# Adaptive Waveform Inversion – FWI Without Cycle Skipping

*Michael Warner, Imperial College London; Llu ś Guasch, Gang Yao*, Sub Salt Solutions Limited*

**Summary**

We present a newly proposed method for performing full-waveform inversion that appears to be immune to the effects of cycle skipping – Adaptive Waveform Inversion (AWI). The method uses Wiener filters to match observed and predicted data. The inversion is formulated so that the model is updated in the direction that drives these Wiener filters towards delta functions at zero lag, at which point the true model has been recovered. The method is computationally efficient, it appears to be universally applicable, and it recovers the correct model when conventional FWI fails entirely.

**Introduction**

Full-waveform inversion (FWI) is a technique for building highly accurate models of physical properties in the subsurface, and especially for building high-resolution p-wave velocity models. FWI has had some spectacular successes in improving the quality of subsequent depth migrations (Warner *et al.*, 2013), and in providing directly interpretable images of physical properties in both the overburden and at the reservoir. Despite its growing success, FWI suffers from a fundamental problem of cycle skipping because seismic data are oscillatory and FWI is a local minimization scheme.

In this paper, we present a newly proposed method for performing waveform inversion that appears to be entirely immune to the effects of such cycle skipping. Since the method works by adapting one dataset to another using Wiener filters, we refer to it as Adaptive Wavefield Inversion or AWI. So far as we have been able to discover, AWI seems to have no disadvantages that are not also suffered by conventional FWI, it requires a similar total compute effort and similar computer codes, and it has some additional advantages beyond its ability to avoid the effects of cycle skipping.

Here we explain the method and the rationale behind it, and present three synthetic data examples.

**Methodology**

Conventional FWI seeks to minimize an objective function of the least-squares difference between observed and predicted datasets given by

$$f = \frac{1}{2}\|\mathbf{p}(\mathbf{m}) - \mathbf{d}\|^2 \qquad (1)$$

where the column vector $\mathbf{d}$ represents the field data, and $\mathbf{p}$ represents the predicted data by a model $\mathbf{m}$. Because seismic data are oscillatory, the objective function typically has many local minima. In this case, local inversion methods will fail to converge to the global minimum unless the starting model is close to the true model.

The alternative approach that we develop here, employs Wiener filters rather than direct subtraction as the means to compare the two datasets. To do this, we perform the inversion in two stages. In the first stage, we design a Wiener filter $\mathbf{w}$ that matches a observed trace $\mathbf{d}$ to an predicted trace $\mathbf{p}$ in a least squares sense, that is $\mathbf{d} * \mathbf{w} \approx \mathbf{p}$. Thus, the Wiener filter coefficients $\mathbf{w}$ can be obtained by minimizing the least-squares objective function $f_1$, where

$$f_1 = \frac{1}{2}\|\mathbf{D}\mathbf{w} - \mathbf{P}\|^2 \qquad (2)$$

Here $\mathbf{D}$ is the convolutional Toeplitz matrix that contains the vector $\mathbf{d}$ in each column. Equation (2) is applied trace-by-trace. It is linear; it has only one minimum, and provides a unique solution for $\mathbf{w}$.

Now, if the filter $\mathbf{w}$, determined by equation (2), was simply a delta function at zero lag, then the predicted and observed data would be identical (apart from a scalar). Thus, in the second stage, we minimise (or maximise) a second normalized least-squares objective function $f_2$, given by

$$f_2 = \frac{1}{2}\frac{\|\mathbf{T}\mathbf{w}\|^2}{\|\mathbf{w}\|^2} \qquad (3)$$

where T is a simple diagonal matrix that acts to weight the filter coefficients as a function of the magnitude of the temporal lag. If the weights increase with the magnitude of the lag, then $f_2$ should be minimized, and if the weights decrease with lag, then $f_2$ should be maximized. The most important feature of $f_2$ is that it has only one minimal or maximal point. Consequently it is immune to cycle skipping (Warner and Guasch, 2014).

With this formulation, the adjoint source $\delta\mathbf{s}$ for one source-receiver pair is given by

$$\delta\mathbf{s} = \frac{\partial f_2}{\partial \mathbf{p}} = \mathbf{D}\,(\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}\left(\frac{\mathbf{T}^2 - 2f_2\mathbf{I}}{\mathbf{w}^{\mathrm{T}}\mathbf{w}}\right)\mathbf{w}. \qquad (4)$$

Reading this expression from right to left, in order to compute the adjoint source for one receiver, we must first find the Wiener filter $\mathbf{w}$, normalize it by its inner product $\mathbf{w}^{\mathrm{T}}\mathbf{w}$, weight the coefficients using a function of temporal lag $(\mathbf{T}^2 - 2f\mathbf{I})$, deconvolve this using the auto-correlation $(\mathbf{D}^{\mathrm{T}}\mathbf{D})$ of the observed data, and finally convolve the result

with the original data **D**. This operation is performed trace-by-trace, and it is computationally efficient.

**Application to the Marmousi model**

To demonstrate the method, we apply it to the well-studied Marmousi model (Figure 1(a)). Data from this model are easy to invert using most FWI schemes provided that the starting model is a smoothed version of the true model, that the inversion begins at very low frequencies, and that the data are noise free. To provide a realistic test therefore, we here use a simple one-dimensional starting model (Figure 1(b)) that provides only a poor match to the true model, we run the inversion using data that have a 10 Hz Ricker wavelet, we run the inversion using the full data bandwidth without beginning at lower frequencies, and we add noise to the synthesized observed data.

For the demonstration, we apply trace-by-trace amplitude normalization to both the observed and predicted data, and we spatially precondition the gradient by dividing it locally by the energy in the incident wavefield averaged over all times and all sources. Beyond that, we use no enhancements or additional features.

The experiment includes free-surface multiples, primary reflections and wide-angle turning arrivals in the data to a maximum offset of 7 km. At 10 Hz, the majority of the data predicted by the starting model are cycle skipped; at later travel times this mismatch can be more than one cycle.

Figure 1(c) shows that conventional FWI is unable to recover the correct model at all under these circumstances. Cycle skipping dominates the data, and the recovered model is a poor match to the true model. The background model is not recovered correctly, velocities are adjusted in the wrong direction within the shallow high-velocity fault blocks, and the deeper structure cannot be properly focused by the starting model.

In contrast, Figure 1(d) shows the results of applying AWI to these data using an identical inversion scheme. AWI is evidently not affected by the cycle skipping, and it iterates successfully to recover the true model. Its accuracy is affected only by the finite bandwidth, finite aperture, acquisition geometry, and finite noise levels of the incident wavefield. When AWI runs, deep features are initially misplaced in depth, but as the inversion proceeds to improve the model, these features move smoothly and continuously towards their correct depths. AWI does not follow a convoluted path from starting to final model.

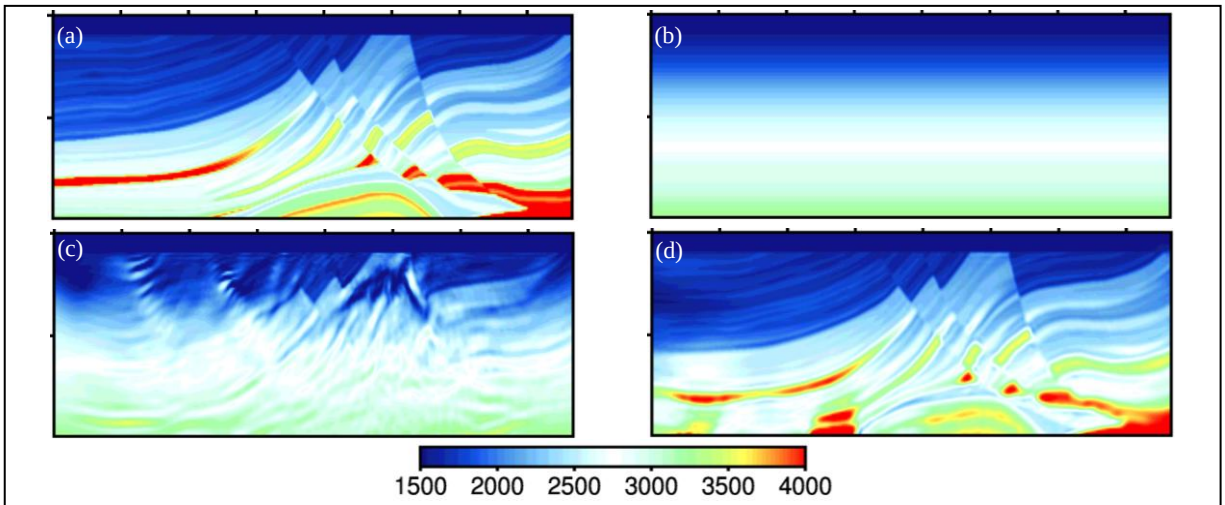


Figure 1: (a) true model; (b) starting model; (c) FWI result; (d) AWI result.

**Incorrect Source**

According to our study on AWI, we found AWI has an extraordinary feature: it is very resilient to source errors. To demonstrate this, two low-frequency source wavelets determined from an OBC field dataset using two different methods are applied to the Marmousi model. The bandwidth of both sources is similar, and their nominal zero times and polarities are the same, but their waveforms are different. We have used source A (Figure 2(a)) to generate synthetic data, and then attempted to invert those data using both the source A and B (Figure 2(b)). In the test, both FWI and AWI were carried out under these

conditions. The forward and inverse modeling were identical, and there was no noise in the data.

In an effort to make these tests as realistic as possible, we limit the temporal bandwidth to 4–12 Hz, and the maximum source-receiver offset to 7 km. The inversion does not honour absolute amplitudes, and we retain the various heuristics and stabilizations within the code that we would use to invert field data. We use a simple approximate diagonal Hessian to precondition the gradient.

In the test, an accurate starting model (Figure 2(c)) is chosen to ensure that the results from FWI are not cycle skipped. If the correct source wavelet shown in Figure 2(a), is used for both FWI and AWI, the results are not identical, but are extremely similar; they differ most at the edges and base of the model where there is less-than-ideal data coverage. This implies that FWI and AWI differ minimally under these ideal circumstances. However, FWI and AWI with the wrong source wavelet shown in Figure 2(b), give total different results in Figures 2(d) and 2(e). FWI failed to recover the model as expected, however AWI unexpectedly managed to recover the model. This appears to be an additional benefit provided by AWI, which can recover a high-quality final result despite the large errors in the assumed source wavelet.

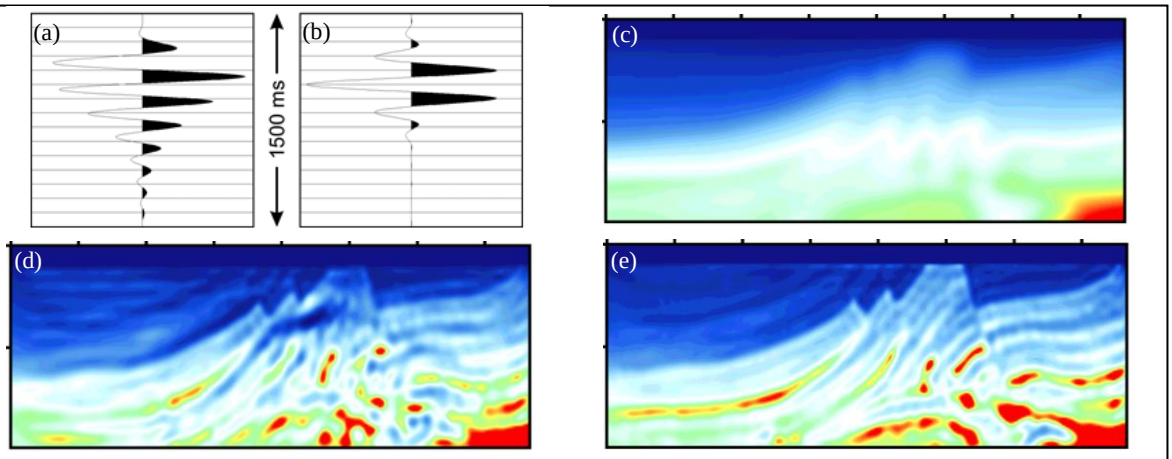


Figure 2: (a) and (b) are two source wavelets determined from an OBS dataset using different approaches. (c) Starting model; (d) FWI result by inverting the data of the source wavelet shown in (a) with the source wavelet shown in (b) ; (e) the counterpart result of (d) by using AWI. This demonstrates AWI is very resilient to the source wavelet errors. The color scale is the same as that in Figure 1.

**Chevron model**

In 2014, Chevron designed a highly realistic seismic model, based upon field seismic data, well logs and rock physics. Using this model, they generated fully visco-elastic synthetic data, simulating a conventional 2D towed-streamer marine acquisition geometry with a maximum offset of 8000 m. The model has a non-deterministic, realistic relationship between velocity and density, it is locally isotropic, it has limited bandwidth, ambient noise, surface ghosts and multiples, and is modelled using the full visco-elastic two-way wave equation. Chevron released this dataset as a benchmark to test FWI and related technology. Chevron has not released the true model. In addition to the synthetic data, the source signature, a sonic profile in one well, a smooth 1D starting p-wave velocity model (Figure 3(a)), the velocity of the water column and the geometry of the seabed, were released.

Figure 3(b) shows the result of applying a conventional FWI algorithm to the refracted portion of these data, starting at 3 Hz, and increasing the bandwidth by stages to 10 Hz. The algorithm and parameterisation used here are identical to those which we would normally use to invert analogous field data. The FWI result shows three features. Firstly, it is cycle skipped at both ends of the model, producing the non-physical structures seen at the ends of the profile. Secondly, it produces useful and sensible velocity updates over the top 2 km of the model everywhere except where the results are affected by cycle skipping. Thirdly, conventional FWI produces only minor updates below the region that is penetrated by the refractions.

Figure 3(c) shows the final result of applying AWI, from 3 to 10 Hz, successively in three stages, followed by conventional FWI from 10 to 27 Hz. Both reflections and refractions were included throughout; and multiples, ghosts and other non-primary phases were included. For the first pass of AWI, the model was smoothed mildly from the surface to 2 km depth, and smoothed aggressively below 2 km. Following this, the total accumulated model update was increased by 50% before beginning a second phase of AWI. Scaling of the total update in this way acts to

accelerate convergence in those portions of the model where convergence is slow, and can also act to increase spatial resolution where convergence is rapid.

During phase two, the smoothing was less severe below 2 km. Rescaling of the update accumulated during phase two was again applied, but only below a depth of 2 km. During a third and final phase of AWI, only minimal smoothing regularisation was applied at all depths, and the accumulated update was not scaled at the end of this phase. Following AWI, a final phase of conventional FWI was applied to the full bandwidth of the data. Although AWI overcomes cycle skipping and deals well with reflections, FWI ultimately has greater absolute accuracy when the starting model is everywhere close to the true model. Typically then, we will obtain the best outcome by applying AWI early and FWI late in the inversion.

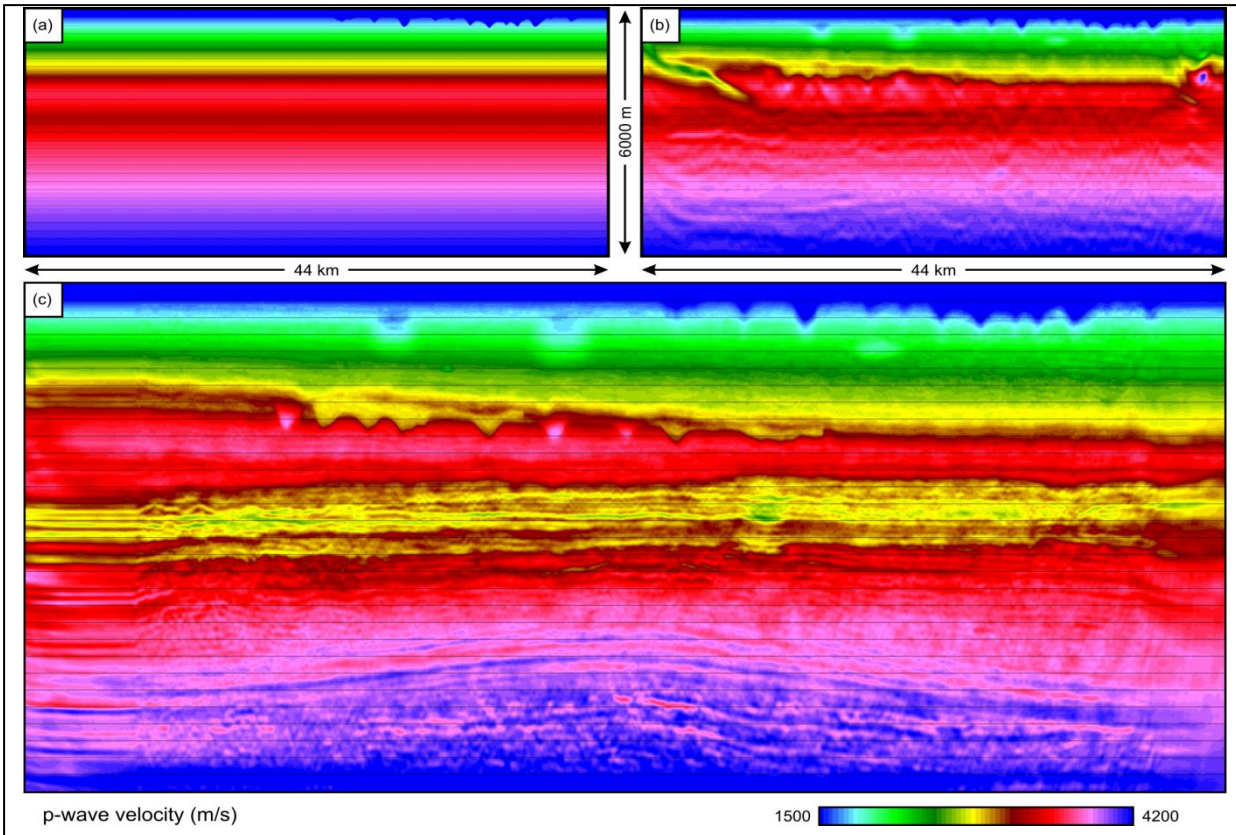


Figure 3: The velocity model. Each panel measures 6 × 44 km. (a) 1D starting model showing seabed topography. (b) After conventional FWI using refractions only. (c) Final model recovered by AWI using reflections and refractions. Unlike FWI, the AWI result is not cycle skipped, and there has been a significant long-wavelength velocity update below 2 km.

## Conclusions

AWI appears to be capable of recovering the correct global solution in circumstances where conventional FWI fails entirely. The method works by adopting some of the characteristics of WEMVA while retaining the essential elements of an FWI scheme, and by doing so it becomes effective while remaining computationally efficient.

## Acknowledgements

We thank Chevron provides the synthetic dataset and Sub Salt Solutions Ltd for permission to present this paper. The methodology described here is the subject of GB patent application number GB1319095.4.

## References

Warner & Guasch [2014] Adaptive waveform inversion – FWI without cycle skipping: Theory. EAGE Extended Abstracts, Amsterdam, 2014.