Adaptive Waveform Inversion: Theory

Mike Warner, Imperial College London, and Lluís Guasch, Sub Salt Solutions Limited*

Summary

We present a new 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 operates by attempting to find an earth model that minimizes the difference between a predicted and an observed pre-stack dataset, or less-commonly that maximizes the zero-lag of the temporal cross-correlation of the two datasets. FWI has had some spectacular successes in improving the quality of subsequent depth migrations, 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. This occurs when the inversion produces a match between predicted and observed data that is shifted in time by one or more wave cycles. Since FWI is a local minimization scheme, these cycle-skipped models represent local minima in the objective function into which the inversion can easily fall. Practical prevention of cycle skipping involves the use of low frequencies in the field data, starting the inversion from an accurate initial model following intensive model building using reflection travel-time tomography, and applying rigorous quality control during FWI.

In this paper, we present a new 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 a simple synthetic demonstration. In companion papers, we examine the performance of AWI on reflection-dominated data, and apply it to a 3D field dataset.

Method

Conventional FWI compares observed and predicted data by subtracting one from the other to generate a residual dataset, and then minimizes the sum of the squares of these residuals. Because seismic data are oscillatory, this approach necessarily leads to an objective function that has many secondary local minima. AWI uses a different means to compare two datasets. It operates trace-by-trace, and it uses a least-squares convolutional filter, a Wiener filter, to match the predicted to the observed data. If this filter is acausal and significantly longer than the trace length, then the match that it provides will be extremely close. As we will see, it is the closeness of this matching that removes cycle skipping from the AWI formulation.

AWI begins by generating a dataset of predicted traces. For each of these traces, we design a separate Wiener filter that converts a predicted trace into its corresponding observed trace. The coefficients that form these filters are necessarily functions of the current model and of the field data. Now, if a particular model provides a perfect match to the field data, then each of these Wiener filters will necessarily consist of a delta function of unit amplitude at zero lag. We can therefore set up an objective function that measures the deviation of each these filters from a zero-lag delta function, and use a conventional adjoint formulation to find the earth model that minimizes this functional.

The mathematics of the method are relatively straight-forward, and it leads to formulations that are similar to conventional FWI. Its essential features are these: For each source in the field data, a forward wavefield is calculated. The predicted and observed data are matched at each receiver, and the resultant Wiener filter is used to generate an adjoint source which is back-propagated into the model. The forward and backward wavefields are cross correlated at every point in the subsurface, and the zero spatial and temporal lag of this cross-correlation is stacked over all sources to form the gradient of the objective function. The only significant difference in this formulation between AWI and FWI is in the formation of the adjoint source which is outlined below for AWI.

There are various objective functions that can be minimized or maximized in AWI. It is possible to use filters that

DOI http://dx.doi.org/10.1190/segam2014-0371.1

match the predicted to the observed data, or that do the reverse. In the analysis below, we consider the reverse scheme since it leads to simpler final expressions; the forward scheme though is closely analogous. The simplest useable form of AWI employs the objective function:

$$f = \frac{1}{2} \frac{||\mathbf{T}\mathbf{w}||^2}{||\mathbf{w}||^2} \tag{1}$$

where $\mathbf{w}$ is an acausal Wiener filter that transforms a single observed trace into its equivalent predicted trace, $\mathbf{T}$ weights the coefficients of this filter by the absolute value of their temporal lag, and the summation implied by the L_2 norm is over all temporal lags and all traces.

This objective function is designed to penalize Wiener coefficients that lie far from zero-lag, and so it should be minimized. This form for the objective function does not seek to match absolute amplitudes between pairs of traces; other forms for the objective function can be devised that achieve this, and more sophisticated forms of weighting and normalization are possible. The normalization term in the denominator is important; without it, the functional will tend to act to minimize all the coefficients which is equivalent to minimizing the predicted data by suppressing all possible arrivals, a result that will not lead towards the true earth model.

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

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

where $\mathbf{D}$ is a matrix that represents convolution by the observed data, and $\mathbf{p}$ is the predicted data. 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 $\mathbf{D}$. This operation is performed trace-by-trace, and it is computationally efficient.

Immunity to cycle skipping

Analyzing the adjoint source in equation (2), suggests the reason why the AWI formulation is immune to the effects of cycle skipping. Without the weighting, amplitude normalization and deconvolution, this expression is simply a Wiener filter convolved with an observed trace, and so it closely matches the predicted data. The adjoint and the predicted data are consequently always in phase, and there can be no cycle skipping. The deconvolution and the normalization of the adjoint do not change its phase, and the weighting simply adjusts the adjoint according to how well the original predicted and observed data match.

Although it is a data-domain method, AWI has important features in common with image-domain approaches such as wave-equation migration velocity analysis (WEMVA); it is these features that make it immune to cycle skipping.

WEMVA methods typically extend the model in some non-physical way, setting up the inversion so that models evolve towards physical outcomes. In conventional WEMVA, the model is extended by introducing sub-surface offset. This represents a non-physical scattering in the subsurface, whereby an incident wavefield at one location generates a coeval scattered wavefield at another. The inversion is then formulated to focus the energy to zero sub-surface offset, thus producing a physical outcome in which the incident and scattered wavefields are coincident.

AWI also extends the model in an analogous non-physical way. The Wiener filter can be regarded as a means of redistributing energy, non-physically, in time. In this case, energy arriving at a receiver at a particular time produces a signal at that receiver that is distributed across earlier and later times. These non-physical arrivals disappear when the filter becomes a zero-lag delta function, corresponding to a physical outcome. WEMVA involves non-physical action at a distance in the sub-surface, whereas AWI involves non-physical interaction across time at the receivers.

This insight also reveals why AWI is immune to cycle skipping. The data $\mathbf{p}$ predicted by a physical model $\mathbf{m}$ can be cycle skipped with respect to $\mathbf{d}$. Considering the forward formulation of AWI, the data predicted by the extended model ($\mathbf{m}$ and $\mathbf{w}$ combined) is instead the convolution of $\mathbf{w}$ with $\mathbf{p}$ which is not cycle skipped since it is always a close match to the observed data $\mathbf{d}$. We can therefore regard AWI as an analogue of WEMVA that seeks to focus energy – all energy and not only primary reflections – at zero temporal lag just as WEMVA seeks to focus primary reflections at zero sub-surface offset.

A different view of the way that AWI operates can be obtained by recalling that the wave equation represents a linear relationship between the source and the wavefield. Thus any observed seismic trace, generated by a point source, can always be predicted exactly by any non-pathological subsurface model simply by adjusting the effective source that is used to model it. The required source may be acausal, and it may be of infinite duration.

The Wiener filter used in AWI then simply provides the best least-squares estimate of a finite-duration version of this source. The effective sources required for every trace in the observed dataset will normally be different. AWI is

DOI http://dx.doi.org/10.1190/segam2014-0371.1

configured so that the inversion drives all these sources towards the same source – the actual source that was used to acquire the physical data. AWI is then not cycle skipped because the effective sources are always adjusted by the Wiener filters so that the observed data are well matched by the predicted data; there is no phase difference between observed and predicted, and so no cycle skipping occurs.

Application to a synthetic model

To demonstrate the method, we apply it to the well-studied Marmousi model, Figure 1. 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 that provides only a poor match to the true model, we run the inversion using data that have a dominant frequency of 10 Hz (inset in Figure 2), we run the inversion using the full data bandwidth without beginning at lower frequencies, and we add noise to the initial data.

For the demonstration, we use a vanilla version of steepest-descent FWI with no model regularization. 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.

Figure 1(a) shows the true model and Figure 1(b) shows the starting model. We include 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 (Figure 2); 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, and this property contributes to its computational efficiency.

Provided that there are refracted arrivals at the target depth, the final outcome of AWI is affected not at all by the quality of the starting model. In a companion paper we show that the method also has advantages for reflection FWI. Although the final recovered model is not affected by the starting model, the number of iterations required to reach that final model does increase as the distance that the method has to travel from its starting point increases. It will always therefore be computationally advantageous to start as close to the true model as possible.

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.

Unlike most WEMVA-based schemes, adaptive FWI does not require computationally expensive non-zero-lag convolutions in the subsurface, it does not require the solution of a computationally expensive extended wave equation, nor does it require primary-only input data. AWI can operate with refracted energy, with primary reflections, and with multiples. Unlike other formulations for overcoming cycle skipping, AWI does not degrade in complex models that involve many interfering events.

Wiener filters are well understood, and the simple scheme that we have outlined here can, in practice, be extended in obvious and productive ways to take advantage of that understanding. The filters can for example be varied with time, be regularized across receivers, or be multi-dimensional. It is possible to match predicted to observed, observed to predicted, or both datasets to each other.

So far as we have been able to tell, AWI has no dis-advantages that are not also features of conventional FWI. It has a range of advantages over conventional FWI of which the avoidance of the detrimental effects of cycle skipping is the most significant.

Acknowledgements

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

DOI http://dx.doi.org/10.1190/segam2014-0371.1

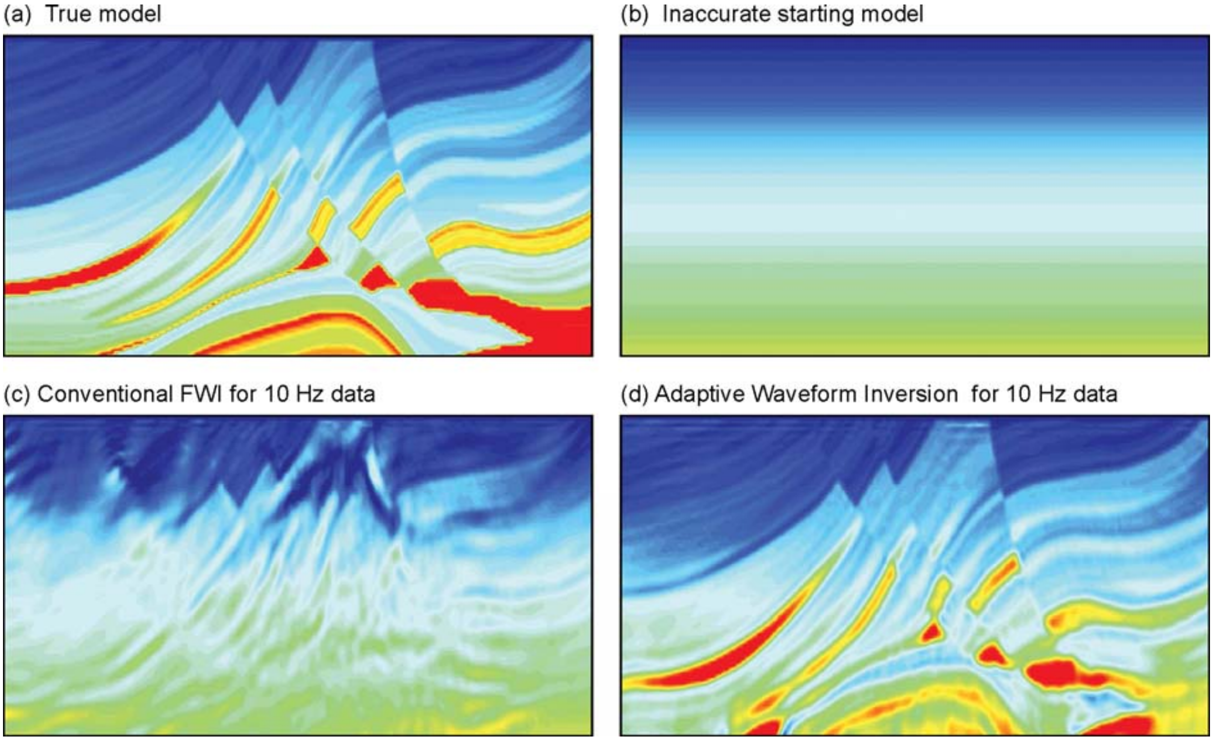

Figure 1: Application of Full-Waveform Inversion and Adaptive Waveform Inversion to the 2D Marmousi model for a dataset with a dominant frequency of 10 Hz, a poor one-dimensional starting model, and free-surface multiples, ghosts, and random noise in the input data. At 10 Hz, the data generated by this starting model are badly affected by cycle skipping so that conventional FWI fails entirely while AWI continues to be accurate and effective.

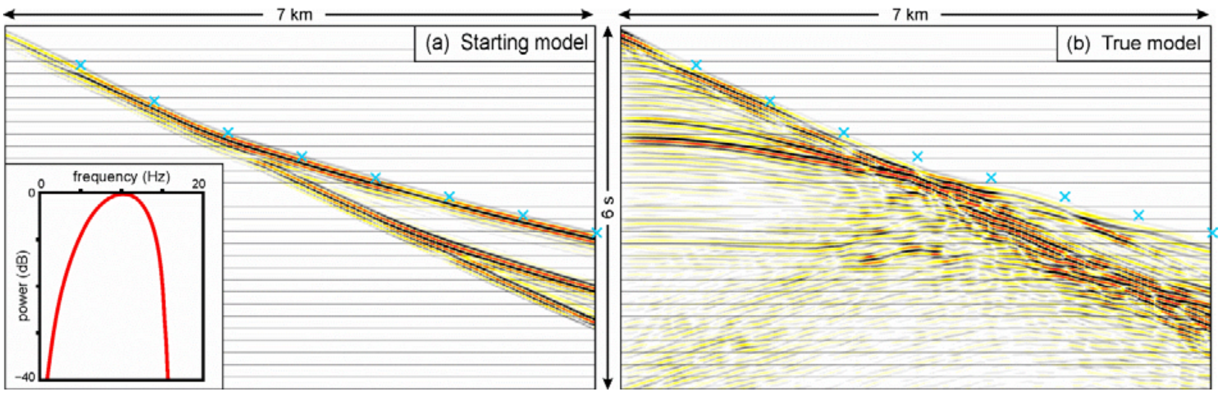

Figure 2: A synthetic shot record generated in (a) the starting and (b) the true model. The crosses on both records are the same. The starting data are cycle skipped with respect to the true mode. The inset shows the power spectrum of the true data in (b).

DOI http://dx.doi.org/10.1190/segam2014-0371.1

http://dx.doi.org/10.1190/segam2014-0371.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

No references