Full waveform inversion with wave equation migration and well control

Gary F. Margrave, Robert J. Ferguson and Chad M. Hogan

ABSTRACT

We examine the key concepts in full waveform inversion (FWI) and relate them to processes familiar to the practicing geophysicists. The central theoretical result behind FWI is presented as a mathematical theorem, the Fundamental Theorem of FWI. This theorem says that a linear update to a migration velocity model can be obtained from a reverse-time migration of the data residual (the difference between the actual data and data predicted by the model). Critically, this migration is only proportional to the required update and the proportionality must be estimated. We argue that in many cases this proportionality factor will be complex-valued and frequency dependent, or in the time domain, a wavelet. The estimation of the velocity update from the migrated section is closely related to the common process of impedance inversion. Then we argue that FWI can be viewed as a cycle of data modelling, migration of the data residual, and calibration of this migration to deduce the velocity update. We present an extended example using the Marmousi model in which we use wave-equation migration of the data residual and we calibrate the migration by matching it to the velocity residual (the difference between actual velocity and migration velocity) at a well. Our example produces an encouragingly detailed inversion but raises many questions.

INTRODUCTION

Full waveform inversion (FWI) was introduced to the exploration seismology community by Lailly (1983) and Tarantola (1984), although its use in whole-earth seismology is much older (Virieux and Operto, 2009). As originally proposed by Lailly (1983) the procedure required a sequence of pre-stack migrations each one using a velocity model improved upon by the previous migration. More recently, FWI is formulated as a generalized inverse problem and a numerical iterative solver (usually a gradient method) is used, can proceed using only a forward modelling code and its adjoint. While general and efficient, this loses the perspective of FWI as a sequence of migrations, and we return to that perspective here. We suggest that FWI can be viewed as an iterative cycle involving forward modelling, pre-stack migration, impedance inversion, and velocity model updating in each iteration.

Seismic migration has been used in exploration seismology as an imaging technique since at least the 1950's and was cast on firm theoretical grounds beginning with the work of Claerbout (1971 and 1976) and proceeding to the present. The computation of impedance traces from a merging of migrated traces and well information was introduced by Lindseth (1979) and has subsequently been called *impedance inversion* (e.g. Cooke and Schneider, 1983, Oldenburg et al., 1983). Migration methods have been the key process in seismic imaging since Claerbout's groundbreaking work and have evolved into methods of ever increasing physical fidelity and algorithmic complexity. Typically, a migration is followed by an inversion for rock properties under the assumption that the migration estimates reflection coefficients. Thus, practically, migration followed by

impedance inversion can be viewed as a two-step more general inversion. Twenty years ago, the migrations were almost always post-stack and the subsequent inversions estimated p-wave impedance only. Today, the migrations are usually pre-stack, giving angle-dependent reflectivity estimates, and the inversions estimate p-wave and s-wave impedance and sometimes density.

More recently, FWI has seen increasing use. Although originally posed in the time domain (e.g. Tarantola, 1984), effective frequency-domain implementations, especially with wide-angle refracted data, are demonstrated by Pratt and Worthington (1990), and Pratt (1990 and 1999). In either domain FWI uses an iteration where each linearized step reduces the objective function (or data misfit function: the sum-squared error between the actual data and forward modelled synthetic data) by deriving a model update; and the updated model is used in the subsequent step. An initial model of earth properties, in the simplest context this is a p-wave velocity model, is used to start the process and is updated with each iteration. Ideally the initial model should be very smooth with detail being added during the iterations. While each step in the iteration proceeds from a linear approximation, the entire process is a nonlinear inversion (Tarantola, 1984). The nature of the assumed physical model determines what is inverted for, for example a constant-density acoustic model allows inversion for p-wave velocity only while a variable-density, isotropic, elastic model allows inversion for p-wave and s-wave velocities and possibly density. Many other inversions are possible.

Viewed as a "black-box", or generalized process characterized only by its input and output, FWI is functionally a replacement for pre-stack migration plus inversion. The original proposal by Lailly (1983) was for a sequence of migrations although the connection with impedance inversion was not drawn. The link between FWI and migration is very strong and follows from the fundamental result that the model update is deduced directly from a migrated image by a simple scaling (Tarantola, 1984). However, in contrast with standard practice, only the data residual (the difference between actual data and that predicted by forward modelling) is migrated while standard practice is to migrate the data itself. Thus, if the velocity model is almost correct, then it will predict most features in the data and the data residual will be "small". Migration of this residual will predict a "small" update. In impedance inversion, an impedance model is deduced directly from the migrated section, typically with assist from well control, which has an analogy to the scaling step in FWI. However, the impedance model deduced from impedance inversion is almost never compared with the migration velocity model and no attempt is made to make them consistent. We could say that standard migration and impedance inversion are a mutually inconsistent set. FWI improves upon this in that a model update is deduced from the migration of the data residual and this update is then added to the existing model. Thus the velocity model is incrementally updated and is always consistent with both migration and the data. However, there is usually no attempt to harmonize FWI with well control as is commonly achieved with impedance inversion. In another parallel, FWI requires very low frequency signal in the seismic data, while impedance inversion obtains these low frequencies from well control.

In this paper we investigate the link between pre-stack migration, impedance inversion, and FWI. All of these methods are related by a common theme which is to

estimate subsurface properties from seismic data. However, they are algorithmically diverse and the conceptual and theoretical linkages between them are not always apparent. We begin by first summarizing important results from FWI theory and making a conjecture about their possible generalization. In particular, we conjecture that any standard depth migration and impedance inversion methods can be used in an FWI iteration with the proviso that the data residual is migrated and then matched at well control locations to the velocity residual. We have already defined data residual and by velocity residual we mean a similar thing: the difference between the exact velocity and the migration velocity. Since this velocity residual can only be known at wells, that is where we attempt a match and we call this process *calibration* of the migration. Then we present an extended exploration of an iterative migration scheme, using pre-stack wave-equation migration (WEM) and matching to well control.

THEORY AND METHOD

FTFWI: Fundamental Theorem of FWI (Lailly, Tarantola): Given real data from a constant-density acoustic medium and an approximate velocity model, $v_k(x,z)$, at the k^{th} step of an iteration, then a linearized update to the velocity model, δv, is given by

$$\delta v_k(x,z) = \lambda \int \sum_{s,r} \omega^2 \hat{\psi}_s(x,z,\omega) \delta \hat{\psi}^*_{r(s),k}(x,z,\omega) d\omega . \qquad (1)$$

where λ is a scalar constant, the hat (^) over a variable indicates its temporal Fourier transform, $\hat{\psi}_s(x,z,\omega)$ is a model of the source wavefield for source s propagated to all (x,z), $\delta\hat{\psi}_{r(s),k}(x,z,\omega)$ is the k^{th} data residual for source s back propagated to all (x,z), $*$ is complex conjugation, and ω is temporal frequency. Specifically $\delta\hat{\psi}_{r(s),k}(x,z,\omega) = \hat{\psi}_{r(s)}(x,z,\omega) - \hat{\psi}_{r(s),k}(x,z,\omega)$ where $\hat{\psi}_{r(s)}(x,z,\omega)$ is the real data at receivers $r(s)$ as back propagated into the medium and $\hat{\psi}_{r(s),k}(x,z,\omega)$ is the k^{th} data model for the same.

FWI (full-waveform inversion) can be viewed as an iterative forward modelling technique where at each stage of the iteration the earth model is updated and the updated model is used in the next iteration. In this paper we will only address the simplest FWI scheme which assumes constant-density acoustic waves. Therefore, the earth model is simply the velocity model. We let $v_k, k \in [0,N]$ denote the velocity model after the k^{th} iteration, with the initial model v_0 assumed to be very smooth. Then let $\psi_k(s,r) = M_{s,r}(v_{k-1}), k \in [1,N]$ be synthetic data for the k^{th} iteration where $M_{s,r}$ is a forward-modelling operator and (s,r) denotes the complete set of source and receiver locations for our seismic survey. Then the FWI objective function (also called the misfit function) is

$$\phi_k = \sum_{s,r} \left(\psi - \psi_k\right)^2 , \tag{2}$$

where the sum is taken over all sources and receivers. Thus the objective function measures the difference between the recorded data and modelled data at the k^{th} iteration. FWI seeks the optimal velocity model v_{opt} which minimizes equation 2, that is an iteration is sought such that $\phi_k \to \min$ as $v_k \to v_{opt}$. This iteration is driven by a famous result (equation 1) which we formulate here as a theorem in two spatial dimensions. This theorem, which we term the *fundamental theorem of FWI* (FTFWI) was first proven by Lailly (1983) and Tarantola (1984) and has since been demonstrated many times. A particularly transparent derivation may be found in Hammad (2010, see section 2.4.2 and equation2.35). The derivation proceeds by calculating the gradient of the objective function with respect to the velocity model. This remarkable calculation has the result that, for a monochromatic wavefield,

$$\nabla_v \phi_k \left(x,z,\omega\right) = \omega^2 \sum_{s,r} \hat{\psi}_s(x,z,\omega)\delta\hat{\psi}^*_{r(s),k}(x,z,\omega) \tag{3}$$

where ∇_v denotes the gradient with respect to the velocity model, and the other various symbols are defined in the box containing the FTFWI. If equation 3 gives the gradient of the objective function, then it follows that we know the direction to step in a gradient search for the minimum of equation 2, but we do not know how far to step. This "step length" is essentially the constant λ in equation 1. To derive equation 1 from equation 3 the step length is symbolized a $\lambda(\omega)$ and we then integrate over all frequencies. If the step length is assumed to be independent of frequency, the result is equation 1. The derivation further assumes that the source wavelet is known, and if this is not the case then the step length becomes frequency dependent with $\lambda(\omega)$ possibly complex.

Interpretation of equation 1 is very important at this stage. Essentially, equation 3 is a monochromatic pre-stack migration where a cross-correlation imaging condition is used and both the source and receiver fields have been time-differentiated resulting in the factor ω^2. Letting T denote record length, equation 3 is often written in the time-domain as

$$\nabla_v \phi_k \left(x,z\right) = -\sum_{s,r} \int \partial_t \psi_s(x,z,t)\partial_t \delta\psi_{r(s),k}(x,z,T-t)dt \tag{4}$$

where it is recognizable as a pre-stack reverse-time migration (RTM) with a cross correlation imaging condition between time-differentiated wavefields, which is then stacked over all sources and receivers. Thus we can interpret equation 1 as saying that the k^{th} update to the migration velocity model is proportional to an RTM of the k^{th} data residual $\delta\psi_{r(s),k}$. This is a very beautiful result that makes a great deal of sense in the light of experience with seismic migration. A "successful" depth migration is usually morphologically similar to the velocity model. Moreover, the many years of demonstrated success of impedance inversion of migrated sections is consistent with this. If a migrated section can be inverted for impedance, then it should be possible to deduce

a velocity model from the impedance estimate and re-migrate the data with improved results (this is rarely done). These observations suggest the direction of investigation taken in this study that any migration, not just the double-differentiated RTM can be used to update the velocity model.

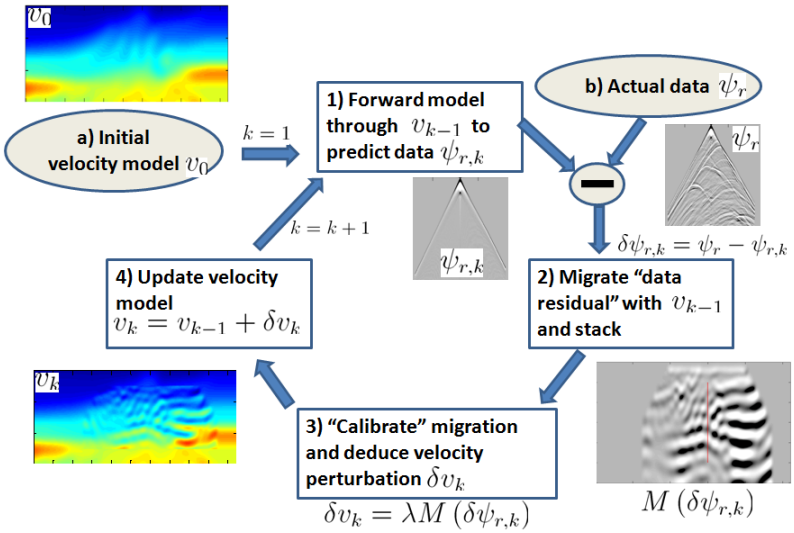

Fig. 1. The cycle of acoustic FWI (full waveform inversion). The cycle has external inputs at a) the initial velocity model, and b) the actual (or recorded) seismic data. The counter k begins at 1 and increments to N when some stopping criterion (not shown) is met. Step 1) Velocity model v_{k-1} is used to predict synthetic seismic data matching the acquisition geometry. Step 2) The data residual (real data - synthetic) is pre-stack migrated and stacked. Step 3) The pre-stack migration is "calibrated" to estimate a velocity perturbation δv_k. Step 4) The velocity model is updated by adding the perturbation to v_{k-1} to estimate v_k.

Figure 1 shows the cycle of FWI in the acoustic, constant-density context considered here. The cycle consists of four steps that are iterated until completion. There are two external inputs: the initial velocity model and the real seismic data. An ordinary migration does a single iteration of step 2) and bypasses step 1). Also, migration commonly uses a highly detailed initial model. In contrast, FWI uses a very smooth initial model but one which contains at least regional gradients. Step 1) uses velocity model v_{k-1} to predict synthetic data using the same recording geometry as the real data. Then, in step 2), the data residual, $\delta\psi_{r,k}=\psi_r-\psi_{r,k}$, is pre-stack migrated, again using velocity model v_{k-1}, and stacked. The FTFWI states that this migrated stack can be used to predict an update to the velocity model. Step 3) shows the *calibration* process which turns the migrated stack into a velocity perturbation. Under the ideal circumstances of the FTFWI, this is a simple scaling but under more realistic conditions it involves estimation of frequency-dependent residual amplitude and phase adjustments. In industry practice, this step is approximated by post-migration impedance inversion. Finally, step 4) adds the velocity perturbation to velocity model v_{k-1} to produce the updated velocity model v_k. There is no guarantee that this cycle converges to the correct answer mostly because the seismic inverse problem is nonlinear. This means that equation 2 has

multiple minima and, if the initial model is not sufficiently close to the true model, then the process may find a local minimum at best.

To help clarify matters, it has been useful for us to ask, and answer, the question:

When can an ordinary industry depth migration be expected to yield a useful velocity model update?

Since an ordinary depth migration migrates the data directly and not the data residual, the answer is:

When the velocity model is so smooth that the predicted data from it is essentially negligible, then the ordinary depth migration can possibly be used to update the velocity model.

This condition is often not met in practice since it is common to use initial velocity models that have abrupt discontinuities. However, the best practice in FWI uses very smooth initial models. Furthermore, it is common to migrate only very low frequencies first and then proceed to higher frequencies using a velocity model updated by the low-frequency result (Pratt, 1999). The smoothness of the initial model is directly linked to the need for low frequency signal in the data. Generally, the very lowest frequencies become low wavenumbers after migration and give low detail velocity perturbations. Pratt's scheme of progressing through the frequency band from low to high means that the velocity model gets progressively more detail as the FWI iteration proceeds. Pratt (1999) suggests that this allows the nonlinear inversion to avoid local minima.

The potential use of a conventional depth migration in an FWI iteration raises several questions. Can we expect the simple proportionality of equation 1 to hold? Can a deconvolution imaging condition be used? How should the migration be converted to a velocity model update? Can we easily iterate from low to high frequencies?

It seems unlikely that the simple proportionality of a frequency-independent scalar will hold in this generalized context. There is the issue already mentioned of an unknown source waveform giving rise to a complex-valued, frequency-dependent $\lambda(\omega)$. Furthermore, it may be that a code is available that does depth migration with a cross correlation imaging condition but without the double time differentiation of equation 4 or the ω^2 factor of equation 3.

If a deconvolution imaging condition is used (which is a better estimate of reflectivity) then further corrections can be expected. The general relation between a reflectivity estimate and an impedance perturbation is complicated and is explored in detail in Bleistein et al (2000). However, it is common practice to compare migrated seismic sections to synthetic seismograms created from well control and then to estimate acoustic impedance. In the constant-density, acoustic context an impedance estimate is essentially a velocity estimate. The common algorithm involves wavelet estimation by matching the seismic data to a synthetic seismogram, computation of bandlimited impedance from the matched data, and merging of low frequency information from the well log. The main

point here is that there is a long and successful history of estimating impedance (or velocity when density is known or constant) from depth migrated seismic data.

While the FTFWI gives a very simple formula for converting a specific kind of migration to a velocity perturbation, there is every reason to believe that this can be generalized to other migration algorithms. Rather than say that we are estimating a "step length" we prefer to call this process *calibrating the migration*. We propose to calibrate the migration by some sort of matching to well control. More specifically, we will match the stacked migrated data residual to the *velocity residual* at the well. In detail, we will match the stack of pre-stack migrated data residuals, after smoothing with a selected size Gaussian smoother, to the velocity residual at the well. By velocity residual we mean the difference between the migration velocity at the well and the velocity measured in the well. That being said, there are many ways to do this matching and we have only begun this process. Whatever the matching process, it must be designed at the well over the limited range of the well log and then applied to the entire migrated stack. We have investigated Wiener match filters and found that they produce a very good estimate at the well but tend to produce artefacts when applied away from the well. A more stable method has been to estimate a least-squares amplitude scalar and a least-squares constant-phase rotation. This can be done on narrow frequency bands to accommodate the expected frequency dependence.

Our choice of migration algorithm is constrained by the codes that we have readily available. However, we are confident that any depth migration can be made to work in this context. In this paper we have chosen a pre-stack PSPI (Gazdag and Squazerro, 1984) shot-record migration (hereafter simply called WEM (wave-equation migration)) because we know it works well on the Marmousi model and because it migrates each frequency independently. This last point is a nice advantage over RTM because it allows the easy production of migrations limited to any desired frequency band. Pratt (1999) prefers frequency domain FWI because it allows the velocity model to be updated first with low frequencies and then with progressively higher frequencies. Using WEM allows us to do something similar while still doing time-domain modelling.

In deriving the FTFWI it is assumed that the source wavelet is known. This wavelet is then used in the forward modelling which then optimizes the cross correlations in equation 1. In the next section we make this same assumption and so bypass any such difficulties. In practice, we anticipate that industry standard deconvolution methods will be useful to estimate and compress the natural wavelets in the raw data and also to remove trace-to-trace variations. Our final matching to well-control should then be able to cope with the residual wavelet.

TESTING

We have conducted a numerical experiment using the Marmousi model (Figure 2), where we created our own finite-difference (acoustic, constant density) data so that there would be no doubt about either the source waveform or the details of the modelling algorithm. We created 40 shot records using finite-difference modelling (second order in time and space), with sources at the surface beginning at 4000 m, incrementing by 100m, and extending to 7900m. Each source record has a split-spread receiver pattern with

offsets ranging from -2000m to +2000m at increments of 8.333m. We show sample source records in Figure 3. The source wavelet was chosen to be minimum phase with a 5 Hz dominant frequency, even so the dominant frequency of the recorded data is higher at about 10 Hz (Figure 4).

A starting velocity model was created by convolving the exact velocity model with a Gaussian smoother whose width was taken to be the nominal wavelength at 5 Hz, computed as $\lambda_{nom} = f_{dom} / v_{mean}$ which works out to be about 580 m. This starting model is shown in Figure 5. In Figure 6 is a comparison of shot 20 (see Figure 3) as modelled through the true Marmousi model and through the smooth model. It is apparent that the main feature in the record from the smoothed model is the first breaks and there is very little apparent reflection energy. However, extremely low frequency reflection events are very difficult to identify visually.

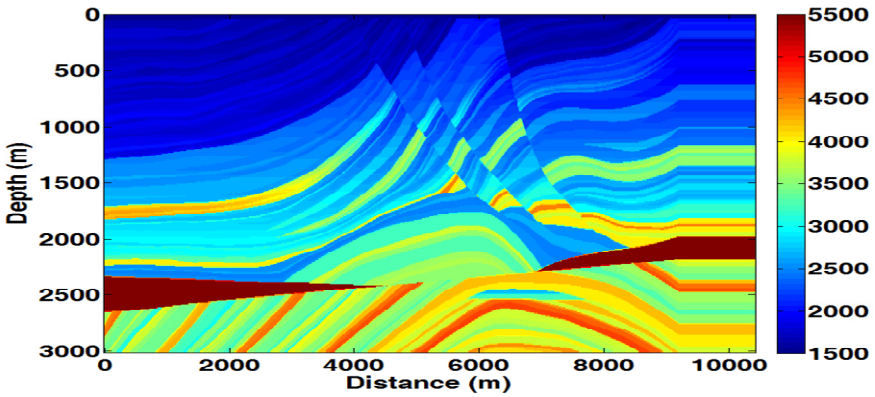

FIG. 2. The Marmousi 2D model is shown with colour indicating acoustic wavespeed in meters/sec. There is a 2:1 vertical exaggeration.

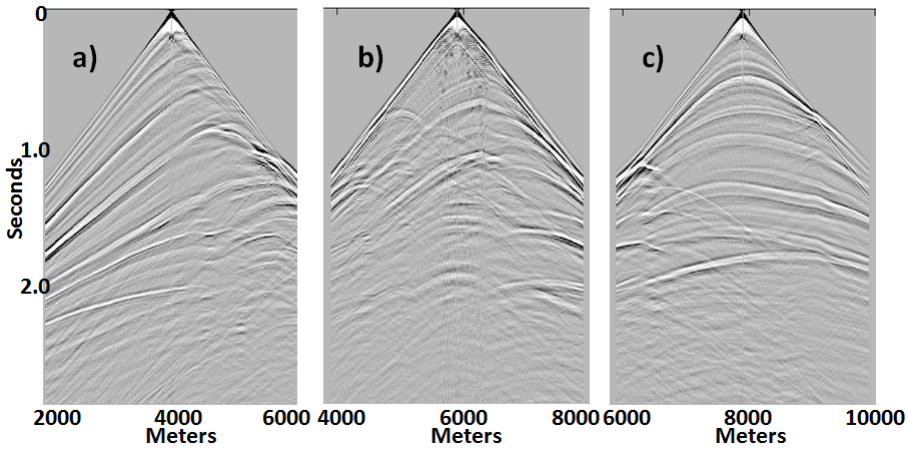

FIG. 3: Three of the 40 modelled shots used in this experiment. a) Shot 1, b) Shot 20, c) Shot 40.

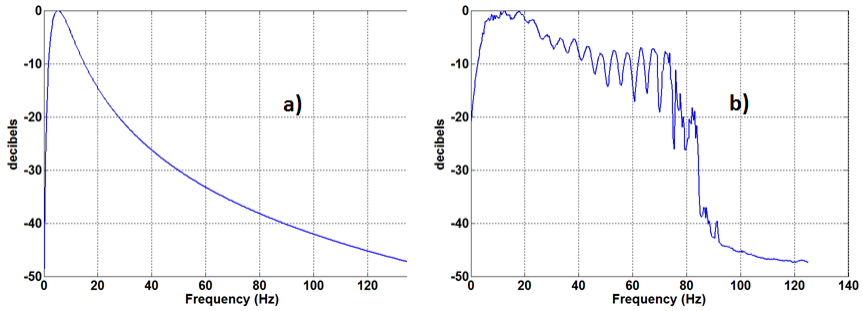

FIG. 4: a) Amplitude spectrum of the source wavelet used in the modelling. b) Average amplitude spectrum of shot 20 (see Figure 3).

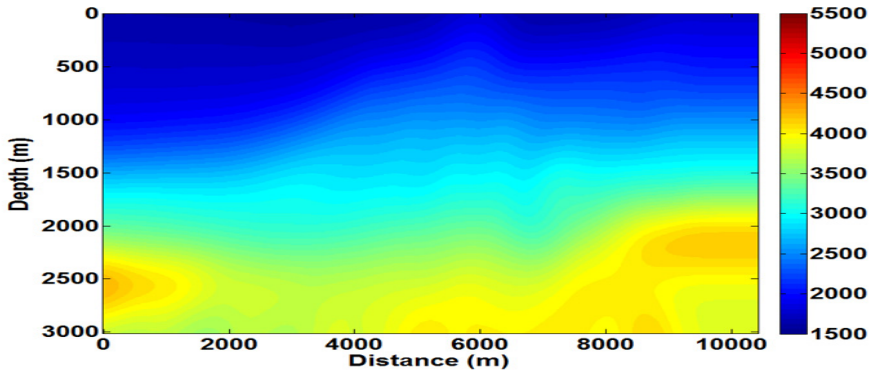

FIG. 5: The Marmousi velocity model of Figure 1 after convolution with a Gaussian smoother of width 580 m. This is the initial model in the FWI simulation.

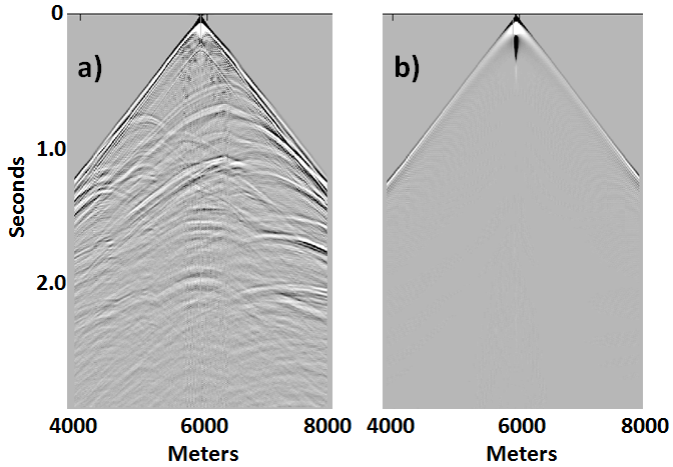

FIG. 6: a) Shot 20 of the test dataset (same as Figure 3b). b) A shot at the same location but modelled through the smoothed model of Figure 5.

As a migration algorithm, in this study we use a pre-stack PSPI (phase-shift-plus-interpolation) algorithm (Gazdag and Squazerro 1984) implemented by us (Ferguson and Margrave, 2005). PSPI is a space-frequency algorithm which means that each frequency is migrated independently. This has proven to be a very useful property because it facilitates an inversion process that is progressive in frequency by which we mean the low frequencies are inverted first and higher frequencies are inverted using the model as updated by the lower frequencies (as done by Pratt 1999).

PSPI is also a one-way depth-stepping algorithm and images primaries only whereas RTM, as is typically used in FWI, images all multiples. However, the cross-correlation imaging condition used in RTM produces many artefacts which can be generally described as events which correlate in space but cannot possibly be reflections. An example of such events would be a wave moving in a certain direction on the forward propagated source model and an event moving in nearly the same direction on the back-propagated data. Such events will correlate, often over a large areal extent, but one cannot possibly be the reflection of the other as they are moving in the same direction. The character of such an artefact is a low-frequency (i.e. very smooth) blurring of the image. On the other hand, a one-way depth-stepping method such as PSPI will only correlate events moving down against those moving up and so will not produce similar artefacts. So, while we cannot migrate multiples, we expect cleaner images of primaries than RTM.

We have investigated both cross-correlation and stabilized deconvolution imaging conditions in our PSPI migrations. If $\psi_{s\downarrow}(x,z,\omega)$ is the frequency-domain, downward-traveling shot model at depth z and $\psi_{r(s)\uparrow}(x,z,\omega)$ is the upward-traveling receiver data from that same shot at the same depth, then the cross-correlation image is given by

$$I_{cc,s}(x,z) = \int \psi^*_{s\downarrow}(x,z,\omega)\psi_{r(s)\uparrow}(x,z,\omega)d\omega \tag{5}$$

and the stabilized deconvolution image is given by

$$I_{dec,s}(x,z) = \int \frac{\psi^*_{s\downarrow}(x,z,\omega)\psi_{r(s)\uparrow}(x,z,\omega)}{\psi_{s\downarrow}(x,z,\omega)\psi^*_{s\downarrow}(x,z,\omega) + \mu I_{\max}(z)}d\omega \tag{6}$$

where $I_{\max}(z) = \max_{x,\omega}\left[\psi_{s\downarrow}(x,z,\omega)\psi^*_{s\downarrow}(x,z,\omega)\right]$. In both of these equations, integration is over all migrated frequencies. In equation 6, μ is a small non-negative constant which we have taken to be 10^{-4}. Note that equation 6 has the two limiting forms

$$\begin{aligned} I_{dec,s}(x,z) &\approx \int \frac{\psi_{r(s)\uparrow}}{\psi_{s\downarrow}}d\omega, \quad \mu I_{\max} \ll \left|\psi_{s\downarrow}\psi^*_{s\downarrow}\right| \\ I_{dec,s}(x,z) &\approx \frac{1}{\mu I_{\max}}\int \psi_{r(s)\uparrow}\psi^*_{s\downarrow}d\omega, \quad \mu I_{\max} \gg \left|\psi_{s\downarrow}\psi^*_{s\downarrow}\right| \end{aligned} \tag{7}$$

so that when $\mu I_{\max}$ is negligible we get the ratio of reflected to incident fields which is an estimate of a reflection coefficient while in other extreme we get a simple, stable cross-correlation.

When used in a shot-record migration as in the present context, the cross-correlation imaging condition results in an image that must be gained to approximate a reflection coefficient. Consider the simple case of normal incidence on a planar reflector beneath a

homogeneous medium of velocity v. Then we can approximate both the incident and reflected fields at the image point as 2D Green's functions with the asymptotic forms $\psi_{s\downarrow}(x,z,\omega) \approx \rho^{-1/2} e^{ik\rho + i\pi/4}$ and $\psi_{r(s)\uparrow}(x,z,\omega) \approx R\rho^{-1/2} e^{ik\rho + i\pi/4}$, where $k = \omega / v$, ρ is the 2D distance from source to image point, and R is the normal incidence reflection coefficient. Then, in this special case, we find $I_{cc,s}(x,z) \sim \dfrac{aR}{\rho}$ and $I_{dec,s}(x,z) \sim a'R$ where $a = \int d\omega$ and $a' \approx \mu^{-1} I_{\max}^{-1} \int d\omega$ are constants. So the deconvolution image is proportional to R while the cross-correlation image must be gained by ρ to achieve simple proportionality.

While it is theoretically possible to calibrate individual migrated shots into velocity perturbations, we have elected to stack the shots and calibrate the migrated stack. Thus we ignore any AVO effects and also any misalignment in common image gathers. Therefore, the pre-stack migrations discussed here will all be stacked into a final image.

It is interesting to compare migrations of the data only (standard practice) with migrations of the data residual as required by the FTFWI. Figure 7 shows the migrated stacks using frequencies 0-5 Hz and the velocity model of Figure 5. In panel 7a), is the image $I_{dec} = \sum_s I_{dec,s}$ formed using the data only while panel 7b) is a similar image formed by migrating the data residual, i.e. the difference between the actual data and that predicted by the migration model (Figure 6). The large horizontal black event at the top of 7a) is an artefact formed by the first breaks. In panel 7b) this artefact is suppressed because the data residual reduces the first breaks (see Figure 6). In panel 7c) is the image $I_{cc} = \sum_s I_{cc,s}$ formed from the data only while panel 7d) is a similar image formed with the data residual is in 7d). The I_{cc} image in 7c) shows the first-brek artefact mentioned previously but is also clouded by a very-low frequency, white, pervasive artefact of unknown origin. Interestingly, this artefact is also not found in the data-difference migration. While I_{cc} and I_{dec} are quite different when only the data is migrated, when migrating the data residual they are quite similar. This observation may well be specific to the particular frequency band used in this test, nevertheless, for the remainder of this test, we will only show I_{dec} images of the data residual.

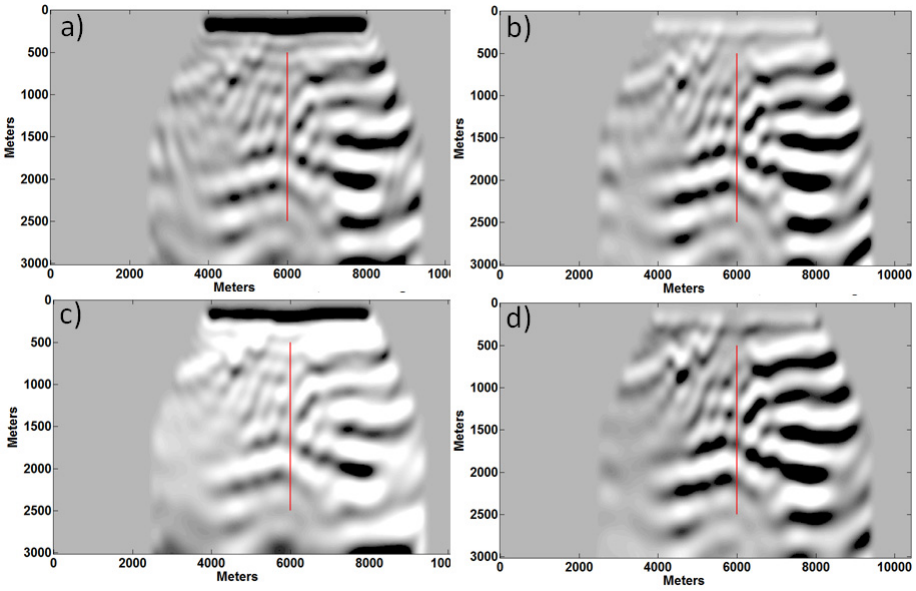

FIG. 7: Comparison of migration of the data versus migration of the data residual. All migrations used only 0-5 Hz. a) Deconvolution imaging condition migration of data only, b) As in a) except migration of the data residual, c) Cross-correlation imaging condition migration of data only, d) As in c) except migration of the data residual. The vertical red line on each panel is the location of the assumed well and log.

As discussed previously, standard FWI uses a "line search" to find a single scalar to calibrate the migrated image into a velocity perturbation, and we feel that the WEM stack produced by PSPI will require a more detailed procedure. In particular, since we feel this step is closely approximated in practice by tying to well control and inverting for impedance, we simulate this process using a single column of velocity values from $x = 6000$ m in the exact velocity model to simulate a well log. Moreover, we limit the depth span of this log to the interval $[500, 2500]$ m (Figure 7).

Figure 8 shows the result of the calibration process at the well. In detail, we first smooth the migrated stack with a 2D Gaussian smoother whose half-width is related to the shortest wavelength. Secondly we determine the best (least squares) scalar required to match the migrated trace in panel 8a) to the velocity residual at the well, which is the difference between the blue and red curves in panel 8b). Thirdly, we determine the best (least squares) constant phase rotation to match the scaled, migrated trace to the velocity residual. Finally, this amplitude scalar and phase rotation are applied to the entire smoothed migrated stack of Figure 7b) to create the velocity perturbation. This perturbation is the added to the migration model to form the velocity model for the second iteration, as shown in Figure 9.

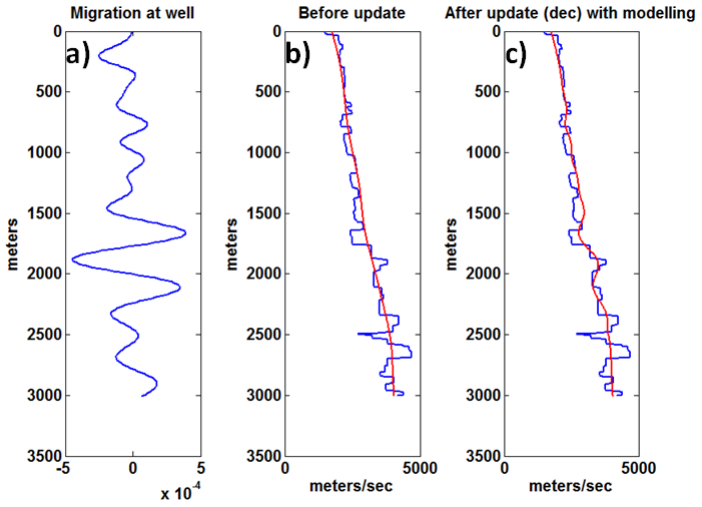

FIG. 8: The calibration process at the well location (red line in Figure 7) is shown. a) The migrated trace from panel 7b at the well. b) The well velocity (blue) and the migration velocity at the well (red). c) the calibrated migration trace has been added to the migration velocity in the interval [500,2500] m over which the well was "logged".

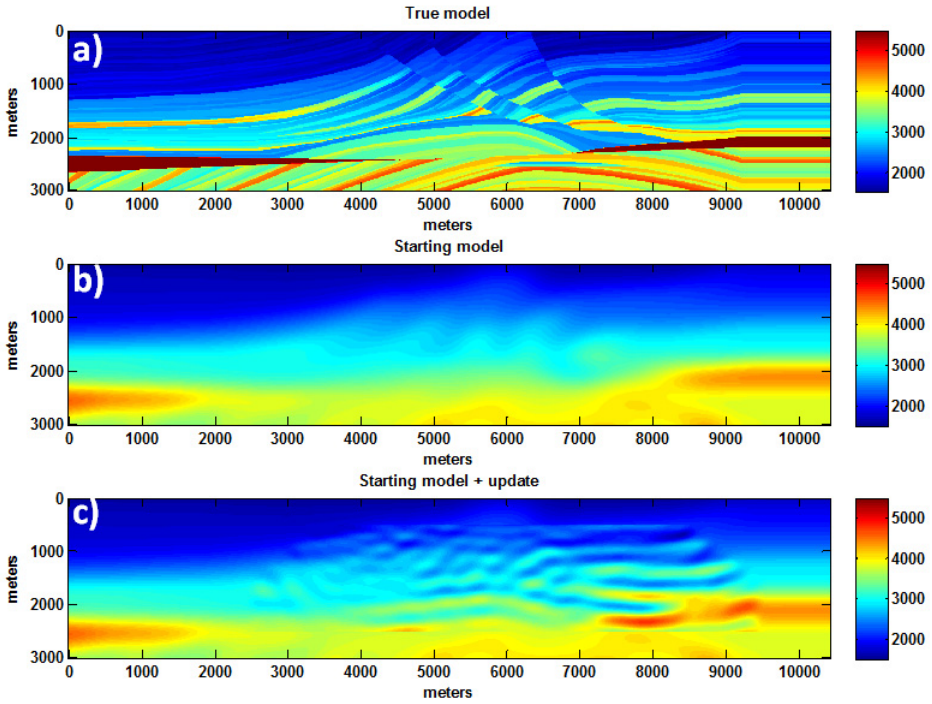

FIG. 9: a) The exact Marmousi velocity model (same as Figure 2). b) The smoothed starting model used in the migrations of Figure 7 (same as Figure 5). c) The velocity model after updating with the calibrated migration stack. The update was a scaled and phase rotated version of that in Figure 7b. The phase rotation and scaled factor were constant for the entire stack and are illustrated at the well in Figure 8.

There are many ways that an iterative scheme can be constructed from these basic ideas. Pratt (1999) inverts single frequency data using a number of iterations for each frequency. Beginning with a very low frequency, in the range 2-5 Hz, Pratt then

proceeds to a higher frequency only after the inversion at a lower frequency has stabilized. Pratt and co-workers have demonstrated very high resolution results using as few as 3 frequencies. We have not yet been able to obtain results with so few frequencies but we have used the basic idea of beginning with low frequencies and moving up the spectrum. Figure 10 illustrates one such iteration where we begin with the frequency band 0-4 Hz and progress towards higher frequencies in non-uniform steps. In each case only one iteration was taken at each frequency band and the starting model for each frequency band was the result from the previous band. Using a space-frequency migration algorithm like PSPI facilitates the selection of any desired frequency band, although in the modelling stage we are still modelling all frequencies.

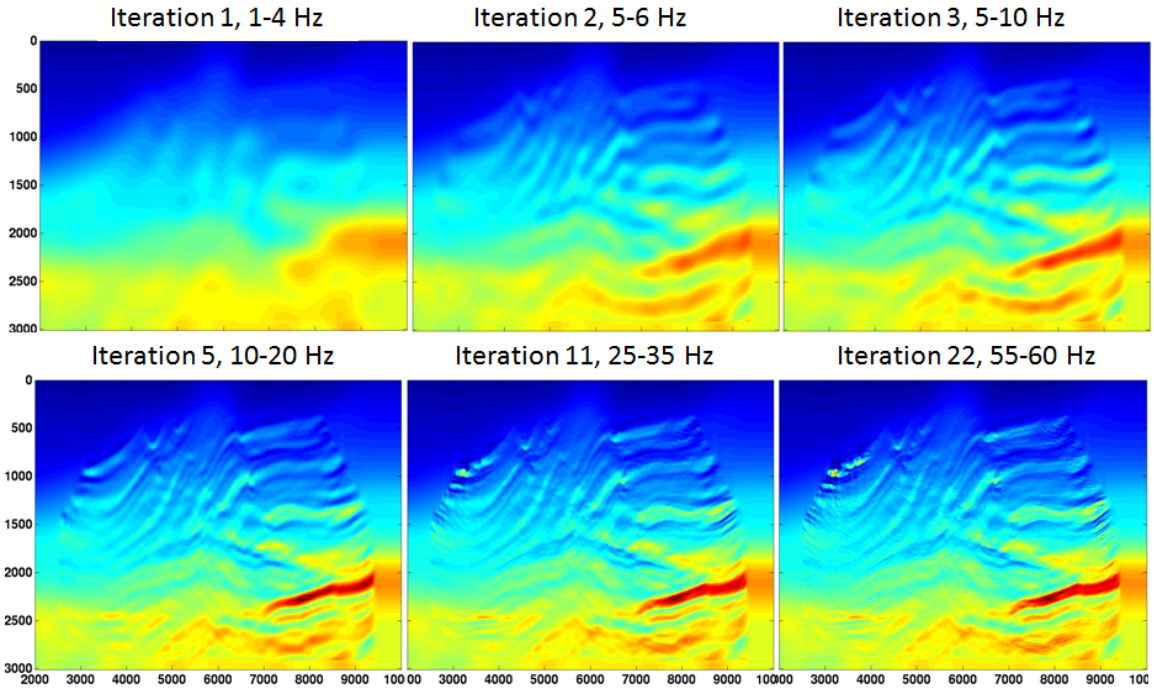

FIG. 10: A sequence of velocity models resulting from a frequency-dependent iteration. Narrow frequency bands were inverted sequentially beginning with the lowest frequencies. The starting model for each iteration was that deduced by the previous iteration. Most of the changes occur in the first few iterations but changes are still apparent on the last iteration (e.g. near (x,z)=(7200,500).

In our calibration step, we smooth the migrated stack with a 2D Gaussian smoother of a selected size as the initial action. In Figure 10, the smoother half-width was programmed to steadily decrease according to $h_w = v_{mean} / (6 f_{\max})$ where $v_{mean} = 2850$ m/s is the mean velocity of the Marmousi model and $f_{\max}$ is the maximum frequency in each iteration. An investigation of the effect of this smoother half-width is shown in Figure 11. Here we fixed the frequency band at 1-40 Hz and varied the smoother on each iteration. As before, each iteration provides an update to the the model determined by the previous iteration. On each the twelve iterations, the smoother half-width was [1000 750 500 400 300 200 100 50 40 30 20 10] in meters. Figure 11 only shows the results from 6 selected iterations. Through iteration 6, very little has happened, then at iteration 7, with a smoother half-width of 100 m, the result comes sharply into focus. Further

improvements through iteration 12 are present but subtle. Using v_{mean} to estimate the mean wavelength, we find that the 200m smoother width is comparable to the mean wavelength at 15 Hz, which is essentially the dominant frequency (Figure 4b). Figure 12 shows the L_2 norm of the data residual corresponding to the experiment in Figure 11. The data residual is seen to decrease sharply on the iteration after the model sharpens.

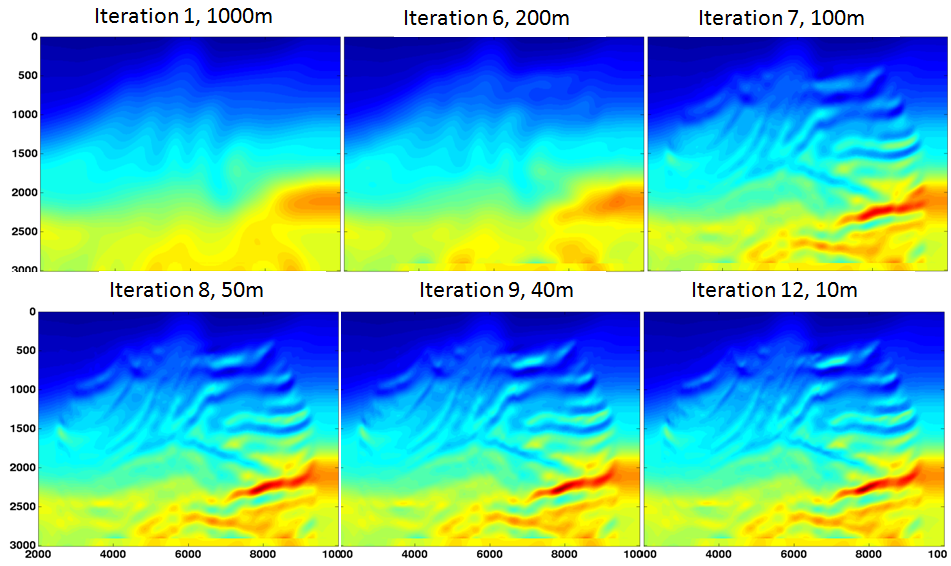

FIG. 11: A sequence of velocity models resulting from an iteration in which the frequency band is kept constant (1-40Hz) but the size of the 2D Gaussian convolutional smoother applied to the migrated stack is varied. The smoother size is noted above each image.

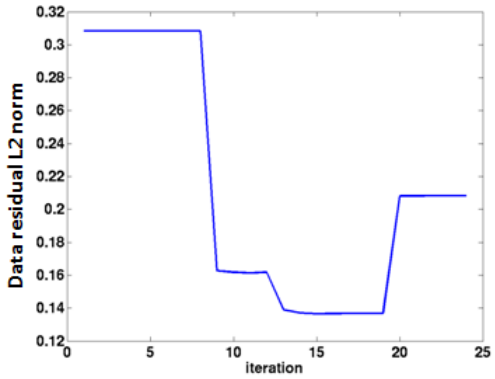

FIG. 12. The L_2 norm of the data residual of the experiment in Figure 11.

The experiments shown so far use a programmed mute in the migration stack. Figure 12 shows a sample CIG (common image gather) at the location of shot 40. This gather was formed on the first iteration and hence used the velocity model of Figure 5. It is apparent that stacking all offsets will not be optimal and three different mute trajectories are shown. The normal mute represents a 1:1 offset:depth curve truncated by a maximum offset of 1500 m (here offset means the distance from the image point to a given source location). The narrow mute is a 1:2 offset:depth curve and the wide mute is a 2:1 curve. Both are also limited by the maximum of 1500m. The normal mute was used so far and

thus specifically in Figure 11. We repeat the experiment of Figure 11 with the wide mute in Figure 14 and with the narrow mute in Figure 15. The wide mute has allowed some clear artefacts into the solution but has also resulted in slightly higher spatial resolution. The narrow mute produces a smoother image with fewer artefacts and overall lower resolution.

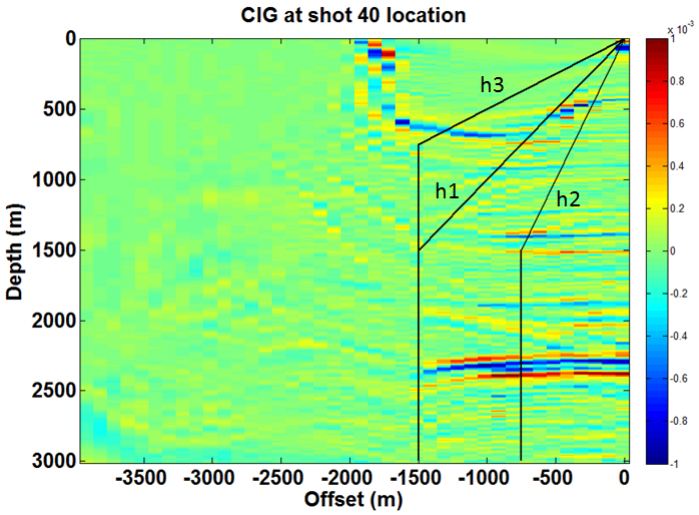

FIG. 13. A CIG (common image gather) formed at the location of shot 40 on the first iteration of the experiment of Figure 11. Three mute trajectories are shown: normal (h1), narrow (h2), and wide (h3). For any depth, only samples at offsets less than the mute trajectory are allowed into the stack.

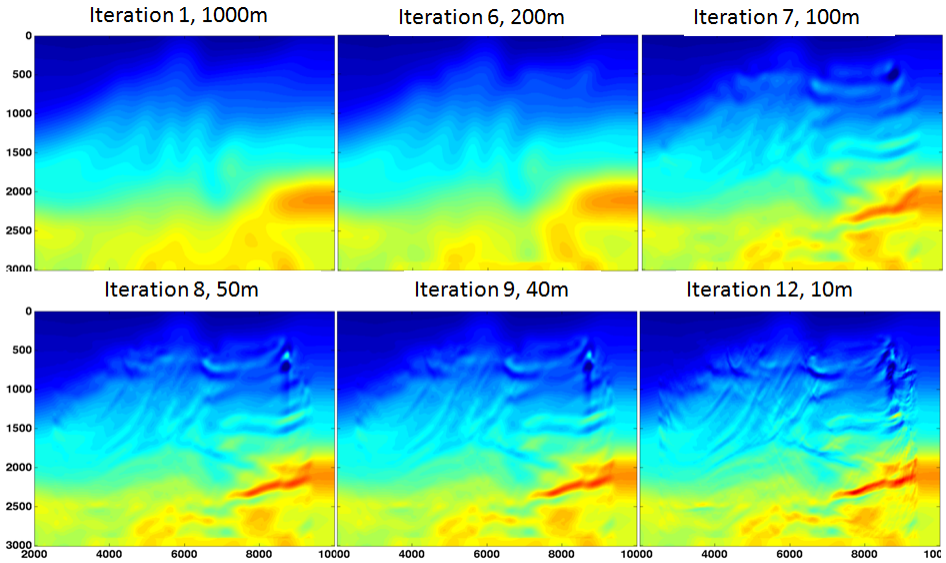

FIG. 14: Similar to Figure 11 except that the wide mute (Figure 13) was used.

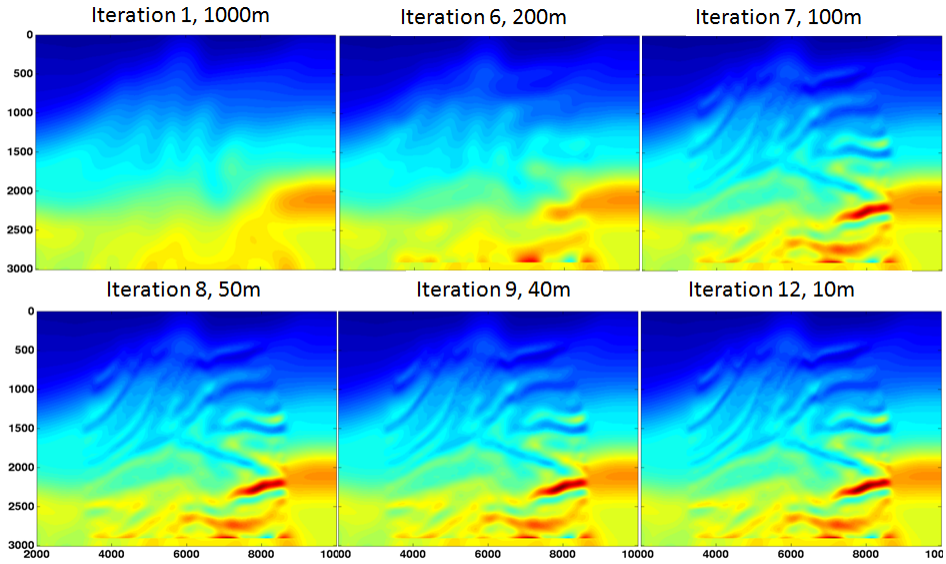

FIG. 15: Similar to Figure 11 except that the narrow mute (Figure 13) was used.

As a final experiment, in Figure 16 we show an experiment similar to the previous ones except that the smoother was fixed at 12m for all iterations and the wide mute was used. Here the image sharpens right away and is essentially stable after 3 iterations. The resolution seems quite high. Figure 17 shows the L_2 norm of the data residual for this experiment and we see that it does reach a minimum at the third iteration. In Figure 18 we compare the original shot 20 with the prediction of shot 20 after the third iteration. While these shots have many similar features, there are also clear differences.

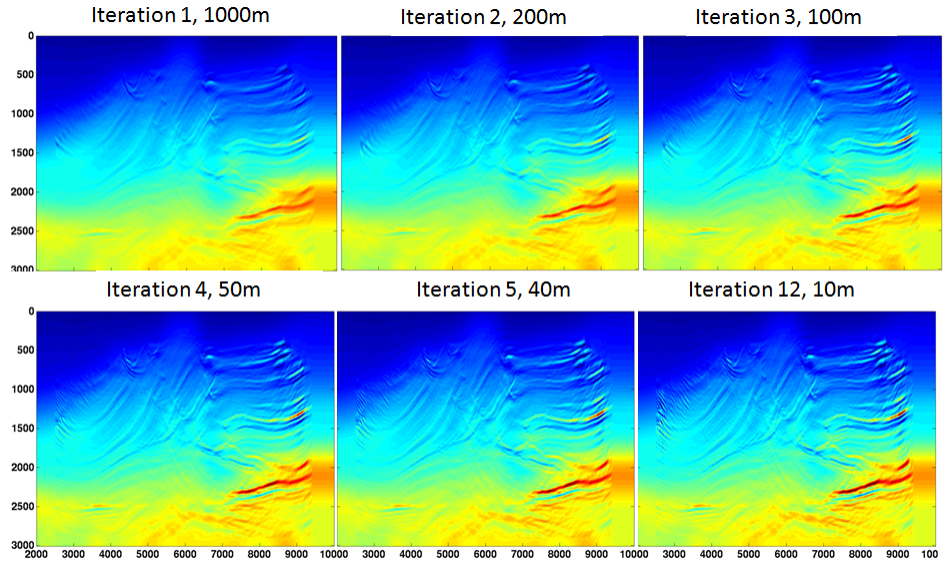

FIG. 16: An experiment similar to those of Figures 11, 14, and 15 except that the smoother was held constant at 12m and the wide mute was used.

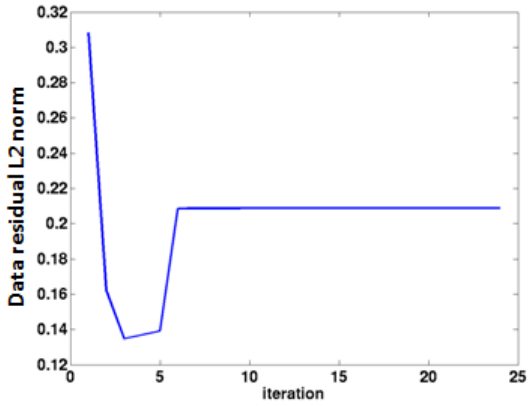

FIG. 17: The L_2 norm of the data residual for the experiment of Figure 16 is plotted versus iteration number.

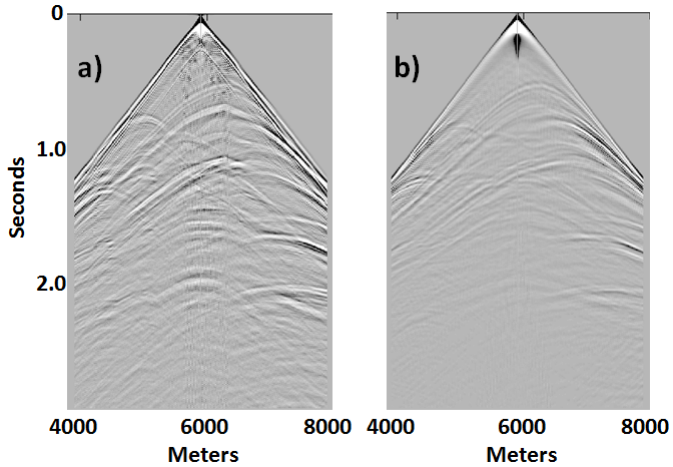

FIG. 18. a) Shot 20 of the test dataset (same as Figure 6a). b) A shot at the same location but modelled through the velocity model Figure 16 (iteration 3).

DISCUSSION AND CONCLUSIONS

We have presented an investigation into the foundations of FWI (full waveform inversion) in which we have isolated the four essential steps in the FWI iteration and compared them to standard practice. External inputs to the iteration are the recorded data and an initial velocity model, which should be very smooth. The first step in the iteration is a forward modelling step to predict the recorded data given the initial velocity model. Then the data residual, the difference between the recorded and predicted data is migrated and stacked. Theory seems to call for a reverse-time migration but we have demonstrated the use of a standard wave-equation migration (PSPI) is possible. The third step in the cycle, which we call calibration, is designed to convert the migrated stack into an estimate of a velocity model update. Theoretically, this is expected to be a simple frequency-independent scaling that is accomplished by a simple forward modelling iteration called a line-search. However, we have argued that this simple scaling is not likely true when the source waveform is unknown and when a wave-equation migration is used. We have argued that this calibration step is comparable to the standard process of impedance inversion by matching to well control. We have tested the implementation

of calibration using well control where, at each step in the iteration, we calibrate to the velocity residual which is the difference between the well velocity and the migration velocity. We implemented a three-step calibration procedure: (i) smooth the migration with a selected convolutional smoother, (ii) determine the best (least-squares) scalar to match the smoothed migration to the velocity residual, and (iii) determine the best (least-squares) phase rotation to match the scaled and smoothed migration to the velocity residual.

We demonstrated these concepts using 40 shots created from the Marmousi model using time-domain, acoustic finite-difference modelling. Our starting velocity model was created from the exact one by 2D smoother with a 590m wide Gaussian. We show a variety of iterations in which we varied the frequency band from low to high, varied the smoother length, and varied the mute used in the migration stack. In all cases, we were able to get reasonable results after only a few iterations. Examination of the L_2 norm of the data residual showed that it tends to decrease sharply and then either level off or increase. The number of useful iterations was generally less than 12.

This study raises a number of questions including:

- What is the best way to calibrate the migration when well control is available?
- How does calibration to well control compare to calibration without well control using a line search?
- How does the type of migration algorithm influence the calibration step?
- What limitations are imposed by using one-way migration instead of reverse time migration? What advantages are gained?
- What is the optimal iteration procedure? Should the iteration proceed from low to high frequency or should the smoother be varied? Or perhaps another process altogether.

We hope to report soon on further investigations into these issues.

ACKNOWLEDGEMENTS

We thank the Sponsors of CREWES for their continued support and also thank NSERC, MITACS, and the sponsors of POTSI. We thank Hussain Hammad for an excellent thesis and informative discussions.

REFERENCES

Bleistein, N., J. K. Cohen, and J. W. Stockwell, 2000, Mathematics of multidimensional seismic imaging, migration, and inversion: Springer, Interdisciplinary Applied Mathematics Series, **13**.

Claerbout, J., 1971, Towards a unified theory of reflector mapping: GEOPHYSICS, **36**, 467-481.

Claerbout, J., 1976, Fundamentals of geophysical data processing: McGraw-Hill.

Cooke, D.A., and W. Schneider, 1983, Generalized linear inversion of reflection seismic data: Geophysics, 48, 665-676.

Ferguson, R. J., and Margrave, G. F., 2005, Planned seismic imaging using explicit, one-way operators: Geophysics, Society of Exploration Geophysicists, 70, S101 - S109.

Gazdag, J., and Squazerro, P., 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, **49**, 124-131.

Hammad, H. I., 2010, Waveform inversion for areas with complex near surface, MSc. thesis, Department of Geoscience, The University of Calgary.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering, Theory and Application, Society of Industrial and Applied Mathematics, Expanded Abstracts, 206-220.

Lindseth, R. O., 1979, Synthetic sonic logs - a process for stratigraphic interpretation: GEOPHYSICS, **44**, 3-26.

Oldenburg, D., T. Scheuer, and S. Levy, 1983, Recovery of the acoustic impedance from reflection seismograms: Geophysics, 48, 1318-1337.

Pratt, R. G., 1990, Inverse theory applied to multi-source cross-hole tomography, Part II: Elastic wave-equation method: Geophysical Prospecting, **38**, 311-330.

Pratt, R. G., and M. H. Worthington, 1990, Inverse theory applied to multi-source cross-hole tomography, Part I: Acoustic wave-equation method: Geophysical Prospecting, **38**, 287-310.

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.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: GEOPHYSICS, **49**, 1259-1256.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: GEOPHYSICS, **74**, WCC1-WCC26.