A practical approach to handle land data: full-waveform inversion case study

Introduction

Full-waveform inversion (FWI) can be formed as a data fitting optimization problem, in which we minimize the difference between the observed field data and the modeled data by solving partial differential equations (Pratt et al., 1998; Li et al., 2016). The evaluation of the difference between observed data and modeled data is the so-called cost function or objective function according to the optimization theory. While we bring the objective function down to the desired tolerance, we believe the iterated velocity model can represent the actual subsurface geology.

A successful application of FWI technology to the real land dataset can be very challenging because of the following reasons. The first one is the low signal-to-noise ratio—especially at the low-frequency part—of the real data, which requires us to start with a good initial model to avoid the "cycle skipping" phenomenon (Plessix et al., 2001). The second reason is that an accurate source wavelet of land data is usually very hard to obtain. The conventional least-squares solution could lead to a very inaccurate wavelet due to the high noise level as well as a lot of components that can not be modeled in the data. An inaccurate wavelet can cause an unstable optimization process (Pratt et al., 1998). The third reason is that land data usually has a lot of components that can not be modeled accurately, such as surface waves and elastic waves, and so on (Li et al., 2013). Therefore, to invert a land dataset, we have to come up with a more reasonable way to invert for components that can be simulated with our modeling method first.

In this abstract, we propose a new workflow to solve the issues listed in the previous paragraph. Firstly, we calibrate our initial source wavelet with a matching filter, which is calculated from the phase and amplitude errors between the modeling data and the real data at near offset. And we repeat this process every time when we change the frequency band. Secondly, we utilize the first break information to drive FWI. We also add strong attenuation at the top of the velocity model to mitigate waveforms that travel through unrealistic areas above the ground surface. After we extract a good initial model from the first break, we run a full FWI with all the actual data. Our final result shows this workflow can help land data FWI converge better than just directly running a conventional FWI will all the actual data.

Theory

Full-waveform inversion (FWI) involves the solution of an optimization problem. There are many advanced ways to form the corresponding objective function of the optimization problem that can mitigate the cycle-skipping phenomenon (Warner and Guasch, 2016; Biondi and Almomin, 2013). But in this project, we use the most popular objective function, which is the $\ell_2$ norm measurement of the distance between observed data and modeled data, as shown in equation 1.

$$\min_{m} \frac{1}{2n_s} \sum_{i=1}^{n_s} ||b_i - \mathcal{F}[m, q_i]||^2_2,$$

(1)

where $b_i$ is $i$th shot, while $\mathcal{F}[m, q_i] = P_r A^{-1}m q_i$ is the modeling of velocity model $m$ and the $i$th source signature $q_i$. $P_r$ is the detection operator, restricting the data to the receiver positions. $A$ is the discrete wave equation operator, while $n_s$ is the total shot number.

There are many numerical optimization methods to solve equation 1, such as steepest descent, nonlinear conjugate gradient, quasi-newton, and so on. Theoretically, different optimization methods can only affect the convergence in terms of total iteration, but not necessarily help to avoid the local minima. According to our experience, a quadratic method such as the Gauss-Newton method does not speed up the inversion in terms of computational cost, evaluated by the total number of wave-equation solved (Plessix et al., 2001). So in this project, we use the first-order gradient-based method to drive the inversion.

Obtaining the source wavelet

One of the most important ingredients of a successful application of FWI to real data is an accurate source wavelet. Theoretically, we can estimate the source wavelet while doing inversion by solving
a deconvolution problem (Pratt et al., 1998). But in practice, this process can cause instability of the inversion due to the low signal-to-noise ratio, especially for the land data. In our approach, we firstly extract an initial wavelet (blue in Fig.1b) for the input data itself as shown in Fig. 1a, followed by a quality control (QC) modeling of a small selected subset of shots with the initial model (Fig. 2a). Then, we compute the cross-correlation of the modeled data and observed data at carefully selected near offset and early time as shown in Fig. 2b. We can observe the time lag in Fig. 2b with the extracted wavelet from the data. Based on the phase and time shift error in Fig. 2b, we can adjust our wavelet accordingly. Red wavelet in Fig. 1b shows the adjusted wavelet. The modeled data with adjusted wavelet and the observed data match with each other very well as shown in Fig. 2c and Fig. 2d.

Figure 1: (a) One raw 3D land shot gather (three receiver lines). (b) Wavelet (blue is initial wavelet; red is adjusted wavelet).

Figure 2: (a) Comparison between real data and modeled data with initial wavelet. (b) Cross-correlation of real data and modeled data in Fig. 2a. (c) Comparison between real data and modeled data with adjusted wavelet. (d) Cross-correlation of real data and modeled data in Fig. 2c with adjusted wavelet;

Real 3D land data example

The 3D dataset used in this project (Fig. 1a) was acquired in Kedong area, Xingjiang province of China. Shot line and receiver line distances are both 300 meters. The shot interval and receiver interval of each line are 60 meters and 30 meters respectively. The maximum offset is 9000 meters approximately.

The FWI for this project is carried out in two steps. The first step is first-break FWI up to 7Hz, followed by an optional reflection tomography step to build up the velocity below 3km. For every tomography run, we preserve the near surface model result from the first break FWI down to 3km and let tomography update the deeper section of the model. The second step is FWI with all the data up to 9Hz.

First break FWI with Q

Inverting a land dataset is very challenging, because of the complexity of the data. As shown in Fig. 1a, the raw shot has ground roll and guided wave components that are very hard to simulate accurately with widely used wave equation modeling methods (Plessix et al., 2010; Wang et al., 2020). Guided wave is refraction energy that travels multiple times between ground surface and reflectors. To play safe, one of the most practical approaches to run FWI on land data is to start inversion with first break information as shown in Fig. 3, which can be obtained freely from the static correction processing step.

The topography is an inevitable issue for land data. To simulate shots with accurate source and receiver locations, we use a regular velocity grid and replace the property above the ground surface with a filling velocity. In that way, we can use the sources and receivers at their right location spatially. However, we
also observe a lot of spurious events that are caused by the filled velocity as shown in Fig. 3a. To solve this issue, we introduce a strong attenuation in the filled velocity area above the ground surface. As we can observe in Fig. 3b, the spurious events are suppressed greatly, providing better data fitting during inversion.

The first break FWI was carried out with an initial velocity (Fig. 5a) that is converted from time migration velocity (RMS-interval), followed by a reflection tomography. We run FWI up to 7Hz with a wavelet depicted in Fig. 1b. Fig. 5b is the inverted result from the first break. We can see a significant change in the shallow part above 2km, where the refraction waves can reach. To better illustrate the convergence of this FWI step, Fig. 4 shows the data fitting before and after FWI. Fig. 4a is the overlap between real data (grayscale) and synthetic data (blue-red scale) with the initial model. There is a lot of mismatch between white and blue areas which are supposed to be lined up. In Fig. 4b, we can observe that the real data (white area in the background) match with the synthetic data (blue area in the foreground) after inversion, which indicates that our inversion converges towards the right direction.

![Figure 3: (a) Synthetic shot record without Q. (b) Synthetic shot record with Q added above the ground surface.](image)

![Figure 4: (a) Data comparison with initial velocity model. (b) Data comparison after first-break FWI](image)

**FWI with the full real data**

To further improve our FWI result, we continued our FWI process from the velocity model obtained with the first break FWI in Fig. 5b. The second round of FWI uses all the data and goes up to 9hz. Fig. 5c is the detailed final FWI result with all the data. To verify our final FWI velocity, we compute the cross-correlation between real data and modeled data before and after FWI as depicted in Fig. 6. The red line in Fig. 6 illustrate the offset of the shot record, the max offset of which is about 9km. The inconsistency of the cross-correlation phase along offset in Fig. 6a gets noticeably improved in Fig. 6b. The final depth imaging with initial model and FWI result in Fig. 7 also proves the convergence of our FWI. Our FWI process brings much more details in Fig. 7b compared to Fig. 7a. The perfect match between dip structure in the final RTM image and well-log dips information (Fig. 7c) also further confirmed our FWI workflow.

![Figure 5: (a) Initial velocity obtained from time-to-depth conversion. (b) Velocity after first break FWI. (c) Final FWI velocity model.](image)
Conclusions
In this abstract, we proposed a practical workflow to invert land dataset. According to our findings, starting with first-break can notably improve the convergence behavior of the whole inversion process. The introduction of a strong attenuation on top of the velocity model can help reduce the spurious events caused by the topography issue, yielding better data fitting.

Acknowledgements
We deeply thank BGP for the permission to publish this data. We also would like to thank Pengyuan Sun, Xudong Zhang without whose help none of this paper could be possible.

References