

Anisotropic 3D full-waveform inversion

Michael Warner¹, Andrew Ratcliffe², Tenice Nangoo¹, Joanna Morgan¹, Adrian Umpleby¹, Nikhil Shah¹, Vetle Vinje², Ivan Štekl¹, Lluís Guasch¹, Caroline Win², Graham Conroy², and Alexandre Bertrand³

ABSTRACT

We have developed and implemented a robust and practical scheme for anisotropic 3D acoustic full-waveform inversion (FWI). We demonstrate this scheme on a field data set, applying it to a 4C ocean-bottom survey over the Tommeliten Alpha field in the North Sea. This shallow-water data set provides good azimuthal coverage to offsets of 7 km, with reduced coverage to a maximum offset of about 11 km. The reservoir lies at the crest of a high-velocity antiformal chalk section, overlain by about 3000 m of clastics within which a low-velocity gas cloud produces a seismic obscured area. We inverted only the hydrophone data, and we retained free-surface multiples and ghosts within the field data. We invert in six narrow frequency bands, in the range 3 to 6.5 Hz. At each iteration, we selected only a subset of sources, using a different subset at each iteration; this strategy is more efficient than inverting all the data every iteration. Our

starting velocity model was obtained using standard PSDM model building including anisotropic reflection tomography, and contained epsilon values as high as 20%. The final FWI velocity model shows a network of shallow high-velocity channels that match similar features in the reflection data. Deeper in the section, the FWI velocity model reveals a sharper and more-intense low-velocity region associated with the gas cloud in which low-velocity fingers match the location of gas-filled faults visible in the reflection data. The resulting velocity model provides a better match to well logs, and better flattens common-image gathers, than does the starting model. Reverse-time migration, using the FWI velocity model, provides significant uplift to the migrated image, simplifying the planform of the reservoir section at depth. The workflows, inversion strategy, and algorithms that we have used have broad application to invert a wide-range of analogous data sets.

INTRODUCTION

Full-waveform inversion (FWI) is a technique that seeks to find a high-resolution high-fidelity model of the subsurface that is capable of matching individual seismic waveforms, within an original raw field data set, trace by trace. The method begins from a best-guess starting model which is then iteratively improved using a sequence of linearized local inversions to solve a fully nonlinear problem. In principle, FWI can be used to recover any physical property that has an influence upon the seismic wavefield, but in practice the technique has been used predominantly to recover P-wave velocity.

The underlying theory is well established, but its practical application to 3D field data sets in a form that is simultaneously efficient, effective, and robust, is still a subject of intense research. [Virieux and Operto \(2009\)](#) present a recent review of waveform inversion with an extensive bibliography, and [Pratt et al. \(1998\)](#) and [Pratt \(1999\)](#) provide an overview of the development of the underlying theory using a formulation similar to that presented here. [Pratt and Shipp \(1999\)](#) presented an early application on crosshole field data in 2D with modeling and inversion carried out in the frequency domain, and [Shipp and Singh \(2002\)](#) presented an application to surface data in 2D implemented in the time domain.

Parts of this paper were first presented at: 81st Annual International Meeting, SEG, as “Full Waveform Inversion: A North Sea OBC case study.” Manuscript received by the Editor 20 August 2012; revised manuscript received 16 November 2012; published online 1 March 2013.

¹Imperial College London, Department of Earth Science and Engineering, South Kensington campus, London, U. K. E-mail: m.warner@imperial.ac.uk; tenice.nangoo08@imperial.ac.uk; j.morgan@imperial.ac.uk; a.umpleby@imperial.ac.uk; nikhil.shah09@imperial.ac.uk; i.stekl@imperial.ac.uk; l.guasch08@imperial.ac.uk.

²CGGVeritas Services (UK) Limited, Crawley, West Sussex, U. K. E-mail: andrew.ratcliffe@cggveritas.com; vetle.vinje@cggveritas.com; caroline.win@cggveritas.com; graham.conroy@cggveritas.com.

³ConocoPhillips Norge, Tananger, Norway. E-mail: alexandre.bertrand@conocophillips.com.

© 2013 Society of Exploration Geophysicists. All rights reserved.

Ben-Hadj-Ali et al. (2008), Sirgue et al. (2008), and Warner et al. (2008) demonstrated the first practical applications of FWI in 3D, and Sirgue et al. (2010) and Plessix and Perkins (2010) have presented recent applications of FWI to 3D field data sets.

Because FWI honors the physics of finite-frequency wave propagation, its spatial resolution is limited only by the source and receiver distribution, the noise level, and the local seismic wavelength. In contrast, methods that are based upon traveltimes and therefore implicitly upon simplified wave propagation have a spatial resolution that is limited also by the size of the local Fresnel zone (Williamson, 1991). In practice, this means that FWI models are almost always better spatially resolved than are equivalent models generated by more-conventional methods.

In a commercial context, 3D FWI is most commonly used to recover high-resolution P-wave velocity models within heterogeneous overburden above deeper reservoirs. These shallow velocity models are then combined with conventional prestack depth migration (PSDM) to improve the imaging of the underlying reservoir using subcritical reflection data. Often, the spatial resolution and complexity of the FWI-recovered velocity model require that a high-fidelity PSDM scheme is used for the migration — typically this will be a reverse-time migration (RTM) scheme based upon the full two-way wave equation. Unlike conventional reflection imaging, FWI typically uses wide-angle refracted arrivals to build its velocity model (Pratt et al., 1996). Because FWI uses a purely local and not a global inversion scheme, low frequencies in the field data are normally essential for robust and effective inversion (Sirgue, 2006).

Because FWI seeks to match the observed data in detail, it must, as a minimum, be capable of generating models that can match the observed kinematics of the major seismic arrivals. In many petroleum-related data sets, some form of anisotropy is required to match the kinematics exactly, especially when the field data include near-normal-incidence reflections and wide-angle refracted arrivals, and when a wide range of azimuths and offsets are contained within the data set. Prieux et al. (2011), and Plessix and Cao (2011), discuss some of the difficulties associated with anisotropic FWI.

Here, we demonstrate the application of FWI to a 3D, high-density, full-azimuth, long-offset, field data set that displays significant P-wave anisotropy (Ratcliffe et al., 2011). We show that the resulting anisotropic FWI velocity model successfully predicts the raw field data, improves the match to available wells, improves the flatness of common image gathers, and successfully migrates the field data leading to improved continuity and significantly simplified images at reservoir depths.

Within this paper, we explain our workflow and inversion strategy, we describe how we have preprocessed and selected the field data, and we demonstrate how to design the starting model and the source wavelet. We also explain how we are able to speed up the computation without compromising the quality of the inversion, and we discuss some of the pitfalls that can arise during the application of 3D FWI to field data. Elsewhere, in related publications, we discuss the technical implementation of the anisotropic modeling and inversion code that we have used (Umpleby et al., 2010), we demonstrate the methodology that we used to provide quality assurance during FWI (Shah et al., 2012), we show how FWI can be extended so that it is able to provide velocity models directly at reservoir depths from this data set (Nangoo et al., 2012), and we use FWI over an extended frequency range to improve the velocity

model used for PS reflection imaging within a seismic obscured area (Vinje et al., 2012).

FULL-WAVEFORM INVERSION

Theory

The theory that underpins FWI has been derived many times using a variety of formulations. We present here only the results that are required to understand how our computer codes operate and to demonstrate the approximations that we have made. For a more complete development of the theory, the reader is referred to Pratt et al. (1998) and Virieux and Operto (2009), and references therein.

FWI begins from the assumption that we can solve a numerical wave equation of the form

$$\mathbf{G}(\mathbf{m}) = \mathbf{p}, \quad (1)$$

where \mathbf{m} is a column vector that contains the set of parameters that describe the subsurface model, \mathbf{p} is a column vector that contains the predicted seismic wavefield at all points within the model, and \mathbf{G} is a function that describes how to calculate the seismic data given the model. For the data set described in this paper, \mathbf{m} contains about seven million slowness values that define the velocity model on a regular Cartesian grid, \mathbf{p} represents about 50 GB of data for each source, or about 70 TB for the full data set after decimation, and \mathbf{G} is a time-domain finite-difference computer program that implements the 3D anisotropic acoustic wave equation, and that also contains a definition of the acquisition geometry and the source wavefield.

FWI is an algorithmic approach that uses the repeated application of the forward problem expressed in equation 1 to solve the nonlinear inverse problem that can be expressed as

$$\mathbf{G}^{-1}(\mathbf{d}) = \mathbf{m}', \quad (2)$$

where now \mathbf{d} contains the observed field data and \mathbf{m}' is the model that we are trying to discover. The inverse of \mathbf{G} in equation 2 expresses the idea that, if the resulting model \mathbf{m}' is placed back into equation 1, then the data set \mathbf{p}' that equation 1 predicts at the receivers will be the original observed data set \mathbf{d} . In practice, although computer codes can be written to provide accurate and explicit solutions to equation 1, it is not possible to write an explicit solver for equation 2. In part, this is because \mathbf{G} is not a matrix; it is a nonlinear function for which there is no formal inverse. Instead, we proceed by using an approximate solution to equation 2, and then seek to improve upon this solution by iteration.

We begin from a starting model \mathbf{m}_0 , and assume that this is reasonably close to the true model. We define the residual data set $\delta\mathbf{d}$ to be the difference between the predicted and observed data, that is $\delta\mathbf{d} = \mathbf{p}' - \mathbf{d}$; clearly the residual data set will be a function of the starting model. The problem is now to find a correction $\delta\mathbf{m}$ to the starting model, to generate a new model $\mathbf{m} = \mathbf{m}_0 + \delta\mathbf{m}$, which reduces the size of the residual data set $\delta\mathbf{d}$ toward zero.

We define a scalar quantity f , variously called the “misfit,” “objective function,” or “functional,” which is equal to the sum of the squares of the residual data set $\delta\mathbf{d}$. The inversion problem then is to find the $\delta\mathbf{m}$ that minimizes the value of f . If the problem is linearized by assuming a linear relationship between small changes in the model and corresponding changes in the data residuals, then the solution for $\delta\mathbf{m}$, that is the change that must be made to the starting model, is well known to be

$$\delta\mathbf{m} = -\mathbf{H}^{-1} \frac{\partial f}{\partial \mathbf{m}}. \quad (3)$$

In equation 3, \mathbf{H} is the square symmetric Hessian matrix containing all the second-order differentials of the functional f with respect to all combinations of the model parameters, and $\partial f / \partial \mathbf{m}$ is the gradient of the functional with respect to each of the model parameters, commonly just called the “gradient.”

The gradient is a vector of the same size as the model \mathbf{m} , so for this study it has about seven million elements. Using the adjoint method first introduced in this context by Tarantola (1984), the gradient is straightforward to compute using code that can solve equation 1. The Hessian, however, is much larger; for this study it contains about 5×10^{13} elements, and although it might be within computational reach to calculate these, it is not normally feasible to find the inverse which is what is required to solve equation 3 directly. Consequently, here we retain only an approximation to the diagonal elements of \mathbf{H} which then becomes trivial to invert.

It is straightforward to calculate the approximate relative magnitude of the diagonal elements of \mathbf{H} using spatial preconditioning. This leads to a final expression for the update that must be applied to each of the model parameters m_i

$$\delta m_i = -\frac{\alpha}{h_{ii}} \frac{\partial f}{\partial m_i}. \quad (4)$$

Here, α is a global scaling factor, commonly called the “step length,” h_{ii} is the approximate relative magnitude of the diagonal element of the Hessian that corresponds to the model parameter m_i , and $\partial f / \partial m_i$ is the corresponding component of the gradient for that model parameter.

To apply equation 4 to a real data set, we need a starting model, a source wavelet, and a computational implementation of equation 1. We use the method of Tarantola (1984) to calculate the gradient. This involves the forward propagation of the source wavefield for each source within the starting model to produce a predicted data set, the formation of the residual data by subtracting the predicted and observed data sets sample-by-sample at each receiver, the back propagation of this residual wavefield from the receivers, and the crosscorrelation of these forward and backward wavefields in time at each point within the interior of the model to form the gradient for each source. These individual gradients are then stacked together to form the global gradient required in equation 4.

Those familiar with RTM will recognize similarities between this portion of FWI and RTM. The key differences are that it is the residual wavefield and not the entire wavefield that is back propagated in FWI, and that the crosscorrelation imaging condition is different in detail between the two methods — in FWI the imaging is typically into velocity space (or more correctly in our case into slowness), whereas RTM is designed to generate a reflectivity image.

We use the method of Pratt et al. (1998, equation 12) to compute the step length α . This involves perturbing the starting model by a small amount in the appropriate direction, and observing the effect that this perturbation has upon the residual wavefield. By assuming a linear relationship between the model and data perturbations, we can then calculate the optimal amount by which to perturb the model to minimize the residual data set.

We use spatial preconditioning to approximate the diagonal elements h_{ii} of the Hessian matrix using the pseudo-Hessian following

the approach of Shin et al. (2001). The observed seismic data are typically more strongly influenced by some parameters within a model than by others. That is, the magnitudes of $\partial \mathbf{d} / \partial m_i$ tend to be much larger for model parameters that, for example, represent parts of the model that are close to sources and receivers or where the incident wavefield is large, than for parameters that represent regions of the model that are far from sources and receivers or where the incident wavefield is small. The diagonal of the Hessian compensates for these effects; we approximate it using the sum of the squares of the amplitude of the forward wavefield suitably stabilized to avoid dividing by small values. Spatial preconditioning then acts to boost the gradient where the incident wavefield is small, and reduce it where the wavefield is large.

Use of equation 4 involves some approximations. We therefore iterate the process, seeking to make a series of stepwise linearized updates to the starting model which move ever closer to the true model. Although we make use of a linearized relationship in the inversion, because the full nonlinear forward modeling expressed by equation 1 is used throughout, the iterated inversion scheme is able to solve the full nonlinear inversion problem correctly.

In practice, various enhancements are additionally incorporated into the basic inversion scheme described above. These include processing, selection, and weighting of the field data, processing of the predicted data, processing of the residual data, processing of the gradient, modification of the functional, and varying the approximation to the Hessian. Various steps must also be taken to stabilize, regularize, and otherwise make the inversion scheme more robust. Where such modifications have important implications for the quality of the FWI result, we outline them in later sections.

In the far-field, that is, at distances of more than a wavelength or two from sources and receivers, the maximum spatial resolution that can be obtained is locally around half the shortest seismic wavelength. This maximum is achievable in practice in FWI provided that the model is locally well sampled by the wavefield in three dimensions. For the data analyzed below, the maximum frequency used during FWI is 6.5 Hz, so that we should be able to resolve the velocity model locally to a resolution of about 115 m when the background velocity is around 1500 m/s, and proportionately less well as the velocity rises. In the near-field however, where evanescent waves become important, the resolution can be better than this. Here, the resolution depends not upon the wavelength directly, but upon the distance to the nearest source or receiver, a property that is used in high-resolution optical microscopy (Betzig et al., 1991). In the present context, this means that resolution may improve close to the seafloor where it will be controlled by the acquisition geometry rather than the local seismic wavelength.

At all stages, FWI only solves a local inversion problem. Its successful application therefore requires that the starting model is sufficiently close to the true model that the nearest minimum in the objective function is the global minimum rather than some local minimum. An inversion that leads to a local minimum will result in a model that will often be no better, and may be worse, than the starting model, but it will nevertheless lead to a reduction in the objective function and it will generate synthetic data that match the field data more closely than did the starting model. Effective and robust FWI therefore requires a means to ensure that the starting model is adequate given the data that are available, and that inversion does indeed proceed toward the global minimum. We have applied a robust quality-assurance process during this study

which is described more fully in [Shah et al. \(2012\)](#). We do this by extracting the wrapped phase of the amplitude-normalized residual wavefield, trace by trace, windowed in time on the dominant arrivals, for the lowest usable frequency in the field data. The spatial variation of this phase residual in common-receiver gathers is diagnostic of the existence of cycle skipping in the starting model, and its evolution between iterations is directly diagnostic of the detrimental effects of cycle skipping during the inversion.

Implementation

At the heart of any FWI code lies the efficient solution of equation 1; that is, given a model and a source, calculate the resultant seismic wavefield everywhere within the model. We have developed 3D finite-difference codes, and associated FWI infrastructure, that can solve equation 1 in the time domain or the frequency domain ([Umpleby et al., 2010](#)), using the acoustic or elastic wave equation ([Guasch et al., 2012](#)), with or without anisotropy ([Štekl et al., 2010](#)), with or without attenuation, although we cannot yet combine all these aspects into the same calculation.

In this study, we have used an explicit time-domain finite-difference time-stepping algorithm to solve the 3D anisotropic acoustic wave equation on a regular cubic mesh. The code allows for spatially variable, tilted transversely isotropic (TTI) anisotropy. To do this, we solve two coupled second-order differential equations; these differ in their exact details, but are similar in principle to the equations derived by [Alkhalifah \(2000\)](#) and [Duvaneck et al. \(2008\)](#). The equations describe two wavefields — one defines the observable acoustic pressure, and the other defines a nonphysical wavefield that acts to allow TTI kinematics within the pressure wavefield to be simulated correctly. Note that acoustic TTI media are a nonphysical abstraction, and so there is no dynamically correct formulation for them. Because we solve two scalar wave equations, the CPU and memory requirements of the TTI anisotropic code are about twice that of a simple isotropic acoustic FWI code.

Within this scheme, anisotropy is parameterized using five parameters. These are: the P-wave velocity parallel to the local symmetry axis, the two angles that define the local orientation of the symmetry axis in space, and the two Thomsen parameters ([Thomsen, 1986](#)), delta and epsilon, that control the variation of P-wave velocity away from the symmetry axis. Provided that the last four anisotropy parameters do not vary significantly over small distances, and provided that sources and receivers are located only within isotropic (or strictly only within elliptically anisotropic) regions, the code is robust, stable, and accurate.

The code is nominally second-order in time and fourth-order in space, but it uses an optimized 53-point finite-difference stencil within a $5 \times 5 \times 5$ cube that uses all face-centered, edge-centered, and corner grid cells to give a performance that is approximately equivalent to fourth-order in time and sixth-order in space. The code is accurate at five grid points per wavelength and above, and its maximum phase-velocity errors remain below 0.5% in all directions at four grid points per wavelength. In practical applications for FWI, we run the code such that the coverage is at least five grid points per wavelength over most of the model, but may be reduced to 4.5 grid points per wavelength in localized low-velocity regions; for example, in a shallow water layer or within the local intense center of a gas cloud.

Sources and receivers can be located at their true locations anywhere within the model, not only at grid cells. We use a 3D version

of the 2D interpolation scheme from [Hicks \(2002\)](#) to do this. We used a free-surface boundary condition on the top of the model in this study, and absorbing boundaries elsewhere. For 3D FWI, the principal requirement of absorbing boundary conditions is that they should be computationally efficient, and especially that they should not require significant extension of the computational domain outside the area of direct interest. In this code, therefore, we use a non-reflective boundary condition that is only two cells wide. This boundary condition applies a simplified one-way wave equation to predict the wavefield beyond the boundary, using this predicted wavefield to populate a conventional two-way finite-difference stencil at the boundary. FWI is robust against back-scattering from less-than-perfect boundaries. At worst, such imperfections lead only to a modest increase in incoherent noise within the interior of the model and to spurious structure close to the boundary itself where the inversion should in any case be regarded as unreliable because of poor data coverage.

The inversion code is designed to run in parallel on a large cluster of networked multicore compute nodes under the control of a single master node. In the simplest configuration, one source runs on each compute node, and P-threads are used to spread the computation across all available local cores. Parallelization on the cluster is across sources — that is, different sources run on different nodes — and we do not distribute the model as subdomains across nodes. If there are too few nodes available, then several sources can be run in parallel on a single node. When this is also insufficient because of limited memory or limited numbers of cores per node, then the individual nodes can be reused to compute multiple sources sequentially, although the latter will necessarily lead to increased disk or network traffic.

So far as is possible, everything is held in memory during FWI, and where that is not possible, intermediate results are either written to local disk on each node, or recomputed when required depending upon which is most efficient. Typically, a node will hold the field data for one source in local memory; it will compute, compress, and store the forward wavefield, then compute, store, and back-propagate the residual wavefield, crosscorrelating this with the regenerated forward field to form the local gradient as it does so. The wavefields are discarded as the crosscorrelation proceeds. During forward propagation, the diagonal of the approximate Hessian is also calculated using spatial preconditioning.

Once the gradient has been computed for a single source on a compute node, that gradient and the spatial preconditioner are sent via the network to the master node. Here, the local gradients from all sources from all nodes are stacked together to form a global gradient. The local preconditioners are also stacked to form a global preconditioner which is used to scale the global gradient. The preconditioned global gradient is then sent back to each of the compute nodes. The compute nodes use this to calculate an optimal step length for their respective sources, and they pass these values back to the master node which combines them to form a global step length. The global step length is then sent back to the compute nodes, which use this to scale the global gradient and update the local model; each node then moves to the next iteration. In this way, once the initial data and model have been distributed, the only significant information that needs to be passed across the cluster is the gradient and the preconditioner, which for the data set inverted here represent only about 60 MB. When there is sufficient hardware available then, the FWI code involves little network communication

beyond that required for the definition of the problem and the initial distribution of the field data to be inverted.

Some enhancements help to speed up the code. The computational domain grows and shrinks during the modeling so that wavefields are only computed within those regions in which they are required. We do not require all nodes to complete their computations to move forward, so that one or two tardy nodes will not compromise the efficiency of the whole cluster. We have found that the time taken to access local memory, rather than the number of floating-point operations, is often the key factor that controls compute efficiency within these memory-intensive codes. Consequently, we access and store data so far as is possible in a way that maps efficiently onto the local memory cache. We also find that it is often faster to recompute earlier results than it is to store them and recover them from memory — this applies to the density model and to the finite-difference stencil, both of which we recompute as we need them.

TOMMELITEN FIELD DATA SET

Acquisition

The field data set for this study was taken from the Tommeliten Alpha field in the North Sea. Tommeliten Alpha is a gas condensate discovery located 25 km southwest of the Ekofisk field in the Norwegian North Sea, Block 1/9. The reservoir consists of two fractured chalk formations, Ekofisk and Tor, situated at the crest of a broad anticline, approximately 3000 m below the surface. A large part of the reservoir is located in a seismic-obscured area, caused by the presence of gas in the overlying section of interbedded silt and sandstone within the 1000–2000 m depth range (Granli et al., 1999).

A high-density, full-azimuth, 3D, 4C, ocean-bottom-cable survey was acquired in 2005 with the aim of improving images of the reservoir beneath the gas cloud. The data were acquired using three swaths, each composed of eight parallel cables, in water depths of around 75 m (Figure 1). The cables were 6 km in length; the inline receiver spacing was 25 m, and the crossline spacing between cables was 300 m (Figure 2). Flip-flop shooting using two air-gun arrays, each of 3930 cubic inches towed at 6 m depth, was orthogonal to the cables, and used a 75 m cross-track and 25 m along-track separation (Figure 2). For each receiver swath, the shooting patch measured 10 × 12 km, and together the three patches covered a survey area of about 180 km². In total, the survey employed 5760 4C receivers and about 96,000 sources.

This survey provides high fold and good azimuthal coverage to offsets of about 7000 m. The maximum offset available in the data set is in excess of 11,000 m, but there is reduced fold, reduced azimuth, and only partial spatial coverage available as the offset increases beyond 7000 m. The gas cloud and corresponding seismic obscured area lies close to the center of the survey area, Figure 1. Four wells lie within the survey area; of these, one lies entirely outside the gas cloud, two lie on its periphery, and one lies within it.

PSDM reflection sections

The reflection portion of the OBC data set was processed by the original contractor to generate the P-wave PSDM reflection images shown in Figure 3. Line locations are indicated in Figure 1. In this conventional reflection processing, the hydrophone and

vertical geophone were matched and summed to remove the down-going receiver ghost. Wide-angle refracted arrivals, wide-angle postcritical reflections, and surface-related multiples were removed from these PZ-summed data as far as possible, and the source wavelet was debubbled, shaped, and filtered to provide a final broadband zero-phase wavelet. The data were imaged using 3D pre-stack Kirchhoff depth migration, and the final section was stacked, band-pass filtered, and balanced. The lowest frequencies available in the hydrophone data were not retained through this sequence. As we will see, this reflection-processing sequence is almost exactly the opposite of that which will subsequently be required for FWI.

Figure 3a shows the structure where it is not obscured by the gas. Bright reflectors at about 3000 m depth indicate the chalk section; the reservoir is located near the top of this section within a broad anticline. There are almost no bright reflectors in the clastic section within the upper 3000 m of the section. On a large scale, the section

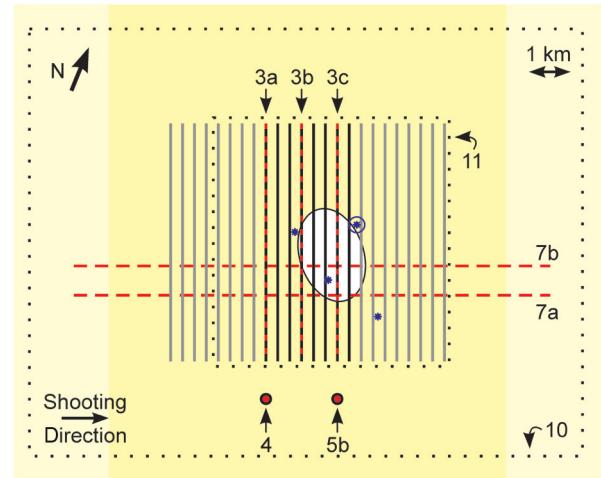


Figure 1. Experimental geometry for the OBC survey across the Tommeliten Alpha field. Black lines show the eight ocean-bottom cables that form the central swath; these receivers recorded sources located throughout the central yellow rectangle. A similar source-receiver geometry was used to acquire an additional swath to either side of the central swath; their cables are indicated by grey lines, and their additional source coverage is indicated in pale yellow. The outer box shows the limits of the FWI velocity model. The white area shows the approximate location of the gas cloud and the seismic obscured region. Numbers refer to the locations of data shown in later figures. Well locations are shown in blue; the circled well corresponds to that shown in Figure 15.

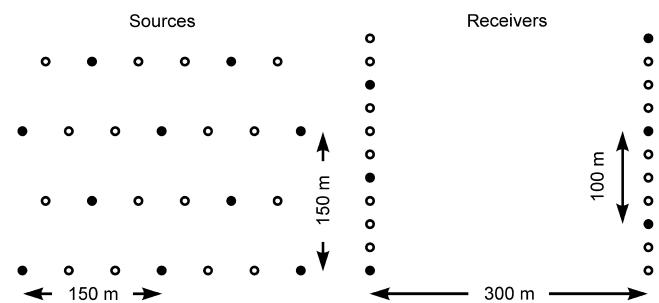


Figure 2. Detail of the source and receiver geometry. Circles show the nominal location of sources and receivers in the original data set. Solid circles show the subset of data that was used during FWI.

is not structurally complicated, and at this scale it is broadly 1D. On a finer scale, not readily visible at the resolution of Figure 3, there are shallow channels within the upper 300 m, and many subvertical, small-offset faults disturb the clastic section.

Figure 3b shows the same structure on a line that passes through the periphery of the gas cloud. The seismic-obscured area is now beginning to appear, and bright reflections are now visible within the middle of the previously less-reflective clastic section. This brightening is presumably caused by gas preferentially occupying the more-sandy layers, and consequently significantly increasing the seismic contrast between the sand and silt. Two small vertical fingers of brightening reflectivity also extend upward from the main reflective gas cloud, and there is some indication of localized brightening above 1000 m. These fingers presumably represent faults or fractures up which the gas has percolated.

Figure 3c shows a line through the center of the gas cloud. This shows the seismic obscured area at its maximum extent. It shows strong reflections in the clastic section starting at about 1000 m

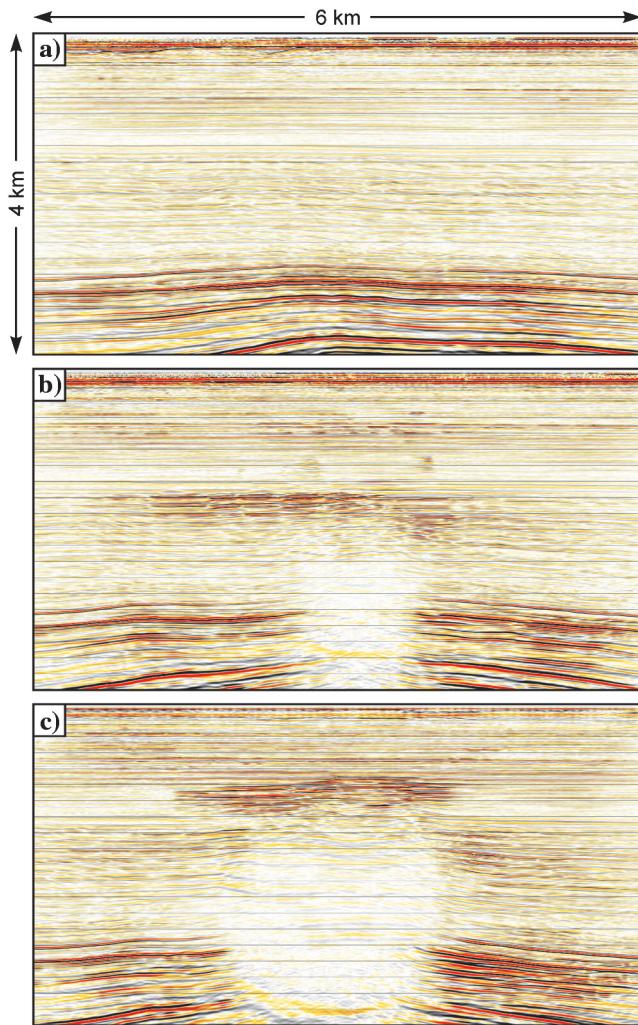


Figure 3. Prestack depth migrated sections across the Tommeliten Alpha structure: (a) outside the gas cloud, (b) on the periphery of the gas cloud, and (c) within the central region of the gas cloud. Bright reflections below about 3000 m represent the chalk section; the reservoir is located near the top of chalk. Migration bandwidth is approximately 7–56 Hz.

depth, with deeper moderately bright reflectors extending laterally away from the gas cloud at about 2000 m depth. These latter reflections suggest that the gas is able to penetrate the clastic section partly by moving sideways along the stratigraphy as well as penetrating upward along faults and fractures.

Although the central portion of Figure 3c is significantly obscured, the effect is not complete. There are low-amplitude, low-frequency, coherent events visible within the obscured area, and these appear to represent the same events that are visible outside this area within the reservoir section and above. However, there is no meaningful depth control on these weak events, and so they do not help to define the geometry of the reservoir. A better velocity model through and beneath the gas cloud potentially could provide sensible depth control for the reservoir, even if it remained difficult to improve the quality of the reflection image in the central obscured area because of the irreversible effects of attenuation within the gas cloud. Nangoo et al. (2012) demonstrate the applicability of FWI using such an approach, and Vinje et al. (2012) demonstrate the utility of using an FWI velocity model for PS reflection imaging beneath the gas cloud.

Shot records

Figure 4a shows a raw hydrophone shot record acquired along a single cable, and Figure 4b shows the same record after it has been preprocessed for FWI — the details and rationale for the latter are explained later in the paper. Figure 5a shows the same record after low-pass filtering; Figure 1 shows the location of the source and cable. This record is not significantly affected by the presence of the gas cloud; it lies along the section shown in Figure 3a.

In Figures 4a and 5a, the data are shown with no additional processing and no temporal gain. They are trace equalized to show near and far traces sensibly at the same scale. It can be seen immediately that the raw data are dominated by wide-angle refracted arrivals at all offsets. These arrivals are a mixture of turning and head waves, and postcritical reflections, together with their ghosts and surface multiples. It is principally these events that we will use to drive FWI.

There are also subcritical reflections visible in Figure 4a. At a travelt ime of about 2000 ms, at the shortest offsets, high-frequency reflections from the clastic section are visible. At about 3000 ms and below, brighter more-continuous reflections from the chalk section can be seen.

Figure 5b shows a second shot record, also low-pass filtered; it is located on Figure 1, and lies along the section shown in Figure 3c. This record is strongly affected by the gas cloud. Comparing Figure 5a and 5b should reveal how the gas cloud manifests in the prestack data. It produces some variations in arrival times, relative amplitudes, and waveforms within the main package of refracted arrivals, but these are subtle and are not obvious without detailed analysis of the records. These subtle changes are nonetheless sufficient to drive FWI, and in most regions of the model they are responsible for the features that we will see in later figures.

However, the dramatic difference between Figure 5a and 5b is provided by the bright subhorizontal arrival that appears in Figure 5b at offsets beyond about 4500 m and times after about 4300 ms. This event is also associated with weak diffractions that extend it to shorter offsets and later times, and with a disruption and delay to the shorter-offset top-chalk reflections. The anomalous bright arrival in Figure 5b is a postcritical reflection from the top of the chalk. Outside the gas cloud, this arrival normally appears at longer offsets and

earlier times, where it is partly obscured by the slower shallower refracted arrivals. However, beneath the gas cloud, low velocities within the cloud change the ray paths such that the critical distance for the top chalk reflector/refractor is reduced. These bright, wide-angle, postcritical arrivals then become visible at shorter offsets.

Because the gas cloud is limited in lateral extent, the anomalous postcritical arrivals are also limited in their spatial extent. Where these arrivals become truncated, they produce the weak diffractions seen in Figure 5b. Traveltimes associated with wide-angle reflections and refractions from the chalk are also anomalous where they have passed through the low-velocity gas cloud. These wide-angle postcritical arrivals are sensitive to the detailed geometry and velocity structure within the gas cloud. They provide a means for FWI to image within and beneath the cloud that is not available to conventional subcritical reflection-based velocity analysis such as reflection tomography, nor to early arrival inversion schemes that use traveltimes or that undertake FWI based upon only the earliest arrivals.

It is interesting to note that these anomalous arrivals have traveled through the nominally obscuring gas cloud, and that they reflect strongly from the reservoir section immediately beneath the gas cloud. Using an FWI velocity model, together with a full two-way RTM algorithm that can deal correctly with postcritical arrivals, it is possible to image the section below the seismic obscured area. That approach is explored in Nangoo et al. (2012), and it is

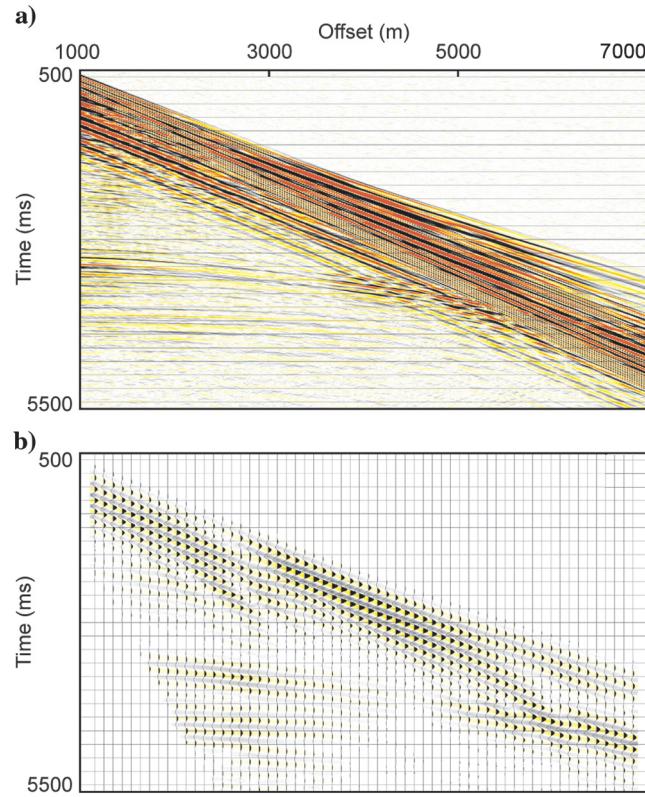


Figure 4. Shot record lying outside the zone of influence of the gas cloud. (a) Raw record, trace equalized, no gain recovery. (b) The same record preprocessed immediately prior to FWI. During preprocessing, the data have been low-pass filtered using an Ormsby filter rolling off from 5.0 to 7.5 Hz, three in four receivers have been removed, and short-offsets containing low-frequency Scholte waves have been muted.

complementary to more-commonly employed PS imaging (Granli et al., 1999; Vinje et al., 2012).

FWI WORKFLOW

Workflow summary

The workflow that we have applied to this data set is largely generic, and with minor modifications it is applicable to a wide-range of data sets where FWI is applied to long-offset data to obtain a high-resolution velocity model for subsequent PSDM. We have inverted the Tommeliten data set many times, exploring a wide range of experimental parameter settings and strategies. The workflow outlined here represents a synthesis of our conclusions from that testing, and is the final workflow that we used to derive the results shown in the figures.

Our workflow takes the following steps:

- 1) Choose an appropriate problem: FWI is not a panacea — as we use it here, it is a means to obtain a high-resolution velocity model. It uses principally transmitted arrivals to perform tomography of the target region. It can only resolve the model to about half the seismic wavelength, and it needs an accurate starting model. It works well for shallow targets, that are

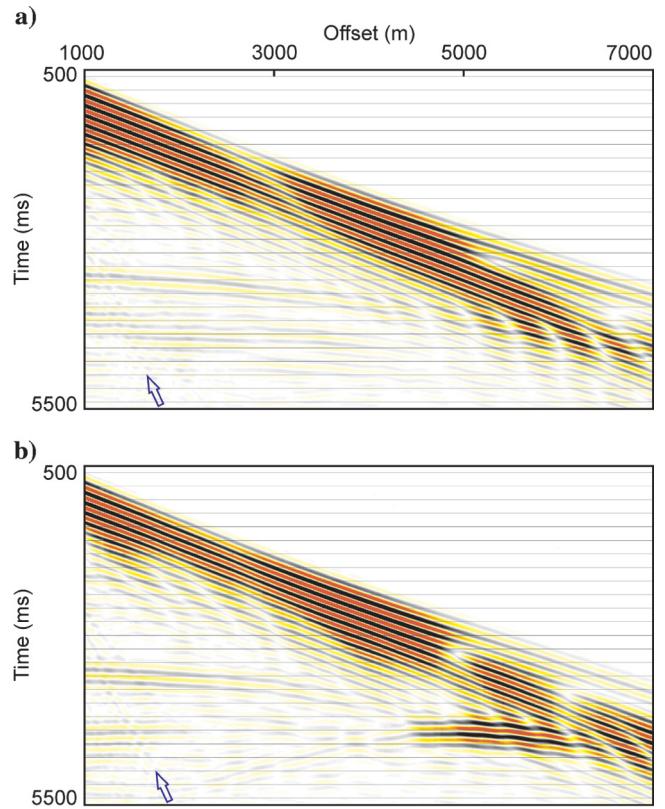


Figure 5. Low-pass filtered shot records: (a) outside the gas cloud — this is the same record as shown in Figure 4 — and (b) inside the gas cloud. The low-pass filter here rolls off over the interval of 6–9 Hz, which is less harsh than that applied prior to FWI. Arrows mark the main train of Scholte waves that is just visible at this bandwidth but which dominates the inner traces at 3 Hz. Note the anomalous postcritical reflection/refraction from the chalk that appears inside the gas cloud.

- adequately covered by refracted arrivals, using long-offset data, containing (very) low frequencies, and ideally many azimuths. Most affordable 3D algorithms do not deal properly with elastic effects, attenuation, or complicated velocity-density models, and so will not deal well with problems and data sets where such effects are dominant.
- 2) Obtain an appropriate field data set: transmission FWI is not normally possible without low-frequency, long-offset, turning arrivals that penetrate to target depths. If these data have not been acquired, or have been lost during subsequent processing, then FWI will not be able to modify significantly the macrovelocity model. In this case, FWI will become little more than an iterated least-squares RTM algorithm — this may provide a sensible outcome, but it will not normally be able to provide the uplift to the macrovelocity model and subsequent PSDM that we illustrate here.
 - 3) Determine the starting frequency: discussed below.
 - 4) Build the starting model, including anisotropy if required: discussed below.
 - 5) Build the source wavelet: discussed below.
 - 6) Check the adequacy of the starting model, source wavelet, and starting frequency: This is potentially the most important stage required to ensure a favorable outcome. The requirement here is that the initial predicted synthetic data must match the field data to within better than half a cycle at the starting frequency. In large or complicated data sets, it can be difficult to assess this effectively by manual examination of the data in the time domain, especially when pushing the inversion into regions of marginal signal-to-noise ratio (S/N) at the lowest frequencies. We have developed a rigorous means to determine whether the starting model is adequate, which is explained fully in [Shah et al. \(2012\)](#). We have applied that approach to this data set, but do not discuss it further in this paper.
 - 7) Preprocess and reduce the field data volume: discussed below.
 - 8) Devise modeling strategy: discussed below.
 - 9) Devise inversion strategy: discussed below.
 - 10) Invert the data with continued quality assurance: The inverted data are discussed below; quality assurance is discussed in [Shah et al. \(2012\)](#).
 - 11) Check the accuracy of the final synthetic data against the field data: discussed below.
 - 12) Check for consistency with reflection geometry, wells, and other a priori information, and check migrated image gathers: discussed below.
 - 13) Use the recovered FWI velocity model for RTM of broadband reflection data: discussed below.

At the current stage of development, quality assurance before, during, and after FWI is normally essential for a successful and validated outcome. FWI is not yet a particularly robust procedure, and it does not normally fail elegantly. Without careful and rigorous quality assurance at all stages, it can lead the unwary practitioner, sometimes with undue confidence, in an entirely spurious direction.

Choosing the starting frequency

A key requirement for successful wide-angle FWI is the presence of low frequencies within the field data. Because FWI is a local inversion scheme, it can only reach the vicinity of the global minimum of the objective function if the starting model predicts data that

differ by no more than half a cycle from the field data, at least for the vast majority of dominant arrivals within the data set. The condition that the starting data should not be cycle skipped with respect to the field data is clearly more easy to meet the lower is the dominant frequency present in the field data.

Consequently, for successful FWI, we require low, and ideally very low, frequencies in the field data, and we begin the inversion using only the lowest frequencies present, following the strategy first suggested by [Bunks et al. \(1995\)](#), and developed by [Sirgue and Pratt \(2004\)](#). In time-domain implementations, which we are using here, this means low-pass filtering the data to preserve only the very lowest frequencies. The question then arises for a particular data set as to how low in frequency can we go? If we begin too low in frequency, then we will introduce unnecessary noise into the results, and if we begin too high, then parts of the data may be cycle skipped and the inversion will likely head to entirely the wrong model.

Notwithstanding source and receiver ghosts, many marine hydrophone data sets have significantly lower frequencies present than might otherwise be suspected, provided that these have not been removed in the field by unnecessary analog or digital filters. These low frequencies are often not easy to identify on time-domain displays, and Fourier amplitude spectra cannot easily distinguish between signal and noise. Indeed, it is not the absolute amplitude of low-frequency data that is of interest, instead it is the signal-to-noise ratio at low frequencies that matters even if the absolute amplitudes are much lower than at higher frequencies.

Consequently, we use plots of the form shown in Figure 6 to chose the starting frequency. These show data from a single common-receiver gather. The raw data have been Fourier transformed, and their phase extracted at a single frequency for every source. The figure shows these phase data plotted for three frequencies at the location of each source. Where such plots contain coherent structure, this indicates source-generated signal.

In Figure 6, at 3.6 Hz, there is clearly good S/N as evidenced by the concentric circular structure in the phase plot. At 3.0 Hz, the data are becoming noisier, but there is still clear source-generated coherent signal at the longer offsets in the outer portions of the receiver gather. At shorter offsets, the phase data at 3.0 Hz are still coherent, but they have a different appearance. The horizontal wavelength is much reduced, and the phase appears to have a four-fold symmetry forming a cross-like pattern centered on the receiver. The short offsets at 3 Hz are dominated by Scholte waves; these are surface waves, or more correctly boundary waves, that are localized near the seafloor, and that couple strongly to the source in shallow water at very low frequencies.

These Scholte waves are not visible in the raw data in Figure 4, but they are visible in Figure 5, which has been low-pass filtered at about 7 Hz, as the low-velocity wave-trains appearing at short-offset. Although these boundary waves are quite weak at 7 Hz, they are dominant at the lowest frequencies, and they dominate the phase plot at 3 Hz at the shorter offsets. The fourfold symmetry visible in Figure 6 results from the azimuthally variable affects of source and receiver arrays. Although the Scholte waves are low frequency, their low velocities mean that the source and receiver arrays are partially effective in suppressing them at 3 Hz. These source and receiver arrays are mutually perpendicular and have approximately the same dimensions. Consequently, their combined 3D response has a fourfold symmetry in a horizontal

plane, and so the suppression of the Scholte waves varies azimuthally, giving the appearance seen in Figure 6.

At 2.4 Hz, the Scholte waves are still in evidence at the shortest offsets; at longer offsets the S/N is poor, and there is only weakly coherent energy visible. The raw field data were acquired using a low-cut acquisition filter set to roll off at 18 dB per octave below 3 Hz, so it is not surprising that the data at 2.4 Hz are limited. Figure 6 therefore suggests that FWI can sensibly start at about 3.0 Hz provided that the Scholte waves are suppressed at the shortest offsets where they dominate the records at these low frequencies.

This is not the conclusion that would have been reached by examining only filtered time-domain data or amplitude spectra. Even at 3.6 Hz, the coherent energy clearly visible in Figure 6 is not readily apparent on filtered traces. In the raw data, at 3 Hz, the power level is about 40 dB down on the power at 10 Hz. Nonetheless, these low frequencies are coherent, they have good signal to noise ratio, and they can be used to drive FWI if they are processed appropriately. We speculate that even lower frequencies would have been usable in this data set had a field filter not been applied to remove them. The useful low-frequency limit for OBC hydrophones is ultimately likely to be controlled by pressure changes associated with changing water depth as swell passes over the receivers at around 1 Hz, and not by the characteristics of the air-gun source and its ghost.

Starting velocity and anisotropy model

Obtaining an adequate starting velocity model is a necessary requirement for successful wavefield tomography. Using the simple inversion scheme that we apply here, all major arrivals in the synthetic data, generated using the starting model, must match the real data to better than half a cycle at the lowest inversion frequency to avoid cycle-skipped local minimums. In this survey, we used an anisotropic starting model that was originally generated for PSDM by the original processing contractor. We were not involved in generating this model, but the process is a familiar one.

An initial stacking velocity model was used as the starting point, taken from an earlier surface-streamer data set. This model was further refined by picking residual moveout, over a 600 m grid, on the prestack time-migrated PZ-summed reflection volume. Following this, the time-migrated data were stacked and matched to the four wells to obtain an initial estimate of anisotropy. The velocity and anisotropy model was then further refined by reflection travel-time tomography with the anisotropy model constrained to follow stratigraphy. The tomography was applied using residual moveout picked on depth-migrated common image gathers, and the scheme was run in a layer-stripping mode with reference made to the well ties at each iteration to constrain depth correctly via adjustments to the anisotropy model. In the absence of strong evidence to the contrary, a vertical axis of symmetry was assumed, and no azimuthal anisotropy and no lateral changes in anisotropy other than tracking stratigraphy were introduced — in particular, the anisotropy inside and outside the gas cloud were kept the same.

For this data set, it is not possible to fit accurately and simultaneously the short-offset and long-offset reflection travel times, nor to fit near-normal-incidence reflections and horizontally traveling refractions, without incorporating some form of anisotropy into the velocity model. The surface data, however, do not completely define the anisotropy, and a match to wells is essential to define the

problem fully. If we apply isotropic FWI to these data, we invariably find that we are unable to match the arrival times accurately, and we also often find that spurious horizontal layering is introduced into the velocity model by the inversion as it attempts to fit anisotropy using heterogeneity.

Within and beneath the gas cloud, there is limited reflection coverage. Here, the tomography model was constrained manually to match generic velocities from the wells, with the geometry and intensity of the low-velocity gas cloud estimated using its apparent effect in obscuring underlying reflectivity and in producing the enhanced shallow reflectivity that is seen in Figure 3. Such interpretative velocity model building is common for PSDM when reflection tomography proves otherwise inadequate; its limitations are clear, and it is one of the approaches on which FWI seeks to improve and ultimately to supplant.

Vertical slices through the resultant velocity model are shown in Figure 7; all velocity sections shown in this paper show vertical velocity. The corresponding anisotropy model is shown in Figure 8. This is the model that was used to depth migrate the data shown in Figure 3, modified in three ways for FWI. We have added a sharp seafloor at the top, we have extrapolated the model laterally to cover the full extent of the survey, and we have smoothed the model with a horizontal wavelength of about 300 m in both horizontal directions. This smoothing had limited effect over most of the model, but in a few areas the original contractor's model had relatively sharp internal boundaries that would have been likely to compromise the early iterations of FWI.

Using a smooth starting model is important. The starting velocity model in general should not normally contain any structure that is sharper than about half a wavelength at the starting frequency unless the location, in depth, in three dimensions, of that structure is certain. In practice, when inversion begins, the only structure of which we can normally be certain is the seafloor. In the section here, top chalk is in reality likely to be a sharp interface at which velocity increases rapidly. However, even though we know its location in depth at a few points where there are wells, its true 3D shape in depth is not known accurately prior to FWI, and so it should not appear as a sharp interface in the initial model. Similar considerations apply to salt, to basalts, and to other layers with

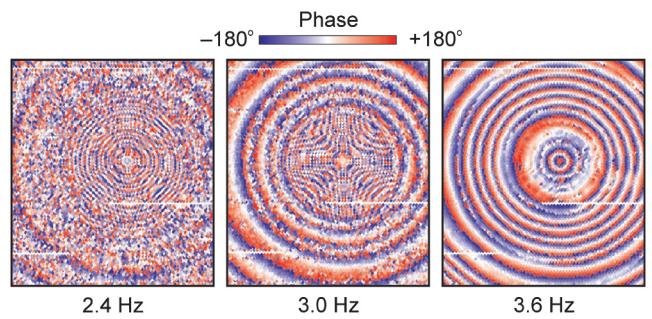


Figure 6. Phase variation within a common receiver gather at 2.4, 3.0, and 3.6 Hz. Data are shown for all sources recorded on a single hydrophone located at the centre of the circles; the graininess of the plots represents the true graininess of the source distribution. The raw field data have been Fourier transformed, and the resultant phase for a single frequency is plotted at the physical location of the corresponding source. Spatial coherence indicates source-generated signal.

strong sharp contrasts — these should not normally be present as sharp interfaces in a starting model for FWI.

The resulting starting model contains a rather 1D clastic section overlying a broad antiformal high-velocity chalk section. Within the clastics, a rather poorly constrained low-velocity gas cloud appears with a somewhat blocky structure. This is an artifact of how the velocity model had been built. We note that the velocity structure below the gas is not strongly constrained by the subcritical reflection data.

Anisotropy is high in this model; below the top few hundred meters, epsilon values are consistently above 10%, and in parts of the section have values of 20%. Delta values are also quite high, but are generally only half or less of those of epsilon. The anisotropy model would not be well-approximated by elliptical anisotropy. There is no strong evidence for a tilted symmetry axis, for lateral changes in anisotropy, or for significant azimuthal P-wave anisotropy. Given the lack of constraint upon the anisotropy model, the inversions were performed using the single 1D anisotropy profile shown in Figure 8.

The original contractor's model maintained high epsilon and moderate delta values into the chalk section as shown by the dashed lines in Figure 8. We think that such high anisotropy values are unlikely to be correct within a chalk section. Consequently, we

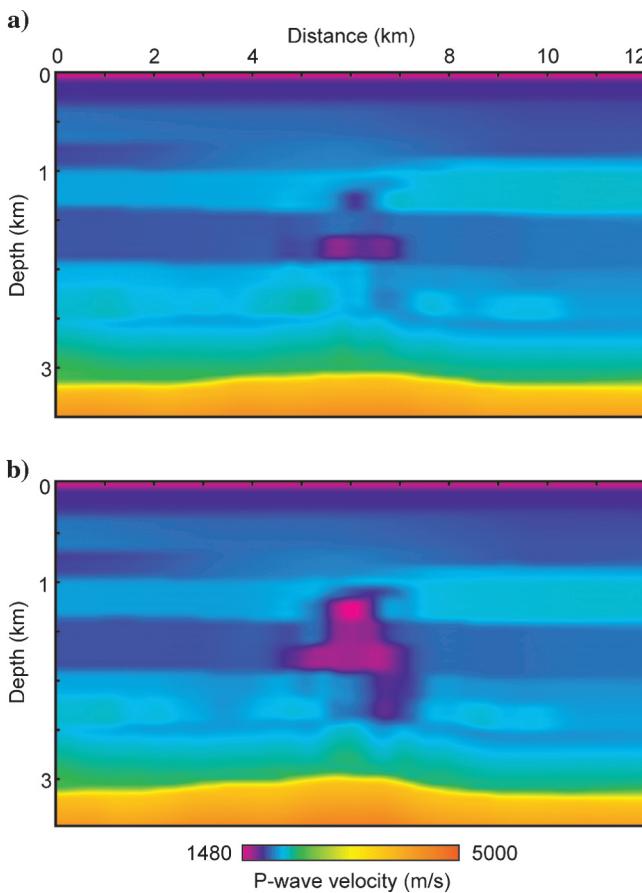


Figure 7. Vertical slices through the starting velocity model: (a) through the periphery of the gas cloud, and (b) through the central portion of the gas cloud. The gas cloud is represented by the central low-velocity feature. The antiformal yellow-red layer toward the bottom of the model represents the chalk.

arbitrarily reduced delta and epsilon values to zero within the chalk as shown. This has minimal effect upon the inversion above 3000 m, and it does not change the qualitative structure below 3000 m. However, it does affect the match to the wells below 3000 m. Although during FWI we used the 1D model of anisotropy shown in Figure 8, all the migrations presented here, including those that used the FWI velocity model, were performed using the original contractor's anisotropy model; this follows stratigraphy and retains the high values of delta and epsilon within the chalk.

Source wavefield

Obtaining an accurate source wavelet is important for FWI. For conventional reflection processing, obtaining a wavelet is also important, but there are some significant differences. For FWI, we require a wavelet that is accurate at the lowest frequencies, and we care not at all about the wavelet at moderate and high frequencies. In conventional processing, if the wavelet estimation is poor, then this will compromise the accuracy of the final result, but it will not normally be catastrophic — well ties will be less good, velocity picks will be a little wrong, multiple suppression may be a little less accurate, and so forth. However, in FWI, if

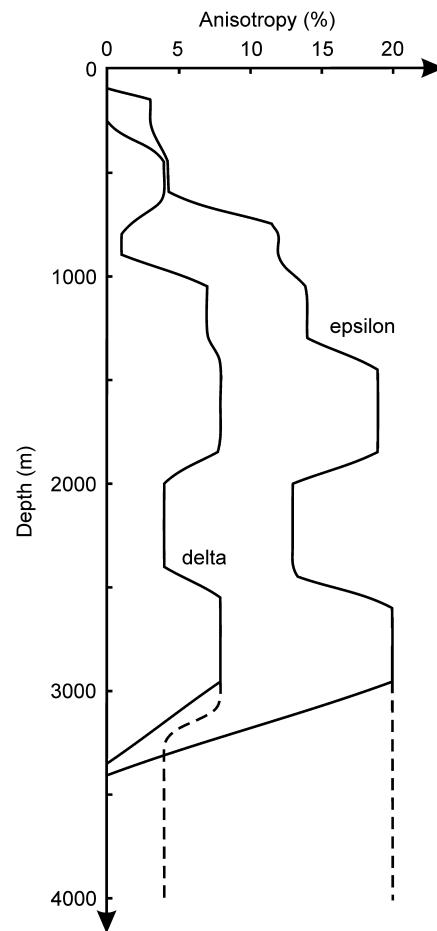


Figure 8. Anisotropy parameters used for FWI. We used a 1D model of delta and epsilon with a vertical axis of symmetry. The dashed profiles represent the anisotropy model used in the original PSDM; the solid line, which drops to zero within the chalk, represents the anisotropy model used during FWI.

the wavelet is significantly incorrect, then this may be sufficient to push the inversion toward a local minimum, severely compromising the inversion, and leading potentially to significant artifacts in the resultant velocity model.

There are many ways to estimate the source wavelet; Figure 9a shows the basis for two of these. Here, the left trace shows the acquisition contractor's estimate of the source wavelet including the source ghost. It has been derived using an heuristic mixture of numerical simulation of the appropriate physics, matched to direct observation using deep-towed hydrophones in deep water, and controlled by near-field hydrophone measurements made among the air guns. Such estimates are well-established, and generally work well over the normal bandwidth of reflection seismic data; however, as we will see, they can be less effective at the lowest frequencies.

The right trace in Figure 9a shows the direct arrival as recorded on an ocean-bottom hydrophone at a lateral offset of about 25 m in 75 m of water, shifted to zero time. In addition to the genuine direct arrival and its source ghost, this trace contains the effects of shallow subseabed reflectivity and of free-surface multiples. It is also affected by spherical divergence in that the source ghost originates effectively further away than does the direct arrival. In principle, if we could deghost, demultiple, and deconvolve the shallow reflections, then we could build a source wavelet from this trace.

Figure 9b shows the same two traces after low-pass filtering. Note that the time axes in Figure 9a and 9b differ so that the latter is even lower frequency than it might otherwise appear. Figure 9b shows that there is a significant phase shift between the contractor's wavelet and the observed direct arrival. This shift is not explained by multiples, by subseafloor reflectivity, or by other obvious signals contaminating the direct arrival. It is consistent throughout the data set, over which there are variations in seafloor and subseafloor reflectivity. It indicates that the contractor's wavelet is not accurate for this data set at the lowest frequencies. We speculate that this may be because there are phase shifts within the acquisition system at low

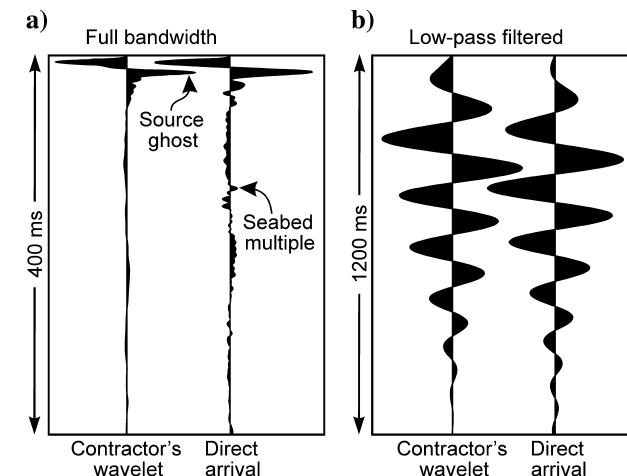


Figure 9. Source signatures. (a) Full-bandwidth waveforms. (b) The same waveforms low-pass filtered using an Ormsby filter rolling off from 5.0 to 7.5 Hz — note the different time scale. On each figure, the leftmost waveform shows the contractor-supplied source estimate, and the rightmost signature shows the direct arrival recorded by a near-source ocean-bottom hydrophone. At low frequency, the phase of the two waveforms is significantly different.

frequency that are not properly accounted for in the contractor's modeling. Whatever the reason, the contractor's wavelet was not considered to be sufficiently accurate for FWI at low frequency. We note that it is likely, as FWI grows in importance, that conventional low-frequency wavelet estimates generated by contractors will correspondingly improve.

We consequently did not use the contractor's wavelet for the inversions shown below. Rather, we used the direct arrival, and deconvolved from this the source ghost and the first and second seafloor multiples, correctly taking into account the finite offset and the effects of divergence. We did not attempt to remove the effects of subseafloor reflectivity, but we do not see any variation in waveforms at low frequency when we apply this process at different receivers around the survey, so we do not believe that this is likely to be a significant omission. We require a deghosted source wavelet for FWI because we use an explicit free surface in the modeling, and this will reapply the source ghost into the modeled data. Finally, having obtained an estimate of the source wavelet, we low-pass filtered it using the same filter that we applied to the field data. This step is essential, and it is important that no additional or alternative filters are applied to the raw source estimate that may change either its phase or amplitude spectra.

Another approach to source estimation during FWI is to allow the inversion itself to estimate the source (Pratt, 1999). Although there are advantages to this, especially if the source varies from shot to shot, there is necessarily a trade-off between model and source, with the potential that systematic errors in the velocity model may map into consistent time shifts in the source. We therefore did not use that approach here.

Preprocessing

We have tried inverting pure hydrophone and PZ-summed data. We have tried inverting field data that contain surface multiples and ghosts by including the free surface in the modeling, inverting field data from which the multiples but not the ghosts had been removed while using an absorbing upper boundary and explicit ghost sources and receivers in the modeling, and inverting field data from which all multiples and ghosts had been nominally removed. It is clear, for this data set, that leaving the multiples and ghosts in the data, and inverting the pure hydrophones, while using a free surface in the modeling, gave the most reliable and stable results.

The hydrophones consistently record lower frequencies than do the geophones, and these low frequencies are lost during PZ summation. The summation is designed to suppress the receiver ghost for near-normal-incidence arrivals, and it performs less than adequately for wide-angle arrivals. The summation leads to a nonphysical data set in which the free surface is partially suppressed for receivers, but present for sources, and it is not straightforward to simulate such a system with a forward modeling code. The only really successful way to model and invert PZ-summed data, is to model hydrophone and geophone responses, and to match and sum these for the synthetic data as was done for the field data. This is not straightforward, and we do not recommend it.

Although it is relatively easy to suppress surface-related multiples in these data for near-normal-incidence reflections, it is not at all easy for wide-angle turning rays. Typical surface-related-multiple elimination algorithms, parabolic radon filters, and deconvolutional approaches do not deal adequately with multiples in turning-wave data, and the best that can normally

be achieved for refracted arrivals in shallow water is to produce an admixture of remnant multiple energy and damaged primaries.

As a consequence, in the results presented below, we use the hydrophone data only, and we leave the source ghost, receiver ghost, and all multiples in the field data. We use a free surface in the modeling, we place sources and receivers at their correct depth below the free surface, and we use a deghosted source wavelet. With this approach, we are able to match the field data accurately during FWI, and we see no evidence of multiple or ghost contamination of our resulting velocity models.

Figure 4b shows the preprocessed data prior to FWI; Figure 4a shows the corresponding original data. We have applied only minimal processing to these data — we have applied a short-offset bottom mute as shown in Figure 4b to remove the Scholte waves, we have applied a top mute ahead of the first arrivals, we have selected only every fourth receiver and every third source, we have deleted bad traces, we have low-pass filtered the data using an Ormsby filter that rolls off from 5.0 to 7.5 Hz, and we have truncated the record length to 7000 ms. We have done nothing else. We do not shape, debubble, change the phase, or otherwise alter the wavelet, we do not apply any multidimensional or data adaptive filtering, we do not apply any form of multiple suppression, PZ-summation, or deghosting, and we do not apply any form of time-variant or data-dependant gain. We are especially careful not to apply any process that may damage the lowest frequencies, or the wide-angle arrivals — note that almost everything that is done conventionally to process reflection data will have one or both of these as an undesirable consequence for FWI. It is therefore almost always necessary to return to the raw, unadulterated field data to apply FWI to a previously processed data set.

The rationale behind our approach is to do as little as possible to the field data, and to put all the resulting phenomena into the forward modeling and inversion code. We remove only those aspects of the field data that our forward code is not intended to simulate — in this case, our forward code will not simulate Scholte waves, and so we remove them from the field data. We could remove these using a multichannel filter, but we prefer here to use a simple mute because that allows us to be certain that we apply identical processes to field and synthetic data. Many practical multichannel filters have an element of data adaptivity — for example, an AGC that is applied and removed after the filter — which will not perform identically for field and synthetic data. As FWI proceeds, additional filtering, spectral shaping, and amplitude normalization is undertaken, changing through successive iterations; here, we regard these as part of the inversion strategy rather than as a form of preprocessing, and we discuss them in that context below.

Modeling strategy

FWI can be computationally expensive. For a 3D model with a linear dimension of n grid cells, the runtime for FWI scales as $\sim n^4$ because the total number of cells is $\sim n^3$, and $\sim n$ time steps are required to cross the model. For a time-domain algorithm such as we use here, the runtime also scales in proportion to the number of sources. To reduce the computational cost, we therefore want to minimize the number of grid cells; that is, to maximize the cell spacing, to minimize the number of sources used per iteration, and to minimize the number of iterations.

For this data set, we have to include within our model a water layer with a depth of around 75 m, principally so that water-bottom

multiples are correctly modeled. We therefore chose an initial grid spacing of 50 m because this is about the coarsest that we can use to capture the water layer effectively. Our modeling code is accurate at five grid points per wavelength, and the errors remain small down to four cells per wavelength. We inverted these data using a starting frequency of 3 Hz, increasing this frequency by stages as the model improved, to a maximum of 6.5 Hz. At this maximum frequency, a 50 m grid spacing provides more than 4.5 grid points per wavelength in the thin water layer, and provides more than five grid points per wavelength everywhere that the velocity is above 1625 m/s. For the modeling code that we used here, this generates minimal numerical dispersion, and the code is stable everywhere in the model with a time-step of 4 ms. A finer grid than this would allow higher frequencies to be modeled, but at a cost proportional to n^4 ; a coarser grid than this would be possible using a higher-order finite-difference stencil, but it would not easily allow the shallow water layer to be properly incorporated. Allowing for boundaries, and for room to distribute sources and receivers that are not located at integer positions, we arrived at a final model size of $16 \times 13 \times 4$ km, or $321 \times 261 \times 81$ cells.

During preprocessing, we selected every third source and every fourth receiver for FWI, generating the geometry shown by solid circles in Figure 2. This provided inline and crossline receiver separations of 100 m and 300 m, respectively, and sources sampled on a square mesh rotated by 45° with a spacing of about 106 m. We are therefore sampling the wavefield horizontally at a density that is similar to the expected far-field resolution except in the crossline direction where we are constrained by the original acquisition. Provided that the subsurface is properly spatially sampled in at least one domain, sparsity in the acquisition will not normally compromise FWI. In addition, FWI is rather robust against most sources of noise, and so we do not require high fold to aid noise suppression. Including additional sources and receivers in the inversion will not therefore improve the resolution except close to the seafloor, and they will increase run times and data volumes.

Following subsampling, we applied source-receiver reciprocity, labeling sources as receivers and receivers as sources. We do this to reduce the total number of effective sources that we must model. This is a step that is normally appropriate for OBC and some land data sets, but it is not normally required for towed-streamers where it confers no advantage. After data selection and reciprocity, we have a data set that contains 1440 (reciprocal) sources, each recorded on about 10,000 (reciprocal) single-component receivers. This represents around 100 Gbytes of field data which it is straightforward to hold permanently in memory during FWI when distributed source-by-source across a cluster.

Inversion strategy

It is possible to invert these data using all 1440 sources at every iteration. However, for a given compute cost, this is not an efficient strategy. Because we have good coverage in the receiver domain, we can obtain a reasonably good velocity update with only sparse coverage in the source domain. A sparse iteration will generate an improved velocity model such that a second sparse iteration, using a different subset of sources, is able to generate a further improvement that is better than would have been arrived at by using both subsets together in one iteration. Consequently, for a fixed compute effort, it is normally better, indeed much better, to use more iterations and fewer sources per iteration than it is to use

all the sources together in few iterations (van Leeuwen and Herrmann, 2012).

For this data set, we used just 80 sources per iteration, chosen randomly so as not to produce a regular interference pattern (Díaz and Guitton, 2011) with 18 iterations per frequency band, so that each of the 1440 sources is used just once per frequency band. This ratio was chosen partly to map the problem neatly onto 40 compute nodes running two sources per node, and partly because extensive testing with similar problems has shown that such a ratio is close to optimal for dense full-azimuth data sets. If we instead use every source at every iteration, we find that we require around 5–10 iterations per frequency to obtain an equivalent model, a strategy that requires 5–10 times the computer resources. Our approach is similar in purpose to the use of composite sources with phase encoding, and similar techniques, that have been used successfully in RTM and FWI (Ben-Hadj-Ali et al., 2011; Schuster et al., 2011), but our approach is more simple; it does not require large volumes of data to be accessed to build the composite shots, it does not suffer problems when there are missing or moving receivers, and it does not produce cross-talk noise in the final result.

In total we used six frequency bands, at 3.0, 3.5, 4.1, 4.8, 5.6, and 6.5 Hz, where these are the cut-off frequencies of a low-pass filter that is applied to field data and source wavelet during FWI. This filter is in addition to that applied to all data during preprocessing. The filter rolls off rapidly above the cut off. We also apply a more gradual low-cut filter that reduces amplitudes below this frequency. The filters are data-adaptive such that they ensure that the peak frequency in the data being inverted is close to the nominal frequency even if the source or data amplitude spectra have reduced amplitude at the nominal frequency; identical filters are applied to field and model data.

Our lowest frequency is a function of the field data. Our highest frequency is partly a function of the computing cost that we are willing to spend, and is partly a recognition that, at least for subsequent depth migration of the full-bandwidth data, additional resolution of the velocity model is unlikely to be of significant practical benefit much beyond that which we can expect to achieve at 6.5 Hz. Conventional RTM separates the earth-model into a macrovelocity model and a reflectivity model, with the implicit assumption that it is the reflectivity model and not the macrovelocity model that produces reflections. The lowest effective frequency in the PZ-summed reflection data that we use in the final RTM is around 7 Hz. Terminating FWI at about this frequency therefore ensures that the resolutions of the macrovelocity and reflectivity models approximately coincide, and this provides an optimal starting point for broad-band RTM.

Running FWI to higher frequencies will therefore significantly increase the cost for minimal additional benefit. Higher frequencies also increase the risk of introducing and magnifying an acquisition footprint as the seismic wavelength becomes shorter than the cable separation. Running FWI to higher frequencies can, of course, be beneficial if the intention is to interpret the FWI image directly, but at these higher frequencies, FWI comes to resemble iterated least-squares migration although the former properly deals with multiple scattering while, at least as it is typically implemented, the latter generally does not.

We use six frequencies because past experience suggests that around 100 sparse iterations in total provides a good balance between cost and effectiveness — at 18 iterations per frequency,

we are using 108 iterations in total. If we instead use around 1000 iterations in total, we obtain only a modest additional improvement in our fit to the field data, with little objective improvement in the velocity model. This is presumably because there are aspects of the data that our forward model is unable to match even in principle, for example more complicated anisotropy, attenuation, elastic, and density effects. We note that this behavior is quite different from that which is seen in synthetic tests when an identical algorithm is used for forward and inverse modeling without realistic source-generated noise. In this unreal situation, huge numbers of iterations will often continue to improve the model and the fit to the data, almost without limit; however, they provide no realistic guide to performance on noisy and inadequately modeled field data sets which we have always to contend with in reality.

Our modeling is acoustic rather than elastic, it takes no account of anelastic attenuation, it maintains a deterministic relationship between velocity and density as the inversion proceeds, it matches only the kinematics of TTI anisotropy correctly not its dynamics, and our model of anisotropy is smoothly varying. Our modeling therefore will not correctly predict the absolute amplitude behavior of the field data. Typically it is attenuation rather than elastic effects that dominate the amplitude mismatch, with unmodeled density complications providing an additional source of amplitude mismatch that can be similar in magnitude to that produced by elastic effects.

Although it is possible in principle to invert for a full elastic model, with attenuation, with density as an independent parameter, and with dynamically correct elastic and anelastic anisotropy, it is not often productive to do so for field data sets. For most field data, such an approach generates too many unknown parameters that are ill-constrained by the too-few available data. Although we are experimenting with 3D codes that are able to deal with each of the effects listed above (Guasch et al., 2012), these are not yet routinely effective on conventional seismic field data sets. We consequently proceed here using an anisotropic acoustic algorithm, we hold the anisotropy fixed during the inversion, and we normalize amplitudes so that the inversion fits amplitude only locally.

For this data set, the inversion first equalizes the band-pass-filtered field data so that the rms value of every trace is the same. During the modeling, we normalize the amplitude of the modeled data, trace by trace, such that the resultant residual data are minimized. This normalization uses a broad sliding time-window, so that different phases that are widely separated in time are normalized somewhat independently. The net effect of this approach, validated on synthetic anelastic data, is that we are successful in minimizing the influence of attenuation, elasticity and unknown density upon the inversion at the cost of some marginal loss of resolution in regions where the data coverage is poor (Warner et al., 2012).

This approach to amplitude normalization retains and uses in the inversion the decay of amplitude within a single waveform — produced, for example, by short-period free-surface and interbed multiples. It also retains the amplitude effects of the direct interference of coexistent phases, and it retains and uses the gross amplitude variation between different arrivals within the same trace. It does not directly use the AVO response for single phases, but if AVO effects cause one phase to change amplitude relative to another within the same trace, then that effect will be at least partially retained during the inversion. We have found this to be a robust and accurate approach for synthetic data (Warner et al., 2012) — it effectively tries

to use all those aspects of the data that our modeling is able to reproduce, and it largely normalizes away those amplitude effects that require a more complete description of the physics and a more complete description of the subsurface.

We also use the following parameterizations. We invert for slowness rather than velocity because we have found in synthetic tests that this marginally improves resolution and convergence rate. We do not invert separately for density; we use Gardner's law (Gardner et al., 1974) to calculate density from velocity, modified to match the density of sea water below a velocity of 1510 m/s. Within the gas cloud, Gardner's law is unlikely to be adequate, but we impose it anyway. We keep anisotropy fixed throughout the inversion. We invert using the same 50-m cubic mesh on which we model. We use conjugate gradients rather than simple steepest descent. We do not attempt any form of quasi-Newton inversion, in part because we are using a different subset of the field data at every iteration.

We approximate the diagonal of the Hessian matrix using spatial preconditioning, effectively dividing the gradient by a (stabilized) measure of the local total energy within the ambient wavefield, source by source. It is possible in principle to apply a global Hessian to the global gradient, or to apply a local Hessian to the local gradient determined using only a single source. Here, we adopt both approaches, applying a strongly stabilized portion of the local Hessian locally, with the remainder of the Hessian, less strongly stabilized, applied globally. We find that this hybrid approach helps to minimize the acquisition footprint without introducing additional artifacts or slowing the rate of convergence.

We use a straightforward least-squares functional, and we have not here applied any explicit regularization, for example by penalizing model roughness using a modified functional. The latter is perhaps surprising, but we find that the raw gradient is normally sufficiently smooth on scale lengths of less than half a seismic wavelength, at least away from the vicinity of sources and receivers.

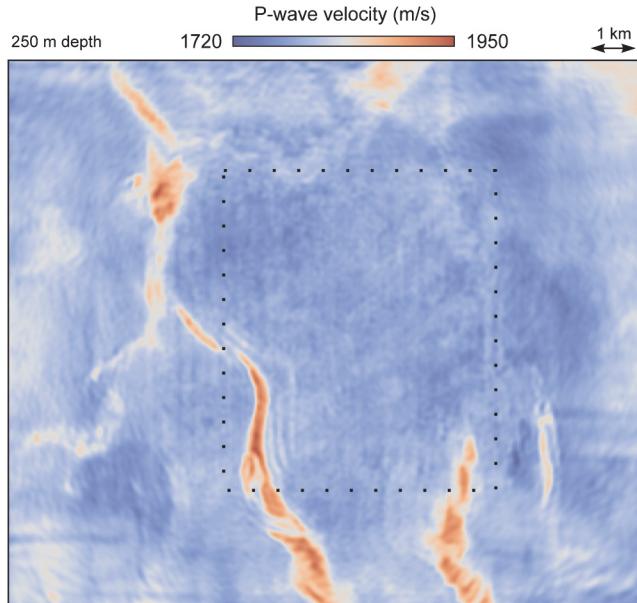


Figure 10. Shallow channels shown on a horizontal slice through the FWI final velocity model. The location of the slice is indicated in Figure 1; the dashed line shows the location of Figure 11.

FWI is effectively self-regularizing because the means by which the gradient is generated is necessarily smooth over most of the model at scale lengths below half a seismic wavelength, and with appropriately stabilized spatial preconditioning it will not tend to invent spurious structure where the wavefield amplitudes are tiny.

For the results presented here, we ran the inversion on a cluster of Intel Xeon 2.67 GHz Westmere-based nodes, each with 12 cores. The cluster was composed of a master node and 40 compute nodes, each with 24 Gbytes of RAM. The network between the nodes was simple gigabit Ethernet. FWI inversion ran in about 62 hours elapsed time, of which the majority was absorbed by computing and memory access rather than by disk or network traffic.

RESULTS AND DISCUSSION

Shallow section

Figure 10 shows a horizontal slice through the final FWI velocity model, at a depth below sea-surface of 250 m; the location of this slice is indicated in Figure 1. The starting model at this depth was a single constant velocity. Note that only the central region of this figure is covered by the cables, and that much of the region is covered only by sources.

The most obvious features in Figure 10 are the high-velocity channels that meander across the section. Figure 11 shows the same depth slice, extracted from the original PSDM volume; note that Figure 11 occupies only the central region of Figure 10. The PSDM section was migrated with the original contractor's velocity model which did not contain channels. At this shallow depth, the PSDM coverage is low fold, restricted only to the region where there are sources and receivers, and contains a strong acquisition footprint. For this study, we did not have access to the original PSDM volume covering the four westernmost ocean-bottom cables.

Although the PSDM image at 250 m is incomplete, it shows the same channel system as does the FWI velocity model. The gross features match, but so too do many of the fine details. In the bottom right corner, the channel is divided into two segments with matching geometries in the two figures. At bottom left, the larger channel widens to the south and forks marginally, and the same structure is evident in the PSDM slice. The channel on the PSDM volume also forks, though this can be seen most clearly on slices that are about 20 m deeper than the one shown.

On the velocity model in Figure 10, within the region covered by the PSDM, the left channel shows, to its right and above it in the orientation of this figure, what at first sight might appear to be shadows of the main channel produced perhaps as a result of the finite bandwidth of the FWI modeling, misplaced multiple scattering, or some similar artifact. However, comparing this to Figure 11, which is a broad-bandwidth migration, we can see similar features with similar geometry in similar locations. The separation of the features is less than the cable separation, so that they are only poorly captured by the PSDM image, but they can nonetheless be seen. There is no reason to believe that these are artifacts in the PSDM image which was migrated using a simple velocity model, and we therefore conclude that they are real within the FWI velocity model. Presumably these are levees or perhaps positions where the main channel was previously located. At their closest, these small features are about 200 m apart; at 6.5 Hz, we expect to be able to resolve features that are separated by about 150 m here, so they should be resolvable by FWI.

It is worth noting that the two methods — PSDM of subcritical reflections and FWI applied to all the data but dominated by transmitted energy — are imaging these channels largely independently. In particular, the FWI image is not constructed predominately using reflections from the channels. This can be seen in that the channels continue to be imaged outside the region where there are local receivers, and where there are therefore no recorded near-normal-incidence reflections from the channels. Instead, the FWI imaging of the channels comes principally from a mixture of shallow turning and deeper transmitted arrivals probably supplemented by interbed and/or free-surface multiples and back scattering of these wide-angle arrivals. It is therefore particularly significant that the PSDM and FWI images agree in structure because they are largely independent images.

As well as containing geologically meaningful channels, Figure 10 shows various acquisition footprints, as indeed does the PSDM image. Some vertical striping is visible within the central region; this is a residual acquisition footprint produced by the cable separation which is here larger than the model depth. Also visible is some horizontal banding toward the left and right edges of the image. This is an acquisition footprint produced by missing source lines; it does not generally affect the central region where multiple source lines were acquired into the three swaths. The short horizontal line seen bottom left is also visible as missing data in the bottom left of Figure 6 — a shot line here was not acquired at its left end for the left swath. Finally, there are four long lines visible that are vertical in the orientation of Figure 10, seen most clearly toward top right. These are produced by the edges of each of the three patches of sources used in the acquisition. Below about 350 m depth, these acquisition artifacts are no longer visible in the resultant velocity models.

In addition to this footprint, FWI has edge effects that are not readily visible within Figure 10. These occur beyond the boundary of Figure 10; they are produced predominantly by the finite aperture of the acquisition system, and they are similar in cause and appearance to the artifacts that appear in migrated images at the edge of the acquisition. As is the case for migration, these edge effects extend further into the interior of the model with increasing depth. They can be suppressed during FWI, but only at the necessary cost of incompletely capturing the velocity model toward its edges. If the edges of the model are important, then the proper solution is to acquire data that is better capable of imaging the edges by extending the acquisition area.

Although the acquisition footprint in FWI is not intense, it can be important to remove it before an FWI velocity model is used to depth migrate or depth convert reflection data. This is true of the weak shallow footprint seen in Figure 10, and the more significant effects that can occur at the edges of the acquisition region. If these are not removed, they can introduce an undesirable footprint into the migrated reflection section. In practice, the footprint in the velocity model is relatively easy to remove post-FWI using the array of heuristic tools that contractors already commonly deploy to solve similar problems in conventional velocity models and PSDM images.

Deep section

Figure 12 shows vertical depth slices through the FWI velocity model; these are the same slices as those shown for the starting model in Figure 7. Figure 12 has been truncated laterally and

vertically so that edge effects from the finite extent of the acquisition do not impinge significantly into the region shown, and the model degrades beyond the extent of Figure 12.

The resulting model is broadly similar to the starting model, as indeed it must be, but there are many differences of detail, particularly associated with the gas cloud where the original model was poor. The background velocity in parts of the model has changed, so that the depth to particular reflection horizons will also change. In principle, this should then be used to adjust the anisotropy model which has been designed in part to ensure that the depth-migrated reflection data tie accurately to the wells. In this study, we have not done that — we keep anisotropy fixed throughout the analysis. The inversion and update of anisotropy during FWI is an important topic, but we have not explored it in this study.

The inverted gas cloud is now much less blocky in its structure. It can be seen that the gas appears to have migrated laterally along particular stratigraphic intervals, and that, toward the top of Figure 12a, it appears to have moved vertically within a narrow zone, presumably a faulted or fractured region. In Figure 12a, the lateral extent of the gas-charged layer is significantly extended in the depth range 1500–2000 m, but vertically it is confined principally to just two layers. In Figure 12b, which runs through the heart of the gas cloud, low velocities extend deeper into the section than for the start model.

In Figure 12b, a region at the top of the chalk has lower than normal velocities, and this may have implications for the reservoir. We explore this further in Nangoo et al. (2012), but we note here that similar behavior is seen within the wells. Referring to Figure 1, for the well furthest outside the gas-charged zone, top-chalk appears as an abrupt increase in velocity over an interval of about 120 m from 2700 to 4600 m/s; in the well, this increase occurs as two simple steps, with no significant intervening decrease. For the well that is closest to the center of the gas, the same velocity increase occurs over a 520-m interval, and it does so as a sequence of four

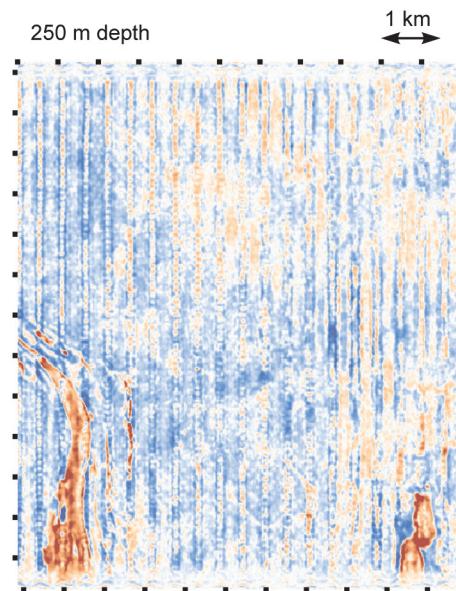


Figure 11. Shallow horizontal slice through the PSDM reflection image coincident with the central region of Figure 10. The fold of the reflection coverage is low at this depth. Color represents relative reflection amplitude on a conventional red-white-blue scale.

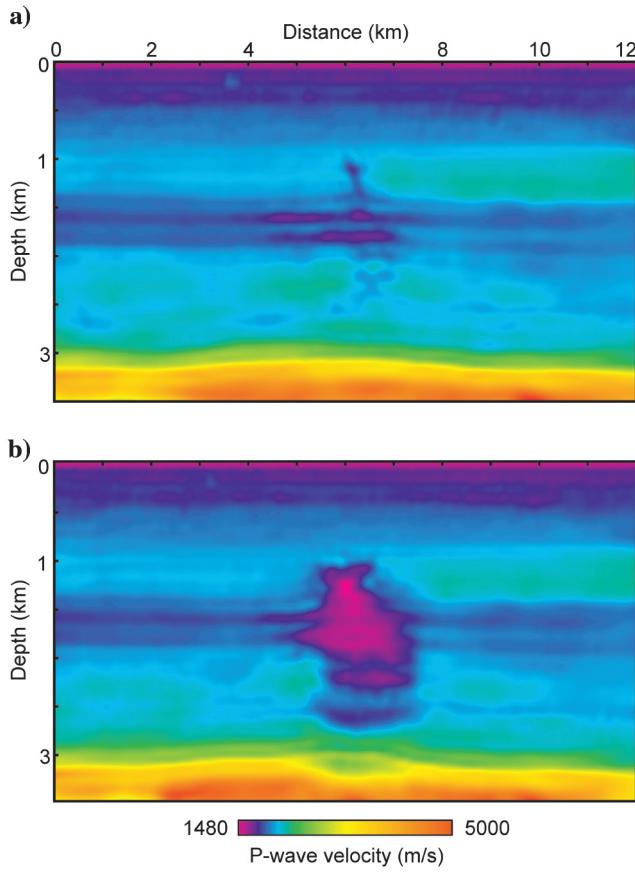


Figure 12. Vertical slices through the final FWI velocity models coincident with those shown in Figure 7. The low velocities recovered by FWI within the top of the chalk anticline are also seen in the central well.

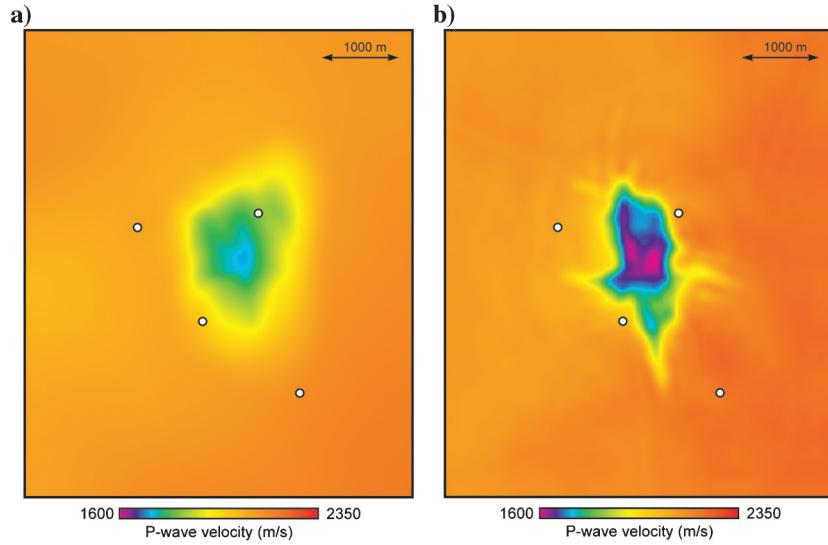


Figure 13. Horizontal slice through velocity model at 1200 m depth: (a) starting model; (b) FWI-recovered model. Circles show well locations from Figure 1.

local increases each of which is followed by a significant subsequent decrease. In addition, although the central well lies closer to the crest of the anticline, the depth at which velocity first increases above 4600 m/s lies 400 m deeper in this central well than it does in the nominally deeper outermost well. The FWI results therefore show qualitative agreement with the wells at top chalk; their direct quantitative comparison requires further refinement of the anisotropy model at depth.

Figure 13a shows a horizontal depth slice through the upper gas cloud at a depth of 1200 m for the starting velocity model. Figure 13b shows the same slice through the FWI-recovered model. Figure 14 shows the corresponding depth slice through the original PSDM volume — note that this was migrated with the starting model. All three figures show the four wells that are present in Figure 1.

Comparing Figure 13a and 13b, we see that the inversion has tightened the boundaries of the gas cloud, it has intensified the magnitude of the lowest velocities within it which now drop to below 1600 m/s, and it has introduced several sharp prominences where previously the outline of the gas cloud was smooth. The well at top right has moved from within to outside the main region of the gas cloud at this level as the boundary has sharpened; we compare velocities within this well below. We note that Figure 12a passes through the southernmost low-velocity prominence seen in Figure 13b, and this coincides with the narrow vertical low-velocity feature seen in Figure 12a.

Comparing Figures 13b and 14, we see that the geometry of the two are clearly related, and especially that three of the four prominences that appear in Figure 13b coincide with similar geometric features in Figure 14. This coincidence occurs not only at this level, but is consistent at the full depth range over which these prominences are seen. We note that the PSDM image and the FWI velocity model are generated predominantly by different subsets of the field data; the former employs only back-scattered reflected energy, while the latter relies most heavily upon forward-scattered transmitted energy. The close agreement between Figure 13b and

Figure 14 is therefore particularly significant. We conclude that the FWI velocity model is revealing genuine structure around the gas cloud — on the reflection data, these prominences are seen to be fault zones, and the gas is clearly invading the faults.

Figure 15 shows a comparison of the starting and FWI velocity models with wire-line sonic measurements made within the well circled on Figure 1, spanning the depth interval of Figures 13 and 14. The sonic velocity is shown in gray, the starting model in red, and the FWI model in blue. When interpreting Figure 15, it should be remembered that the minimum seismic wavelength used for FWI at this depth is about 300 m, so that we do not expect to resolve velocity structure below about 150 m; the velocity models are themselves only defined at 50-m grid spacing. All three traces show vertical velocity.

It is clear from Figure 15 that both models follow the long-wavelength trend of the well except in the depth interval around 1200 m. Here, the starting model is significantly in error — it is

clear that the manual model building used to insert the gas cloud into the tomographic velocity model was not accurate at this point. FWI has correctly removed this erroneous region of low velocity, and provides a much closer match to the well. Figure 15 therefore provides further support for the hypothesis that the FWI model better represents the subsurface than does the starting model. There are no wells that penetrate the heart of the gas cloud, and so we cannot confirm directly that the very-low velocities seen there are quantitatively correct.

Match of synthetic to field data

In addition to the match to the geometry of the depth migrated reflection data and to the absolute velocities in wells, the match of synthetic data calculated using the final velocity model to the raw field data is an important tool for quality assurance. Figure 16 shows a shot record acquired on a single cable that passes through the central region of the gas cloud. Figure 16a shows the field data, Figure 16b shows the equivalent shot record generated using the starting model, and Figure 16c shows the equivalent record obtained using the final FWI velocity model. The dashed line delineates the region within which the field data were muted prior to FWI.

There are two key regions of these figures that are relevant to assessing the performance of FWI. The most obvious of these is the bright postcritical reflection from the top of the chalk, seen at late times and longer offsets. As discussed earlier, the amplitude, timing, and appearance of this event is strongly influenced by the low-velocity gas cloud. This event is apparent in the starting model, but its location, absolute amplitude, and spatial variation are not correct, and it does not reproduce the weak diffracted event seen in the field data that slopes toward shorter offsets. After FWI, the agreement between the synthetic and field data is much improved — the event is correctly located, it reproduces the high associated amplitudes, it reproduces their decay in time and broadly their variation laterally, and it reproduces the diffracted event. Although it is hard to judge from a single printed figure alone, the quantitative match of traveltimes is also close.

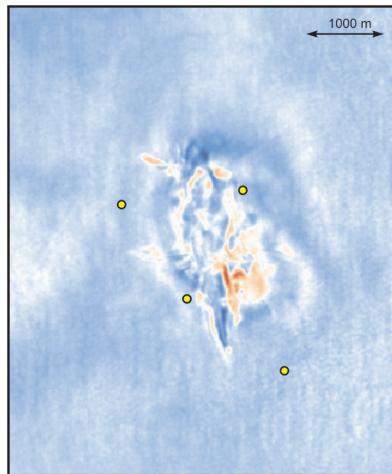


Figure 14. Horizontal slice through the PSDM reflection image coincident with Figure 13. Circles show well locations from Figure 1. Color represents relative reflection amplitude on a conventional red-white-blue scale.

The second feature that is significant in Figure 16 is the detailed structure of the early refracted arrivals and their ghosts and multiples that form the major suite of arrivals. These look superficially similar in all three figures. Careful inspection, however, shows that many of the fine details that are produced by the interference of coexistent phases are properly reproduced in the FWI record but not the starting model. These are easiest to see in the printed figure where there are local drops in amplitude, or jumps in local phase velocity, visible in the field data. Most of these detailed features are reproduced accurately by the FWI model. These features mostly represent small changes to the macrovelocity model that are required to match the arrival times of particular phases that are hidden within the larger suite of arrivals.

In Figure 16, it is difficult to judge the match quantitatively. Figure 17 shows the same shot record displayed in a format that makes this easier. Here, the data are plotted every 250 m, with a trace from the field data plotted (left) next to an equivalent trace from the synthetic FWI model (plotted right). The traces have been normalized — that is, the absolute values of amplitude are not meaningful, but the relative amplitude within each trace is preserved. Unlike Figure 16, Figure 17 has had the FWI bottom mute applied.

Examining Figure 17 trace by trace, we see that the quantitative match between field and model data is close. We match the traveltimes of every phase, we match their waveforms, we match their

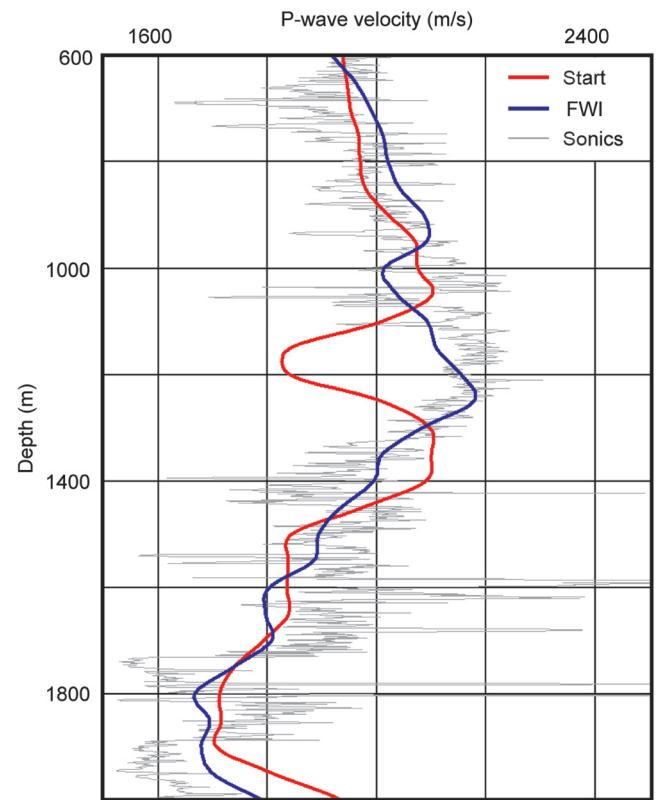


Figure 15. Comparison between well data, and starting and FWI velocity models. The well is shown circled in Figure 1, and lies on the periphery of the gas cloud. The FWI model more closely matches the well. Note that the minimum seismic wavelength used during FWI was about 300 m, and that the starting and FWI velocity models are defined on a grid with a spacing of 50 m.

relative amplitudes within individual traces, and we match the decay of their coda which is different for different phases. It is interesting to note that, as explained above, although we do not pay undue attention to the amplitudes of the field data, our FWI model does reproduce with quite a high degree of accuracy the amplitudes of all the main phases in the field data. In particular, it can reproduce the high amplitudes of the posterical reflection from top chalk where these are affected by the gas cloud; fitting these high amplitudes, even though they have been partially suppressed by normalization, is clearly helping to drive the inversion in the right direction.

Close examination of Figure 17 does reveal a systematic mismatch between model and data. In all cases, the data predicted by the FWI model is delayed slightly with respect to the field data. Careful examination of large numbers of traces also reveals that this

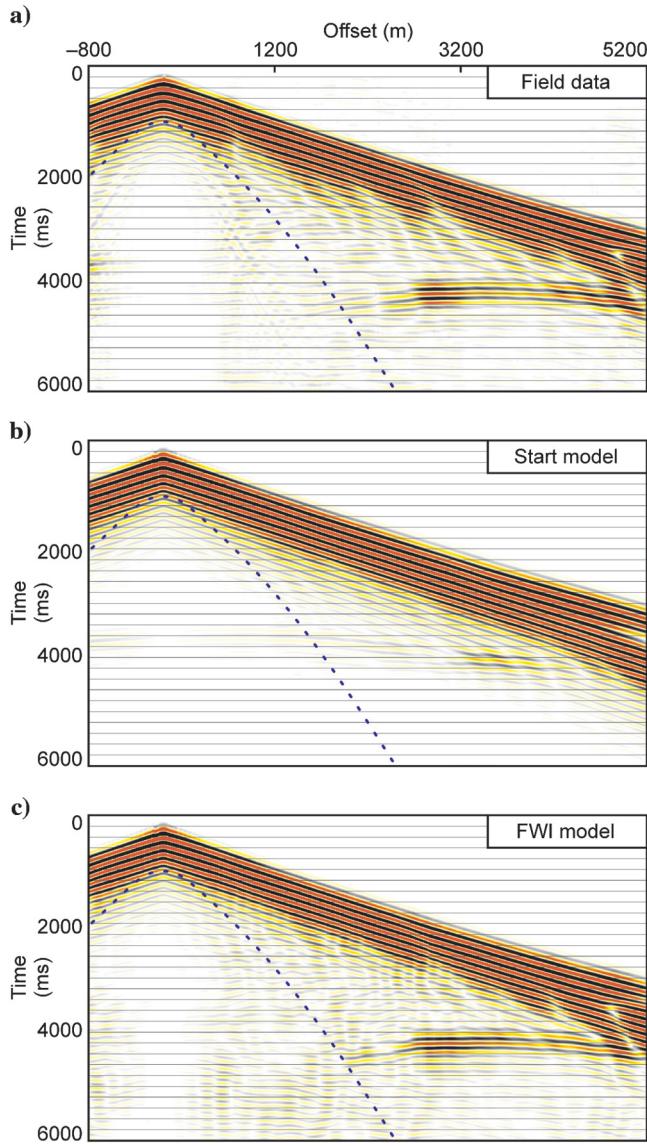


Figure 16. A single shot record in the region affected by the gas cloud: (a) original field data; (b) synthetics generated using the starting model; (c) synthetics generated using the FWI model. The dashed line marks the position of the bottom mute applied prior to FWI.

delay is not obviously a function of offset, azimuth, or position within the model. It also reveals that the delay is slightly longer later within a long wave-train, which suggests that it may be related to the accuracy with which water-bottom multiples are reproduced because it is these that dominate the coda. The effects are not large, but they are consistent throughout the model. For the earlier portion of the wave-train, the mismatch averages about 12 ms, equivalent to about 30° of phase at the maximum frequency of 6.5 Hz.

At first glance, it is not clear how such a systematic mismatch could arise because FWI is designed specifically to remove such effects. It is possible that the mismatch originates because we have used an inappropriate source wavelet. We note that if we use the contractor's wavelet to generate the synthetics, then the mismatch becomes larger. We also note that our wavelet does match the absolute time of the first arrival at short offset when this is modeled using only a water layer, a free-surface, and the source and receiver in their correct positions. We therefore think that it is unlikely that the explanation lies with the source wavelet, though that remains a possibility. We note that, if we invert for the source wavelet as part of the inversion, or arbitrarily change it, then this reduces the effect on the final synthetics, but this does not imply that modification of the source is necessarily the correct explanation.

We suspect that the explanation is related to the problem of properly capturing the water depth when using a coarse mesh. We model and invert on a 50 m cubic mesh. The water layer is about 75 m deep. We do not allow the inversion to alter the velocity or the density in the water layer which are held at their physical values of 1480 m/s and 1000 kg m⁻³. We do allow the inversion to change the velocity at the first grid point below the seabed, and this also changes the density here via Gardner's law. Changing the velocity below the seabed also implicitly affects the water depth which is not otherwise specified explicitly, and it changes the seafloor reflection coefficient.

The finite-difference codes are running on this coarse mesh, and they implicitly assume that velocity and density vary smoothly (and minimally) between nodes. At the seabed, and wherever there is rapid change in velocity and density, there is implicitly then a question of how to upscale correctly from a fine-scale model that properly describes the sharp sea floor, to the coarser model that most closely reproduces the wavefield that a finer model would have generated. This is not a trivial problem — it depends upon propagation

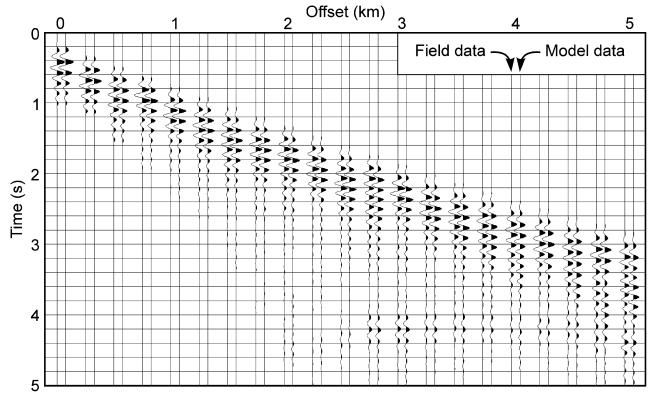


Figure 17. Trace-by-trace comparison between the original field data and the final FWI synthetics for a shot passing beneath the gas cloud.

angle and frequency, it depends upon the specific details of the numerical method that is being used to solve the wave equation, it is different for velocity and density, and it affects transmitted and reflected waves differently.

Consequently, there is no single coarse model that can properly describe the response of the water layer and the sea floor, at least not one that maintains physically meaningful values of velocity and density immediately above and below the seabed. We therefore think that the residual mismatch that we see may represent the best least-squares compromise that the inversion can find between fitting the absolute traveltimes, the water-multiple periodicity, and the sea-floor reflectivity which in turn controls multiple amplitudes and hence the temporal decay of the wave-train. We are investigating this further, and we surmise that it may be possible to deal with it by introducing nonphysical velocities and densities into the cells immediately adjacent to the seafloor. These velocities would be those that most accurately upscale the sharp seafloor onto a coarser mesh. If we are correct in this explanation, then we note that the effect will not be seen in synthetic studies that use the same mesh for initial data generation and FWI, and indeed we do not see it in our own such studies. It will also likely be a common problem for FWI schemes on coarse meshes applied to marine field data, though we have not seen it previously reported.

Migration

The final test for an FWI model is to use it to depth migrate the reflection data. The migrated images that we show below were generated using the original 3D contractor's model of anisotropy including high values of delta and epsilon within the chalk rather than the 1D version that we employed during FWI; this change makes minimal qualitative difference to the migrations. In addition, prior to migration, the final velocity model was processed to remove the edge effects that are introduced during FWI. This processing ramped the FWI velocity model smoothly back into the starting model, in depth over the range 3250–3750 m, and laterally around the periphery of the model. For these migrations, we did not apply any smoothing to the FWI model to remove the minor shallow acquisition footprint.

Figure 18 shows common-image gathers from a Kirchhoff depth migration using the starting model and the FWI model, centered on the top chalk, and passing through the periphery of the gas cloud — in the center of the gas cloud, there are no deep reflections to migrate. A good model should flatten these gathers. There is residual multiple energy within the gathers, which complicates their analysis. It is clear, however, that the FWI model does a significantly better job at flattening the image gathers than does the starting model. We note that the starting model was generated using traveltime tomography, and that process is designed to produce flattened gathers using ray-based traveltime corrections; these are evidently not adequate to deal with this data set close to the gas cloud.

Figure 19a shows a vertical depth slice through the PSDM volume, that has been migrated using the starting model, and Figure 19b shows the same slice migrated using the FWI model. Both sections were migrated using a high-fidelity RTM algorithm because the velocity model produced by FWI is of greater resolution than can be properly accommodated by ray-based migration methods. Note that Figure 19a is not the same processing and migration as was shown in Figure 3 — it is a new version designed to be directly comparable with the FWI-migrated version.

Within the clastic section in Figure 19, above about 2000 m, there are few significant differences between the two migrations although there are some differences of detail within the gas-affected central region. In the depth range 2000–2400 m, reflector continuity is similar in the two migrations, but their geometry is different. On the FWI section, subhorizontal events here remain approximately subhorizontal as they approach the gas-charged region, whereas on the original migration there is significant pull up beneath the gas cloud; the latter is presumably spurious overcorrection for the low-velocity cloud. The most significant differences between the two migrations though appear within the deeper target reservoir section below about 2800 m. Here there are many differences, with the FWI section being superior in all respects. The FWI-migrated section is less complicated, it removes conflicting cross-dips, it improves continuity on a large and a fine scale, and it is clearly better focused as events are brighter, sharper and less noisy. It is worth noting that the migration uplift seen between Figure 19a and 19b is produced almost entirely by improving the shallow velocity

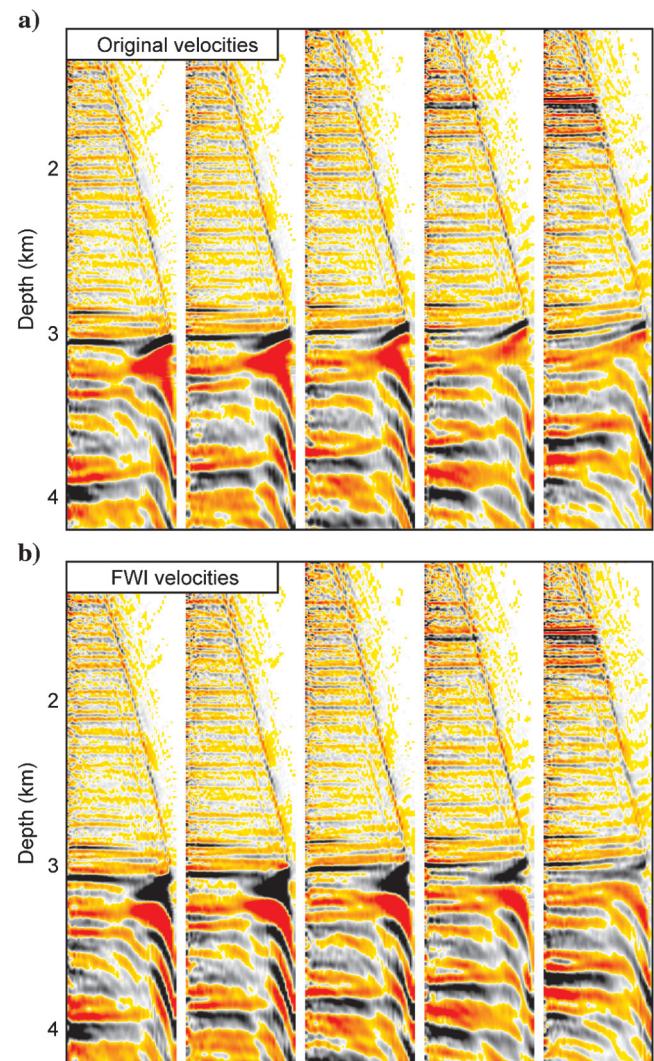


Figure 18. Kirchhoff-migrated common-image gathers: (a) migrated using the starting model; (b) migrated using the final FWI model. The maximum offset shown is 5000 m.

model down to about 2000 m, and that the improvement in migration is seen predominantly below this depth.

Figure 20 shows an even more dramatic example of the uplift obtained by using FWI-based migration. This shows a horizontal slice at a depth of 3600 m, within the chalk section. Figure 20a shows the RTM using the starting model, and Figure 20b using the FWI model. In the former, the structure shown represents an elongate dome, with significant deformation, and a moderately complicated planform. In contrast, the FWI migration in Figure 20b

shows a much simpler elongate dome, with minimal apparent deformation, and a simple elliptical planform.

The velocity model used to migrate Figure 20b was significantly more complicated than that used to migrate Figure 20a. The result of this more complicated migration, however, is a much-simplified final structure. There is only one way realistically that a more-complicated velocity model can produce a simpler outcome, and that is if the structure at depth is simple and if the more-complicated velocity model is more correct. Figure 20 therefore provides strong

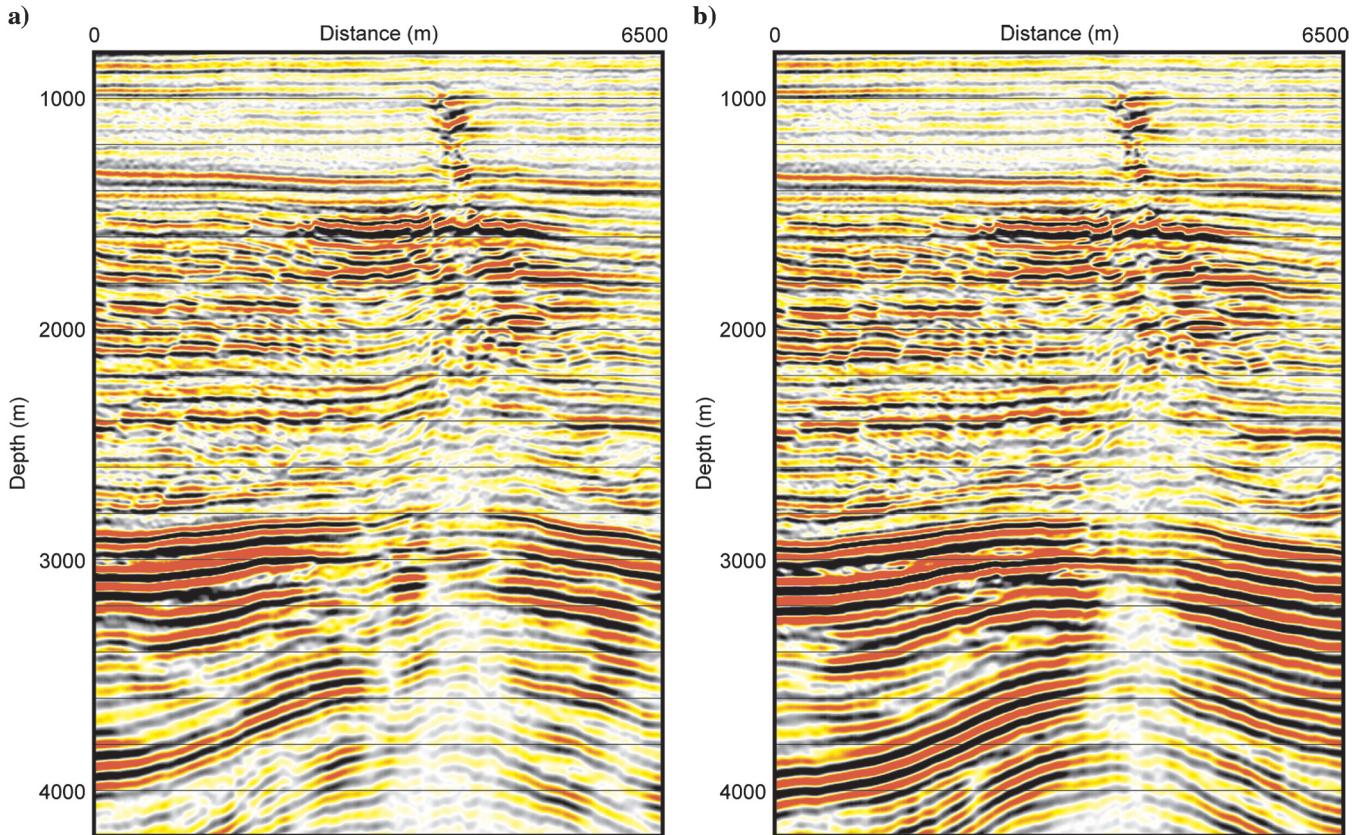
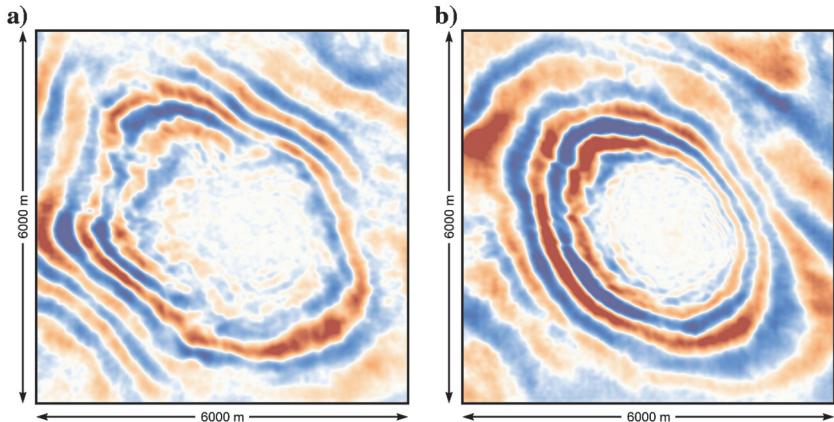


Figure 19. Reverse-time migrated vertical sections: (a) using the starting model; (b) using the final FWI model. The section lies approximately along the line shown in Figure 7a. There is no amplitude equalization; both sections are displayed at the same gain. Use of the FWI velocity model provides a significant uplift in final image quality below about 2800 m depth.

Figure 20. Reverse-time migrated horizontal depth slices at 3600 m depth. (a) Using the starting model. (b) Using the final FWI model. Both sections are displayed at the same gain. Use of the FWI velocity model removes major structural artifacts from the migrated image revealing a more-regular broad dome. These changes significantly influence predicted reservoir geometry.



support for the veracity of the FWI velocity model, and for its utility for deeper depth migration. The outcome of the improved migration in Figure 20 is to change the geometry of the reservoir section, and in principle there will be consequent implications for the evaluation of the resource.

CONCLUSION

We have successfully applied acoustic full-waveform seismic inversion to a 3D anisotropic field data set. We have demonstrated an algorithmic approach, a strategy, and a workflow that moves from the original raw field data to a final inverted high-resolution high-fidelity velocity model. We have shown that this model depth migrates the reflected portion of the data set better than does a conventional model. We have shown that the model better matches well data, generates synthetic data that match the field data at all offsets trace-by-trace, matches geometric features seen in the reflection data at high resolution, and better flattens common-image gathers. We have explained our parameterization and rationale in detail, both of which are based upon extensive prior testing, and we have provided a work flow that is largely generic and that others can follow to process and invert analogous data sets.

Two elements are particularly important: (1) ensuring that the combination of starting model, source wavelet and preprocessed field data are not cycle skipped at the lowest frequencies used to begin the inversion, and (2) processing and inverting the data in a way such that maximum information is retained and used by the inversion while suppressing the effect of those elements of the data that a particular algorithm will not in principle be able to model — in this case, these are principally the effects of attenuation, elasticity, and anomalous density.

We have demonstrated the utility of using few sources per iteration, and shown that this leads to a more than fivefold reduction in compute time. We have outlined some of the principal pitfalls that may be encountered in applying practical FWI on real data, and explained how to identify and circumvent these. We stress the differences that may be encountered between inverting real data and inverting synthetic data especially when the latter have been generated using simplifying assumptions that match those that are also used in the inversion.

It is our intention that this paper will help to move FWI from an experimental and specialist technique, that can on occasion be difficult, expensive, and unreliable, into a practical and routine process that is straightforward, robust, cost-effective, and properly quality-assured.

ACKNOWLEDGMENTS

The authors would like to thank the PL044 partnership: ConocoPhillips Skandinavia AS, Total E&P Norge AS, ENI Norge AS, Statoil Petroleum AS for their kind permission to use their data and publish this work. We would like to thank Paul Valasek, formerly of ConocoPhillips and now of Maersk, for originally identifying this data set as suitable for FWI analysis, and making it available to us. Imperial College London wishes to thank the sponsors of the FULLWAVE consortia consisting of: BG, BP, Chevron, CGGVeritas, ConocoPhillips, DONG, ENI, HESS, Maersk, Nexen, OHM, Rio Tinto, Shell, TOTAL, and Tullow Oil, under the U. K. ITF program with additional support from the DTI, BEER, and BIS. Richard Wombell, Chris Purcell, Clare Grice, Shirine Swan, Yamen Belhassen, Thomas Hellmann, and Grunde

Rønholt of CGGVeritas are thanked for their assistance in supporting the project and in preparing the data. The 3D ProMAX/SeisSpace package, supplied by Halliburton under a university software grant, was used to preprocess and analyze the field data within Imperial College. The inversion software used in this study is operating commercially within CGGVeritas; access to the software is also available for application to academic problems through collaboration with Imperial College London, and to commercial partners through membership of the FULLWAVE consortium.

REFERENCES

- Alkhalifah, T., 2000, An acoustic wave equation for anisotropic media: *Geophysics*, **65**, 1239–1250, doi: [10.1190/1.1444815](https://doi.org/10.1190/1.1444815).
- Ben-Hadj-Ali, H., S. Operto, and J. Virieux, 2008, Velocity model building by 3D frequency-domain, full-waveform inversion of wide-aperture seismic data: *Geophysics*, **73**, no. 5, VE101–VE117, doi: [10.1190/1.2957948](https://doi.org/10.1190/1.2957948).
- Ben-Hadj-Ali, H., S. Operto, and J. Virieux, 2011, An efficient frequency-domain full waveform inversion method using simultaneous encoded sources: *Geophysics*, **76**, no. 4, R109–R124, doi: [10.1190/1.3581357](https://doi.org/10.1190/1.3581357).
- Betzig, E., J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak, 1991, Breaking the diffraction barrier — optical microscopy on a nanometric scale: *Science*, **251**, 1468–1470, doi: [10.1126/science.251.5000.1468](https://doi.org/10.1126/science.251.5000.1468).
- Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: *Geophysics*, **60**, 1457–1473, doi: [10.1190/1.1443880](https://doi.org/10.1190/1.1443880).
- Díaz, E., and A. Guittot, 2011, Fast full-waveform inversion with random shot decimation: 81st Annual International Meeting, SEG, Expanded Abstracts, 2804–2808.
- Duvaneck, E., P. Milcik, P. Bakker, and C. Perkins, 2008, Acoustic VTI wave equation and their application for anisotropic reverse-time migration: 78th Annual International Meeting, SEG, Expanded Abstracts, 2186–2190.
- Gardner, G. H. F., L. W. Gardner, and A. R. Gregory, 1974, Formation velocity and density — the diagnostic basics for stratigraphic traps: *Geophysics*, **39**, 770–780, doi: [10.1190/1.1440465](https://doi.org/10.1190/1.1440465).
- Granli, J. R., B. Arntsen, A. Sollid, and E. Hilde, 1999, Imaging through gas-filled sediments using marine shear-wave data: *Geophysics*, **64**, 668–677, doi: [10.1190/1.1444576](https://doi.org/10.1190/1.1444576).
- Guasch, L., M. R. Warner, T. Nangoo, J. Morgan, A. P. Umpleby, I. Štekl, and N. Shah, 2012, Elastic 3D full-waveform inversion: 82nd Annual International Meeting, SEG, Expanded Abstracts, SI6.8.
- Hicks, G. J., 2002, Arbitrary source and receiver positioning in finite-difference schemes using Kaiser windowed sinc functions: *Geophysics*, **67**, 156–166, doi: [10.1190/1.1451454](https://doi.org/10.1190/1.1451454).
- Nangoo, T., M. Warner, J. Morgan, A. Umpleby, I. Štekl, and A. Bertrand, 2012, Full-waveform seismic inversion at reservoir depths: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, W015.
- Plessix, R.-E., and Q. Cao, 2011, A parameterisation study for surface seismic full-waveform inversion in a acoustic vertical transversely isotropic medium: *Geophysical Journal International*, **185**, 539–556, doi: [10.1111/gji.2011.185.issue-1](https://doi.org/10.1111/gji.2011.185.issue-1).
- Plessix, R.-E., and C. Perkins, 2010, Full waveform inversion of a deep water ocean bottom seismometer data set: *First Break*, **28**, 71–78.
- Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part I: Theory and verification in a physical scale model: *Geophysics*, **64**, 888–901, doi: [10.1190/1.1444597](https://doi.org/10.1190/1.1444597).
- Pratt, R. G., C. Shin, and G. H. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: *Geophysical Journal International*, **133**, 341–362, doi: [10.1046/j.1365-246X.1998.00498.x](https://doi.org/10.1046/j.1365-246X.1998.00498.x).
- Pratt, R. G., and R. M. Shipp, 1999, Seismic waveform inversion in the frequency domain, Part 2: Fault delineation in sediments using crosshole data: *Geophysics*, **64**, 901–913, doi: [10.1190/1.1444598](https://doi.org/10.1190/1.1444598).
- Pratt, R. G., Z.-M. Song, P. Williamson, and M. Warner, 1996, Two-dimensional velocity models from wide-angle seismic data by wavefield inversion: *Geophysical Journal International*, **124**, 323–340, doi: [10.1111/gji.1996.124.issue-2](https://doi.org/10.1111/gji.1996.124.issue-2).
- Prioux, V., R. Brossier, Y. Gholami, S. Operto, J. Virieux, O. I. Barkved, and J. H. Kommedal, 2011, On the footprint of anisotropy on isotropic full waveform inversion: The Valhall case study: *Geophysical Journal International*, **187**, 1495–1515, doi: [10.1111/gji.2011.187.issue-3](https://doi.org/10.1111/gji.2011.187.issue-3).

- Ratcliffe, A., C. Win, V. Vinje, G. Conroy, M. Warner, A. Umpleby, I. Štekl, T. Nangoo, and A. Bertrand, 2011, Full waveform inversion: A North Sea OBS case study: 81st Annual International Meeting, SEG, Expanded Abstracts, 2384–2388.
- Schuster, G. T., X. Wang, Y. Huang, W. Dai, and C. Boonyasiriwat, 2011, Theory of multi-source cross-talk reduction by phase-encoded statics: *Geophysical Journal International*, **184**, 1289–1303, doi: [10.1111/gji.2011.184.issue-3](https://doi.org/10.1111/gji.2011.184.issue-3).
- Shah, N., M. Warner, T. Nangoo, A. Umpleby, I. Štekl, J. Morgan, and L. Guasch, 2012, Quality assured full-waveform inversion: Ensuring starting model adequacy: 81st Annual International Meeting, SEG, Expanded Abstracts, SI3.5.
- Shin, C., S. Jang, and D.-J. Min, 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: *Geophysical Prospecting*, **49**, 592–606, doi: [10.1046/j.1365-2478.2001.00279.x](https://doi.org/10.1046/j.1365-2478.2001.00279.x).
- Sirgue, L., 2006, The importance of low frequency and large offset in waveform inversion: 68th Annual International Conference and Exhibition, EAGE, Extended Abstracts, A037.
- Sirgue, L., O. I. Barkved, J. Dellinger, J. Etgen, U. Albertin, and J. H. Kommedal, 2010, Full waveform inversion: The next leap forward in imaging at Valhall: *First Break*, **28**, 65–70.
- Sirgue, L., J. T. Etgen, and U. Albertin, 2008, 3D frequency-domain waveform inversion using time-domain finite-difference methods: 70th Annual International Conference and Exhibition, EAGE, Extended Abstracts, F022.
- Sirgue, L., and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: *Geophysics*, **69**, 231–248, doi: [10.1190/1.1649391](https://doi.org/10.1190/1.1649391).
- Štekl, I., A. Umpleby, and M. Warner, 2010, Seismic anisotropy effects in 3D full-wavefield tomography: 72nd Annual International Conference and Exhibition, EAGE, Extended Abstracts, Workshop WS6.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266, doi: [10.1190/1.1441754](https://doi.org/10.1190/1.1441754).
- Thomsen, L., 1986, Weak elastic anisotropy: *Geophysics*, **51**, 1954–1966, doi: [10.1190/1.1442051](https://doi.org/10.1190/1.1442051).
- Umpleby, A., M. Warner, and I. Štekl, 2010, Time vs frequency for 3D wavefield tomography: 72nd Annual International Conference and Exhibition, EAGE, Extended Abstracts, Workshop WS6.
- van Leeuwen, T., and F. J. Herrmann, 2012, Fast waveform inversion without source-encoding: *Geophysical Prospecting*, ; <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2478.2012.01096.x/abstract>, accessed 31 October 2012.
- Shipp, R. M., and S. C. Singh, 2002, Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data: *Geophysical Journal International*, **151**, 325–344, doi: [10.1046/j.1365-246X.2002.01645.x](https://doi.org/10.1046/j.1365-246X.2002.01645.x).
- Vinje, V., A. Ratcliffe, A. Bertrand, T. Hellmann, and Y. Belhassen, 2012, Improved PP and PS imaging using full-waveform inversion — Tommeliten OBC case study: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, W010.
- Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, no. 6, WCC1–WCC26, doi: [10.1190/1.3238367](https://doi.org/10.1190/1.3238367).
- Warner, M., I. Štekl, and A. Umpleby, 2008, Efficient and effective 3D wavefield tomography: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, F023.
- Warner, M. R., J. V. Morgan, A. Umpleby, I. Štekl, and L. Guasch, 2012, Which physics for full-wavefield inversion: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, W029.
- Williamson, P. R., 1991, A guide to the limits of resolution imposed by scattering in ray tomography: *Geophysics*, **56**, 202–207, doi: [10.1190/1.1443032](https://doi.org/10.1190/1.1443032).