Tu P1 02

Convergence Regions for AWI and FWI

J. Yao* (Imperial College London), L. Guasch (S-Cube, London), M. Warner (Imperial College London)

Summary

Cycle-skipping is the most significant local minimum FWI suffers in practice, while adaptive waveform inversion (AWI) provides a new waveform-inversion scheme which is robust against cycle-skipping. In this paper, we present an extensive test exploring the convergence properties of both FWI and AWI against cycle-skipping. A set of 1300 initial models are designed by progressively smoothing the Marmousi model and by bulk shifting its mean slowness. The convergence regions of FWI and AWI are mapped based on the recovered models of both approaches. AWI shows a convergence region broader than FWI. It succeeds refining the initial models to the global minimum which FWI cannot.

Introduction

Full-waveform inversion (FWI) is a local optimisation algorithm that is driven by the difference between a given set of data points (the observed data) and an equivalent dataset numerically generated by solving a wave equation in a known, discretised model (the predicted data). These differences are then used to modify the model in order to reduce the data misfit. There are numerous factors that contribute to the presence of local minima in FWI, but perhaps the most significant in practice is that caused by data cycle-skipping – this occurs when observed and predicted data differ by more than half a cycle.

Warner and Guasch (2016) proposed Adaptive Waveform Inversion (AWI) as a new waveform-inversion scheme that is insensitive to cycle-skipping; it defines the functional to be minimised as a lag-weighted Wiener filter. The inversion strategy is then designed to focus the energy of such filters at zero lag in order to create a delta-function-shaped Wiener filter. In this paper, we explore the practical convergence limits of both conventional FWI and AWI systematically for the widely-used Marmousi model, generating a large suite of starting models by applying various degrees of smoothing and by bulk shifting the mean slowness of the true model.

Theory

Conventional FWI measures the mismatch between the two datasets using the L2 norm squared of the data differences by defining the following objective function:

$$\text{FWI data error} = f_{FWI} = \|\mathbf{p} - \mathbf{d}\|_2^2 \tag{1}$$

where $\mathbf{p}$ is the predicted data and $\mathbf{d}$ is the observed data. Since seismic data are oscillatory, FWI is likely to become trapped within a local minimum when the phase mismatch between $\mathbf{p}$ and $\mathbf{d}$ is over half a cycle. In contrast, rather than minimising the differences between the two datasets directly, AWI seeks to minimise the non-zero lag components of the Wiener filter that matches $\mathbf{d}$ into $\mathbf{p}$, thus

$$\text{AWI data error} = f_{AWI} = \frac{\|\mathbf{Tw}\|_2^2}{\|\mathbf{w}\|_2^2} \tag{2}$$

where $\mathbf{w}$ denotes the Wiener filter obtained from the least-squares solution of $\mathbf{d} * \mathbf{w} = \mathbf{p}$, and $\mathbf{T}$ is a diagonal matrix that monotonically weights the non-zero lag components of $\mathbf{w}$. Due to the monotonic character of this penalty function $\mathbf{T}$, the AWI objective function is insensitive to cycle-skipping unlike its FWI counterpart. Despite this insensitivity, AWI may still fail for extreme models since there are other causes of local minima in waveform inversion unrelated to cycle skipping.

Experiment Design

To introduce cycle-skipping progressively in the predicted data, we designed a set of 1300 initial velocity models by progressively smoothing the Marmousi model and by bulk shifting its mean slowness. The models are 9.6 km wide and 2.5 km deep, and have a free surface at the top. Synthetic data were generated using 91 sources and 187 receivers at a depth of 25 m. A schematic map of the distribution of the starting models is shown top left in Figure 1, where the horizontal axis indicates the velocity decrease or increase of the true model as a percentage perturbation ranging from -15% to $+15\%$, and the vertical axis indicates the radius of the gaussian-smoothing operator applied which ranges from 0 to 1200 m. The true model is in the centre of the bottom axis on this diagram.

This suite of models was then inverted with both FWI and AWI. The observed data were generateded using a Ricker wavelet with a peak frequency at 9Hz; the total number of iterations was 300 for all runs, and all iterations are run at full bandwidth. From the results of the inversions we generated volumes showing model and data mismatch in order to study the practical convergence region of each method. The L1 model error was computed using:

$$\text{model error} = \left\| \frac{(\mathbf{m}_{true} - \mathbf{m_i})}{\mathbf{m}_{true}} \right\|_1 \tag{3}$$
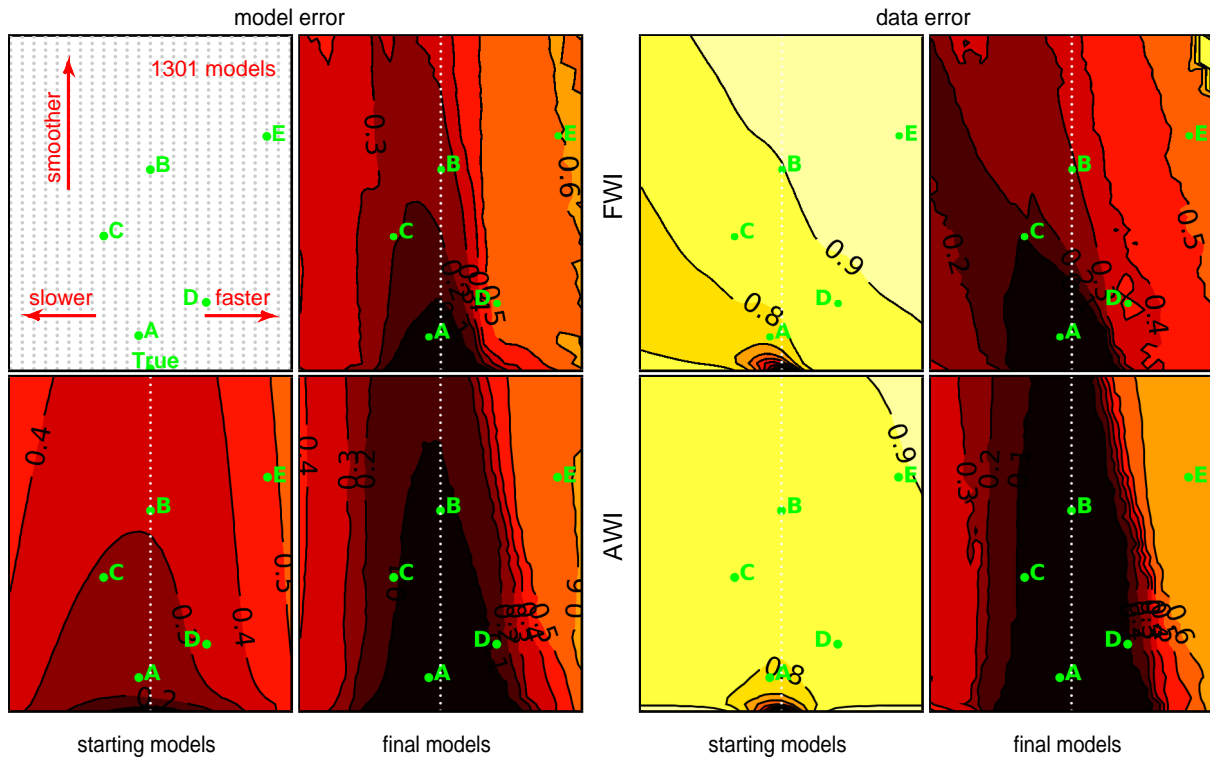
Figure 1: *2D plots of model and data errors. Top left figure illustrates the starting model distribution. Left panels show the model error for the starting and final models (the starting model error is independent of the inversion algorithm). Right panels show the data error, using the FWI or AWI objective function, for the starting and final models. Top row is FWI; bottom row is AWI. Plots are normalised with respect to the maximum model or data error respectively. The positions of the models A to E displayed in Figure 2 are indicated in green.*

Convergence Limits

The contour maps in Figure 1 show how both the model and data errors evolve for the two inversion schemes. The data-errors show a consistent decrease in values for all models both in FWI and AWI, as would be expected since the algorithms are designed to minimise this quantity. The model error, on the other hand, has a clear region of improvement in an area surrounding the true model, but there is also a region where the initial model error increases for less-accurate starting models. The model errors define a region of starting models that can be inverted successfully. The dark-brown areas in the model error (values around 0.1 and below) define the set of starting models that converge towards the global minimum. The subset of the models shown in Figure 2, discussed below, illustrate and demonstrate this choice of threshold value for the convergence region.

For AWI, the region of success covers a wider range of models confirming its robustness against cycle-skipping and its capacity to converge from less-accurate starting models. In addition, the shape of the model and data error surfaces is more consistent for AWI which suggests that, at least within the region of convergence, the functional is a reliable indicator of success.

For the FWI model error, the values are not only higher, that is the fit to the model is less good, but it is also biased to the left of the figure indicating a preference for low-velocity starting models. This is not consistent with the FWI objective-function behaviour (top right image in Figure 1): although both model and data errors show a predilection for slow models, the effect is much more pronounced in the data error to a degree where there does not exist a coherent overlap between the two, suggesting that the FWI objective function is less reliable as an indicator of final model quality than is its AWI counterpart. AWI also has larger model and data errors for fast starting models, but the region of success does not exhibit any significant asymmetry with respect to the sign of the initial bulk-velocity error – it is approximately symmetric around the white dotted line in Figure 1.
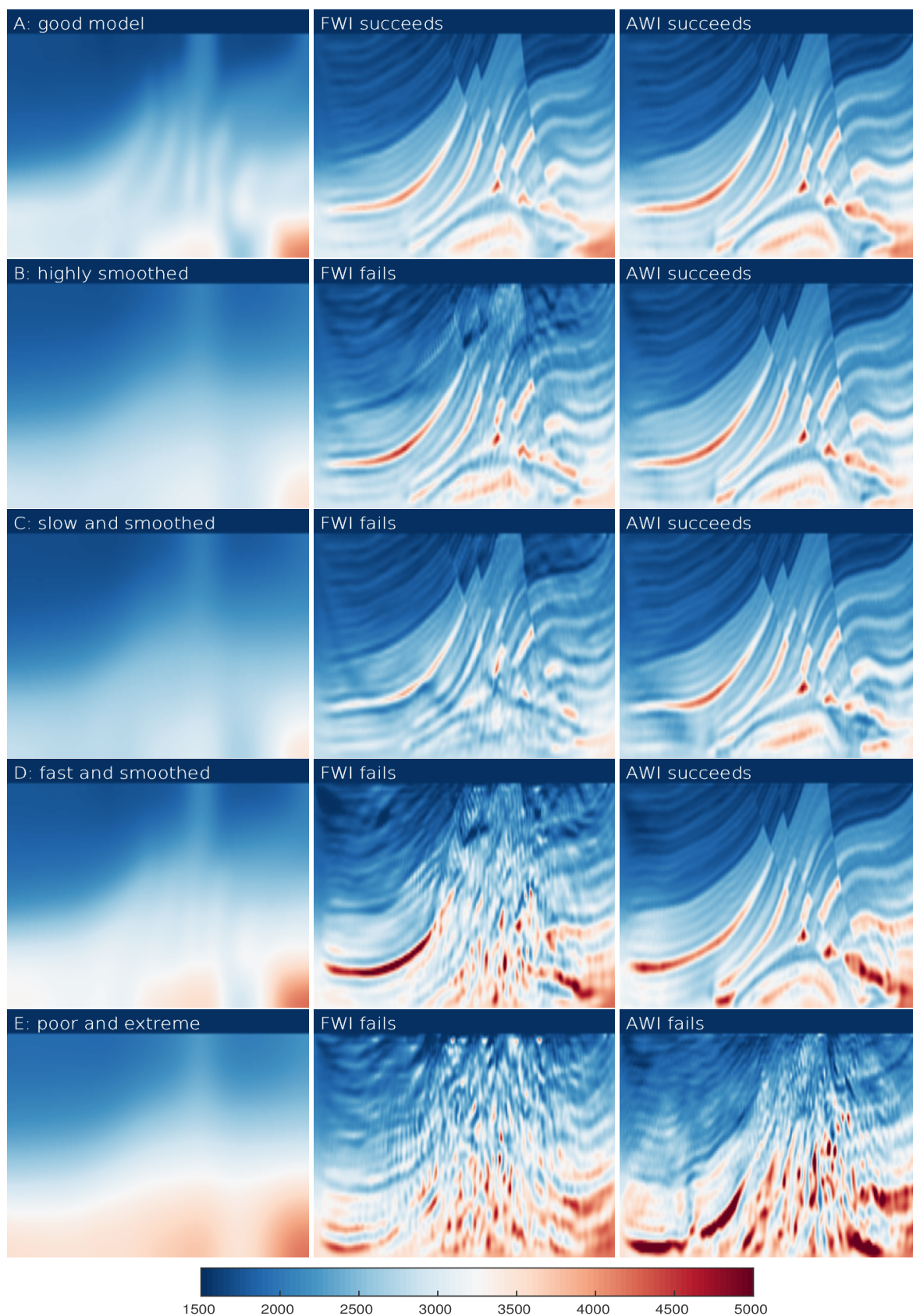
Figure 2: *Each row shows starting and recovered model for one of the five models A to E indicated in Figure 1. Left: starting models. Centre: FWI-recovered models. Right: AWI-recovered models.*

Selected models

We have selected a group of five models (models A to E in Figures 1 and 2) to validate our interpretation of the model-error evolution and its identification with the region of adequate starting models. Model A has a small negative velocity shift and has been modestly smoothed; it is sufficiently close to the true model for both FWI and AWI to converge successfully as shown in the first row of images in Figure 2. Many synthetic FWI studies employ modestly smoothed starting models of this type that have no significant bulk error in the start model. It is not clear that the behaviour of FWI on field data is well represented by the behaviour of FWI on this category of starting model.

Models B to D are outside and inside the area of successful convergence for FWI and AWI respectively. The second to fourth rows in Figure 2 confirm that, while the FWI results show artifacts characteristic of misconvergence, their AWI analogues evolve towards the correct solution without problems. For FWI, it is interesting to note that a large amount of smoothing with no velocity shift has a detrimental effect in the shallow part of the final result (model B), while a less-smoothed starting model with incorrectly low velocities (model C) is unable to recover the deeper parts of the model. When the starting model is slightly smoothed and shifted towards positive velocities (model D), FWI cannot reconstruct any part of it, and its final result is extremely poor. This is in good agreement with the model error from Figure 1.

In the most extreme case, where the model has been heavily smoothed and its velocities shifted by almost 15% (model E), neither FWI or AWI are able to avoid falling in a local minimum. For AWI, this local minimum is not produced by conventional cycle skipping – rather the starting model is sufficiently poor that both algorithms will tend to align the wrong events. In simple models, AWI is able to invert successfully from almost arbitrarily incorrect starting points, but in realistically complicated models that contain many arrivals, neither AWI nor FWI is able easily to proceed when the starting model is sufficiently poor that events of one type in the starting model become misidentified with events of another type in the true model. Both FWI and AWI have some immunity to this effect since other parts of the data volume have the potential to improve the model such that the misidentification becomes resolved as the inversion proceeds. However, in complex models that are seriously in error, the potential for this resolution is severely curtailed; model E appears to occupy such a position.

Discussion

Both FWI and AWI algorithms require starting models that are inside the basin of attraction of their respective global minima in order to converge. However, AWI has a solution space with a broader global minimum, allowing larger errors in initial model estimates. Detailed examination of this suite of synthetic models, and our experience with a range of field datasets, suggests that AWI starting models will normally converge successfully when the initial velocity errors are about twice to three times those that can be endured by conventional FWI. In practice this often means that AWI will succeed when beginning from a simple smoothed pre-stack time-migration stacking-velocity model obtained without using tomography, whereas FWI often cannot proceed successfully from such a model. Equivalently, for the same starting model, AWI can begin using frequencies that are two to three times higher than those required by FWI. In practice, this means that AWI will often be successful on legacy towed-streamer data that does not contain clean low frequencies, whereas FWI will fail on the same dataset unless significant effort is expended in improving the starting model and/or customising the input data.

Acknowledgements

References

Warner & Guasch (2016) Adaptive waveform inversion: Theory. Geophysics, 81, R429-R445.