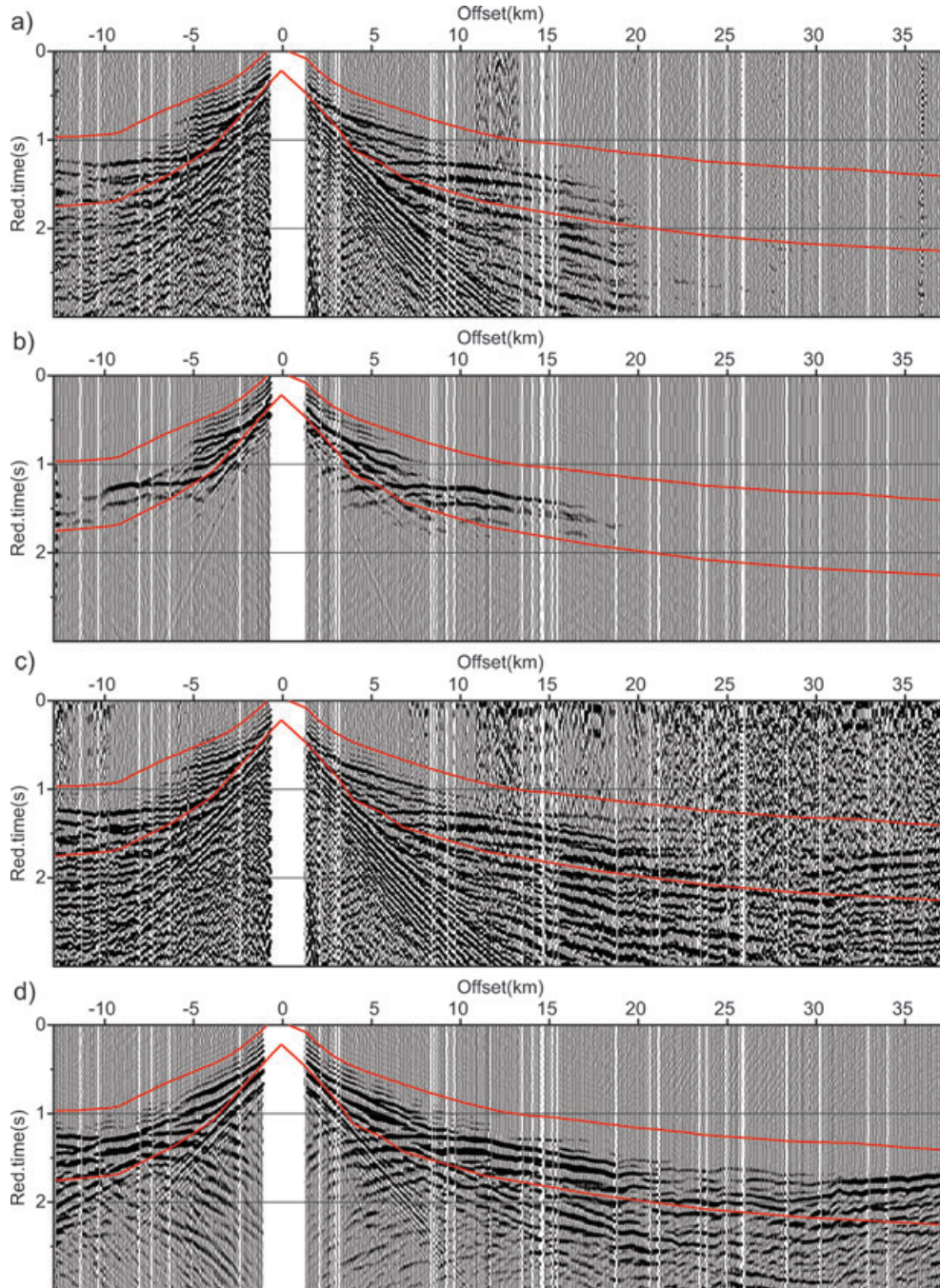

Figure 15. Comparison between real and synthetic seismograms for shot point 7. Data plotted using reduction velocity of 5500 m s^{-1}. (a) Non-normalized real data as pre-processed for visco-acoustic FWI (only the mute is not applied). (b) Visco-acoustic synthetic seismograms computed in the final visco-acoustic FWI model. (c) Normalized real data as pre-processed for acoustic FWI (only the mute is not applied). (d) Acoustic synthetic seismograms computed in the final acoustic FWI model.
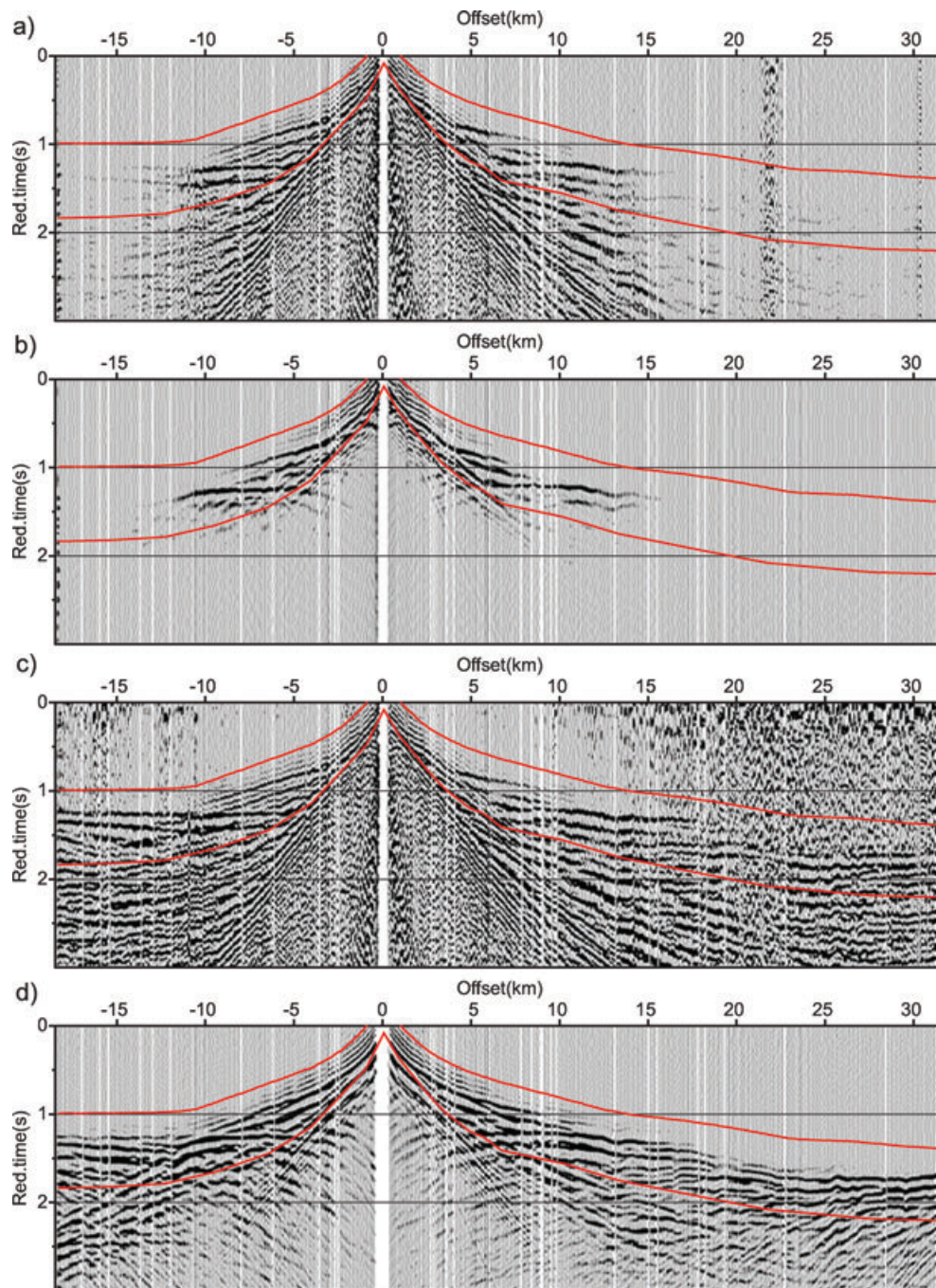
Figure 16. Same as Fig. 14 for shot point 11.

Source wavelets computed with the acoustic seismic modelling from the FAT velocity model and normalized data do not show satisfying focusing and repeatability (Fig. 14a). In contrast, source wavelets computed with the acoustic seismic modelling from the acoustic FWI velocity model and normalized data show improved repeatability, but remain poorly focused in time (Fig. 14b). When non-normalized data are used instead of normalized data to estimate the source wavelets with acoustic modelling, unstable estimation of the source wavelets is obtained (not shown here), that is consistent with the fact that the results of the acoustic FWI of the non-normalized data are quite noisy (Fig. 11c). Source wavelets computed with the visco-acoustic seismic modelling from the final visco-acoustic FWI model and non-normalized data are much more focused and looks like the desired minimum-phase output of the spectral whitening process (Fig. 14c). This supports the relevance of the visco-acoustic FWI model. Of note, the better focusing of the source wavelets inferred from the final visco-acoustic FWI model and non-normalized data (Fig. 14c) compared to that inferred from the acoustic FWI model and normalized data (Fig. 14b) can also be partly explained by the lower frequency content of the normalized data associated with longer-offset data content. We checked also the relevance of the FWI Q model by estimating the source from
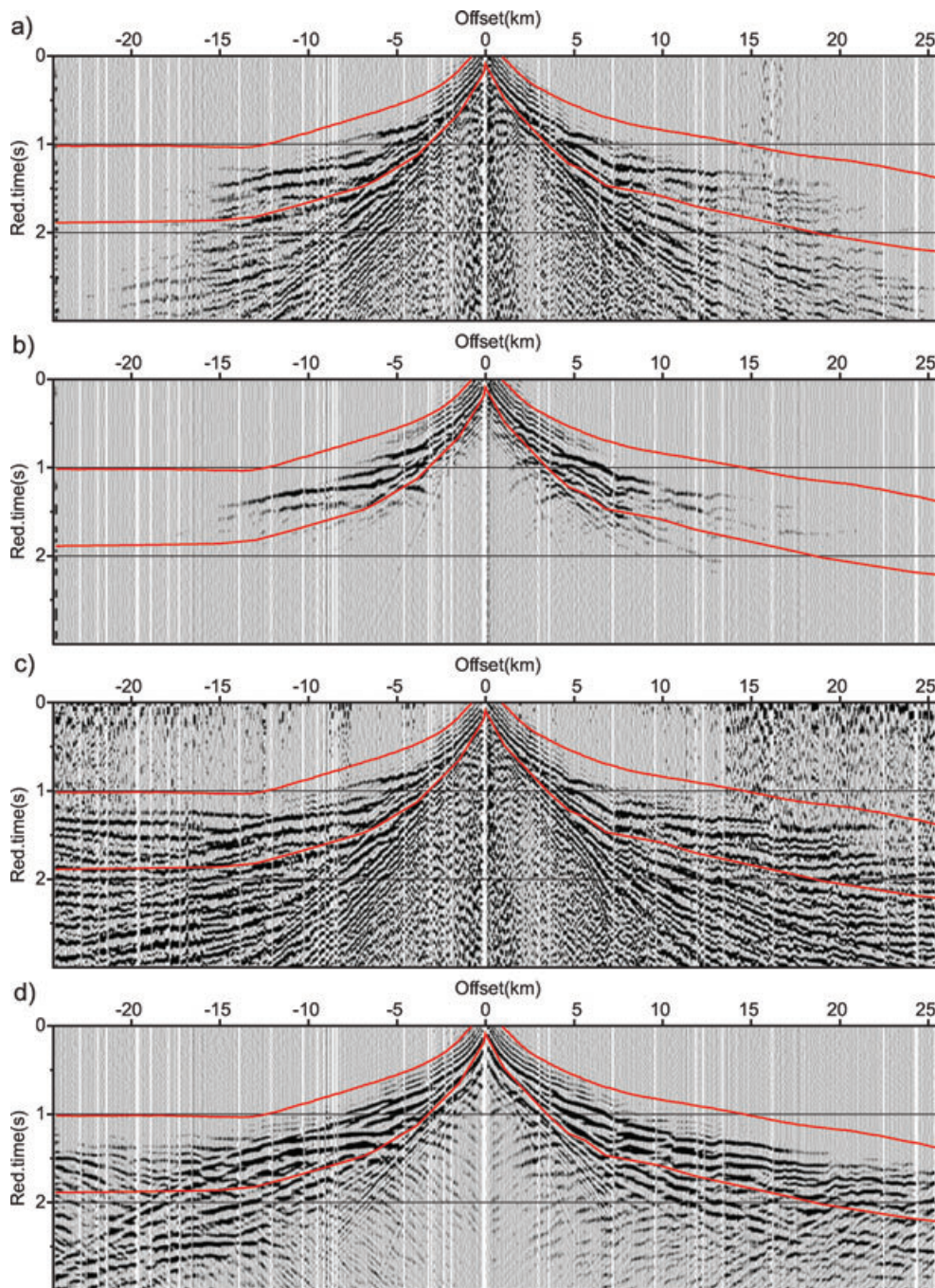
Figure 17. Same as Fig. 14 for shot point 16.

the final velocity model of the visco-acoustic FWI and the constant $Q = 50$ model instead of the recovered one. In such case, we do not obtain a consistent source signature between the shot points.

5.5.3 Time-domain synthetics

In order to test the validity of the obtained velocity and attenuation models, we calculated the time-domain synthetic seismograms using the FD frequency-domain method of Hustedt *et al.* (2004). The synthetic seismograms are computed for the source wavelet inferred from the final FWI models (Figs 14b–c). In Figs 15–18, we compare

the normalized real data with the acoustic synthetic data computed in the final acoustic FWI model (Fig. 11b) and the non-normalized real data with the visco-acoustic synthetic data computed in the final visco-acoustic FWI model (Figs 11e and f). In both cases, the match between the recorded and the modelled data is quite good for the arrivals involved in the inversion by the mute window. However, unlike the acoustic FWI, visco-acoustic FWI clearly allows one to match the true amplitudes of the data. Direct comparison between recorded seismograms and visco-acoustic synthetic seismograms allows for a more quantitative interpretation of the two sets of seismograms (Fig. 19). Both the kinematics and the dynamics of the early arriving phases are quite well matched. Similarly to
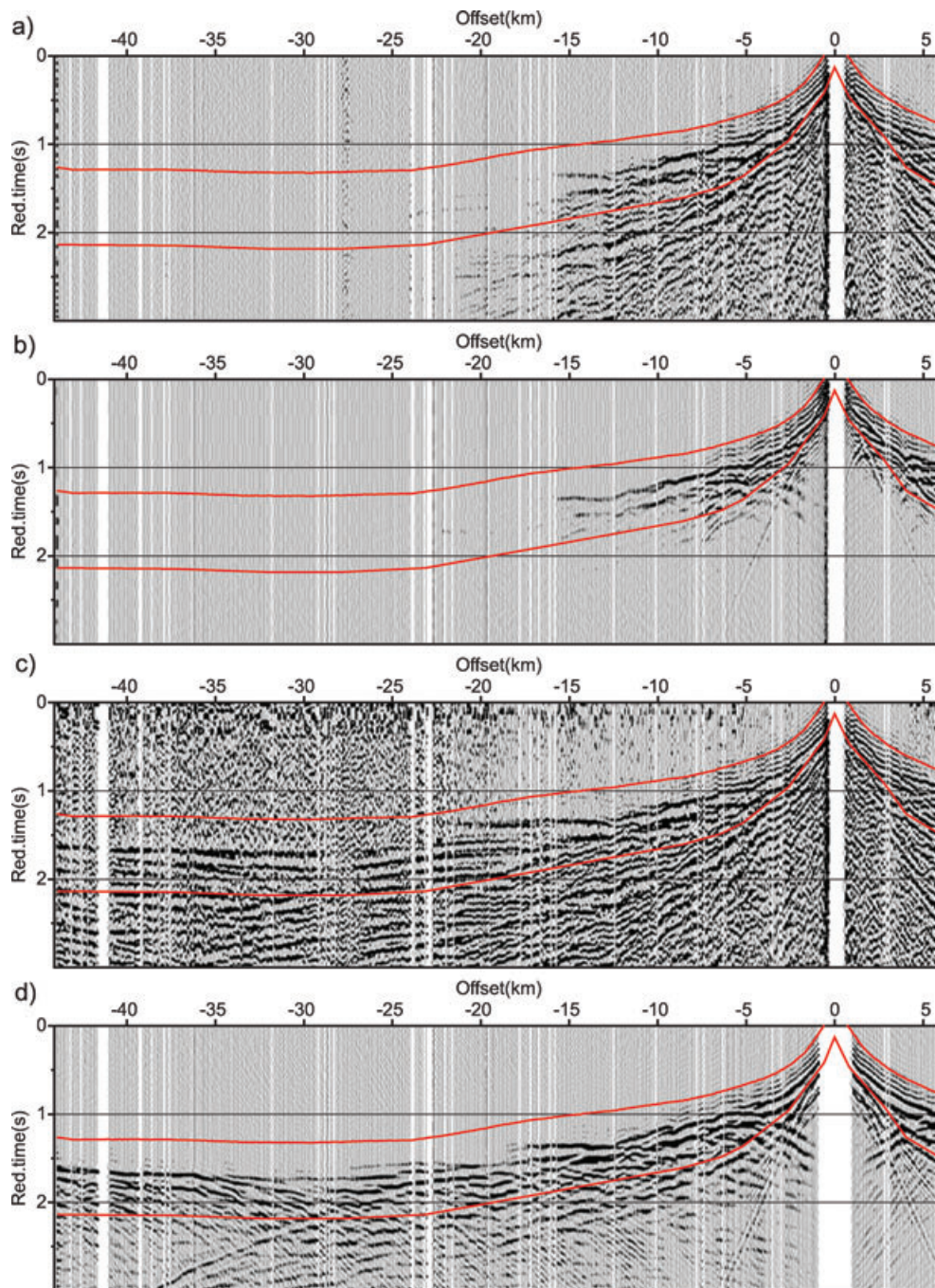
Figure 18. Same as Fig. 14 for shot point 28.

the wavelet estimation in the final visco-acoustic FWT results, we repeated the synthetic modelling replacing the final Q model with a constant $Q = 50$ model. In such case, the time-domain data fit is significantly degraded (not shown here).

5.5.4 Comparison with existing geological data and implications for petrophysical interpretation

Attenuation of seismic waves is potentially a diagnostic parameter to predict petrophysical properties of rock including reservoir properties like permeability. However, for such an interpretation a basic knowledge of the expected lithology is necessary, since

the correlation of attenuation with reservoir rock properties varies much between sandstones and carbonates (e.g. Best *et al.* 1994; Assefa *et al.* 1999; McCann & Sothcott 2009) and it depends on the intra-pore material as well (Klimentos & McCann 1990; Best *et al.* 1994).

Hence, here we bring some geological data in order to verify both the consistency of the obtained Q model with the expected lithology and to link the obtained c and Q models with the rock properties. We use well-log data from the Wilczna-1 borehole (Fig. 20) that is located in the middle of the G01 profile (offset by 3 km to the SE; Fig. 8) and the interpreted stratigraphic horizons provided by the industry (Fig. 20).
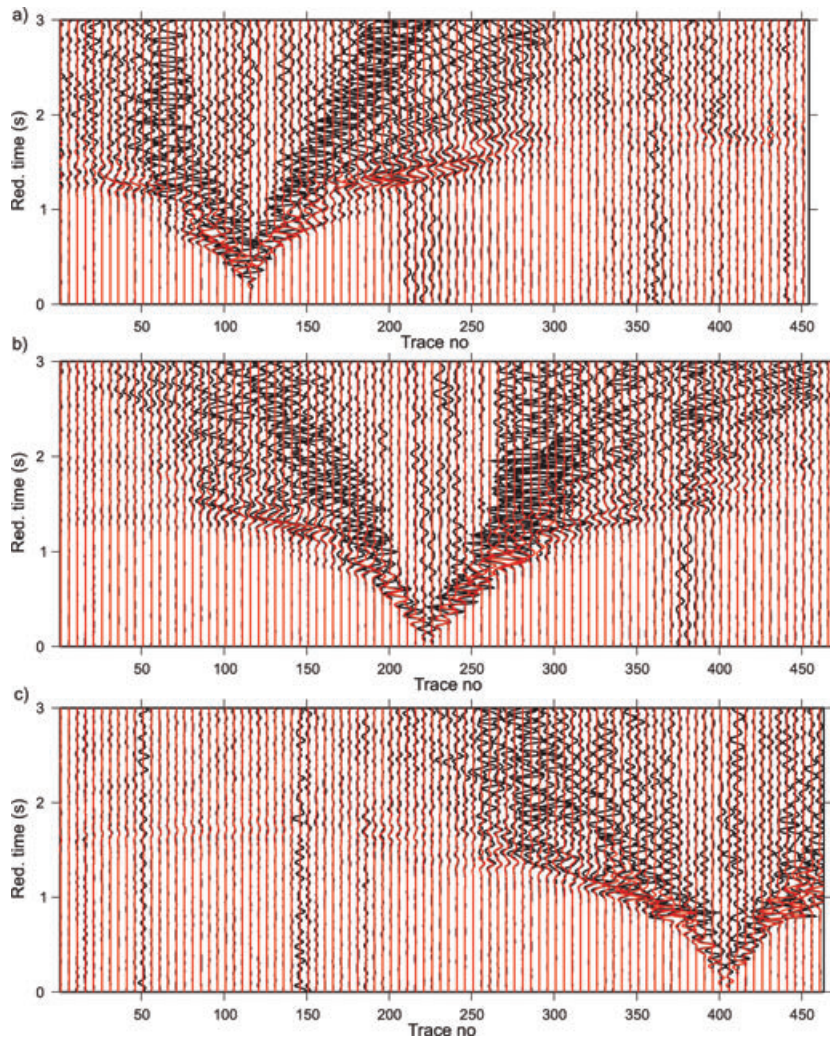
Figure 19. Direct comparison between recorded seismograms (black) and synthetic seismograms (red) computed in the final visco-acoustic FWI model. (a) Shot point 7. (b) Shot point 16. (c) Shot point 28.

Fig. 20 shows a comparison of the well-log data from the Wilczna-1 borehole and 1-D velocity and attenuation logs from the final acoustic and visco-acoustic FWI models (Figs 11b, e and f). There is a quite good match of the recovered velocities with the velocities from the well check-shot survey (Fig. 20a), with a better fit of the velocities from the visco-acoustic FWI than the acoustic one (especially in the shallower part). However, the visco-acoustic FWI failed to recover a low-velocity zone at *ca.* 1750 m depth, where the acoustic FWI performs better. The failure of the visco-acoustic FWI to image the deep part of the target can be related to the small weight of the non-normalized long-offset data in the misfit function resulting from strong attenuation effects. This small weight can make the FWI poorly conditioned for the reconstruction of the deep structure of the target. Again, the acoustic FWI performs slightly better in the deep part because the data normalization leads to a better data preconditioning with offset, as already mentioned. The mispositioning in depth of the low-velocity zone imaged in the acoustic FWI model can be also explained by the 3-km offset of the Wilczna-1 borehole from the seismic profile.

The correlation of the resistivity log (Fig. 20b) and the Q log is not very clear, although the two curves matches well in some places. For example, we show a very good correlation between attenuation and resistivity in the first 300 meters of the profile. A zone of high

resistivity between 600 and 750 m depth correlates well with a zone of weaker attenuation ($Q = 40$). The lowest $Q = 10$ at 1000 m depth close to the top of Upper Jurassic (J3) correlates well with a low resistivity zone. A zone of higher resistivity between 1300 and 1600 m depth correlates well with a zone of weaker attenuation ($Q = 40$). Finally, a resistive layer at 2700–3000 m depth can be correlated with a decrease in attenuation.

We used also the relation of Klimentos & McCann (1990, their eq. 12) that links Q values with the clay content in the pore space. A Q-log converted into the percentage of clay content is shown in Fig. 20(c) together with the clay content calculated from the gamma-ray log. Although the absolute values of the two curves are different, the peak in the clay volume at 1000 m depth are consistently resolved.

There is a good match of the stratigraphy and the velocity model (Fig. 21c) that was already described by Malinowski & Operto (2008). Noticeably, two high velocity layers are observed: in the Upper Jurassic (J3) and in the Middle/Lower Triassic (Tp) and both are due to the presence of carbonates. Transition to Zechstein is marked by a low-velocity zone.

The recovered Q model is also consistent with the geology and could be related to the lithological changes within the stratigraphic units (Figs 20a and b). The highest attenuation (i.e. the lowest Q
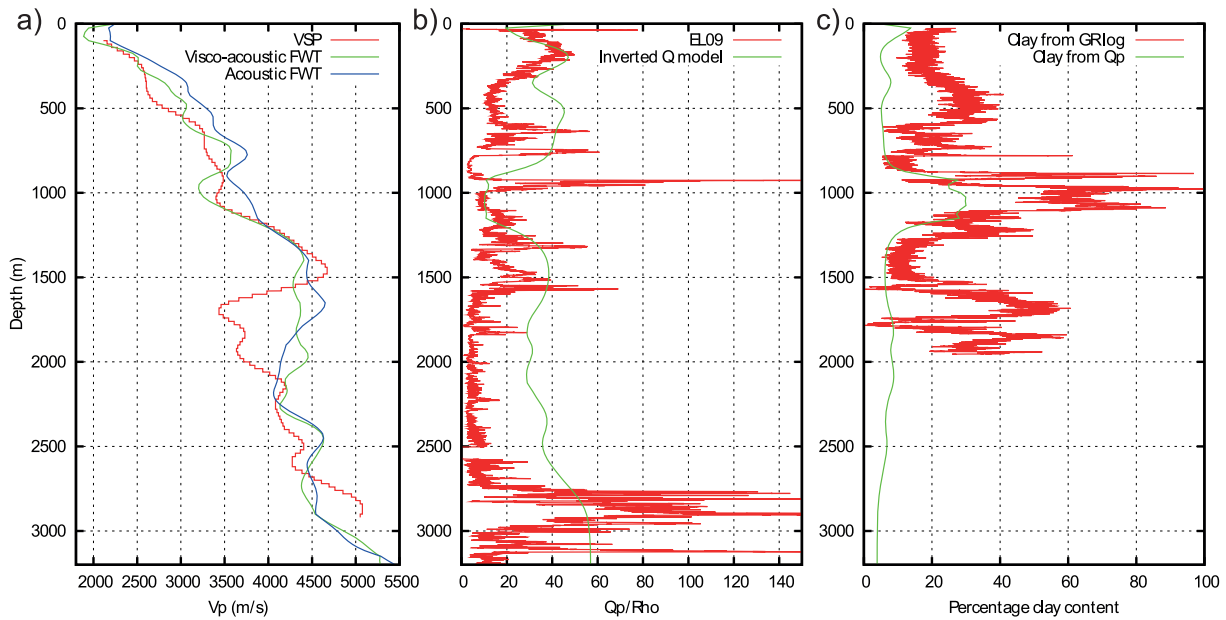
Figure 20. Comparison between the FWI results and well-logs from Wilczna-1 borehole. (a) Velocity profiles. (b) Comparison between resistivity log and *Q* profile. (c) Comparison between the clay volume calculated from the *Q* model according to the formula of Klimentos & McCann (1990) and from the gamma-ray log.

values, $Q < 25$) zones are associated with the transition from Cretaceous to Jurassic (K1/J3). Low Q values are also observed in the Lower Jurassic/Uppermost Triassic (J1/Tre). Those high attenuation zones are likely associated with the increased clay fraction in the clayey-sandstone facies characteristic for those stratigraphic units in the area (Marek & Pajchlowa 1997). The transition from Lower Triassic (Tp) to Zechstein (Z) is marked by a change from lower to higher attenuation that can be attributed to a change from the Buntsandstein sandstones to the clays characteristics for the top of Zechstein.

Relatively higher Q values ($Q > 75$) are observed at shallow depths (<1 km) in the eastern part of the profile within the Upper Cretaceous (K2). According to McCann & Sothcott (2009), $Q > 70$ together with velocity threshold ($3500 < c < 4500$ m s^{-1}) are characteristic for clean sandstones with good permeability. Some of the above high-Q zones fulfil this condition, however it is valid only for the sandstone lithology. In fact, the Upper Cretaceous in this area is developed in the carbonate facies rather than sands (Marek & Pajchlowa 1997). Hence, we associate these anomalies with the transition of the Campanian/Maastrichtian marly facies in the western part of the profile to the lime-silicate Maastrichtian facies to the east (Marek & Pajchlowa 1997).

Within the Triassic, a peak in Q corresponds to the transition from Keuper (Tk) to Muschelkalk (Tp). Again, it can be explained as the effect of the less attenuating carbonates (Muschelkalk limestones) as compared to the rocks with the increased clay content.

We can conclude that most of the Q anomalies present in our model are due to the change in the clay content, since the laboratory measurements provide strong correlation of Q with the volume content of intra-pore clay (Klimentos & McCann 1990; Best *et al.* 1994). No such correlation exists in case of Q versus porosity (Best *et al.* 1994). Provided we have a sandstone lithology, higher Q values ($Q > 70$) are indicating good permeability (Best *et al.* 1994; McCann & Sothcott 2009), but this is not the case here, as we link our low-attenuation zones with carbonate lithology. Shear

wave quality factors together with a shear wave velocity are necessary for determining the fluid content and type in the pore space (Klimentos 1995; Dvorkin & Mavko 2006). These parameters could be estimated in the future, when a visco-elastic FWI is applied to our data set.

Finally, it is worth to note that the stratigraphic horizons used for comparison with the FWI models are projected along the G01 line from a distance greater than 10 km, since no reflection seismic profiles exist in the middle of the G01 line. In fact, based on the FWI results we can refine this stratigraphic interpretation and re-pick these horizons, e.g. top of the Zechstein could be locally uplifted at ca. 30 km of the model distance.

6 CONCLUSIONS

We have presented a sensitivity analysis of the visco-acoustic FWI, where the compressional wave speed and the attenuation factor are jointly reconstructed. The relative sensitivity of the seismic data to the velocity and the attenuation factor is strongly dependent on the values of the attenuation in the medium. We have shown that both the velocity and the attenuation factor can be reliably reconstructed with a comparable resolution and without trade-off for sufficiently attenuating media and sufficiently accurate velocity starting model. Both the amplitude and phase differences of the partial derivative wavefields with respect to c and Q can help to reconstruct the two classes of parameters by non-linear inversion. The imaging of the low wavenumber of the attenuation from the low frequencies can be however unstable. Aggressive smoothing and damping of the gradient of the misfit function with respect to Q parameter is a pragmatical solution to regularize the Q imaging at low frequencies. Realistic synthetic tests have shown how the presence of attenuating near surface weathered layer can lead to erroneous reconstruction of the near surface velocities by acoustic FWI. The erroneous near surface velocities have in turn a strong impact on the reconstruction of the deep velocities. In contrast, visco-acoustic FWI successfully
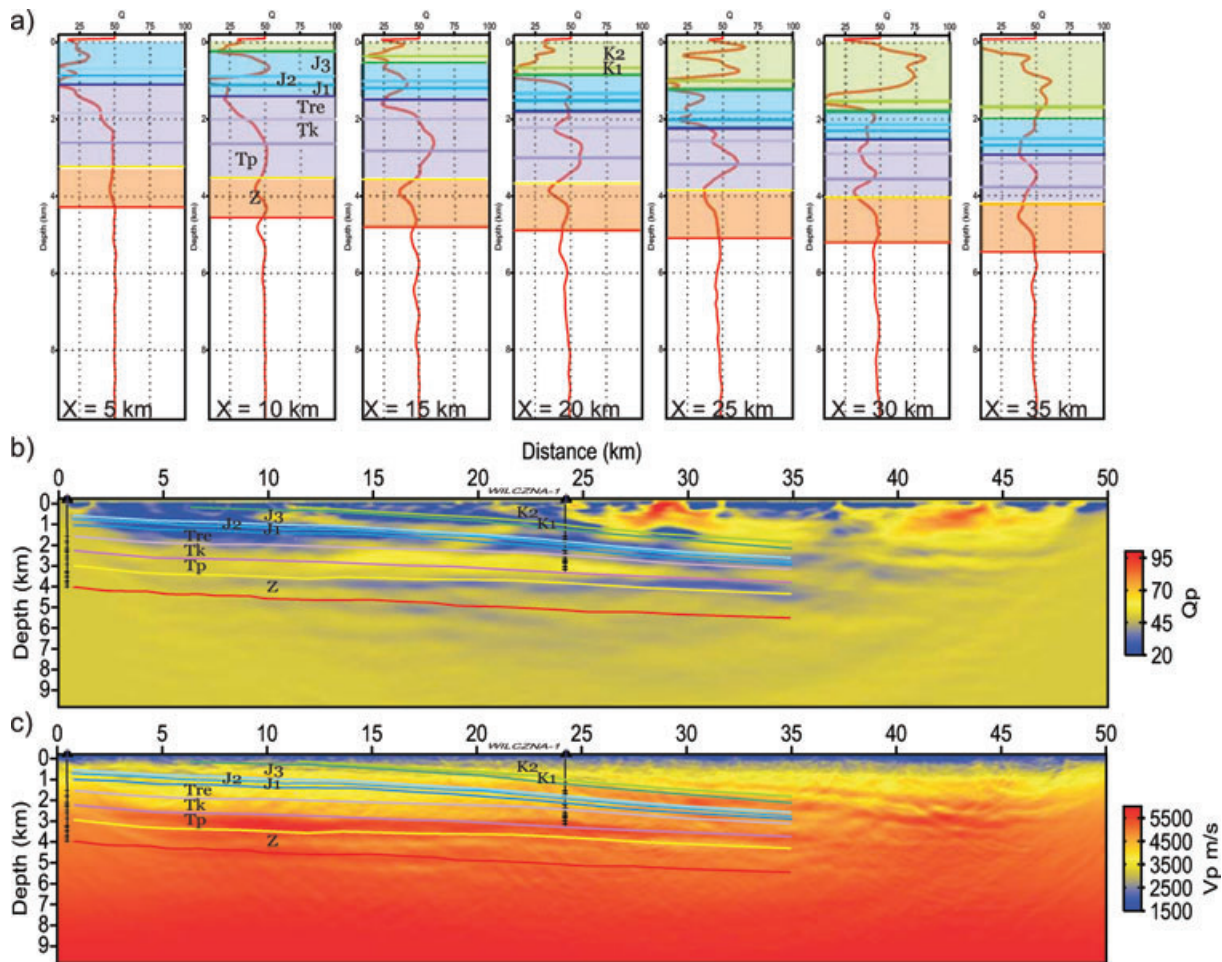
Figure 21. Correlation between stratigraphic horizons and the V_P and Q models. (a) 1-D logs from the recovered Q model extracted at 5 km interval. Stratigraphic horizons from the industry data overlaid on the recovered Q (b) and V_P model (c). Colouring denotes: green – Cretaceous, blue – Jurassic, violet – Triassic, orange – Zechstein; Stratigraphy follows: K2 – Middle Cretaceous, K1 – Lower Cretaceous, J3 – Upper Jurassic, J2 – Middle Jurassic, J1 – Lower Jurassic, Tre – Triassic (Rhaetian), Tk – Triassic (Keuper), Tp – Triassic (Scythian) and Z – Permian (Zechstein).

imaged c and Q in the near surface weathered layer, even when a constant Q starting model is used.

We applied both the acoustic and visco-acoustic FWI to real wide-aperture onshore seismic data recorded in the Polish Basin. This onshore data set shows a strong footprint of attenuation. We have shown how a heuristic normalization of the data with offset allows us to reconstruct a reliable velocity model in the acoustic approximation. Alternatively, visco-acoustic FWI allows us to reconstruct a reliable velocity model and Q model from the true-amplitude data. We have proposed a pragmatical approach based upon seismic modelling and source wavelet estimation to infer the best starting homogeneous Q model for visco-acoustic FWI. We have shown that the source wavelet estimation was quite sensitive to the quality of the velocity and attenuation model used for the estimation. Therefore, the source estimation should provide a good indicator of the relevance of the FWI models. Estimation of the source wavelet also allows us to compute more realistic synthetic seismograms which can be directly compared to the real data in order to check the kinematics and dynamics of the wavefield simulated in the FWI models.

We found the recovered velocity and attenuation models consistent with the expected lithology and stratigraphy in the study area. Most of the high-attenuation zones were interpreted as caused by the increased intra-pore clay content. However, correlation of the Q with the borehole resistivity log suggests that some of the highly attenuating zones could be linked with a presence of the mineralized fluids. Our detailed Q model could be potentially used in reflection seismic processing in order to compensate the attenuation effects by inverse-Q filtering techniques (Wang 2008), hence making the imaging of the sedimentary basins more effective.

REFERENCES

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**(6), 1990–2008.

Assefa, S., McCann, C. & Sothcott, J., 1999. Attenuation of P- and S-waves in limestones, *Geophys. Prospect.,* **47,** 359–392.

Berenger, J.-P., 1994. A perfectly matched layer for absorption of electromagnetic waves, *J. Computat. Phys.,* **114,** 185–200.

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.

Bleibinhaus, F. & Rondenay, S., 2009. Effects of surface scattering in full-waveform inversion, *Geophysics,* **74**(6), WCC87–WCC95.

Bleibinhaus, F., Hole, J.A., Ryberg, T. & Fuis, G.S., 2007. Structure of the California Coast Ranges and San Andreas Fault at SAFOD from seismic waveform inversion and reflection imaging, *J. geophys. Res.,* **112,** B06315, doi:10.1029/2006JB004611.

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., 2010. Two-dimensional frequency-domain visco-elastic full-waveform inversion: parallel algorithms, optimization and performances, *Comput. Geosci.*.

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

Brossier, R., Operto, S. & Virieux, J., 2009. Robust elastic frequency-domain full-waveform inversion using the L_1 norm, *Geophys. Res. Lett.,* **36,** 20 310.

Brossier, R., Operto, S. & Virieux, J., 2009. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion, *Geophysics,* **74**(6), WCC63–WCC76.

Brossier, R., Operto, S. & Virieux, J., 2010. Which data residual norm for robust elastic frequency-domain full waveform inversion? *Geophysics,* **75**(3), R37–R46.

Chavent, G., 1974. Identification of parameter distributed systems, in *Identification of Function Parameters in Partial Differential Equations*, pp. 31–48, eds Goodson, R. & Polis, M., American Society of Mechanical Engineers, New York, NY.

Chavent, G., 2009. *Nonlinear Least Squares for Inverse Problems*, Springer, Dordrecht.

Dvorkin, J.P. & Mavko, G., 2006. Modeling attenuation in reservoir and nonreservoir rock, *Leading Edge,* **25**(2), 194–197.

Gao, F., Levander, A.R., Pratt, R.G., Zelt, C.A. & Fradelizio, G.L., 2006. Waveform tomography at a groundwater contamination site: VSP-surface data set, *Geophysics,* **71**(1), H1–H11.

Hak, B. & Mulder, W.A., 2010. Migration for velocity and attenuation perturbations, *Geophys. Prospect.,* **58,** 939–951.

Hak, B. & Mulder, W.A., 2011. Seismic attenuation imaging with causality, *Geophys. J. Int.,* **184**(1), 439–451.

Hicks, G.J., 2002. Arbitrary source and receiver positioning in finite-difference schemes using kaiser windowed sinc functions, *Geophysics,* **67,** 156–166.

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**(2), 598–612.

Hustedt, B., Operto, S. & Virieux, J., 2004. Mixed-grid and staggered-grid finite difference methods for frequency domain acoustic wave modelling, *Geophys. J. Int.,* **157,** 1269–1296.

Jaiswal, P., Zelt, C., Dasgupta, R. & Nath, K., 2009. Seismic imaging of the Naga Thrust using multiscale waveform inversion, *Geophysics,* **74**(6), WCC129–WCC140.

Kamei, R. & Pratt, R.G., 2008. Waveform tomography strategies for imaging attenuation structure for cross-hole data, in *Rome 2008 – Leveraging Technology*, Proc. 70th EAGE Conf., Extended Abstracts, p. F019, EAGE Publications BV.

Klimentos, T., 1995. Attenuation of p- and s-waves as a method of distinguishing gas and condensate from oil and water, *Geophysics,* **60**(2), 447–458.

Klimentos, T. & McCann, C., 1990. Relationships among compressional wave attenuation, porosity, clay content, and permeability in sandstones, *Geophysics,* **55**(8), 998–1014.

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

Liao, Q. & McMechan, G.A., 1996. Multifrequency viscoacoustic modeling and inversion, *Geophysics,* **61,** 1371–1378.

Malinowski, M. & Operto, S., 2008. Quantitative imaging of the Permo-Mesozoic complex and its basement by frequency domain waveform tomography of wide-aperture seismic data from the Polish Basin, *Geophys. Prospect.,* **56**(6), 805–825.

Malinowski, M. *et al.*, 2007. Effective sub-Zechstein salt imaging using low-frequency seismics—results of the GRUNDY 2003 experiment across the Variscan front in the Polish Basin, *Tectonophysics,* **439,** 89–106.

Marek, S. & Pajchlowa, M., 1997. *The Epicontinental Permian and Mesozoic in Poland, Vol. CLIII of Prace Państwowego Instytutu Geologicznego*, Polish Geological Institute (in Polish with English summary).

Marfurt, K., 1984. Accuracy of finite-difference and finite-elements modeling of the scalar and elastic wave equation, *Geophysics,* **49,** 533–549.

McCann, C. & Sothcott, J., 2009. Sonic to ultrasonic q of sandstones and limestones: laboratory measurements at in situ pressures, *Geophysics,* **74**(2), WA93–WA101.

Mulder, W. & Hak, B., 2009. An ambiguity in attenuation scattering imaging, *Geophys. J. Int.,* **178**(3), 1614–1624.

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

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,** B09306, doi:10.1029/2005JB003835.

Plessix, R.-E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.,* **167**(2), 495–503.

Plessix, R.E., 2009. Three-dimensional frequency-domain full-waveform inversion with an iterative solver, *Geophysics,* **74**(6), WCC53–WCC61.

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., Song, Z.M. & Warner, M., 1996. Two-dimensional velocity models from wide-angle seismic data by wavefield inversion, *Geophys. J. Int.,* **124,** 323–340.

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., Hou, F., Bauer, K. & Weber, M., 2005. Waveform tomography images of velocity and inelastic attenuation from the Mallik 2002 crosshole seismic surveys, *GSC Bulletin,* **585,** 1–14.

Rao, Y. & Wang, Y.H., 2008. The strategies for attenuation inversion with waveform tomography, in *Rome 2008 – Leveraging Technology*, Proc. 70th EAGE Conf., Extended Abstracts, EAGE Publications BV.

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.

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.

Shipp, R.M. & Singh, S.C., 2002. Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data, *Geophys. J. Int.,* **151,** 325–344.

Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics,* **69**(1), 231–248.

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

Smithyman, B., Pratt, R.G., Hayles, J. & Wittebolle, R., 2009. Detecting near-surface objects with seismic waveform tomography, *Geophysics,* **74**(6), WCC119–WCC127.

Song, Z. & Williamson, P., 1995. Frequency-domain acoustic-wave modeling and inversion of crosshole data. Part 1: 2.5-D modeling method, *Geophysics,* **60**(3), 784–795.

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**(3), 786–809.

Sourbier, F., Operto, S., Virieux, J., Amestoy, P. & L'Excellent, J.-Y., 2009a. Fwt2d: a massively parallel program for frequency-domain full-waveform tomography of wide-aperture seismic data—part 1: algorithm, *Comput. Geosci.,* **35**(3), 487–495.

Sourbier, F., Operto, S., Virieux, J., Amestoy, P. & L'Excellent, J.-Y., 2009b. Fwt2d: a massively parallel program for frequency-domain full-waveform tomography of wide-aperture seismic data—part 2: numerical examples and scalability analysis, *Comput. Geosci.,* **35**(3), 496–514.

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

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

Toksöz, M.N. & Johnston, D.H., 1981. *Geophysics Reprint Series, No. 2: Seismic Wave Attenuation*, SEG, Tulsa, OK.

Virieux, J. & Operto, S., 2009. An overview of full waveform inversion in exploration geophysics, *Geophysics,* **74**(6), WCC127–WCC152.

Wang, Y., 2008. *Seismic Inverse Q Filtering*, Blackwell Publishing, Malden, WA.