Regularizing velocity differences in time-lapse FWI using gradient mismatch information

Bram Willemsen, M.I.T and Alison Malcolm, Memorial University of Newfoundland

SUMMARY

We present a method for recovering time-lapse velocity changes using full waveform inversion (FWI). In a preprocessing step we invert for a single intermediate model by simultaneously minimizing the data misfit in the baseline and the monitor surveys. We record the individual FWI gradients corresponding to the baseline and the monitor datasets at each iteration of the inversion. Regions where these gradients consistently have opposing sign are likely to correspond to locations of time-lapse change. This insight is used to generate a spatially varying confidence map for time-lapse change. In a subsequent joint inversion we invert for baseline and monitor models while regularizing the difference between the models with this spatially varying confidence map. Unlike double difference full waveform inversion (DDFWI) we do not require identical source and receiver positions in the baseline and monitor surveys.

INTRODUCTION

FWI (Lailly, 1983; Tarantola, 1984; Pratt, 1999) uses the entire seismic record to invert for subsurface material properties such as seismic velocity and density. When monitoring a reservoir, the primary objective is to recover changes in the material properties due to production changes. A straight-forward application of FWI to this time-lapse problem is to invert for the baseline (i.e., initial) state and the monitor (i.e., new) state independently. Subtracting these models intuitively approximates the time-lapse change in material parameters. However, artifacts are also present due to the nonlinear nature and ill-posedness of the uncoupled inverse problems for the baseline and monitor models.

In an attempt to overcome some of these problems Watanabe et al. (2004) and Denli et al. (2009) introduced DDFWI for monitoring production related changes. In DDFWI the baseline data residual from an initial baseline inversion is subtracted from the monitor dataset. Conceptually this can be seen as removing a part of the data that the initial model could not explain. After the monitor model is inverted from this modified data and the baseline model is subtracted a better time-lapse estimate is obtained. The data subtraction in DDFWI imposes constraints on the seismic acquisition. Both the source and receiver locations have to be the same in the baseline and monitor surveys. Differences are not uncommon in marine surveys as is illustrated for instance by Eggenberger et al. (2014). These differences will result in artifacts in the velocity change estimate.

Several methods have been introduced that avoid the requirement of identical acquisition. Zamanian et al. (2014) invert for the time-lapse change by casting the inverse problem in a hierarchical Bayesian framework. Maharramov et al. (2014)

circumvent some of the problems of subtracting independent baseline and monitor inversions by swapping baseline and monitor datasets a fixed number of times. In a different approach Maharramov and Biondi (2014) introduce a joint inversion in which Total Variation (TV) regularization suppresses the oscillatory model-difference artifacts that arise in joint FWI with noisy datasets. Another approach is to iteratively swap baseline and monitor datasets in FWI and record which regions change consistently as a result. This information can be used to generate a confidence map, which can regularize the model-difference in a subsequent time-lapse inversion. This was done by Yang (2014) in a method called Alternating FWI (AFWI). The idea we introduce in this study is similar to Yang's approach. Instead of performing a potentially long sequence of expensive full waveform inversions for the confidence map, we perform a single joint inversion with a single model. In this joint inversion we minimize both the baseline and monitor data residuals with this single intermediate model. Regions where the gradient of the baseline data term consistently has a different sign than the gradient of the monitor data term are considered to be potential regions of time-lapse change. We quantify this principle and construct a confidence map from the gradient history. Similar to the work of Yang (2014) we use this (different) confidence map to regularize the model difference in a final joint inversion.

METHOD

To obtain a confidence map we first minimize the objective function

$$\widetilde{\chi}(m(x)) = \frac{1}{2}\left|\mathscr{F}_0(m(x)) - d_0\right|^2_{d_0} + \frac{1}{2}\left|\mathscr{F}_1(m(x)) - d_1\right|^2_{d_1}, \quad (1)$$

where the misfit is $\widetilde{\chi}$ and a single model m is used to fit both the baseline data d_0 and the monitor data d_1 through forward model $\mathscr{F}_0(m(x))$ and $\mathscr{F}_1(m(x))$ respectively. The baseline and monitor acquisition are not required to be the same. Both the data norms $|\ |_{d_0}$ and $|\ |_{d_1}$ are implemented as the L_2 norm. Model regularization can be applied to m but this is not done in the examples in this study. The gradient of (1) with respect to the model consists of two parts:

$$\frac{\delta\widetilde{\chi}}{\delta m(x)} = g_0(x) + g_1(x), \quad (2)$$

where $g_0(x)$ and $g_1(x)$ are the standard FWI gradients of the baseline and the monitor data terms. The concepts we introduce in this research are applied to constant density acoustics, but the ideas extend naturally to different physics. Equation 1 can be minimized with any iterative optimization routine, but in this research we use L-BFGS (Nocedal and Wright, 2006). When inverting with the objective function given in (1), we expect to get an intermediate model $m(x)$ that explains both the baseline data, d_0, and monitor data, d_1, as well as possible. At locations x where there is a real time lapse model

Regularizing velocity differences in time-lapse FWI using gradient mismatch information

change we expect that the FWI gradient of the baseline data term, $g_0(x)$, will have the opposite sign to the FWI gradient of the monitor term, $g_1(x)$. Intuitively this means that in regions of true time-lapse change an inversion using just the baseline data would move away from intermediate model m in the opposite direction to that an inversion using just the monitor data would move in. This insight can be used when constructing a confidence map. When inverting for the intermediate model $m(x)$ with an iterative inversion scheme we can keep the gradient history of both the data terms. We now define a confidence map $\beta_g(x)$ based on the gradient history,

$$\beta_g(x) = \sum_i^n \left[\left| sgn(g_{0,i}(x)) - sgn(g_{1,i}(x)) \right| \right.$$
$$\left. \times \left(\frac{1}{2} \left(|g_{0,i}(x)| + |g_{1,i}(x)| \right) \right) \right], \qquad (3)$$

where $g_{0,i}(x)$ and $g_{1,i}(x)$ are the gradients of the baseline and monitor terms at iteration i respectively. $|f(x)|$ is the absolute value of f at each location x and not a norm. To reduce the effects of differences in illumination we precondition the gradients using the inverse diagonal of the pseudo-Hessian proposed by Shin et al. (2001). The confidence map β_g in (3) has high values when the gradient of the baseline and the monitor term consistently have opposing sign and large amplitude. The reason for using the entire gradient history instead of the last iteration is to suppress some of the randomness and the nonlinear effects that take place when minimizing (1).

The confidence map $\beta_g(x)$ is used in a subsequent joint inversion where the following objective function χ is minimized for the baseline model m_0 and the monitor model m_1,

$$\chi(m_0(x), m_1(x)) = \frac{1}{2} |\mathscr{F}_0(m_0(x)) - d_0|_{d_0}^2$$
$$+ \frac{1}{2} |\mathscr{F}_1(m_1(x)) - d_1|_{d_1}^2 + \frac{1}{2}\varepsilon \left| \frac{m_0(x) - m_1(x)}{\beta_g(x)} \right|_m^2. \qquad (4)$$

The data norms $|\ |_{d_0}$ and $|\ |_{d_1}$ and the model norm $|\ |_m$ are implemented as the L_2 norm. The scalar ε weights the model difference regularization term. Depending on the choice of ε, the model regularization term can be used to strongly suppress model differences outside the regions with high beta. Additional regularization can be applied to the individual models, but this is not investigated in this study. Note that there is nothing in this formulation that requires the baseline and monitor acquisitions to be the same. In the remainder of the study we refer to the process of first minimizing (1) and recording the gradients in (2) for confidence map (3) and then solving (4) as Gradient Confidence FWI (GC-FWI).

NUMERICAL INVESTIGATION

The performance of GC-FWI is compared with DDFWI on a synthetic model under noisy conditions. We compare using unrealistically favorable conditions for DDFWI. The source and receiver locations are exactly the same in both surveys and so is the source wavelet. These tests thus give us a conservative estimate of the expected improvements from GC-FWI.
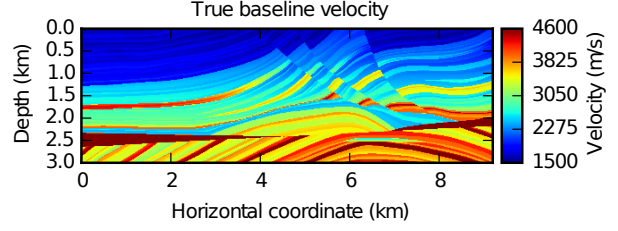

Figure 1: The true baseline model.

Setup of synthetic problem

The model on which the synthetic study is done is the Marmousi model, see Figure 1. The model is discretized using 151 nodes in the z direction and 461 nodes in the x direction. The node spacing is 20 m in both directions. The baseline and monitor surveys have 19 identical shot locations equally distributed over the surface and 461 identical receiver locations. The source wavelet is a 6 Hz ricker wavelet and is exactly the same in the baseline and monitor surveys. The same time-domain solver is used for generating the data and for solving the inverse problem. The model includes a Perfectly Matched Layer (PML) on all sides. No model regularization is used on the individual baseline and monitor models in GC-FWI or in DDFWI. Uncorrelated Gaussian noise is added to the data. We define the noise level r as the energy of the noise relative to the energy of the noiseless signal in the entire seismic survey,

$$r = \frac{\sum_i^{n_s} \sum_j^{n_r} \sum_k^{n_t} n_{i,j,k}^2}{\sum_i^{n_s} \sum_j^{n_r} \sum_k^{n_t} s_{i,j,k}^2}. \qquad (5)$$

In (5), n_s is the number of sources, n_r is the number of receivers, and n_t is the number of time samples in a trace. For a specific noise level we first invert for the baseline model starting from a heavily smoothed initial guess. Figure 2 shows that good reconstruction is achieved for $r = 0.64$. When comparing DDFWI and GC-FWI at a specific noise level r, the corresponding inverted baseline model is used in DDFWI and in GC-FWI as the initial guess for m in (1). Figure 3 shows the true time-lapse velocity change used in this study.
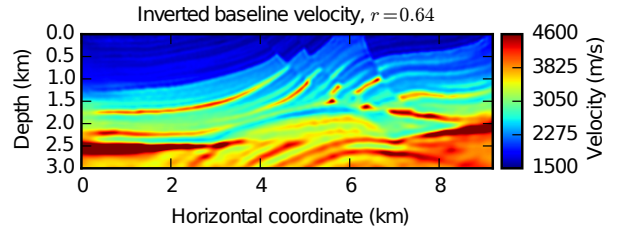

Figure 2: The baseline model inverted from the baseline data using standard FWI for $r = 0.64$.

Calculating the confidence map $\beta_g(x)$

We perform 20 iterations to minimize (1) in a single joint inversion and record the gradients $g_0(x)$ and $g_1(x)$ for noise levels
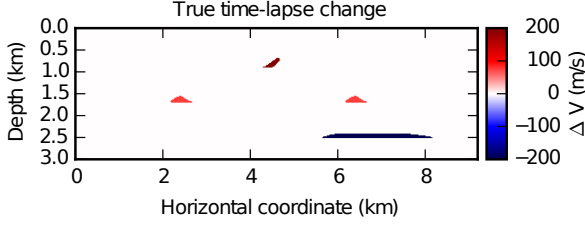
Figure 3: The true time-lapse velocity change investigated in this study.

Figure 5: Evolution of the preconditioned gradients of both data terms of (1) in the inversion, at the locations indicated by arrows of corresponding color in Figure 4. True time-lapse regions correspond with gradients of different signs (black) whereas in regions without timelapse changes the gradients have the same sign (red).

$r = 0.04$ and $r = 0.64$. The gradients are preconditioned (Shin et al., 2001) with the purpose of correcting for differences in illumination. The preconditioned gradient history is then used to obtain the confidence map $\beta_g(x)$ using (3). We normalize the confidence maps by dividing each by its maximum value. The results are compared in Figure 4.

At the locations marked with arrows in Figure 4, we plot the evolution of the preconditioned gradients using matching colors in Figure 5. We see that the preconditioned gradients $g_0(x)$ and $g_1(x)$ have opposite signs and large amplitude at the location marked by to the black arrow. The baseline data term and the monitor data term try to move the intermediate model m, computed by minimizing (1), in opposite directions. This is a sign that the baseline and monitor velocities may differ and thus a true time-lapse change is likely at this location. Equation 3 for the confidence map therefore assigns a high value of $\beta_g(x)$ to the location of the black arrow. At the location of the red arrow the gradients of both data terms are often small and have the same sign. The low amplitude and the agreement in sign makes it unlikely that a true time-lapse change is present at this location, because both data sets agree that the model fits the data well, and if they try to adjust it at all do so in the same direction. In the preconditioned gradient evolution plot of Figure 5 we observe some transient behavior, which is a

reason for involving more than one iteration when computing the confidence map $\beta_g(x)$ in (3).

Comparing GC-FWI and DDFWI

With the confidence map $\beta_g(x)$ we regularize the model difference in the joint inversion (4). We perform 15 nonlinear iterations and minimize the objective function using L-BFGS. We compare the results of GC-FWI at noise level $r = 0.64$ with 20 nonlinear iterations of DDFWI starting from the initial inverted baseline model corresponding to that noise level. Since the weighing of a regularization term is not trivial we show the results for three different values of parameter ε, which weights the model regularization in (4).

We display the results of the comparison between GC-FWI and DDFWI in Figure 6. We see that the level at which GC-FWI suppresses noise in the time-lapse estimate depends on the model regularization weight ε. At the low end of the investigated values (i.e., $\varepsilon = 10.0$) this noise is only slightly less than in DDFWI. At higher values of ε the noise in the recovered GC-FWI time-lapse estimates goes down. At the high end of the investigated values of ε (i.e., $\varepsilon = 1000.0$) we observe the strongest noise suppression, but we also notice that the recovered time-lapse change on the rightmost perturbation is suppressed below the true value of Figure 3.

DISCUSSION

In this study we use identical wavelets and identical source and receiver locations in the baseline and monitor surveys. In reality these conditions are hard to satisfy and deviations will inherently result in artifacts in DDFWI as is documented by for instance Yang (2014). Instead of subtracting data, GC-FWI works in the model space by comparing the stacked gradients of the baseline and monitor survey evaluated at an intermediate model. It is true that differences in source and receiver locations in baseline and monitor surveys will change the gradient contributions of the shots, but the stack of these contributions
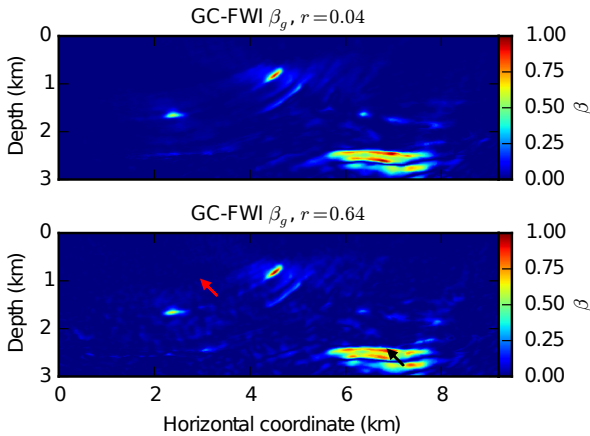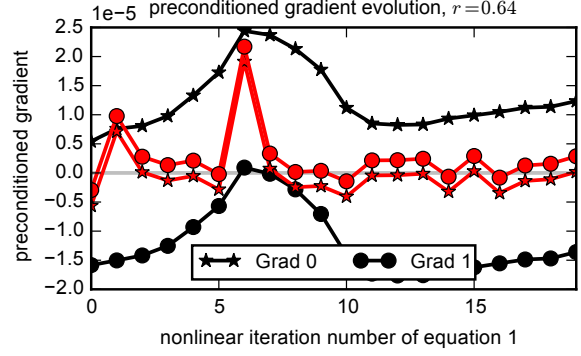

Figure 4: Confidence map $\beta_g(x)$ for GC-FWI for noise levels $r = 0.04$ and $r = 0.64$. The black and red arrows correspond to the preconditioned gradient evolution curves of the same color in Figure 5.

Regularizing velocity differences in time-lapse FWI using gradient mismatch information

seems quite robust to changes in source location. Tests indicate that if the 19 monitor shots are offset from the baseline locations by half the shot spacing, the largest shift possible, the resulting confidence map $\beta_g(x)$ is qualitatively the same. We have not yet investigated how realistic differences in the source wavelet would influence GC-FWI.

For the example in this study, the cost of GC-FWI is approximately 3.5 times that of DDFWI. The cost of $\beta_g(x)$ is 40 standard FWI iterations because each of the 20 iterations of (1) is twice the cost of a standard FWI iteration due to the presence of two data terms. Each of the 15 iterations of the subsequent joint inversion (4) is also twice as expensive as a standard FWI iteration for the same reason. This brings the total cost of GC-FWI to 70 standard FWI iterations compared to the 20 standard FWI iterations of DDFWI in this example. In this calculation we do not include the cost of the initial inverted baseline model that is used in DDFWI and as the initial guess in (1).

Even though GC-FWI significantly reduces the constraints, imposed by DDFWI, on the acquisition of the repeat survey, current time-lapse datasets have very similar source and receiver locations. This knowledge can potentially be used to increase

the accuracy of the confidence map $\beta_g(x)$. Instead of comparing the stacked gradient of the baseline and the monitor data term in (1) we could compare the gradient history of each shot individually. Since the shot and receiver locations are similar in the baseline and monitor survey, the major difference in the gradient contribution of a shot will be due to noise and time-lapse model differences. Generating $\beta_{g,s}(x)$ from the history of each shot s and then stacking the result for the final confidence map $\beta_g(x)$ may give improved results. If the source and receiver locations in the baseline and monitor surveys are similar and DDFWI is possible, $\beta_g(x)$ could also be used to regularize the model difference in DDFWI. A significant portion of the noise in the DDFWI time-lapse estimate in Figure 6 could be suppressed in this way.

CONCLUSION

The method introduced in this study performs an initial inversion where the baseline and the monitor data misfit are minimized with a single model. By comparing the gradients of both data terms at each iteration of the inversion we gain insight about the probability of a time-lapse velocity change in a region. This knowledge is encoded in the confidence map $\beta_g(x)$, which is then used to regularize the model difference in a final joint inversion. This method, which we refer to as GC-FWI, is flexible to differences in acquisition and shows promising results on synthetic examples.
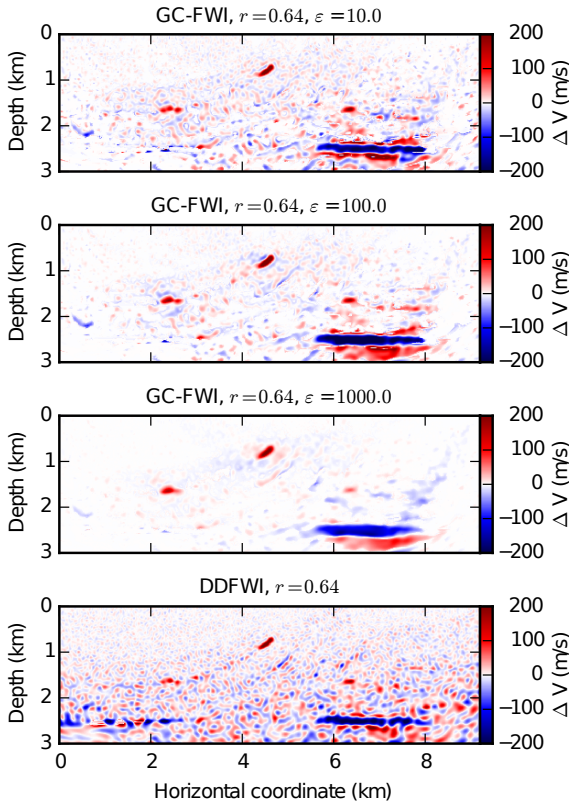
Figure 6: Comparison of GC-FWI using three different values of the model-regularization weighting term ε with DDFWI for noise level $r = 0.64$. Figure 3 shows the true model time-lapse change. At low regularization (top) GC-FWI gives results similar to those of DDFWI, but as ε is increased GC-FWI suppresses spurious changes and results in a more accurate recovery of the time-lapse changes.

Regularizing velocity differences in time-lapse FWI using gradient mismatch information

REFERENCES

Denli, H., L. Huang, et al., 2009, Double-difference elastic waveform tomography in the time domain: Presented at the 2009 SEG Annual Meeting, Society of Exploration Geophysicists.

Eggenberger, K., P. Christie, D.-J. van Manen, and M. Vassallo, 2014, Multisensor streamer recording and its implications for time-lapse seismic and repeatability: The Leading Edge, **33**, 150–162.

Lailly, P., 1983, The seismic inverse problem as a sequence of before-stack migrations: Presented at the Conference on Inverse Scattering: Theory and Application, SIAM.

Maharramov, M., and B. Biondi, 2014, Robust joint full-waveform inversion of time-lapse seismic data sets with total-variation regularization: arXiv preprint arXiv:1408.0645.

Maharramov, M., B. Biondi, et al., 2014, Joint full-waveform inversion of time-lapse seismic data sets: Presented at the 2014 SEG Annual Meeting, Society of Exploration Geophysicists.

Nocedal, J., and S. J. Wright, 2006, Numerical optimization, 2 ed.: Springer.

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

Shin, C., S. Jang, and D.-J. min, 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: Geophysical Prospecting, **49**, 592–606.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Watanabe, T., S. Shimizu, E. Asakawa, T. Matsuoka, et al., 2004, Differential waveform tomography for time-lapse crosswell seismic data with application to gas hydrate production monitoring: Presented at the 2004 SEG Annual Meeting, Society of Exploration Geophysicists.

Yang, D., 2014, Full wavefield inversion methods for monitoring time-lapse subsurface velocity changes: PhD thesis, Massachusetts Institute of Technology.

Zamanian, S. A., D. Yang, W. L. Rodi, and M. C. Fehler, 2014, Hierarchical bayesian inversion of time-lapse seismic data with different acquisition geometries: Presented at the 76th EAGE Conference and Exhibition 2014.