

Adaptive waveform inversion: Practice

Lluís Guasch¹, Michael Warner², and Céline Ravaut³

ABSTRACT

Adaptive waveform inversion (AWI) reformulates the misfit function used to perform full-waveform inversion (FWI), so that it no longer contains local minima related to cycle skipping. It does this by finding a model that drives the ratio of the predicted and observed data sets to unity rather than driving the difference between these two data sets to zero as is the case for conventional FWI. We apply AWI to a 3D field data set acquired over a pervasive gas cloud in the North Sea, comparing its performance with that of conventional FWI in a variety of circumstances. When starting inversion from 3 Hz, and using a good starting model obtained from reflection tomography, FWI and AWI generate similar models although the FWI result contains edge artifacts that are not produced by AWI. However, when the starting frequency is increased to approximately 6 Hz, or when

the starting model is less accurate, FWI fails to recover a good model whereas AWI continues to converge. When both of these conditions apply, FWI fails comprehensively, leading to a model that is significantly worse than the starting model, whereas the AWI result remains largely unaffected. We applied Kirchhoff depth migration to the fully-processed data using the FWI result obtained following reflection tomography, and using the AWI result obtained from a simple one-dimensional starting model. We use the resulting migrated volumes, together with measures of residual moveout throughout the volume, to show that the AWI result from a simple starting model is at least as good as the FWI result obtained following tomography. We conclude that AWI is robust in the presence of cycle skipping on this 3D field data set, and can proceed successfully from a less-accurate starting model, and from a higher starting frequency, in circumstances in which FWI fails completely.

INTRODUCTION

Full-waveform inversion (FWI) is a seismic data inversion technique that seeks to recover a high-resolution model of subsurface physical properties by minimizing the misfit between an observed data set and an equivalent data set simulated using the recovered model. There are many potential ways to measure such a data misfit. In this paper, we explore the application of the measure of misfit that is provided by adaptive waveform inversion (AWI), introduced by Warner and Guasch (2014, 2016), to a 3D anisotropic field data set, and we show that this can have significant advantages over more-established measures. In particular, this new measure appears to allow the relaxation of the requirement to have low-frequency

field data and a high-quality starting model, which are normally prerequisites for successful FWI.

Most current implementations of FWI are based upon the principles and ideas formulated by Lailly (1983) and Tarantola (1984). Computational limitations restricted the initial application of FWI to simple 1D and 2D synthetic examples (Gauthier et al., 1986; Mora, 1987). Subsequent advances in the FWI algorithms and in the computational hardware, enabled the application of the technique to more-realistic synthetic problems (Pratt et al., 1996; Pratt, 1999), to 2D field data (Igel et al., 1996), and ultimately to 3D field data sets (Ben-Hadj-Ali et al., 2008; Sirgue et al., 2008; Warner et al., 2008). Early progress was made in two dimensions in the sparse frequency domain (Pratt et al., 1998; Pratt, 1999).

Manuscript received by the Editor 22 May 2018; revised manuscript received 12 October 2018; published ahead of production 21 February 2019; published online 12 April 2019.

¹S-Cube, 52 Princes Gate, London SW7 2PG, UK. E-mail: lguasch@s-cube.com.

²Imperial College London, Department of Earth Science and Engineering, South Kensington Campus, London SW7 2AZ, UK. E-mail: m.warner@imperial.ac.uk.

³Equinor ASA, Norway Research and Technology, Margrethe Jørgensens vei 4, 9406 Harstad, Norway. E-mail: celr@equinor.com.

© The Authors. Published by the Society of Exploration Geophysicists. All article content, except where otherwise noted (including republished material), is licensed under a Creative Commons Attribution 4.0 Unported License (CC BY). See <http://creativecommons.org/licenses/by/4.0/>. Distribution or reproduction of this work in whole or in part commercially or noncommercially requires full attribution of the original publication, including its digital object identifier (DOI).

Three-dimensional applications to field data (Plessix and Perkins, 2010; Vigh et al., 2010; Warner et al., 2013; Routh et al., 2017) have established time-domain 3D anisotropic acoustic FWI as a standard tool for high-resolution velocity-model building and depth imaging. In parallel with these developments, regional and global seismologists have applied FWI and related techniques to earthquake data to recover elastic models of the earth at larger scale lengths (Fichtner et al., 2008, 2009; Tape et al., 2010).

Under the right circumstances, which include optimal acquisition geometry, availability of low frequencies in the field data, and a good starting model, FWI can resolve subsurface models accurately at all scale lengths longer than about half the local seismic wavelength (Virieux and Operto, 2009). This represents a substantial improvement over methods that are based upon matching traveltimes, for example, traveltime tomography, where the theoretical resolution limit is related to the size of the Fresnel zone rather than to the wavelength. FWI also benefits from minimal data preprocessing when compared to conventional velocity analysis using, for example, residual-moveout reflection traveltime tomography.

Although it can bring improved resolution and fidelity, one of the main limitations of FWI is the phenomenon known as cycle skipping. This occurs when the starting model is unable to predict data that match the observed field data to an accuracy of better than half a cycle at the lowest usable frequencies available within the data set. In these circumstances, conventional FWI is liable to become trapped within a local minimum in the misfit function, for which the recovered model produces predicted data of which some portion is shifted in time with respect to the observed data by approximately an integer number of cycles. Cycle skipping during FWI is normally addressed by using especially low-frequency field data, by using an accurate starting velocity model, and by rigorous quality control using experienced practitioners.

The central purpose of AWI is to reformulate the waveform inversion problem so that it is no longer affected by cycle skipping. Several other authors have also made this attempt. Of these, the methods most closely related to AWI include van Leeuwen and Mulder (2008, 2010), who use an objective function that penalizes energy at nonzero lag in the crosscorrelation of the observed and predicted data sets; the mathematics of this method has similarities to those of AWI, as does the approach of Luo and Sava (2011). Ma and Hale (2013) design an objective function based upon dynamic warping to match the observed and predicted data. AWI uses a similar rationale but replaces nonlinear dynamic warping by linear least-squares matching filters. Biondi and Almomin (2012, 2014) present a method called tomographic FWI that introduces an additional nonphysical dimension into the subsurface velocity model such that the model always predicts the data accurately; the inversion then acts to focus this nonphysical extension out of the final model. This method provided the initial motivation for AWI, which also introduces a nonphysical extension, which must be focused out of the final result. Using wavefield-reconstruction inversion, van Leeuwen and Herrmann (2013, 2015) use a different nonphysical model extension, in which the wavefield no longer exactly obeys the wave equation, and the inversion again acts to find a model for which this nonphysical feature progressively disappears. Huang and Symes (2015) introduce a method similar to AWI, but that use a source that is extended in time and space. Their approach suggests that there is a continuum of methods that extends from that of van Leeuwen and Herrmann (2013, 2015), through Huang and Symes (2015), to AWI as presented here.

Several authors have proposed FWI schemes that seek to overcome cycle skipping by dealing directly with explicit phase or time differences. Among these, Bozdağ et al. (2011) use signal envelopes, which are inherently lower frequency than the raw data and are therefore less subject to cycle skipping, Shah et al. (2012) and Alkhalifah and Choi (2012) attempt to unwrap phase differences at single frequencies in space, and Jiao et al. (2015) present the method of adjustive FWI, which seeks to minimize explicit time shifts between specific arrivals. Métivier et al. (2016) use the concept of optimal transport distance to quantify the misfit between predicted and observed data sets.

Although there is diversity in how each of these methods, including AWI, formulates the waveform-inversion problem, they all share the common feature that they define a new objective function that, under appropriate circumstances, does not pass through a minimum when the predicted and observed data sets differ by a wave cycle. Some of these methods appear unaffordably expensive in three dimensions, some lack robustness when applied to complicated or noisy data sets, and some suffer from new local minima that are unrelated to cycle skipping but that are peculiar to particular methodologies. It remains to be seen which of these methods will prove to be sufficiently robust, efficient, and effective, so that they can be applied with confidence to a wide range of field data sets and problems. The thesis advanced in this paper is that AWI does have those desirable characteristics.

We apply AWI to a 3D ocean-bottom North Sea anisotropic field data set. AWI uses a measure of data misfit that does not have secondary minima associated with cycle-skipped solutions; consequently, the method is immune to the effects of cycle skipping. This immunity does not however mean that AWI will also have immunity to the many other causes of secondary local minima that can affect conventional FWI, and of course any new method may also introduce new problems, or new categories of local minima, that are peculiar to itself when applied to field data. Note that, by *cycle skipping*, we mean an effect that is produced by the finite bandwidth of the observed seismic data, in which there is a misalignment in time of a packet of energy in the predicted seismic data by an approximately integer number of wave cycles such that it produces an apparent fit to the observed data. Cycle skipping in FWI is then necessarily overcome by extending the bandwidth of the observed data to lower frequencies. In contrast, the misalignment of one arrival with another to which it is unrelated, for example, the misidentification of a surface multiple as a primary arrival, will not normally be overcome by extending the bandwidth of the data downward. If such a misalignment produces a local minimum in the conventional FWI objective function that cannot be overcome by a reasonable downward extension of the bandwidth, then this is not *cycle skipping* as that term is conventionally applied and as we use the term here. Warner and Guasch (2016) show that AWI avoids conventional cycle skipping when applied to 2D synthetic data in challenging circumstances when conventional FWI fails entirely. However, the literature of FWI is replete with proposed methods that work in principle and on synthetic inverse-crime data sets, but that fail completely when they are applied to real field data sets. Consequently, it is important to demonstrate for any new method that it is affordable and effective on field data, and that its theoretical advantages can indeed be translated into practical outcomes; testing that hypothesis for AWI is the central purpose of this paper.

We compare the performance of AWI with conventional FWI in a sequence of increasingly challenging scenarios, ranging from the most benign, where a good starting model and low frequencies are available, to the most challenging where the starting model and data are problematic. Our results demonstrate insensitivity to cycle skipping and show that no new problems appear. They also show that AWI is less troubled by edge artifacts associated with the finite extent of the receiver array and the finite extent of the source coverage. We show that AWI can begin from a simple smooth starting model generated without recourse to tomography, and that the model subsequently generated by AWI provides a basis for depth migration that is significantly superior to traveltome tomography.

The structure of the paper is as follows: In a short overview, we outline the mathematics that underpins the AWI algorithm. We describe the field data set, and its preparation for FWI and AWI. We describe a sequence of numerical tests in which we apply FWI and AWI to the same data under increasingly challenging circumstances, and we analyze the resultant velocity models. We use the starting and recovered velocity models to depth-migrate the reflection data, analyzing the residual moveout on migrated common-image gathers and the resultant migrated stacked volumes. The results demonstrate significant uplift in the migrated images following waveform inversion, and they show that AWI, beginning from a simple 1D velocity model, is as good as or superior to, FWI beginning after several rounds of reflection tomography. Finally, we present our conclusions and suggest our interpretation of the results.

AWI THEORY

In general, waveform inversion aims to minimize an objective function f that measures the misfit between two data sets, or alternatively to maximize an objective function that measures the similarity of two data sets. The two data sets are the observed field data \mathbf{d} and the equivalent predicted data \mathbf{p} generated by solving the wave equation for a model \mathbf{m} of the subsurface. The only difference between conventional FWI and AWI is the way in which this data misfit f is defined. This difference leads directly to differences in the respective adjoint sources that must be back propagated for each of the two methods. For FWI, the adjoint source $\delta\mathbf{s}$ is the data residual formed by the difference of the two data sets ($\mathbf{p} - \mathbf{d}$). However, for AWI, the adjoint source is a more complicated entity. A more complete account of the mathematics of AWI is given by Warner and Guasch (2016); here, we present only a summary of the key ideas.

In conventional FWI, the misfit is defined as

$$f_{\text{FWI}} = \frac{1}{2} \|\mathbf{p} - \mathbf{d}\|^2, \quad (1)$$

where \mathbf{p} is a column vector containing the predicted data, \mathbf{d} is a vector containing the observed data, and $\|\cdot\|$ represents the L_2 -norm. When the objective function f_{FWI} is minimized with respect to the model \mathbf{m} , the two data sets \mathbf{p} and \mathbf{d} become more alike, and the model that was used to calculate \mathbf{p} is then assumed to be a better representation of the subsurface. In an idealized case, the two data sets become identical, and the objective function becomes zero, at the global minimum. When applied to field data, the objective function is unlikely ever to reach zero because there will be noise in the observed data, the physics used to simulate the wave equation is likely to be imperfect, and various forms of model regularization are likely to be applied to mitigate the under-determined and ill-

posed aspects of the problem. More significantly, because seismic data are oscillatory and of a finite bandwidth, simple sample-by-sample differencing of two data sets, where parts of one resemble parts of the other shifted in time, will tend to lead to cycle-skipped local minima in the objective function whenever the starting model is too far removed from the true model. In essence, it is the minus sign in equation 1 that causes the problem of cycle skipping in conventional FWI.

AWI defines the misfit between the observed and predicted data differently. Conceptually, rather than attempting to drive the difference of the two data sets to zero, AWI seeks to drive their ratio to unity. In the idealized case, when the two data sets are identical, their difference will be zero and their ratio will be unity, and both approaches will have reached the same global minimum. However, the path that a local, linearized, gradient-descent inversion scheme will follow from a starting model toward this global minimum will depend upon the detailed shape of the objective function; consequently, one approach may become trapped at a particular local minimum that does not appear within the objective function of the other.

The ratio that AWI uses is not a sample-by-sample ratio in the time domain; rather, it is the ratio at each frequency in the temporal Fourier domain. In addition, this ratio must be stabilized so that it does not involve dividing by noise-dominated numbers close to zero. Because division in the frequency domain represents deconvolution in the time domain, and because FWI and AWI use a least-squares formulation, AWI can be described and implemented using damped least-squares convolutional matching filters, that is, Wiener filters. Described this way, AWI designs a suite of Wiener filters that match one data set to the other, and then it seeks to modify the earth model so that these filters tend toward unit-amplitude delta functions at zero lag. In this study, we design and apply these Wiener filters trace-by-trace, using a different filter for every trace in the data set.

The Wiener filter \mathbf{w} that matches an observed trace \mathbf{d} into a predicted trace \mathbf{p} is given by

$$\mathbf{w} = (\mathbf{D}^T\mathbf{D} + \mu\mathbf{I})^{-1}\mathbf{D}^T\mathbf{p}, \quad (2)$$

where \mathbf{D} is a Toeplitz matrix with \mathbf{d} in its columns arranged such that $\mathbf{D}\mathbf{x}$ represents the convolution of \mathbf{d} with a vector \mathbf{x} , and μ is a small positive number that stabilizes the deconvolution represented by the matrix inverse in equation 2. This well-known equation represents the crosscorrelation of the two data sets deconvolved by the autocorrelation of one of them. The exact value of μ is not important for AWI; in this study, we used a value that corresponds to a strengthening of the zero-lag of the autocorrelation matrix by 1%.

The filter \mathbf{w} must now be driven toward a zero-lag delta function. There are several ways to achieve this; the one that we adopt here is to define an objective function of the form

$$f_{\text{AWI}} = \frac{1}{2} \frac{\|\mathbf{T}\mathbf{w}\|^2}{\|\mathbf{w}\|^2}, \quad (3)$$

where \mathbf{T} is a diagonal matrix that acts to weight \mathbf{w} as a monotonic function of the magnitude of lag. If this weighting function has a value of zero at zero lag and increases with the magnitude of lag, then it will act as an annihilator, having a global minimum value of zero when \mathbf{w} is a zero-lag delta function. Minimization of the objective function in equation 3 will drive the coefficients within \mathbf{w} toward zero lag, drive \mathbf{p} toward \mathbf{d} , and, provided that the starting

model is within the global basin of attraction, drive the model \mathbf{m} toward the true earth model. The exact form of the weighting function \mathbf{T} does not appear to be significant for effective AWI — a simple linear weighting with the magnitude of the temporal lag can be sufficient, but we have found that a more-rapid variation with lag tends to speed convergence. Here, we used a Gaussian weighting for \mathbf{T} , centered on zero lag, with a standard deviation of approximately 5% of the trace length, leading to an objective function that should be maximized; \mathbf{T} was not varied during the inversion.

After discretization, for example, with finite differences, the time-domain wave equation can be written as

$$\mathbf{A}\mathbf{u} = \mathbf{s}, \quad (4)$$

where \mathbf{A} is a discrete wave equation operator, and \mathbf{u} is the wavefield generated at all times and all points in the model by a source \mathbf{s} . The predicted data \mathbf{p} are then related to the wavefield by a restriction operator \mathbf{R} that selects the subset of the full wavefield at the receiver positions; thus,

$$\mathbf{R}\mathbf{u} = \mathbf{p}. \quad (5)$$

With this formulation, all FWI-like algorithms lead to an expression for the gradient of the objective function with respect to the model parameters that can be written as

$$\frac{\partial f}{\partial \mathbf{m}} = -\mathbf{u}^T \left(\frac{\partial \mathbf{A}}{\partial \mathbf{m}} \right)^T \mathbf{A}^{-T} \mathbf{R}^{-T} \delta \mathbf{s}, \quad (6)$$

where the adjoint source $\delta \mathbf{s}$ depends upon the observed and predicted data (Warner and Guasch, 2016). In practical terms, reading equation 6 from right to left describes what must be done to compute the gradient for a single shot record. It says: Find the adjoint source for that shot record, inject this into the model, propagate it backward in time through the model, modify this wavefield in a way that depends upon the form of the wave equation and the model parameters, and, finally, crosscorrelate this modified wavefield in space and time with the corresponding forward wavefield generated by the true source and take the zero lag. To find the gradient for the entire data set, these single-shot gradients must be stacked over all sources.

For conventional FWI, the adjoint source is given by the residual wavefield; thus,

$$\delta \mathbf{s}_{\text{FWI}} = \mathbf{p} - \mathbf{d}. \quad (7)$$

But for AWI, the adjoint source is given by a more complicated expression that involves the Wiener filter \mathbf{w} and the lag-weighting \mathbf{T} (Warner and Guasch, 2016); thus,

$$\delta \mathbf{s}_{\text{AWI}} = \mathbf{D}(\mathbf{D}^T \mathbf{D})^{-1} \left[\frac{\mathbf{T}^2 - 2f_{\text{AWI}} \mathbf{I}}{\mathbf{w}^T \mathbf{w}} \right] \mathbf{w}. \quad (8)$$

In practical terms, reading equation 8 from right to left describes how to compute the AWI adjoint source for a single source-receiver pair. It says, use equation 2 to find the Wiener filter that matches the observed data to the predicted data, weight each coefficient in this filter by the expression in square brackets that depends upon the magnitude of its temporal lag and upon the current value of the objective function and the zero-lag of the autocorrelation of the filter,

and then deconvolve this by the autocorrelation of the observed data and convolve it with the observed data. This complicated-sounding operation contains only local 1D operations, and it is cheap to compute. Having obtained the adjoint source for AWI, the inversion proceeds exactly as it would for conventional FWI.

There are two other issues with which we might want to be concerned when implementing AWI in a practical computer code: the AWI Hessian matrix is not the same as the FWI Hessian, and the AWI objective function is more nonlinear than is the FWI function — that is, small changes in the model are approximately linearly related to changes in the FWI residuals, but they are less linearly related to changes in the AWI residuals. The former issue requires more care when preconditioning the AWI gradient with some approximation to the Hessian, and the latter issue requires more care when calculating the magnitude of the model update during the AWI step-length calculation. Another way to state the principle difference between FWI and AWI is that the region of convexity of the objective function is typically wider for AWI, but the region over which the objective function is approximately parabolic is typically wider for FWI. Consequently, AWI will normally converge toward the global minimum starting from less good models than are required for FWI, but driving AWI toward the global minimum will typically require a more carefully designed minimization procedure than does FWI.

In the results presented below, we solve the 3D VTI anisotropic acoustic wave equation using time-domain finite-differences that are 10th-order in space and 4th-order in time. We performed two suites of inversions using a different minimization scheme in each suite. In the first suite, we used a simple diagonal approximation to the Hessian during FWI and AWI to precondition the gradient, and we damped this preconditioning quite strongly during AWI. We used steepest descent on the preconditioned gradient and a simple linear step-length calculation. In the second suite, used to obtain the FWI and AWI models for depth migration, we did not apply a Hessian at all. Instead, we used conjugate gradients on the raw gradient together with a quadratic step-length calculation, using multiple test steps where necessary to ensure that the minimum in the objective function was bracketed properly by these steps. Both minimization schemes produced substantially the same end results.

RESULTS

Data set

The data used in this study come from a field located in the North Sea. The depth of the reservoir is approximately 3 km, overlain by an extensive well-confined gas cloud (Granli et al., 1999). The vertical migration of the gas is thought to be the result of piercement of the antiformal structural high caused by an underlying salt diapir. The presence of gas above the reservoir dramatically decreases the quality of the seismic data below; the gas-charged zone has a low P-wave velocity and a high effective attenuation, and both properties appear to degrade the image. Four wells have been drilled on this site: two outside, one at the edge, and one through the gas cloud. Warner et al. (2013) provide details.

The data were acquired in 2005. Three swaths of eight ocean-bottom cables were deployed (Figure 1). The cable length was 6 km, the spacing between the cables was 300 m, and each contained 4C sensors every 25 m. The source lines were shot perpendicular to the cables using air guns at a depth of 6 m. Dual sources were used

in the flip-flop shooting mode. Each source consisted of a 0.0644 m^3 (3930 in^3) air-gun array. The lateral separation between sources was 75 m, and each fired every 50 m inline. The area covered by each swath was 120 km^2 , and the three swaths combined covered a total of 180 km^2 . The total number of 4C receivers in the final data set is 5760, and the total number of shots is 96,000. The survey provides offsets of approximately 7 km with good azimuthal coverage and maximum offsets of up to 11 km with reduced azimuth and fold, corresponding to the sources in the corners of the shooting area.

For the inversion, only the hydrophone data were used. These data were low-pass filtered, using an Ormsby filter rolling off from 5 to 7.5 Hz. The raw data have an adequate signal-to-noise ratio for FWI down to frequencies less than 3 Hz. The filtered data were muted ahead of the first arrivals, and a bottom mute was applied at a short offset to remove low-velocity, low-frequency Scholte waves. No other preprocessing of the raw field data was applied prior to inversion — no deghosting, no demultiple, no debubble, no signature removal, no denoise, and no pz-summation. Refracted arrivals, postcritical and precritical reflections, and all ghosts, multiples, and the original source waveform were retained within the data to be inverted. One in four of the original receivers were selected giving a 100 m inline and 300 m crossline separation. One in three original sources were selected, giving an approximately 100 m shot spacing in both directions. Source-receiver reciprocity was applied leading to a final data set that contained 1440 reciprocal sources and approximately 20 million source-receiver pairs. A suitable low-frequency source wavelet was extracted from the filtered field data, and it was refined and tested using forward modeling following the procedure described in Warner et al. (2013). Extraction of the wavelet prior to FWI and AWI requires only that the properties of the water column, free surface, and water-bottom reflectivity can be estimated from the raw field data; the wavelet estimate does not use the tomography or other subsurface velocity model. Figure 2 shows the original full-bandwidth data, the muted, subsampled, low-pass filtered data ready for inversion, and the corresponding low-frequency wavelet. The long multicyclic wavelet results from the properties of the low-pass filter used to preprocess the field data. This preprocessed data set and wavelet were used as input for all FWI and AWI inversions described here.

Test inversions

In a suite of tests, we ran AWI and FWI on the same combinations of input data and starting model. Here, we show four combinations in total:

- 1) low starting frequency and a good 3D starting model
- 2) higher starting frequency and a good 3D starting model
- 3) low starting frequency and a simple 1D starting model
- 4) higher starting frequency and a simple 1D starting model.

Our expectation is that AWI and FWI will provide similar high-quality results for the first combination, but that FWI will begin to fail for the more difficult problems because of cycle skipping, whereas AWI should still be able to proceed even as the input data and starting model become less suitable.

We do not, in this paper, explore the importance of, or inversion for, anisotropy; all of the inversions shown used the same 1D VTI model of anisotropy described in Warner et al. (2013), and this

remained fixed throughout the inversion. We did not explicitly correct the data for the effects of attenuation, nor did we include attenuation in the forward modeling during inversion. The inversion code balances amplitudes in various heuristic ways, an approach that serves to mitigate systematic amplitude variations produced by otherwise uncompensated anelastic attenuation, density variations, elastic effects, and short-wavelength changes in anisotropy, and to balance the relative importance of different traveltimes, offsets, and phases in the data. Some form of amplitude matching between the predicted and observed data is normally helpful during FWI of field data to prevent leakage of other parameters into the acoustic velocity model, but the exact form of that matching is not critical; details of the method applied here are given in Warner et al. (2013). The same approach to amplitude balancing and regularization was used during FWI and AWI. Reflections and refractions are included in the data to be inverted, and the amplitude balancing also helps to ensure that both types of arrival contribute to the final inverted model. In practice, however, the low and intermediate wavenumbers in the recovered velocity model are extracted predominantly from the refracted and postcritically reflected

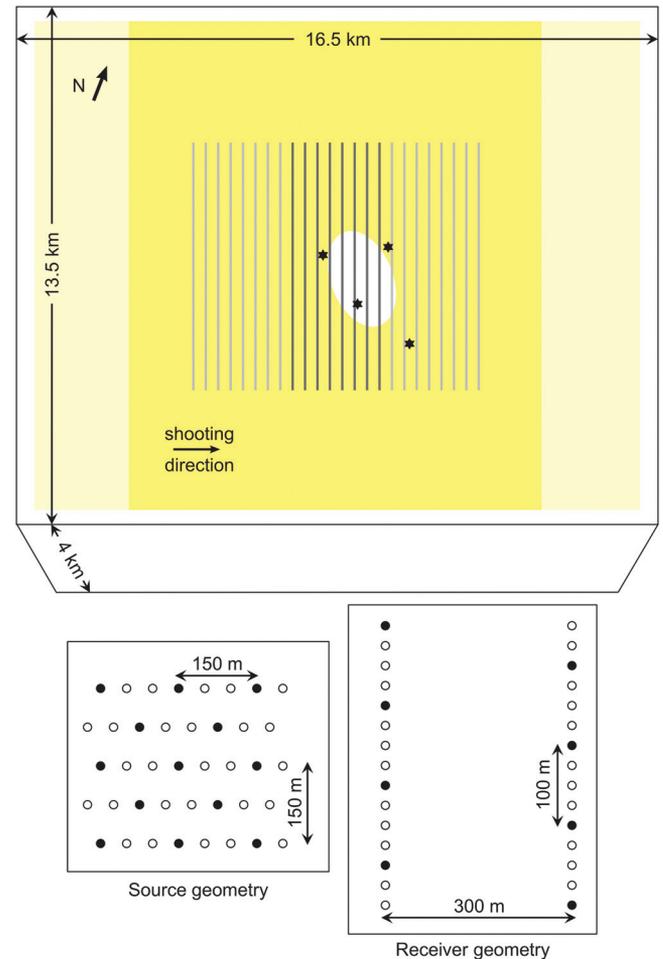


Figure 1. Acquisition geometry and model size. North–south lines show the ocean-bottom cables acquired in three swaths of eight cables each; yellow indicates the three corresponding patches of shots; white ellipse indicates the approximate location of the gas cloud surrounded by four wells. The filled black circles show the subset of original sources and receivers that are used in the inversion.

arrivals with precritical reflections contributing principally only to the high-wavenumber detail. We applied smoothing to the model updates at a scale length of about half the local seismic wavelength, and we used no formal model regularization of any kind beyond that.

We obtained the good 3D starting velocity model from the original processing contractor. It was built using multiple rounds of anisotropic reflection tomography, guided by wells, applied to the p_z-summed, deghosted, and demultiplied data with some manual intervention to delimit the lateral extent of the gas cloud. We obtained the simple 1D starting model by matching the moveout in one corner of the data set and extending this over the entire model in one dimension. The good 3D model contains a reasonable initial representation of the low-velocity gas cloud; the 1D model does not contain any version of this cloud. The good 3D model is the same as that used to start the inversion in Warner et al. (2013). Vertical and horizontal slices through the two models are shown in Figure 3.

For the low starting-frequency inversions, we ran the inversion in six passbands, where the low-pass filter applied increased gradually from 3 to 6.5 Hz. In each case, the nominal frequency represents the corner frequency of a low-pass filter that rolls off at approximately 48 dB per octave; given the sharp roll off of the field data at low frequency, in practice, this means that the inverted data within each passband have a rather narrow bandwidth with a peak frequency coinciding closely with the nominal frequency. Starting at a low frequency and gradually increasing the passband is the conventional way to run time-domain FWI; it begins at a low frequency to mitigate cycle skipping. For the higher starting-frequency inversions, we ran every inversion using the 6.5 Hz filter. Some longer offset portions of the field data appear to be cycle skipped with respect to

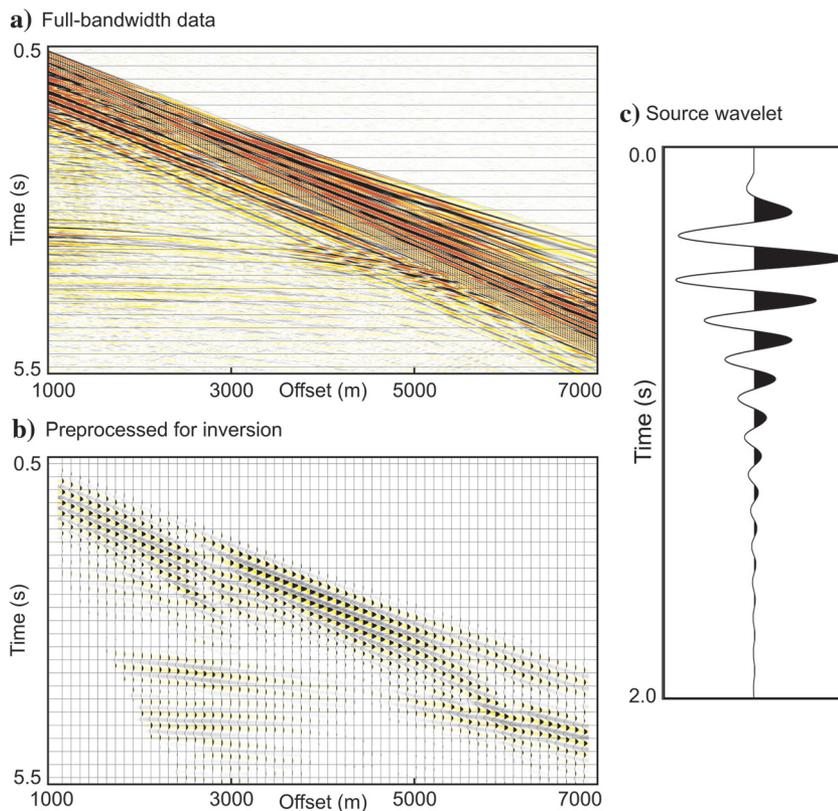
modeled synthetic data at 6.5 Hz, even with the good 3D starting model; therefore, we would not expect FWI to work well in these circumstances. We split the data into 18 random subsets, each containing just 80 reciprocal sources. Within each passband, we inverted each of these subsets in turn before moving to the next passband.

Low starting frequency, and a good 3D starting model

This represents the standard run, serving as a benchmark against which the others can be measured. We do not expect this run to be cycle skipped, and AWI and FWI should prove to be equally successful. For this test, we used one iteration per source at each of six frequency bands, opening up the top of the passband progressively from 3 to 6.5 Hz; we then ran the final passband a second time, giving seven iterations per source in total. As described above, we used 18 subiterations per passband, each run over one-18th of the available sources giving 126 subiterations in total.

In previous studies of this data set, we have found that this level of computational effort represents a sensible commercial compromise between cost and accuracy for conventional FWI; this does not represent the best possible fit to the observed data, nor does it represent a fit to within the estimated noise level of the data. Fitting highly redundant multichannel 3D reflection data to the noise level can require many thousands of iterations of conventional FWI while producing no significant difference in subsequent depth migrations, and is unlikely ever to be commercially justified. In another study, we have run several thousand iterations of FWI and AWI on this data set; when FWI and AWI are successful, this large number of iterations does not improve the practical outcome significantly,

Figure 2. Hydrophone shot record. (a) Full-bandwidth raw data. (b) Preprocessed for inversion. (c) Corresponding low-frequency source wavelet.



and when FWI has become trapped in a local cycle-skipped minimum, using a large number of iterations alone does not enable it to escape. We have observed that AWI often benefits from more iterations than FWI, presumably because the AWI objective function is less parabolic. In practice, however, we would almost never choose to run large numbers of AWI iterations on a field data set; rather, we would run sufficient AWI iterations to ensure that the resultant model was no longer cycle skipped and then switch to conventional FWI to complete the process, moving FWI progressively to a higher frequency. We do not follow that practical workflow in this study because here we are interested in directly exploring the differences between AWI and FWI, rather than devising an optimal practical combination between them.

Figure 4 shows the vertical and horizontal depth slices through the models recovered by FWI and AWI. Above the high-velocity chalk, the results are similar in the central region of the model and the geometry and absolute velocity of the recovered gas cloud appears correct. Warner et al. (2013) show that an FWI model very similar to this one predicts the field data closely; accurately migrates the reflection data and flattens the gathers; and matches the wells, the boundaries of the gas cloud, and the geometry of the migrated reflectors. We therefore believe that this benchmark lies close to the true model.

The recovered FWI and AWI models diverge toward the lateral edges of the model, and within the high-velocity chalk layer below a

depth of about 3000 m. The lateral differences are illustrated in the FWI model by the low-velocity (green) features seen in the bottom left of Figure 4a and 4c; these are edge effects produced by the finite extent of the acquisition geometry and are a common feature of FWI models that can be partially overcome by appropriately designed model regularization. The obvious artifacts seen in the FWI model within the chalk are an extension of these lateral edge effects to greater depths, where they represent an edge to the maximum depth of penetration related in turn to the maximum offset available in the data set. Before migrating with the FWI model, these artifacts must be removed, for example, by regularization toward the starting model in the affected areas.

Interestingly, AWI does not suffer as significantly from these edge effects, either laterally or in the chalk. This means that the AWI-recovered model remains accurate to a greater depth and closer to the edges of the recovered model. Both models also show an imprint of the acquisition geometry at a fine scale; this fine acquisition footprint is stronger in the AWI result than in the FWI-recovered model. The footprint is relatively easy to remove from the model, either in postprocessing or by regularization during the inversion; in the model shown, the footprint is especially marked near top chalk, although it is less intense than the edge effects introduced by FWI in the chalk.

In summary, when the starting model is accurate and the inversion begins at low frequencies, AWI and FWI produce similar

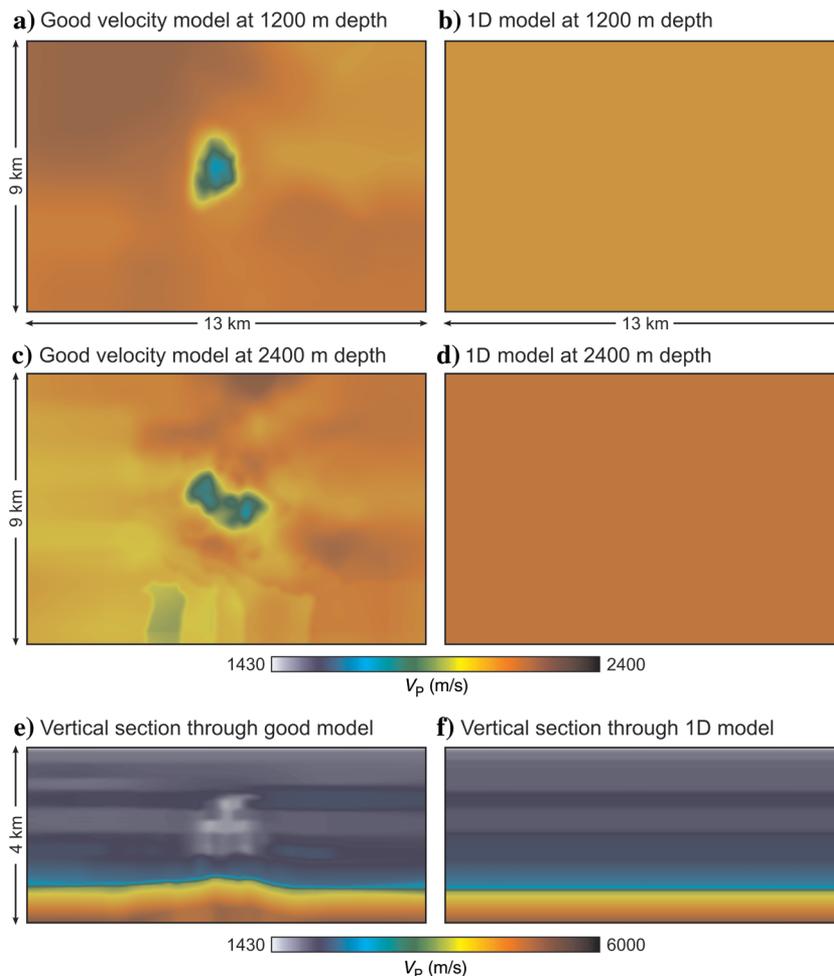


Figure 3. Starting models. Left: good 3D model derived from the tomography. Right: simple 1D model — the gas cloud is not present in the model. Top: horizontal slice at 1200 m depth. Middle: horizontal slice at 2400 m depth. Bottom: vertical slice through the gas cloud. Note that the color-scale limits are different for the horizontal and vertical slices.

outcomes that differ only in their edge effects and acquisition footprint. AWI suffers less from edge effects produced by the finite extent of the source and receiver arrays and by the finite offset; FWI suffers less from a fine-scale acquisition footprint produced by the finite spacing of sources and receivers.

Higher starting frequency, and a good 3D starting model

In this test, the starting model was unchanged, but all iterations were run at the full processed bandwidth of 6.5 Hz; the peak frequency in this case is approximately 5.5 Hz. We expect AWI to proceed satisfactorily in this case, but FWI may begin to be impacted by cycle skipping in some parts of the data set. In this test, we increased the total number of iterations per source from 7 to 14. This increase is desirable because the inversion begins from a model that is effectively further from the true answer than in the previous test. Even though the initial model has not changed in absolute terms, it predicts data that are further from the observed data when that distance is measured in terms of wave cycles at the starting frequency. Consequently, the inversion has a longer effective path to navigate when it begins from a higher frequency, and this typically requires more iterations to reach the same global minimum.

Figure 5 shows the recovered models. The model recovered by AWI is little changed from that shown in Figure 4 demonstrating that AWI is able to proceed successfully from a higher starting frequency. In contrast, the FWI model in Figure 5 has begun to

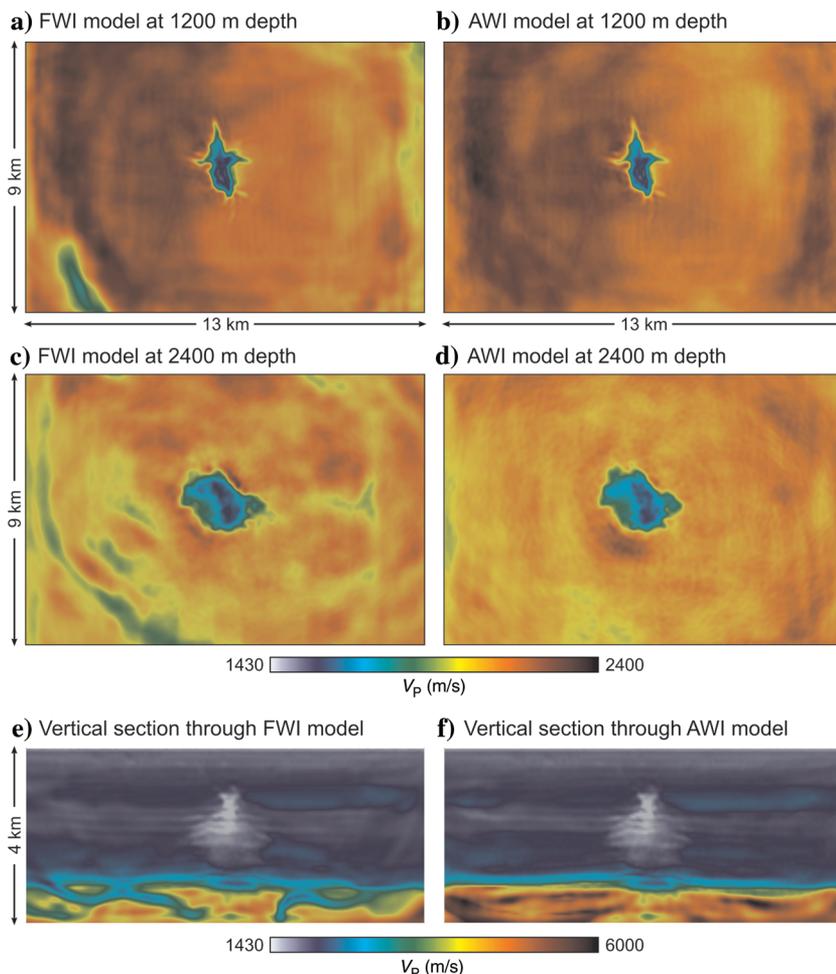
degrade; the edge artifacts have moved closer to the center of the model, and the gas cloud is less saturated and is distorted with respect to the benchmark model. The differences are not great, however, and although some small portion of the data is cycle skipped, the remainder of the data operate to push the model broadly in the desired direction.

Low starting frequency, and a simple 1D starting model

In this test, we retained the low starting frequency, but we used a less-accurate model to begin. This starting model contains no trace of the low-velocity gas cloud; however, the background model still gives reasonably accurate arrival times away from the gas cloud. We ran the inversion twice, iterating from 3 to 6.5 Hz, then continuing a second run as a warm restart, repeating the iterations from 3 to 6.5 Hz, giving 14 iterations per source in total. As in the previous test, this increase is necessary because the inversion begins further from the true model than does the benchmark.

Figure 6 shows the results. The AWI model is still largely unaffected, although it does lose some absolute accuracy with increasing depth through the gas cloud. The geometry and saturation of the gas cloud are well-recovered even though there is no gas cloud in the starting model — the recovered cloud is built entirely by the inversion. As we will see later, a poor starting model, similar to this, does not migrate the field data accurately nor does it result in flat gathers. In production use, we would typically run AWI on a model

Figure 4. Models recovered using a low starting frequency and a good 3D starting model. Left: FWI. Right: AWI. Color scales and slices are the same as Figure 3. Recovery of the gas cloud is similar by both methods, but edge-related artifacts are worse, laterally and in depth, in the FWI result. The AWI result has an enhanced short-wavelength acquisition footprint when compared to FWI.



of this sort until cycle skipping was no longer problematic and then smooth the model and continue the inversion to completion using FWI. We have not followed that enhanced workflow in this example, and we show the pure AWI and pure FWI results in Figure 6.

In contrast to AWI, the FWI result in Figure 6 is badly degraded. The gas cloud is now significantly under-saturated, and FWI is not capable of building the cloud adequately from a model that does not already contain some reasonable representation of it. At greater depth, the gas cloud is barely recovered at all — this is especially clear in the vertical section in Figure 6c. The FWI result is also badly contaminated by edge effects in the shallow section and by shorter-wavelength artifacts in the deeper section. The latter are likely caused by conflicting updates produced by parts of the data that are cycle skipped with respect to other parts that are not cycle skipped. The results of this test clearly demonstrate that AWI is capable of beginning the inversion from a less accurate starting model for which some of the data are cycle skipped and for which the FWI is only able to recover a poor, under-saturated, noisy version of the true model.

Higher starting frequency, and a simple 1D starting model

This is the most challenging case that we examine; all iterations are run at the full 6.5 Hz bandwidth, and the starting model is 1D. Because we start further from the true model in terms of the starting

frequency than in either of the last two tests, we ran a total of 21 iterations per source for this test, all at 6.5 Hz.

Figure 7 shows the results. For AWI, the outcome above the top chalk is similar to the previous result; the geometry of the gas cloud is properly recovered, its absolute velocity is fully saturated in its upper portion, and this reduces only a little with increasing depth. There are now, however, significant artifacts below the top chalk; these develop relatively early in the inversion before the gas cloud is fully formed. In production, we would deal with these by smoothing or constraining the updates strongly within the chalk until the gas cloud had developed, relaxing the constraints as the inversion proceeds when we would also likely switch from AWI to pure FWI. We have not followed that workflow in the results shown here.

In contrast to AWI, FWI goes badly wrong in this test. The gas cloud is poorly recovered, the faster left side of the benchmark model at 1200-m depth is now replaced by a region of spurious lower velocity, and the deeper part of the model contains little structure in the update that is real. The slow region on the left has been produced by pervasive cycle skipping of the predicted data from this area, which generates a gradient contribution that has the wrong sign — the recovered velocity thus decreases when the model actually requires an increase. The FWI-generated model here is significantly worse than the starting model, although of course the objective function has still been driven down as the inversion has moved toward, and has become trapped within, a deep local

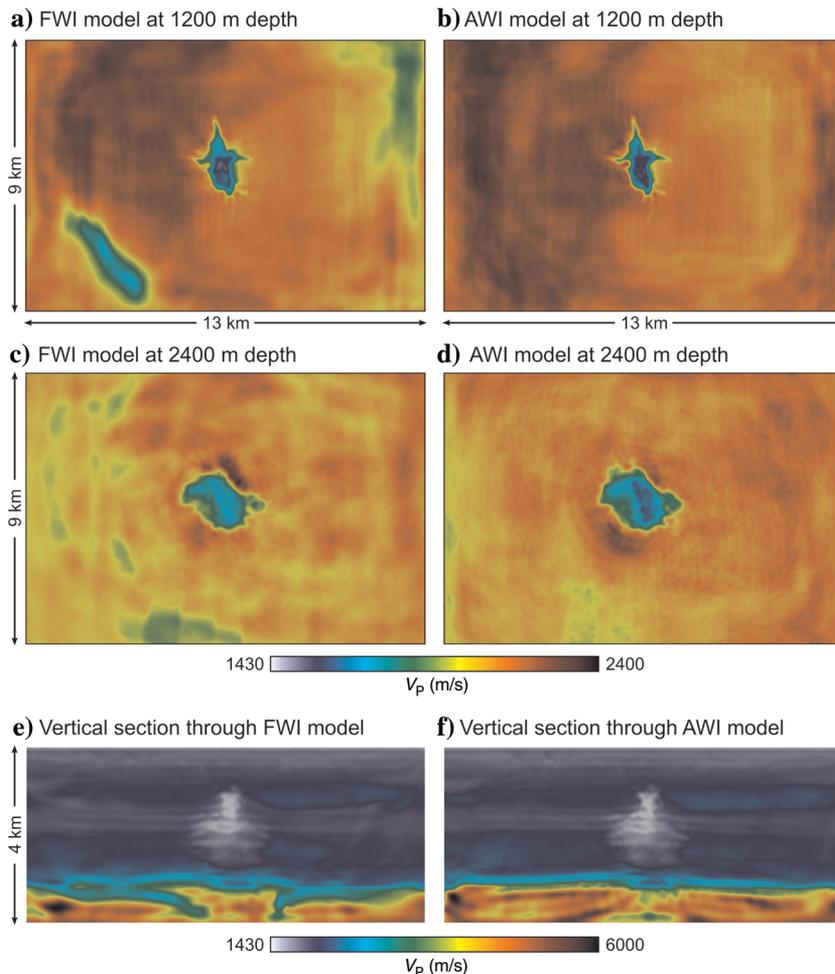


Figure 5. Models recovered using a high starting frequency and a good 3D starting model. Left: FWI. Right: AWI. The AWI result is similar to the benchmark result in Figure 4. However, the FWI result is losing intensity in the gas cloud, especially at depth, and it has enhanced edge artifacts.

minimum. Quality control of FWI that relies principally upon reductions in the objective function, increases in the zero-lag cross-correlation of predicted and observed data, or similar measures of data misfit, will not easily detect that this inversion has gone astray.

Depth migration

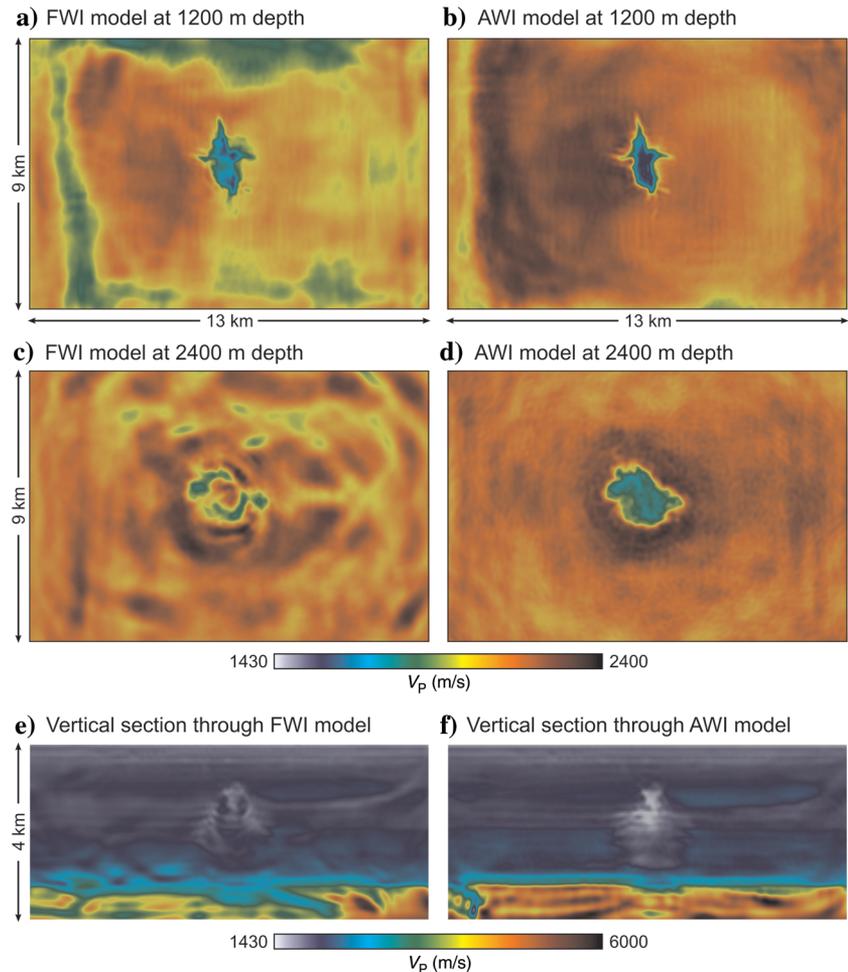
Following the tests described above, we re-ran pure FWI beginning from the tomography model, and we re-ran AWI beginning from an approximately 1D model built by heavily smoothing the tomography model after removing the gas cloud (Ravaut et al., 2017). The anisotropy model was as used for the previous tests. Offsets from 1.2 to 10.5 km were included and late arrivals at nearest offsets were muted. In total, 100 iterations were run for AWI and 50 iterations for FWI, inverting every source at each iteration over a frequency range of 3–6 Hz. These inversions used conjugate gradients and a quadratic step length. Unlike the previous tests, no spatial preconditioning or other approximation to the diagonal Hessian was applied during AWI. The models are lightly smoothed as the inversion proceeds to mitigate the acquisition footprint.

Figure 8 shows vertical slices through the relevant starting and recovered models for FWI and AWI. Although the two methods begin from quite different starting points, the resultant gas clouds that they produce are similar in geometry and magnitude. We then used these four models to migrate the p_z-summed demultiplied deghosted data

using 3D prestack Kirchhoff VTI ray-traced multi-arrival depth migration. The velocity models are not unduly complicated, and more-accurate migration schemes, such as reverse time migration, have not produced noticeably more accurate migrations in our previous studies. Although attenuation is high within the gas cloud, we did not compensate the data for the amplitude or dispersive effects of attenuation, and we did not take attenuation into account during either FWI/AWI or depth migration. We might then expect that the velocity obtained by low-frequency FWI would be slower within the highly attenuating gas cloud than that required for optimal depth migration. We might also expect that anisotropy within the gas cloud would differ from that outside; we did not include any such changes in anisotropy during either FWI/AWI or migration.

Vertical slices through these migrations are shown in Figures 9 and 10, the latter with the corresponding velocity model superimposed. Figure 11 provides a quantitative assessment of the accuracy of these migrations using color-coded measures of automatically picked residual moveout of primary subcritical reflections on common-image gathers. Neither FWI nor AWI seek to flatten gathers explicitly; rather, they seek to minimize the mismatch between the observed and predicted data sets. Consequently, gather flatness provides an independent test of the validity of the velocity model assuming that the anisotropy model is correct. In Figure 11, red picks represent under-migrated gathers indicating that the model velocity is too high, blue picks represent over-migrated gathers

Figure 6. Models recovered using a low starting frequency and a simple 1D starting model. Left: FWI. Right: AWI. The AWI result is similar to Figures 4 and 5, but it is under saturated in the deeper parts of the gas cloud. In contrast, the FWI result is failing badly, and the gas cloud is only partly recovered even at a shallow depth.



indicating that the model velocities are too low, yellow picks represent flat gathers, and white regions indicate that no events were automatically pickable.

Clearly, the migrations using the FWI and AWI results (Figure 8c and 8d) are significantly better than either the tomography or the simple 1D models (Figure 8a and 8b), neither of which is able to focus the image adequately beneath the gas cloud. The gathers are clearly flatter beneath the gas cloud in the tomography model than in the simple 1D model, but despite this, there is no especially clear improvement in the appearance or focusing of the stacked migration in Figure 8a relative to Figure 8b. Comparing the FWI and AWI-based migrations, these are similar. There are fewer conflicting dips in the AWI result beneath the gas cloud than in the FWI result, and the migrations differ with respect to the low-frequency subhorizontal events seen within the anticline beneath the gas cloud; the provenance of these reflections is unknown — they are undrilled and are stronger, more continuous, and more nearly horizontal on the AWI migration.

Compared with FWI, AWI recovers a good gas velocity down to approximately 2.8 km depth giving the same quality of residual moveout (RMO) picks in Figure 11c and 11d. On the FWI model, artifacts associated with the finite aperture of the acquisition are visible on the model shown by the ellipses in Figures 8c and 10c. On the AWI model, this imprint of the acquisition is less pronounced. The artifacts on the AWI results are of a smaller wavelength than those observed on the FWI results.

In Figure 11, all the RMO maps show red picks on the right side of the gas cloud between approximately 600 and 1300 m in depth. Neither FWI nor AWI improve the gather flatness in this region, and, here, the tomography result appears to be somewhat flatter than either of the inversion results. This behavior suggests that these under-migrated gathers are likely related to inaccuracies in the anisotropic model, and that the difference between delta and epsilon should be smaller than we have assumed. Reflection tomography will attempt to map such an anisotropy error into the velocity field in a way that minimizes the overall RMO, whereas waveform inversion will tend to map the error in a way that minimizes the data misfit — both are wrong. The proper way to address this is to improve the anisotropy model, for example, by inverting for the difference between delta and epsilon as part of the waveform inversion. We have not explored that approach here, but have shown it to work elsewhere (Debens et al., 2015).

DISCUSSION

AWI and FWI results are similar when starting from a model and at frequencies that are not cycle skipped (Figure 4). However, when the starting model is less accurate and/or when sufficiently low frequencies are not present in the field data, AWI can still proceed successfully when FWI cannot. Figures 4–7 demonstrate that AWI can recover a good velocity model from this field data set when the

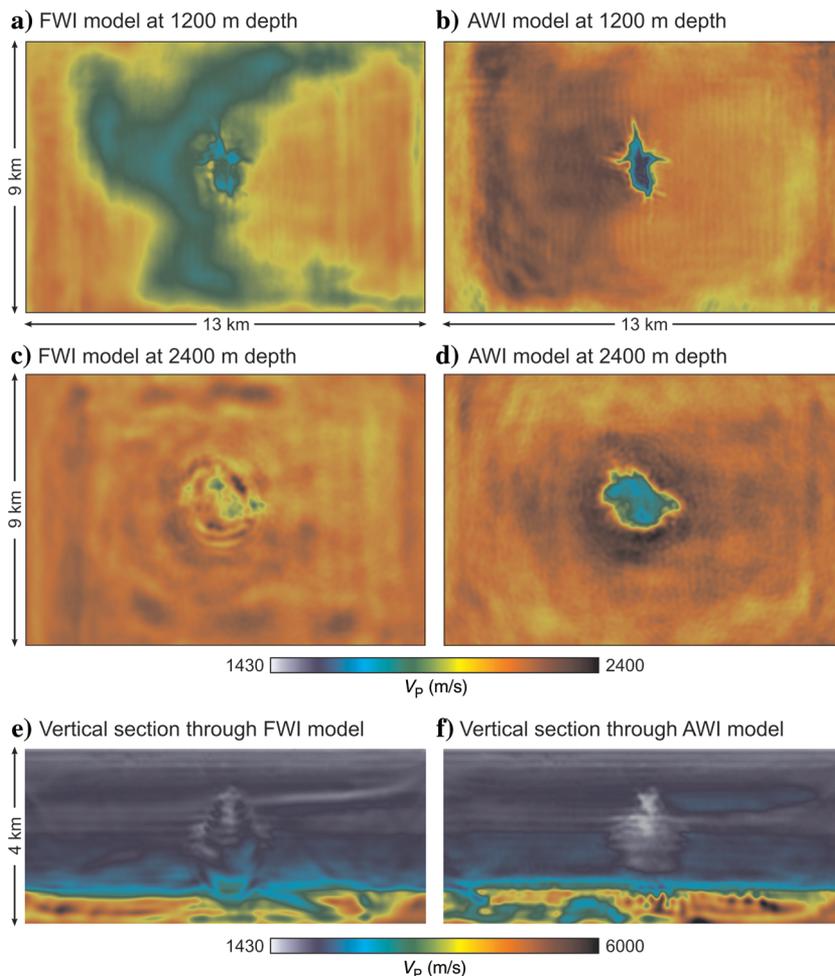


Figure 7. Models recovered using a high starting frequency and a simple 1D starting model. Left: FWI. Right: AWI. Above the chalk, the AWI result is similar to Figure 6, and it is still providing good recovery of the gas cloud. The FWI result is now badly affected by cycle skipping and is significantly worse than the starting model.

effective distance between the starting and true models is about two to three times larger, in terms of the starting frequency, than conventional FWI can tolerate. We have applied FWI and AWI to a range of marine field data sets, ocean-bottom and towed-streamer, and we have found that this quantitative conclusion appears to hold over a wide range of circumstances. In practice, this means that AWI can invert successfully when the usable starting frequency that can be tolerated by FWI is approximately doubled, when the cumulative errors in the start model that can be tolerated by FWI are approximately doubled, or when there is some equivalent combination of both effects. For many marine data sets, this means that AWI can tolerate towed-streamer data starting from time-migrated stacking velocities or following only the simplest reflection tomography, whereas FWI often requires more careful QC, time and/or depth windowing, enhanced low-frequency noise reduction, and multiple passes of structurally guided reflection tomography to proceed successfully. AWI is not a cure for all the problems that can beset waveform inversion, but in real marine data sets, it often spans the critical region that can make an otherwise difficult problem tractable and practical.

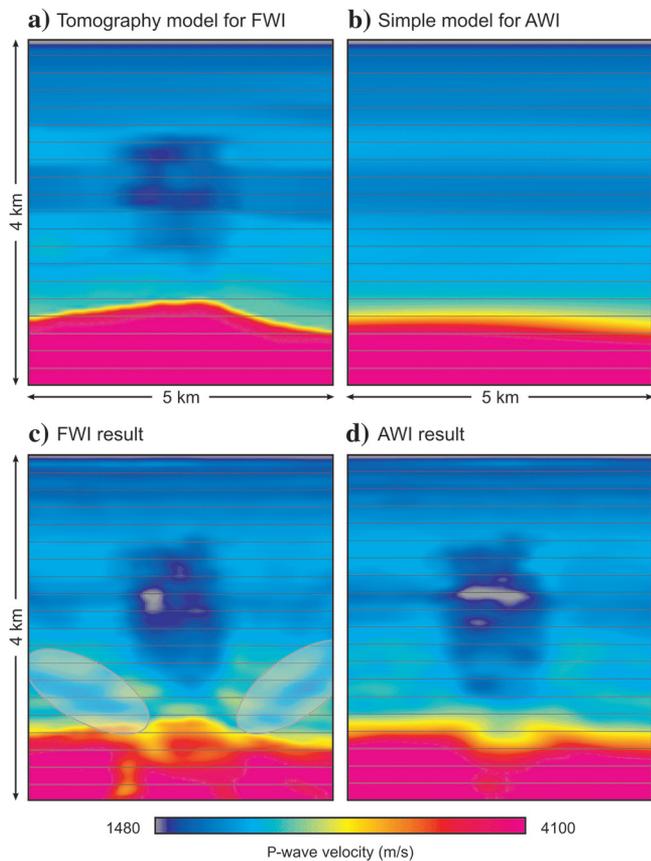


Figure 8. Vertical slices through velocity models used for migration, from the seabed to a depth of 4000 m. (a) Tomography model used as the starting model for FWI. (b) Simple smooth model without a gas cloud used as the starting model for AWI. (c) FWI-recovered model. (d) AWI-recovered model. The FWI result shows edge effects indicated by the ellipses and within the underlying fast chalk layer. Regularization during AWI has removed the previously visible footprint at the level of the top chalk. Lowered AWI-recovered velocities below the gas cloud are confirmed in the wells.

Figure 12 shows a comparison of the impulse responses produced by FWI and AWI in a constant-velocity model. It shows the raw gradient of the objective function with respect to the model parameters for a single source and single receiver in a constant-velocity model. The overall shape of this figure does not depend at all strongly upon either T or μ ; it does depend upon the bandwidth of the source, but not upon its phase spectrum. The gradient generated by a plurality of sources and receivers would be composed of a summation of gradients similar to this one. In Figure 12, the red central region represents updates to the model that have the correct sign; if there were many sources and receivers, then this central region would dominate the gradient and the whole model would be updated appropriately. The blue regions in Figure 12 represent sidebands that have the wrong sign for the model update. If there were many sources and receivers, then these side bands would tend to be suppressed over much of the model by destructive interference, but toward the survey edges, they would tend to interfere constructively to produce edge artifacts in the model having the wrong sign. These edge artifacts will always tend to appear toward the lateral edges of a model, and where the update is dominated by refracted arrivals, they will also tend to appear beneath the survey, where the finite source-receiver offset limits the depth of penetration of the refractions. In a simple model, with turning refracted arrivals and a fixed

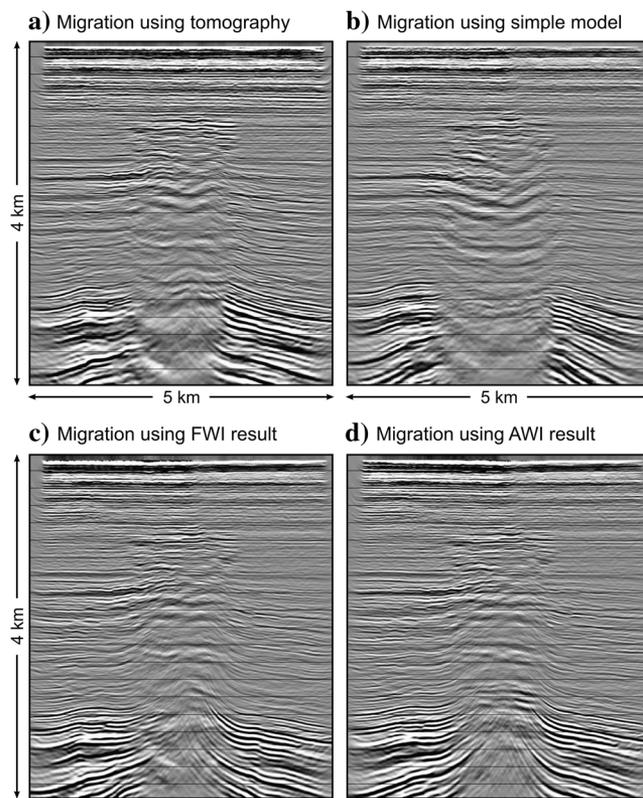


Figure 9. Prestack 3D Kirchhoff depth migrations corresponding to the four models shown in Figure 8. The tomography and simple starting models produce poorly focused migrations. The FWI- and AWI-based migrations are significantly improved. The FWI- and AWI results differ from each other principally with respect to the low-frequency subhorizontal events seen within the anticline beneath the gas cloud; the provenance of these reflections is unknown — they are undrilled and are stronger, more continuous, and more nearly horizontal on the AWI migration.

acquisition geometry, these artifacts will tend to produce a bowl surrounding an otherwise properly updated model.

In Figure 12, the AWI gradient has a tighter central red region than FWI. This will tend to improve the spatial resolution of AWI, while at the same time tending to enhance the local acquisition footprint produced by finite spacing between sources and/or receivers. The AWI gradient also has oscillatory side bands with variable wavelengths; these will tend to interfere destructively even toward the survey edges. In contrast, FWI has a broader central region than AWI, and this will tend to protect against an acquisition footprint produced by sparse sources/receivers, but it will also tend to degrade the spatial resolution with respect to AWI. The FWI gradient also has a single strong reversed-sign blue sideband, which will tend to interfere constructively when there are many source-receiver pairs, and so it will tend to produce stronger artifacts toward the survey edges, and below the deepest refracted arrivals.

In summary, Figure 12 suggests that AWI will tend to display a stronger acquisition footprint produced by finite shot and receiver spacing, and FWI will tend to produce a stronger edge-artifact laterally, and in depth, produced by finite source and receiver array apertures. All of these effects are displayed in Figure 4. The reduction of edge effects at depth means that the effective penetration provided by AWI can often be significantly greater than that of conventional FWI, extending well below the depth of the deepest penetration predicted by ray tracing. The results presented here show this effect in part, but the dramatic velocity contrast present at the top chalk in this survey complicates the outcome.

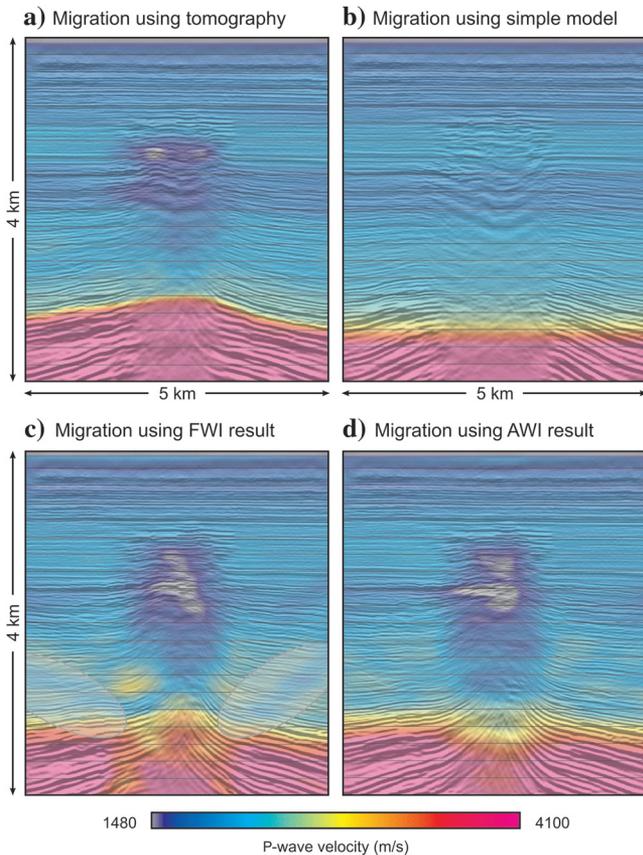


Figure 10. Superposition of depth migrations from Figure 9 and their corresponding velocity models from Figure 8.

In practice, although Figure 12 suggests that there should be an effect, we do not typically see a marked difference in the spatial resolution between FWI and AWI in field data. FWI and AWI use different measures of misfit. The simple difference used by FWI is a purely local measure that compares the predicted and observed values data point by data point. This typically means that,

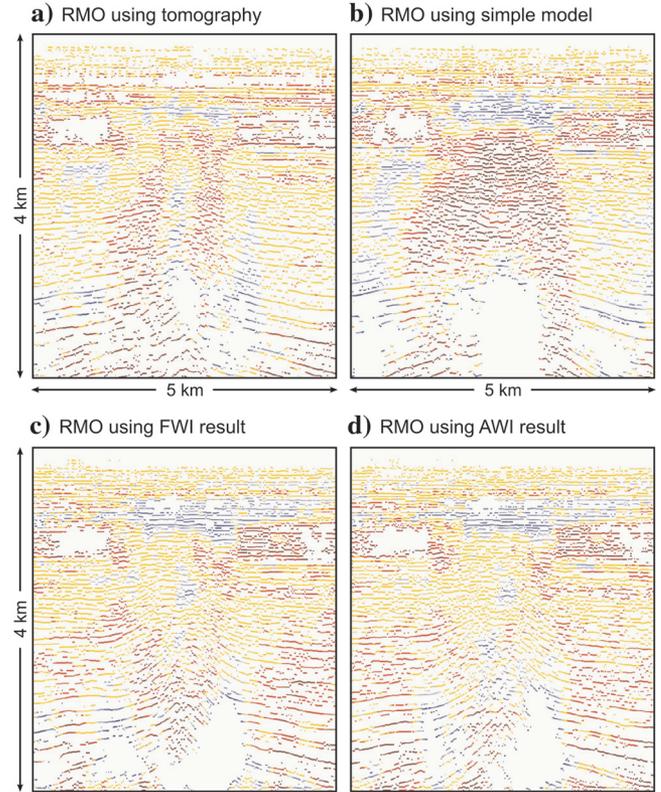


Figure 11. Color-coded RMO on common-image gathers corresponding to the migrations shown in Figure 9. Red represents under-migrated gathers indicating that the model is too fast, blue represents over-migrated gathers indicating that the model is too slow, yellow represents flat gathers, and white has no autopickable events.

Single source-receiver gradient for constant velocity model

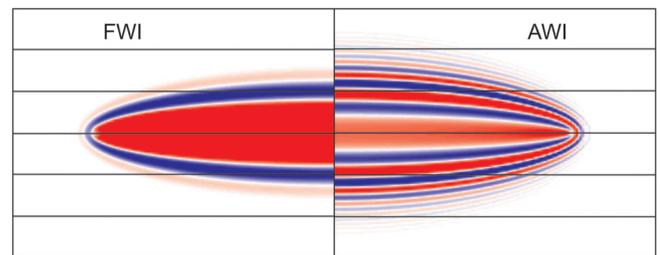
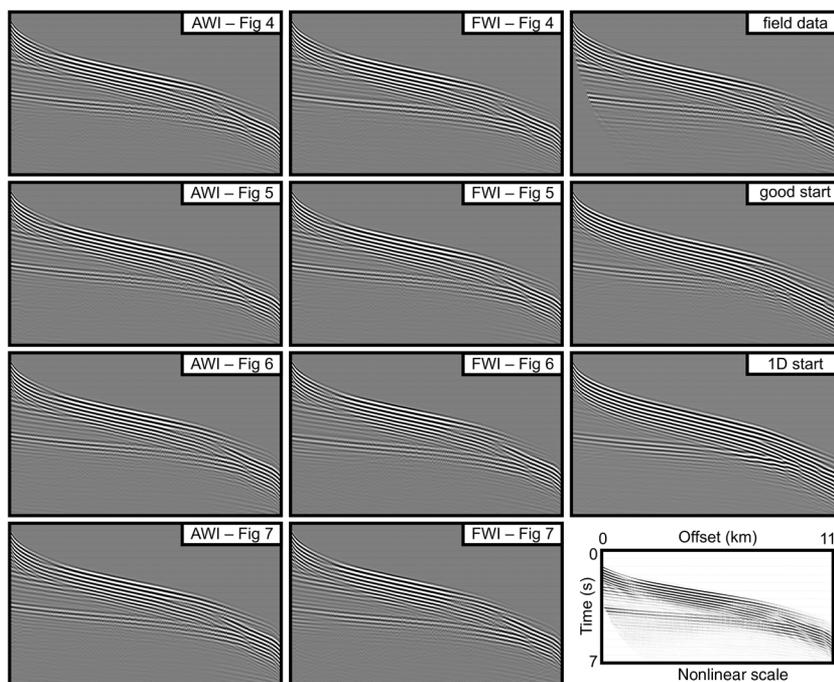


Figure 12. Gradients produced by a single source-receiver pair in a constant-velocity model for FWI (left) and AWI (right). These represent the model updates that would be applied by each of the algorithms, to a constant-velocity model, when starting from a minimally perturbed constant-velocity model, using just one seismic trace, without preconditioning or regularization. The absolute amplitudes of the two gradients, the orientation of the figure, and its horizontal and vertical scale lengths, are all arbitrary. Red regions have the correct sign for the update; blue regions have the reverse sign.

Figure 13. Data comparison for the different inversion runs. The data panels show a single receiver gather. The traces are ordered by offset, hence the ragged nature of the data — nearby traces are not spatially contiguous. Each AWI and FWI panel shows the predicted data computed using the final models recovered for each of the four inversions shown in Figures 4–7. Equivalent panels are also shown for the preprocessed field data and for the tomography and simple 1D starting models. The inset at bottom right shows axis labels appropriate for all panels.



provided the data are not cycle skipped, FWI will tend to phase lock the predicted and observed data tightly together. In contrast, the measure of misfit used by AWI is global within each trace so that every data point influences every other. This provides the ability to overcome cycle skipping, but it will not phase lock the data locally if one trace contains events that are not present in the other. Instead, it will seek the best least-squares compromise that minimizes its Wiener-filter-based misfit function. The AWI result will in consequence lose some of its theoretically higher resolution. In practice, the two effects seem to be of similar importance so that the practical resolution provided by FWI and AWI become rather similar. In typical work flows, we run AWI initially, moving to pure FWI and to higher frequencies as the inversion proceeds.

A common method to assess the quality of waveform inversion results is to compare the final synthetic data with the observed data. Figure 13 shows such a data comparison for all of the recovered models. Here, each panel shows a common-receiver gather over the entire survey, sorted by offset. The synthetics generated by all of the inverted models appear rather similar, and all provide a reasonable fit to the field data even though several of the synthetic data sets are significantly cycle skipped. Waveform inversion is designed to reduce the mismatch between observed and predicted data, and so it will produce models that generate synthetics that resemble the field data. Consequently, matches to the field data must be used with care when assessing model quality. An obvious mismatch is a likely indication of a poor inversion outcome, but a superficial lack of an obvious mismatch to the final synthetic data may not necessarily provide an indication of a good outcome.

CONCLUSION

We have systematically tested the AWI method, introduced by Warner and Guasch (2014, 2016), on a 3D ocean-bottom field data

set. The results demonstrate that AWI here is insensitive to cycle-skipping produced by limited low-frequency bandwidth of the observed data. Consequently, AWI can proceed successfully in circumstances for which conventional FWI performs poorly or fails entirely. AWI will normally require more iterations to converge if the initial model is further from the true solution or when the starting frequency is increased. In the data set examined here, AWI is able to reconstruct a well-resolved and accurate velocity model starting from a simple 1D or 1D-like initial velocity model. AWI has less sensitivity to the finite extent of the acquisition geometry and to the finite maximum source-receiver offset, and it does not show the strong edge artifacts from which conventional FWI can suffer. AWI does, however, have increased sensitivity to finite source and receiver spacing such that a fine-scale acquisition footprint can appear in the recovered model; this footprint is not difficult to remove by regularization during AWI, but in some survey geometries, AWI can require more closely spaced data to be used at each iteration, which may increase the computational cost.

The recovered AWI model migrates the field data more accurately than does the tomographic model, and at least as well as tomography followed by conventional FWI. When using AWI in this way, the velocity model is built, including the starting model, entirely from raw unprocessed field data without recourse to reflection tomography and with little manual intervention required. In practical applications, AWI would normally be followed by conventional FWI to maximize the resolution and ensure that the recovered model is properly phase locked to the field data.

ACKNOWLEDGMENTS

We thank the Norwegian Petroleum Directorate, who made the field data used in this study available on public release from March 2016, and we thank the PL044 partnership of ConocoPhillips Skandinavia AS, Total E&P Norge AS, ENI Norge AS, and Statoil

Petroleum AS for their kind permission to work with this data set prior to its public release. We thank P. Valasek, formerly of ConocoPhillips, for originally identifying the data set as suitable for FWI analysis and making it available to us. We are grateful to S-Cube and Equinor for permission to publish this work. Commercial use of AWI is subject to GB patent number GB1319095.4; international patent protection is pending.

DATA AND MATERIALS AVAILABILITY

The field data associated with this research are available from the Norwegian Petroleum Directorate.

REFERENCES

- Alkhalifah, T., and Y. Choi, 2012, Taming waveform inversion non-linearity through phase unwrapping of the model and objective functions: *Geophysical Journal International*, **191**, 1171–1178.
- Ben-Hadj-Ali, H., S. Operto, and J. Virieux, 2008, 3D frequency-domain full-waveform tomography based on a domain decomposition forward problem: 78th Annual International Meeting, SEG, Expanded Abstracts, 1945–1949, doi: [10.1190/1.3059277](https://doi.org/10.1190/1.3059277).
- Biondi, B., and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: [10.1190/segam2012-0275.1](https://doi.org/10.1190/segam2012-0275.1).
- Biondi, B., and A. Almomin, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: *Geophysics*, **79**, no. 3, WA129–WA140, doi: [10.1190/geo2013-0340.1](https://doi.org/10.1190/geo2013-0340.1).
- Bozdag, E., J. Trampert, and J. Tromp, 2011, Misfit functions for full-waveform inversion based on instantaneous phase and envelope measurements: *Geophysical Journal International*, **185**, 845–870, doi: [10.1111/j.1365-246X.2011.04970.x](https://doi.org/10.1111/j.1365-246X.2011.04970.x).
- Debens, H. A., M. Warner, A. Umpleby, and N. V. da Silva, 2015, Global anisotropic 3D FWI: 85th Annual International Meeting, SEG, Expanded Abstracts, 1193–1197, doi: [10.1190/segam2015-5921944.1](https://doi.org/10.1190/segam2015-5921944.1).
- Fichtner, A., B. L. N. Kennett, H. Igel, and H. P. Bunge, 2008, Theoretical background for continental and global scale full waveform inversion in the time-frequency domain: *Geophysical Journal International*, **175**, 665–685.
- Fichtner, A., B. L. N. Kennett, H. Igel, and H.-P. Bunge, 2009, Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods: *Geophysical Journal International*, **179**, 1703–1725.
- Gauthier, O., J. Virieux, and A. Tarantola, 1986, Two dimensional non-linear inversion of seismic waveforms: Numerical results: *Geophysics*, **51**, 1387–1404, doi: [10.1190/1.1442188](https://doi.org/10.1190/1.1442188).
- 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).
- Huang, G., and W. W. Symes, 2015, Full waveform inversion via matched source extension: 85th Annual International Meeting, SEG, Expanded Abstracts, 1320–1325, doi: [10.1190/segam2015-5872566.1](https://doi.org/10.1190/segam2015-5872566.1).
- Igel, H., H. Djikpesse, and A. Tarantola, 1996, Waveform inversion of marine reflection seismograms for P-impedance and Poisson's ratio: *Geophysical Journal International*, **124**, 363–371.
- Jiao, K., D. Sun, X. Cheng, and D. Vigh, 2015, Matching pursuit full-waveform inversion: 77th Annual International Conference and Exhibition, EAGE, Extended Abstracts, N104 08, doi: [10.3997/2214-4609.201412911](https://doi.org/10.3997/2214-4609.201412911).
- Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering, Theory and Application, Society for Industrial and Applied Mathematics, Expanded Abstracts, 206–220.
- Luo, S., and P. Sava, 2011, A deconvolution-based objective function for wave-equation inversion: 81st Annual International Meeting, SEG, Expanded Abstracts, 2788–2792, doi: [10.1190/1.3627773](https://doi.org/10.1190/1.3627773).
- Ma, Y., and D. Hale, 2013, Wave-equation reflection travelttime inversion with dynamic warping and full-waveform inversion: *Geophysics*, **78**, no. 6, R223–R233, doi: [10.1190/geo2013-0004.1](https://doi.org/10.1190/geo2013-0004.1).
- Métivier, L., R. Brossier, Q. Méridot, E. Oudet, and J. Virieux, 2016, Measuring the misfit between seismograms using an optimal transport distance: Application to full waveform inversion: *Geophysical Journal International*, **205**, 345–377, doi: [10.1093/gji/ggw014](https://doi.org/10.1093/gji/ggw014).
- Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multi-offset seismic data: *Geophysics*, **52**, 1211–1228, doi: [10.1190/1.1442384](https://doi.org/10.1190/1.1442384).
- Plessix, R.-E., and C. Perkins, 2010, Full waveform inversion of a deep water ocean bottom seismometer dataset: *First Break*, **28**, 71–78, doi: [10.3997/1365-2397.2010013](https://doi.org/10.3997/1365-2397.2010013).
- Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain. Part 1: 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.
- 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.
- Ravaut, C., F. A. Maaø, J. Mispel, A. Osen, M. Warner, L. Guasch, and T. Nangoo, 2017, Velocity estimation without conventional tomography for a gas cloud field in the North Sea: 79th Annual International Conference and Exhibition, EAGE, Extended Abstracts, A3 04.
- Routh, P., R. Neelamani, R. Lu, S. Lazaratos, H. Braaksma, S. Hughes, R. Saltzer, J. Stewart, K. Naidu, H. Averill, V. Gottumukkula, P. Homonko, J. Reilly, and D. Leslie, 2017, Impact of high-resolution FWI in the Western Black Sea: Revealing overburden and reservoir complexity: *The Leading Edge*, **36**, 60–66.
- Shah, N. K., M. R. Warner, J. K. Washbourne, L. Guasch, and A. P. Umpleby, 2012, A phase-unwrapped solution for overcoming a poor starting model in full-wavefield inversion: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, P014, doi: [10.3997/2214-4609.20148455](https://doi.org/10.3997/2214-4609.20148455).
- 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, doi: [10.3997/2214-4609.20147683](https://doi.org/10.3997/2214-4609.20147683).
- Tape, C., Q. Liu, A. Maggi, and J. Tromp, 2010, Seismic tomography of the southern California crust based on spectral-element and adjoint methods: *Geophysical Journal International*, **180**, 433–462, doi: [10.1111/j.1365-246X.2009.04429.x](https://doi.org/10.1111/j.1365-246X.2009.04429.x).
- 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).
- van Leeuwen, T., and F. J. Herrmann, 2013, Mitigating local minima in full-waveform inversion by expanding the search space: *Geophysical Journal International*, **195**, 661–667, doi: [10.1093/gji/ggt258](https://doi.org/10.1093/gji/ggt258).
- van Leeuwen, T., and F. J. Herrmann, 2015, A penalty method for PDE-constrained optimization in inverse problems: *Inverse Problems*, **32**, 015007, doi: [10.1088/0266-5611/32/1/015007](https://doi.org/10.1088/0266-5611/32/1/015007).
- van Leeuwen, T., and W. A. Mulder, 2008, Velocity analysis based on data correlation: *Geophysical Prospecting*, **56**, 791–803.
- van Leeuwen, T., and W. A. Mulder, 2010, A correlation-based misfit criterion for wave-equation travel time tomography: *Geophysical Journal International*, **182**, 1383–1394.
- Vigh, D., B. Starr, J. Kapoor, and H. Li, 2010, 3D Full waveform inversion on a Gulf of Mexico WAZ data set: 80th Annual International Meeting, SEG, Expanded Abstracts, 957–961, doi: [10.1190/1.3513935](https://doi.org/10.1190/1.3513935).
- 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., A. Ratcliffe, T. Nangoo, J. V. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Stekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand, 2013, Anisotropic full-waveform inversion: *Geophysics*, **78**, no. 2, R59–R80, doi: [10.1190/geo2012-0338.1](https://doi.org/10.1190/geo2012-0338.1).
- Warner, M., I. Stekl, and A. Umpleby, 2008, Efficient and effective 3D wavefield tomography: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, F023, doi: [10.3997/2214-4609.20147684](https://doi.org/10.3997/2214-4609.20147684).
- Warner, M. R., and L. Guasch, 2014, Adaptive waveform inversion — FWI without cycle skipping — Theory: 76th Annual International Conference and Exhibition, EAGE, Extended Abstracts, E106 13, doi: [10.3997/2214-4609.20141092](https://doi.org/10.3997/2214-4609.20141092).
- Warner, M. R., and L. Guasch, 2016, Adaptive waveform inversion: Theory: *Geophysics*, **81**, no. 6, R429–R445, doi: [10.1190/geo2015-0387.1](https://doi.org/10.1190/geo2015-0387.1).