# The influence of source wavelet estimation error in acoustic time domain full waveform inversion

Paschalia Pavlopoulou[1,2] and Ian F. Jones[1]*

## Abstract

Almost all forms of full waveform inversion use a source wavelet estimate in the extrapolation of the down-going wavefield to produce the forward modelled data and also to create the gradient. This source wavelet can be modelled, extracted from the recorded data during the pre-processing, or sometimes a zero-phase band-limited synthetic wavelet can be used instead. Alternatively, the source information can be regarded as another unknown and the initial source estimate then updated within the inversion.

The importance of reliable source wavelet information during the waveform inversion implementation has been widely implied or briefly mentioned by multiple authors, but little detailed study of the effects of poor wavelets has been presented in the literature. It is the purpose of this study to examine if shape errors in the source wavelet manifest themselves in a significantly damaging way in the velocity model obtained using conventional time-domain acoustic waveform inversion, and also to assess what effect such velocity change has on the subsequent migrated image positions.

We conclude that for modest structural complexity, acoustic time-domain least-squares waveform inversion for refractions with a maximum frequency of 9 Hz, commencing from a non-cycle skipped model, is surprisingly robust with respect to source wavelet error.

## Introduction

Over recent years, full waveform inversion (FWI) has become an important tool for estimating high-resolution model parameters that faithfully describe the properties of the subsurface. Since its introduction in the early 1980s by Lailly (1984) and Tarantola (1984), many authors have introduced multiple techniques to implement waveform inversion through various approximations.

In the context of exploration seismology and more specifically in velocity model building our main goal is to calculate the parameters that describe the physical properties of the subsurface that govern elastic wave propagation and explain the observed seismic field data. The parameters referred to as the Earth model, are subsequently used to produce a realistic representation of the subsurface *image* through the process of migration.

The importance of reliable source wavelet information during the FWI implementation has been widely implied or briefly mentioned by several authors (e.g. Pratt, 1999; Virieux and Operto, 2009; Rickett, 2013; Sun et al., 2014; Wang et al., 2017; Jones, 2018), but little detailed study of the effects of poor wavelets has been presented in the literature (Pavlopoulou, 2019).

After presenting a brief description of inherent source wavelet variability, and of where FWI uses the source wavelet, we will describe some of the mechanisms leading to source wavelet estimation error. The current literature suggests that the source signature plays a crucial role in the inversion, but how crucial this role is, is a matter of question. In this work we quantitatively assess the effect of source wavelet error on both the resulting velocity field and on the depth of horizons in the associated migrated images, firstly for a synthetic 2D data example, and then for a 3D field data example.

## Factors affecting source shape

In a typical marine survey, the source signature is generated by an airgun array which is distributed over a volume of many cubic metres. This gives rise to a directivity dependency in the source wavelet shape, and after interaction with the free surface (producing a 'ghost'), the phase behaviour of the resulting waveform is angle dependent. In addition, the gas bubble produced by an airgun array expands and then contracts repeatedly as it floats towards the surface, producing a quasi-periodic oscillatory bubble pulse. This behaviour constitutes a time-varying source which is difficult to model, so is either incorporated into the source waveform or removed in pre-processing. Variation in gun pressure will also alter the source-signature shape, as will changes in sea-state or even water temperature. For a land survey, where highly variable near-surface elastic effects predominate, each source can be significantly different in terms of emerging wavelet shape, and the directivity of each individual source can be even more pronounced than in the marine case.

Even if we have field recordings of each source, the overall source waveform shape variation can be difficult to model with a finite difference (FD) propagator, given that the FD cell size may be greater than the airgun array size or the distance over which near-surface land conditions vary, and also because directivity is still present. In addition, the FD modelling will usually have a regular or at best curvilinear grid distribution of nodes, which will not in general correspond to actual source locations (although this is a lesser issue). Consequently, the modelling itself also introduces an additional, albeit small, 'numerical' source directivity.

With these observations in mind, we can see how the dependency of the inversion result on wavelet shape is inherently complicated. In addition, the recorded wavelet appearance also depends on the receiver instrument response, its coupling coefficient, and vector fidelity.

## Method

FWI is an indirect seismic parameter estimation technique that searches for a high-fidelity high-resolution Earth model that explains the kinematics of the recorded data. It ideally utilizes the full waveform, which means that the travel-time, phase or amplitude information is involved in the inversion process. Waveform inversion simulates the wavefield by attempting to solve a form of the wave equation: in practice this is usually an acoustic approximation of the wave equation (i.e. shear mode conversion and propagation are not accounted for, and density change is also often ignored). The acoustic FWI approach that inverts only for the anisotropic velocity model is currently the commercial standard (Virieux and Operto, 2009; Warner and Guasch, 2016; Jones, 2018). The high-resolution P-wave velocity model retrieved by FWI is then typically used for subsequent pre-stack depth-migration (preSDM).

Given that changes in the source wavelet will alter the resulting inverted velocities in model $m$, we want to assess the nature and degree of these changes. Here we assess the sensitivity to source wavelet error by considering two norms. Firstly we consider the total residual velocity error $C_v(m)$ in the $m^{th}$ velocity model which we define as being the sum of absolute difference in velocities between some reference velocity model, $v_{ref}$, and the model $v_m$ obtained by FWI using the $m^{th}$ perturbed source wavelet (summed over all: inlines, crosslines, and time-samples ($t$)):

$$C_v\left(m\right) = \sum_{inlines} \sum_{crosslines} \sum_t \left| v_{ref} - v_m \right| \qquad (1)$$

Secondly we assess the vertical depth positioning error $C_z(m)$, with respect to some reference horizon $z_{ref}$, that would result from using the perturbed rather than the reference velocity model in the migration:

$$C_z\left(m\right) = \sum_{inlines} \sum_{crosslines} \left| z_{ref} - z_m \right| \qquad (2)$$

Where $z_m$ is the depth of the reference horizon in the migration using the $m^{th}$ model.

The analysis is conducted initially for a synthetic data set and then for a field data case study example from the Norwegian Sea (Singh et al., 2020). In the synthetic trials, we independently perturb wavelet phase and then wavelet time-shift, and repeat this exercise initially for the exact reference model, and then from using a smoothed version of the reference model.

For the field data, where the 'true' model and source wavelet are unknown, we make do with a best estimate of both, and then derive a series of different, but reasonable, source wavelets to again assess the changes in the velocity model resulting from FWI using these wavelets. In an industrial FWI project, we would process the entire 3D data volume though typically 50-100 iterations of FWI, using different variants of FWI (travel-time norm, data-difference norm, refraction, reflection), and employing staging strategies for differing offset and frequency ranges. The forward modelling would be conducted using parameters for velocity, anisotropy, and perhaps density and Q. Such a project would take a few months, consuming a large amount of CPU-time, and at best employ a source wavelet that was updated occasionally as the model converged (in an attempt to accommodate initial source estimation errors and/or directivity effects). Note that source wavelets are often estimated from near trace stacks whereas the refracted events being used in FWI 'see' wavelets propagating at higher emergence angles. However, due to the practicalities of time and CPU constraints for this study, here we have limited the procedure for the 3D field data analysis as follows:

1) Starting from the final TTI velocity model from a successful commercial FWI project (Singh et al., 2020), we modify this model by inverting a swath of data to a broader bandwidth using the final 'production' source wavelet. The inversion scheme used is that of a conventional data-difference least-squares norm, inverting to a maximum frequency of 9 Hz, employing the refracted (transmitted) wavefield. Hence the data input to the inversion are restricted to a window of data that primarily captures the diving wave energy. The anisotropy parameters were not updated in these tests. The output from this step constitutes our reference model, updated using only the swath of input data employed in this study.

2) As above, starting from the final TTI 'production' model, now rerun the inversion with a suite of alternative source wavelets, and allow the FWI to run for a maximum of ten
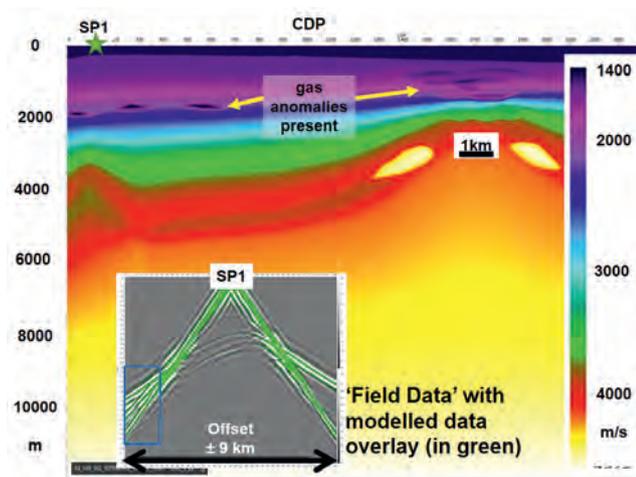


**Figure 1** Synthetic test conducted using the BP 2004 benchmark model with small-scale 'gas' anomalies. Inset: synthetic shot 'field data' (greyscale) with colour overlay of FWI modelled data, from location denoted by SP1. The blue box on the left of the inset indicates the zoom area shown in the next figures.
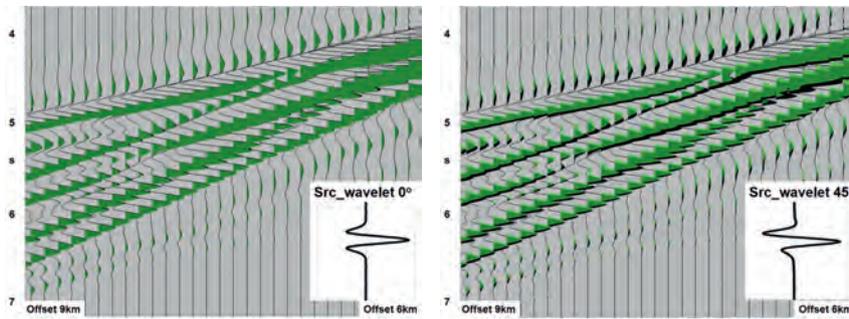
**Figure 2** Zoom of portion of a shot for field data (displayed as a black wiggle trace display) with overlay of modelled data (in green wiggle trace) and inset of the wavelet used. Left: with correct zero-phase wavelet. Right: with erroneous 45⁰ phase rotated wavelet. The waveforms should perfectly overlap, which they do for the correct zero-phase wavelet, but not for the phase-shifted wavelet.
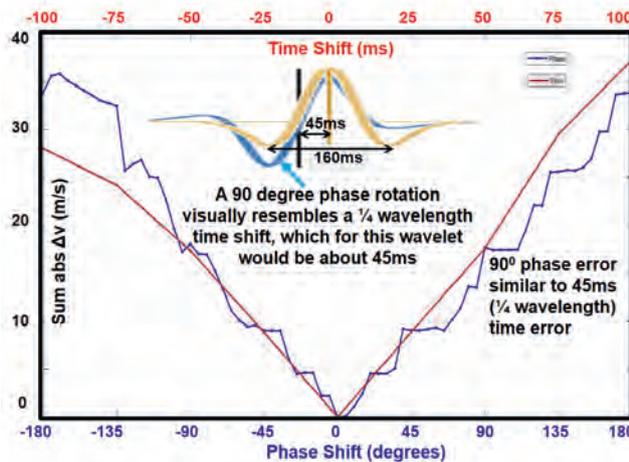


**Figure 3** Plot of velocity errors (from equation 1) versus phase shift (blue curve) and also time shift (red curve). The wavelet is shown, but not at the same scale as the time-shift axis.
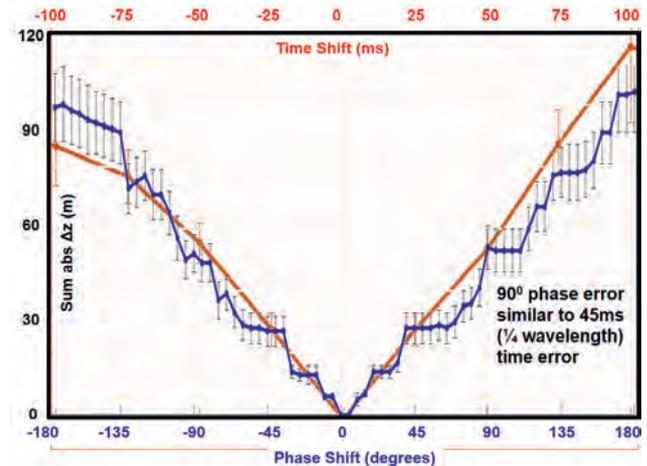


**Figure 4** Plot of depth errors associated with the phase errors (blue curve) and time shift errors (red curve). The error bars are computed for depth variations along the entire seismic line for a given phase error (depth for different CDP locations as indicated in the next figure).

iterations for each alternative source wavelet. Note that we are actually only fine-tuning an already good model, rather than starting from scratch with each new trial source wavelet.

3) Compute metrics comparing the inverted velocities obtained for the reference and trial source wavelets, by assessing the total absolute velocity differences in the region penetrated by the refracted wavefield.

4) Using each of the derived velocity models, perform 3D preSDM and compare the depths for target horizons, between the reference and each of the trial results. This depth error analysis is conducted in the vicinity of a well location, providing a 'ground-truth' reference for both depth and interval velocity.

## Results for the synthetic data

The synthetic data were taken from the non-salt portion of the well-known BP-2004 benchmark model (Billette and Brandsberg-Dahl, 2005). This is used as the reference for assessing overall absolute velocity change and residual absolute depth error, when performing FWI with a perturbed source wavelet. Here, the FWI was performed using a maximum frequency of 12 Hz. Figure 1 shows the portion of the synthetic data used in this study, with a shot record inset from location SP1, and Figure 2 displays a zoomed portion of this shot gather taken from above a 'gas' anomaly showing the 'field data' with an overlay of data modelled with correct zero-phase wavelet, and also with an overlay of data modelled with erroneous 45⁰ phase rotated wavelet. It can be noted in the figure that for the correct zero-phase

wavelet the field and modelled waveforms coincide as expected. However, for the 45⁰ phase rotated wavelet, the modelled data (plotted in green) appear to arrive too early. FWI would interpret this apparent time shift as an indication that velocity should be decreased.

Figure 3 shows the velocity errors (from equation 1) versus phase shift and also versus time shift for a suite of perturbed source wavelets, whilst Figure 4 shows the corresponding depth
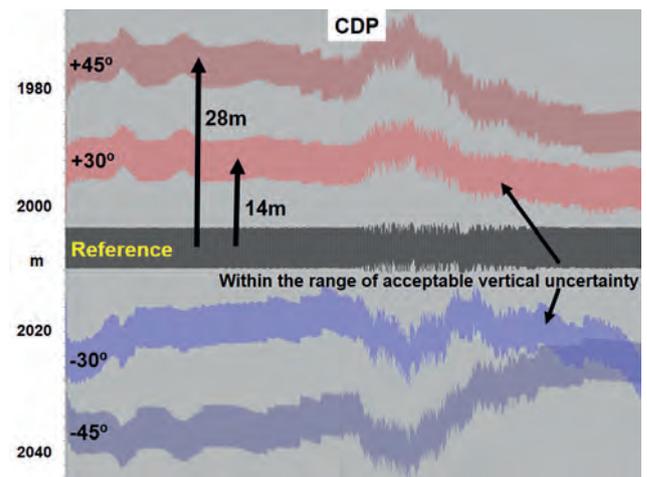


**Figure 5** Depth position of the 2 km-deep reference horizon along the seismic line for the reference, ±30⁰ and ±45⁰ phase shifted wavelets. The vertical depth error is acceptably small (within typical image depth uncertainty range) for the ±30⁰ phase shift.
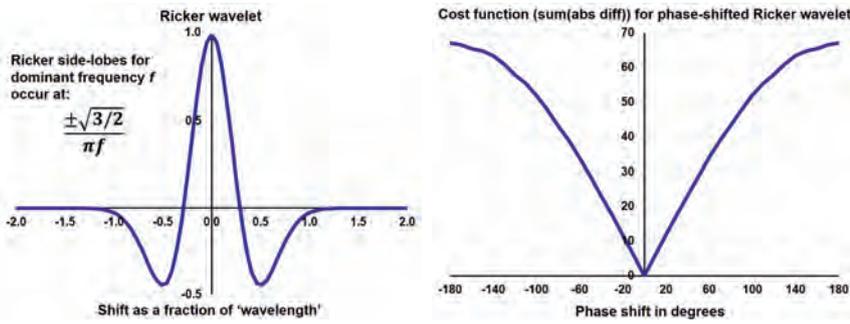
**Figure 6** Left: Ricker wavelet of centre frequency f – the side-lobe separation is √6/πf. Right: cost function as a function of phase shift (in degrees).

errors in the migrated images (from equation 2). The error bars in Figure 4 are computed for the depth scatter along the seismic line for the horizon at 2 km depth. It can be noted in the wavelet inset in this figure that a $+90^0$ phase rotation of the wavelet visually resembles a quarter wavelength time shift, which for this wavelet would be about 45 ms. Consequently, the errors for the $90^0$ phase shifts (blue curve) are similar to the errors for the 45 ms time shifts (red curve). In Figure 5, the depth variation in the migrated images for a reference horizon at 2 km depth (as summarized by the error bars in Figure 4) is shown for the images associated with the $\pm30^0$ and $\pm45^0$ phase shifted source wavelets. For this synthetic case, the depth error is acceptably small for the $\pm30^0$ phase shift (about 14 m: within typical image depth uncertainty range: Vlassopoulou et al. 2019).

## Wavelet cost function for synthetic tests

In the case of a mostly smoothly varying velocity model, as is the case with this portion of the BP_2004 data, the velocity perturbation will for the most part simply be a linearly scaled version of the gradient. And, as the gradient is made from a back-projection of the residual, and in these tests the residual is formed from a simple time or phase shift of the source wavelet, then we can expect the data cost function and the velocity cost function for the modelled data to resemble the corresponding cost functions computed for the wavelet alone.

Figure 6 shows a Ricker wavelet, and the cost function (using the sum of absolute amplitude differences) for phase shifts of this Ricker wavelet. As the results are scale invariant (i.e. the shapes are the same regardless of frequency) we have plotted the

wavelet as a function of wavelength (rather than absolute time). The forms of such cost functions resemble those derived from the BP_2004 synthetic FWI tests, as expected from the above reasoning.

## Field data example

For the field data example shown in Figure 7, comprising relatively flat-lying sediments over dipping unconformable beds, the production FWI revealed a near-seabed low velocity layer, along with other geologically conformable updates (Singh et al., 2020). The velocity change between the smoothed tomographic starting model and the refraction FWI update from the production workflow clearly shows these changes (Figure 8). The original source wavelet used in the production FWI project (derived from a deghosted near trace stack at the sea bed) was perturbed to create several other 'reasonable' source wavelets. Only the refracted wavefield is used in the inversion, the offset range being limited to 2-5.5 km for the early arrivals (Figure 9), and the inversion is performed in frequency bandwidth 2-9 Hz.

The analysis of errors was conducted for events down to the maximum depth of penetration of the diving-wave energy – about 1600 m (as we only considered refraction FWI in this study). Firstly, we analyse the errors in a 4 km x 4 km patch centred on the well position (for two representative horizons), and then for the width of the entire section for a deeper horizontal marker.

Figure 10 displays the total velocity 'error' (change) with respect to the production 'best' FWI model (blue curve), and the
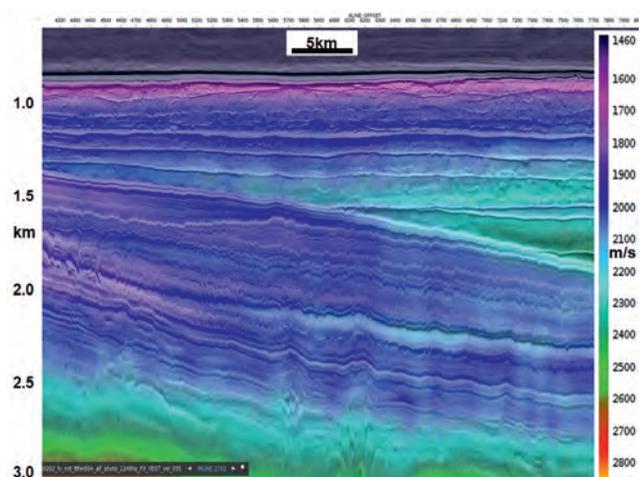


**Figure 7** Seismic image with colour overlay of final production FWI velocity model (starting from a structurally guided non-parametric tomographic inversion).
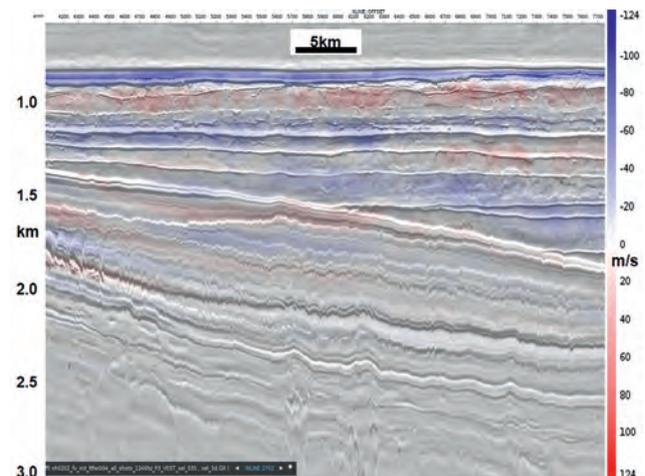


**Figure 8** Seismic image with colour overlay of FWI velocity update (starting from a structurally guided non-parametric tomographic inversion): A near-seabed low velocity layer is introduced, along with deeper updates.
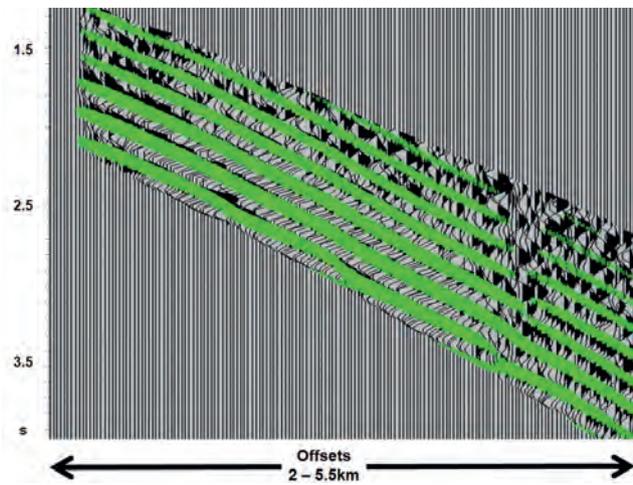
**Figure 9** Portion of shot record showing the peak of the modelled data (in green) overlying the field data (black wiggle trace) following refraction FWI using the best source wavelet estimate, with maximum frequency 9 Hz. A green modelled data peak overlying the black field data peak indicates a good match.
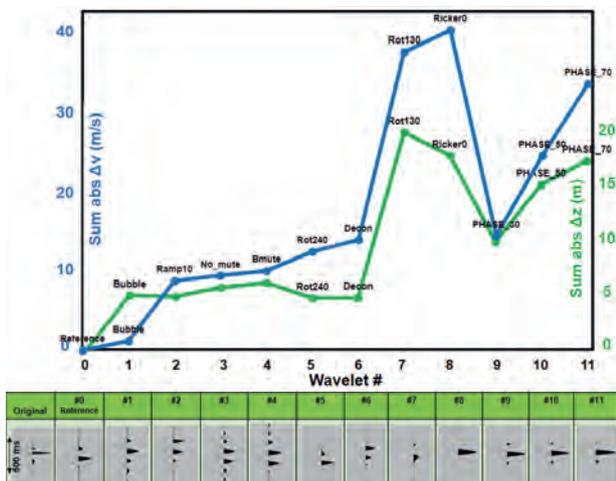


**Figure 10** Sum of absolute velocity 'error' (change) with respect to the 'best' FWI model (blue curve) and associated vertical depth error measured near the well at a depth of 1300 m (green curve). Bottom: display of the original estimated wavelet (1-2-20-30 Hz) and the various source wavelets (1-2-7-9 Hz) used in the study (note: the time-alignment of the wavelets in the display is unimportant).

total vertical depth error corresponding to the differing wavelets, measured for the reference horizon at about 1300 m depth in the vicinity of the well (green curve). The corresponding source wavelets are also shown at the bottom of the figure. The 'original' wavelet (1-2-20-30 Hz) is the source wavelet estimated for the production project (derived from the deghosted near-trace stack at the seabed). This was subsequently filtered for use in the various frequency-stages of FWI in the production workflow. The various (1-2-7-9 Hz) band limited source wavelets used in this study are then shown, with #0 being the 'reference'. Wavelets #1 - #4 are produced by tapering the reference wavelet differently; #5 is a 7 Hz Ricker wavelet with $240^0$ phase rotation: this particular phase rotation created a waveform similar to the data-derived references signature; #6 is derived using a deconvolutional procedure. The worst results (wavelets #7 and #8) are obtained using a 7 Hz Ricker wavelet with $130^0$ and $0^0$ phase rotations, respectively: these waveforms least resemble the reference wavelet. The last three results are obtained using the data-derived
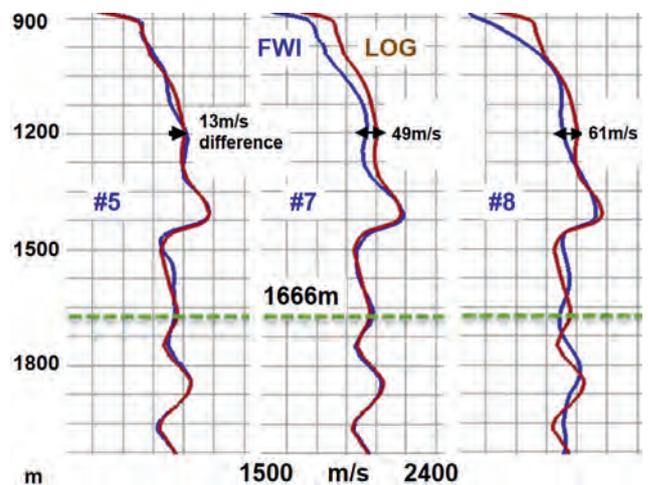


**Figure 11** Well-log (in brown) and FWI velocity profiles (in blue) for wavelet #5 (representative of reasonable errors), and from wavelets #7 and #8 (characteristic of the worst results). The depth of diving wave penetration (1666 m) is indicated, as are the velocity differences at ~1200 m depth.

reference wavelet after $30^0$, $50^0$ and $70^0$ phase rotations for comparison to the synthetic tests.

In Figure 11 are shown comparisons of the well log interval velocity values and the FWI-derived velocity profiles at the well location for three representative trial wavelets. In the depth range where the diving wave penetration was of use for updating the velocity model with refraction FWI (down to about 1666 m), the variations are all comparatively small (the velocity differences at ~1200 m depth are indicated). Below this penetration depth we see no significant model changes, as the data input to the FWI were muted to exclude deeper events (as described earlier).

## Discussion

Perhaps the most surprising observation of this study in the real dataset is the degree to which FWI is 'tolerant' of changes in the source wavelet, at least in the case of a good starting model and limited bandwidth (2-9 Hz). Wavelet changes that appear to be substantial to the eye can result in 'little' difference to the migrated image.

Analysis of the synthetic dataset (which has a smooth background velocity) showed that for small source wavelet phase errors (~<35°) and small source wavelet time errors (~<16 ms), the depth errors associated with the error in the inverted velocity will be acceptable. Expressed in terms of the source wavelet's wavelength, these errors relate to a fractional wavelength error of about 1/10 of wavelength. In addition, the resulting depth errors are less than the typical depth uncertainties expected in migrated image position (e.g. Osypov et al., 2013). This observation is made for FWI with a maximum inverted frequency of 12 Hz.

Results from the real dataset (where the geology consists of relatively flat layering with thermal vents and sill intrusions present in the area) indicated that if the *main peak* of the source wavelet is in *alignment* with the underlying reference wavelet, then in general the subsequent migration depth errors are well within the acceptable image position-uncertainty. Note that the depth error metrics (characterizing deviation from the reference model results) are derived from analysis conducted on relatively flat horizons;

consequently, if the geology was more steeply dipping, the image positioning error would probably have been greater.

The implication of this observation is that for the bandwidth being used (in our case 2-9 Hz), if the modelled data are not cycle skipped with respect to the field data, then the FWI procedure will be tolerant of significant source wavelet-shape error, as it is unlikely to get stuck in a local minimum. Naturally, if we increased the bandwidth to incorporate higher frequencies, then the sensitivity to wavelet error would be more pronounced, but may probably still exhibit similar tolerances when expressed in terms of fractions of the wavelet's wavelength.

Although the velocity field itself might differ in a way that looks significant, with lots of small changes of magnitude of about 30 m/s, the resulting depth change in the associated migrated image is small enough to be considered acceptable (± 15m). This is possibly because small velocity changes, with opposite signs, tend to cancel each other out in terms of the differences they make to the travel times *along ray paths*. If the geology was more steeply dipping, perhaps the image positioning error would have been greater.

## Conclusions

To a certain extent, the velocity and associated image depth perturbations resulting from slight wavelet change can be considered as acceptable, as compared to the depth errors associated with tomographic uncertainty. It is only when we exceed perhaps a $30^0$ source phase error or a very poor source wavelet estimate, that we suffer unacceptable changes in the result. In other words, if the location of the wavelet peak energy in the modelling is similar to that of the field data (not cycle-skipped), then FWI is relatively robust with respect to source error. However, for more complex geological environments with rapid lateral velocity variation and steeply dipping events (e.g. salt), or for FWI run to higher frequencies or with poor starting models, the analysis presented here would need to be repeated in order to draw representative conclusions.

## Acknowledgements

## References

Billette, F.J. and Brandsberg-Dahl, S. [2005]. The 2004 BP Velocity Benchmark. *EAGE 67th Conference & Exhibition*. B035.

Jones, I.F. [2018]. *Velocities, Imaging and Waveform Inversion (The evolution of characterising the Earth's subsurface)*. EAGE, Houten, The Netherlands.

Lailly, P. [1983]. The seismic inverse problem as a sequence of before stack migrations. *Conference on inverse scattering: Theory and application, SIAM*, 206–220.

Osypov, K., Yang, Y., Fournier, A., Ivanova, N., Bachrach, R., Yarman, C.E., You, Y., Nichols, D., Woodward, M. [2013]. Model-uncertainty quantification in seismic tomography: method and Applications. *Geophysical Prospecting,* **61**, 1114–1134.

Pavlopoulou, P. [2019]. *Sensitivity to Source Wavelet Shape in Time Domain Full Waveform Inversion: Applications in 2D synthetic data and field dataset from offshore Norway*. MSc thesis, University of Leeds.

Pratt, R.G. [1999]. Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model. *Geophysics*. **64** (3), 888–901.

Rickett, J. [2013]. The variable projection method for waveform inversion with an unknown source function. *Geophysical Prospecting*, **61** (4), 874–881.

Singh, J., Braidwood, L., Valler, V., Michot, O., Bekkeheien, R., Wang, C., Jones, I.F. [2020]. Waveform inversion methodology for deep structural imaging in offshore Norway. *EAGE 82nd Conference & Exhibition,* Extended Abstracts.

Sun, D., Jiao, K., Cheng, X., Vigh, D. and Coates, R. [2014]. Source wavelet estimation in full waveform inversion *SEG Technical Program Expanded Abstracts 2014*, Expanded Abstracts, 1184–1188.

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

Virieux, J. and Operto, S. [2009]. An overview of full-waveform inversion in exploration geophysics. *Geophysics*, **74** (6), WCC1–WCC26.

Vlassopoulou, A., Felicio, R., Hagen, C., Jones, I., Ackers, M. and Schjelderup, S. [2019]. Bayesian uncertainty estimation in the presence of tomographic model error. *Second EAGE/PESGB Workshop on Velocities*, Extended Abstracts.

Wang, C., Yingst, D., Farmer, P., Jones, I., Martin, G. and Leveille, J. [2017]. Reconstructed full-waveform inversion with the extended source. *SEG Technical Program Expanded Abstracts*, Expanded Abstracts, 1449–1453.

Warner, M. and Guasch, L. [2016]. Adaptive waveform inversion: Theory. *Geophysics*, **81** (6), R429–R445.

**NAME COMPANY**
**CC?????-MA???**