Applications of FWI to the microseismic source problem

Nadine Igonin, Kristopher H. Innanen

ABSTRACT

Microseismic events occur as a result of both natural and man-made processes, and their spatial distribution is important to understanding the mechanisms involved in creating this seismicity. In order to accurately locate microseismic events, a calibrated velocity model is required, but acquiring one is often a challenge. Since both the microseismic source location problem and velocity model inversion problem are wave-based, full waveform inversion (FWI) can be used to solve both problems. In this report, we propose a modified FWI scheme that iteratively solves for both source distribution and the velocity model.

We show numeric examples in a 2D time-domain acoustic framework developed in Python. Examples of the gradients for both source location and velocity model are shown in a variety of scenarios in order to begin to understand their behaviour on their own, as well as the potential cross-talk between them. We see that low frequencies are required for the source-term gradient to produce a correct update. This update has the form of a dipole, with a negative lobe at the incorrect source location and a positive lobe in the direction of the correct source location.

INTRODUCTION

Novel research can come about in many ways, and one approach to searching for gold mines of unexplored research areas is by combining certain fields in a way that has not been done before. For example, extending global seismology methods to microseismic has had tremendous success. Likewise, tailoring novel methods developed for microseismic and applying them to more conventional seismology datasets has also proven fruitful. In this report, we propose a bridge between the FWI method and microseismic data processing. To date, limited work has been done in applying FWI in a proper and rigorous way to microseismic data, and there is still much to be learned about the parametrization of the problem and features of applying the method to both real and synthetic data.

We begin with an overview of microseismic data and its attributes. The term itself refers to seismic events of moment magnitude less than zero. These events have P- and S-wave arrivals whose polarity and amplitude are governed by the source mechanism or moment tensor. Microseismic events occur in any seismogenic setting, such as natural earthquake systems. They can also be triggered by man-made processes, such as geothermal power generation or hydraulic fracturing. These events are studied in terms of their spatial and temporal relationships to the processes being induced in the subsurface. For example, in volcano seismology, microseismic is used to study the movement of gas and magma within a volcano. In the hydraulic fracturing case, microseismic is a proxy of which regions of the reservoir have been stimulated. Therefore, a significant effort in the field of microseismic is toward effectively locating microseismic events.

Figure 1a shows an example of the microseismic events caused by a single stage of hydraulic fracturing. The events show that fractures were activated approximately perpendicular to the injection well and illuminate a region in the shape of an elongated cluster. This sort of clustering is common for conventional hydraulic fracturing. However, in the case of induced seismicity, there is distinctly different clustering behaviour. The microseismic events still tend to cluster close together, but now along a fault, which is independent of the well position, such as shown in Figure 1b.

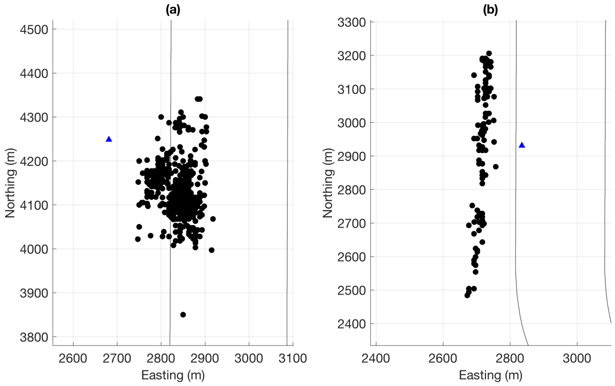

FIG. 1. Example of spatial characteristics of microseismic due to (a) hydraulic fracturing and (b) fault activation. The black dots are microseismic events, and the black line is the trajectory of an example well. (Data is from a proprietary dataset owned by the Microseismic Industry Consortium.)

Figure 2 shows an example microseismic event, with P- and S-waves indicated. The amplitudes and polarities depend on the focal mechanism. For example, during hydraulic fracturing, commonly there are tensile or strike-slip events. These focal mechanisms have nodal planes, such that at a nodal plane the amplitude of the wave is close to zero. The frequency content of the events depends on the magnitude. Higher magnitude events have lower frequencies, and conversely, smaller magnitude events have higher frequencies.

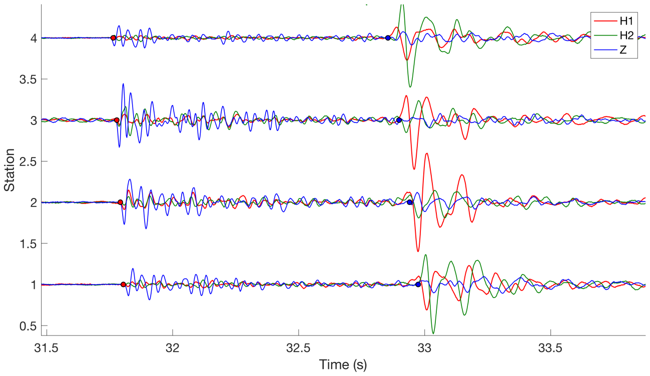

FIG. 2. Example of a microseismic event on four different stations. The H1 and H2 components are the two horizontal components, and Z is the vertical component. The red dots signify the onset of the P-wave and the blue dots signify the onset of the S-wave. (Data is from a proprietary dataset owned by the Microseismic Industry Consortium.)

There are many challenges that arise when working with microseismic data. For one, due to their low magnitudes, microseismic events tend to have low signal to noise ratio, which brings in a whole set of processing challenges. Additionally, reliable absolute locations of events require an accurate velocity model. In most applications, this is a problem that cannot be sufficiently remedied. Typically, crude existing velocity models are calibrated using event data from a known location, such as perforation shots. However, this approach also has caveats and can produce erroneous changes to the original velocity model due to the sparcity of information. Additionally, most applications use a 1D velocity model, which is a gross oversimplification of a potentially complex subsurface. Ignoring the effects of anisotropy can lead to misinterpretations of microseismic activity (Hayles et al., 2011). Therefore, velocity model building in the microseismic field is still an active area of research (Akram and Eaton, 2017).

Next, consider the method of full waveform inversion (FWI) (Tarantola (1984), Virieux and Operto (2009)). The premise of this approach is to find a model that minimizes the discrepancy between true and modelled data. This is done by minimizing an objective function, usually the sum of the square of the difference between the true and modelled data. A starting model is selected and during each iteration, the gradient, which is the derivative of the objective function with respect to the model parameters, guides the direction and magnitude of the update to the model. The updating then occurs iteratively until one of a range of possible stopping criteria are met. Multiparameter FWI is an extension of conventional FWI wherein the goal is to invert for more than one parameter. For example, in an elastic environment, multiparameter FWI can be used to solve for density, P-wave velocity, or S-wave velocity. Likewise, the problem can be formulated in terms of impedance instead, or any other useful parametrization.

One of the challenges with applying FWI is non-uniqueness. For a limited set of sources and receivers, there is more than one velocity model that can accurately reproduce the data. For instance, in a dataset with limited aperture, a high-velocity layer deep below the surface will give a similar response to a shallower lower-velocity layer. These two velocity configurations will therefore be very difficult to distinguish. In many instances of non-uniqueness such as this one, changes to data acquisition can help (such as by increasing the aperture), but non-uniqueness in one form or another is a universal characteristic of geophysical inverse problems.

A related challenge with FWI is that of cross-talk. This is a problem that occurs if one parameter is updated in response to data variations caused in part by a different parameter. For example, given an anomaly in density, the P-wave velocity gradient may attempt to update the region of the density anomaly without being able to determine that it should not do so. For this reason, the Hessian, an operator which measures coupling between different parameters, is often required in order to address and remedy the cross-talk.

The goal of this report is to outline a scheme that combines the microseismic source location problem with the conventional FWI framework in order to simultaneously solve for the source character and the background velocity model. We give this approach the name of *microseismic full waveform inversion*, or MFWI. Based on gradients derived in the companion report (Igonin and Innanen, 2017), we illustrate how both gradients behave in a 2D acoustic environment. The purpose is to examine, within a coherent theoretical and numerical framework, the character of the non-uniqueness and cross-talk occurring during simultaneous updating of the velocity model and the source positions and strengths.

THEORY

FWI starts with the goal of minimizing the difference between observed and modelled data, such that the objective function is

$$\phi = \sum_{r_g, r_s} \frac{1}{2} \int dt \Big(p(\mathbf{r}_g, \mathbf{r}_s, t) - g(\mathbf{r}_g, \mathbf{r}_s, t | s_c, s_s) \Big)^2, \tag{1}$$

where $\mathbf{r}_g$ is the receiver position, $\mathbf{r}_s$ is the source position, s_c is the velocity model, s_s is the source distribution, $p(\mathbf{r}_g, \mathbf{r}_s, t)$ is the observed data and $g(\mathbf{r}_g, \mathbf{r}_s, t|s_c, s_s)$ is the modelled data. This is the time-domain form of the objective function, chosen over frequency domain because of the nature of the microseismic problem. In a Newton scheme, the update to the model would have the form of

$$\delta \mathbf{s} = -\mathbf{H}^{-\mathbf{1}}\mathbf{g}, \tag{2}$$

where

$$\mathbf{g_0} = \frac{\partial \phi}{\partial \mathbf{s_0}}, \tag{3}$$

is the gradient with respect to either s_c or s_s, and

$$\mathbf{H} = \frac{\partial^{\mathbf{2}} \phi}{\partial \mathbf{s_0^2}}, \tag{4}$$

is the Hessian. For a more comprehensive review of the derivations and interpretation of the different components, the reader is referred to (Margrave et al., 2011).

In the companion report, Igonin and Innanen (2017), the sensitivities, gradients and Hessian are derived for the MFWI scheme proposed in this paper. As a summary, the time-domain source-term gradient is

$$\mathbf{g}_s = -\sum_{r_g, r_s} \int dt \delta P(\mathbf{r}_g, \mathbf{r}_s, t | s_c, s_s) g(\mathbf{r}_g, \mathbf{r}, t - t^* | s_c, s_s), \tag{5}$$

where $\delta P(\mathbf{r}_g, \mathbf{r}_s, t|s_c, s_s)$ is the residuals, and t^* is the origin time of the source. Similarly, the velocity-model gradient is

$$\mathbf{g}_c = -\sum_{r_g, r_s} \int dt \delta P^*(\mathbf{r}_g, \mathbf{r}_s, t|s_c, s_s) \int dt' \int dt_s g(\mathbf{r}_g, \mathbf{r}, (t-t^*)-t'|s_{c0}, s_{s0}) \frac{\partial^2 g(\mathbf{r}, \mathbf{r}_s, t'|s_{c0}, s_{s0})}{\partial t'^2} s_{s0}(\mathbf{r}_s, t^*), \tag{6}$$

where $\delta P^*(\mathbf{r}_g, \mathbf{r}_s, t|s_c, s_s)$ is the complex conjugate of the residuals.

A time-domain implementation of FWI was developed by Almuteri and Innanen (2016) in the 2D acoustic environment, and is based on the gradient in equation 6. A series of Python codes were developed, including a finite-difference engine based on the second-order finite difference approximation and Clayton-Engquist absorbing boundary conditions. For the details of the finite-difference engine, boundary conditions, stability, line search and optimization techniques, the reader is referred to Almuteri and Innanen (2016).

In this report, the original codes are validated in order to explore the behaviour of the gradient in equation 6. Then, the gradient in equation 5 is created using the same finite-difference engine as the backbone. The following sections show the results of these experiments, along with a few discussions on the nature of the cross-talk between the two parameters.

MFWI CASE I: UPDATING VELOCITY MODEL ERRORS

In order to begin to understand the new proposed source-term gradient, it is necessary to gain an understanding of the conventional velocity-model gradient. A series of tests were conducted to confirm that the gradient was behaving as expected. Figure 3a shows the geometry and velocity model of one such test. The focus is a circular anomaly in the center of the model space. After 100 iterations, and using a dominant frequency of 20 Hz, Figure 3b shows the inverted result. The size of the anomaly is consistent with the true model, and there are not many artifacts from the geometry.

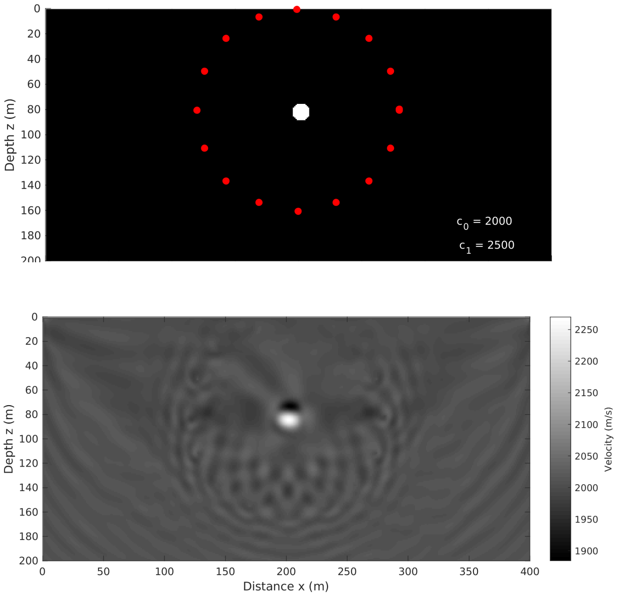

FIG. 3. True (a) and inverted velocities (b). Red circles are sources and there is an array of receivers on the surface.

Of particular interest is how well velocity anomalies can be imaged with microseismic-type geometries. As discussed previously, generally speaking, microseismic events tend to cluster spatially. Thus, the combined group of events provide very similar ray-paths and illuminate only a limited portion of the model space. An example of how this can be problematic is shown in Figure 4. Due to the limited number of unique information, the anomaly is not well resolved.

Another geometry of interest is that due to fault activation. Figure 5 shows a vertical fault illuminated by several events along its length, and the resulting reconstructed velocity field. In this case, because enough of the model space is uniquely sampled, the anomaly is accurately resolved. In practical terms, this means that fault activation is favourable for these applications because of the larger range in depths for recorded events.

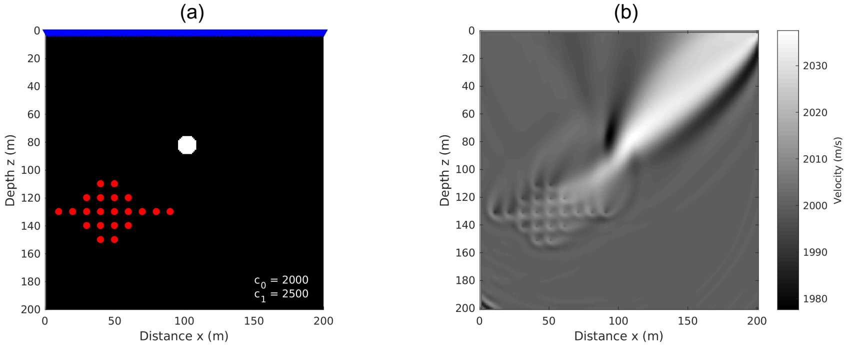

FIG. 4. Microseismic-type geometry (a) and inverted velocity model (b) after 2 iterations.

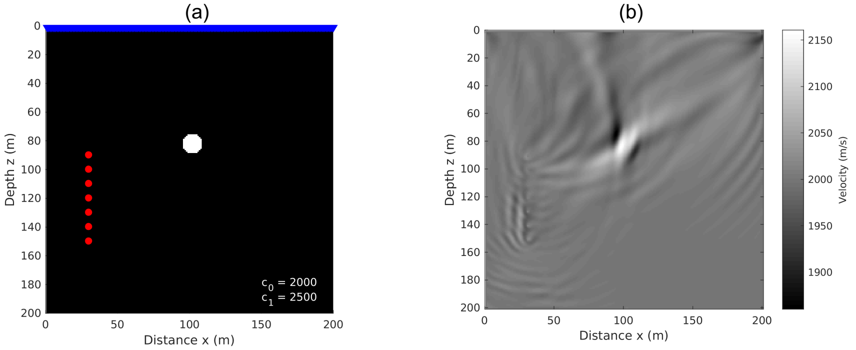

FIG. 5. Fault-activation type geometry (a) and inverted velocity model (b) after 100 iterations.

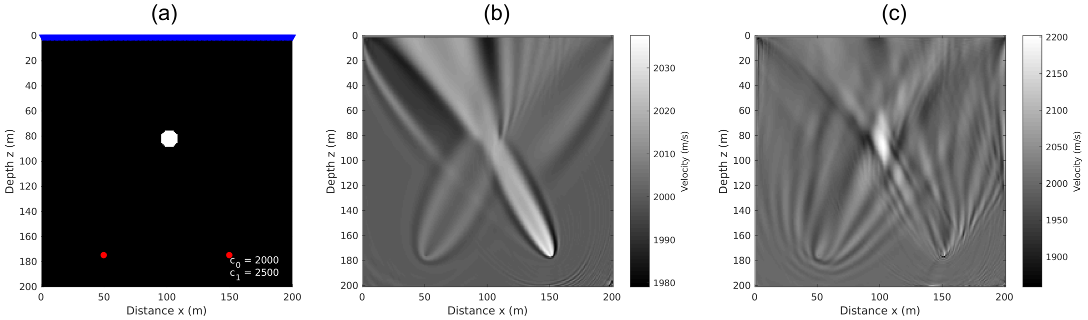

FIG. 6. Effects of varying frequency content. (a) Geometry: left source is 20 Hz and right source is 50 Hz, (b) velocity model after 1st iteration, (c) velocity model after 97th iteration.

Since the frequency of microseismic events depends on the magnitude, it is important to consider the effects of frequency on the inversion. Specifically, the effects of superimposing kernels due to sources of variable frequency. Figure 6 shows the effects of having sources with different frequencies. At the first iteration (Figure 6b), the difference in the update is apparent. However, after 97 iterations (Figure 6c), the anomaly is still correctly resolved. Higher frequency data provides more resolution, but it may be more practical and less computationally expensive to start with the lower frequencies. In practice, this would mean starting with high magnitude events first (due to their lower frequencies), and working down to the lower magnitudes.

Cross talk

With the assurance that the velocity model gradient is working correctly, the next question is the effect of having an incorrect source position on the inversion. This is an indirect way to study the cross-talk in the absence of the Hessian. Figure 7a shows what happens if the modelled source position is 10 m above the true source position, while holding the true and modelled velocity constant. The result is a negative anomaly between the true and modelled source position.

Figure 7b shows what happens if all source positions (originally at 190 m) are moved by 90 m up, such that the model thinks all the sources are 90 m higher than they are in the true model. The effect is an update of constant amplitude, a step, of high magnitude. This shows that the gradient has the tendency to supply a mass update to the velocity model if the source position is off. Therefore, an MFWI implementation with a Hessian may be more fruitful because it may be more able to recognize and address the cross-talk.

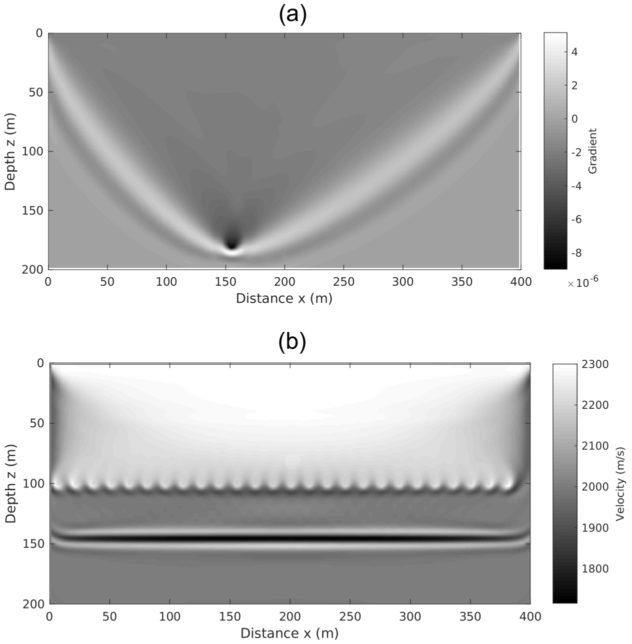

FIG. 7. Impact of introducing error in source position on one source by 10 m (a) and all sources by 90 m (b). The event at 150 m is thought to be a boundary effect.

MFWI CASE II: UPDATING SOURCE ERRORS

Source positioning errors

Consider a universe where the wavefield at all positions is known, that is, there are receivers at every position in the model. At origin time $t^* = 0$, a point source goes off at a certain position (x_t, z_t) and travels through a known homogeneous background velocity field. For this case, we will fix the starting velocity model equal to the true velocity model. However, we will supply an erroneous starting position (x_0, z_0) in order to study how the gradient will attempt to move toward the true location.

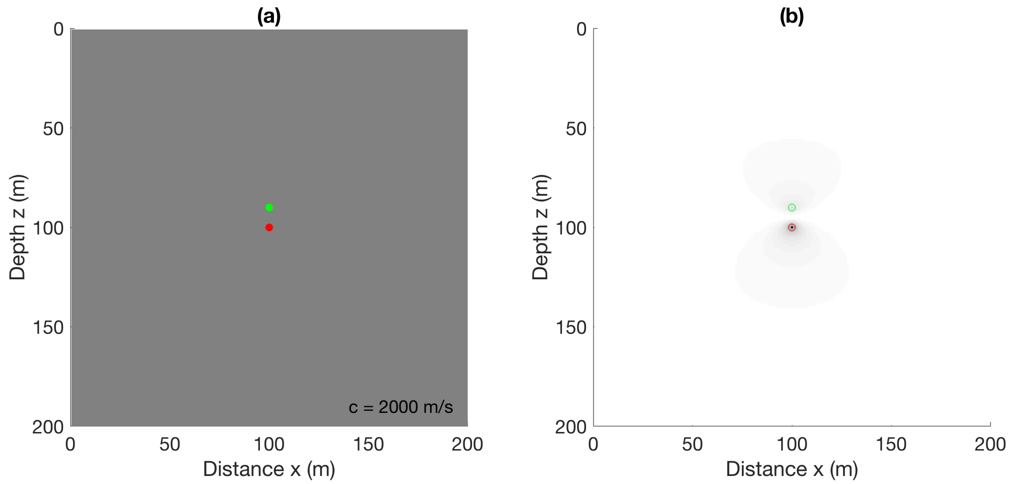

FIG. 8. Model and gradient with a vertical discrepancy. (a) Homogeneous velocity model, with true position in green and starting position in red; (b) gradient, with darker colours being negative.

Using a homogeneous background medium with a velocity of 2000 m/s and a dominant frequency of 20 Hz, Figure 8 shows the gradient after one iteration. There is a 10 m difference between the true and starting depths. The gradient attempts to cancel out the incorrect source position by adding a large negative update at that location, and appears to move in the direction of the true location. The two lobes are not symmetric, and there is a gap between the true and starting location.

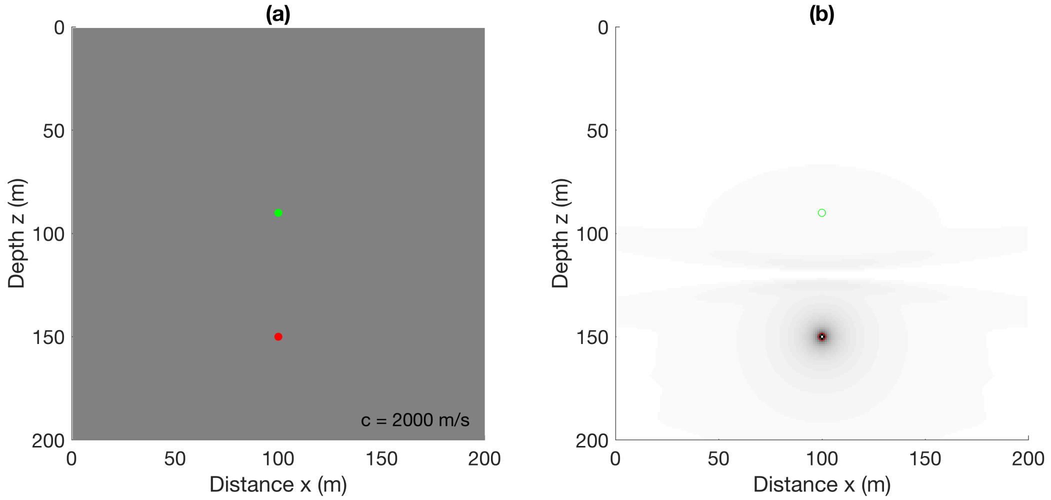

FIG. 9. Model and gradient with a large vertical discrepancy. (a) Homogeneous velocity model, with true position in green and starting position in red; (b) gradient, with darker colours being negative.

As one would expect, this sort of problem suffers from cycle skipping. As an illustration of this, Figure 9 shows the gradient if the starting position is at 150 m, putting it 60 m from the true depth. Again, we see that there is a tendency to cancel out the incorrect location, and there is a directionality of the update in the correct direction. Even so, we see that the center of the upper lobe is about the correct location.

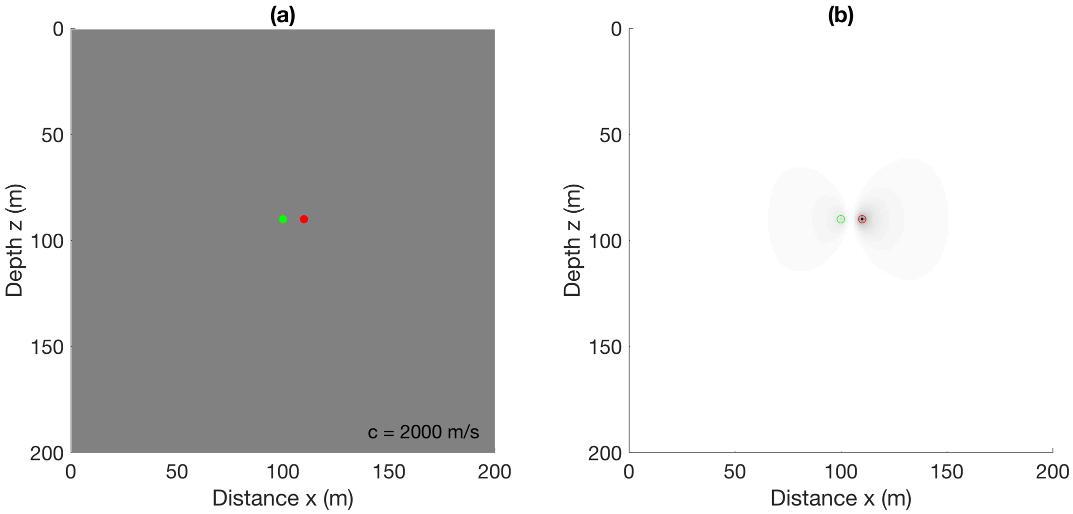

FIG. 10. Model and gradient with a horizontal discrepancy. (a) Homogeneous velocity model, with true position in green and starting position in red; (b) gradient, with darker colours being negative.

The same sort of directionality can be seen if there is a horizontal discrepancy in the source position, as shown in Figure 10.

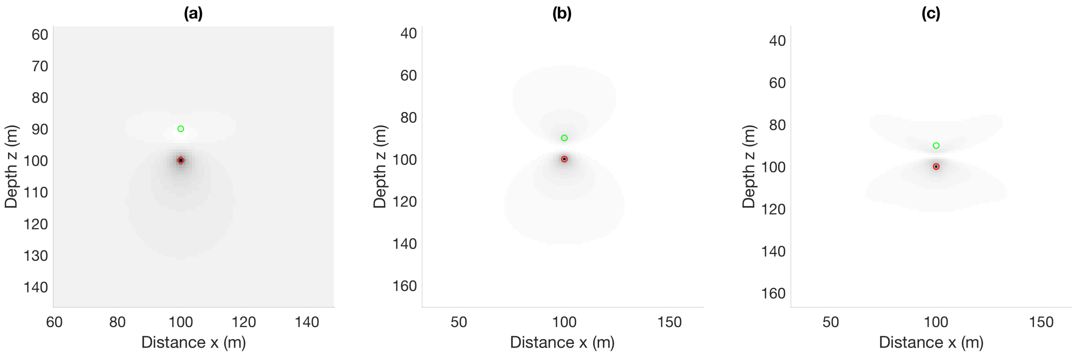

FIG. 11. Gradient as a function of frequency: (a) 5 Hz, (b) 20 Hz, (c) 40 Hz.

Another aspect to consider is the effect of frequency. Commonly, lower frequencies are used to build out the dominant components of velocity models, and higher frequencies are used later on to give higher resolution. To see if this is the case with this gradient, Figure 11 shows a comparison of the update in Figure 8, but with frequencies of 5 Hz, 20 Hz and 40 Hz. Only the 5 Hz update behaves exactly as it should, acting as a dipole by adding negative energy at the incorrect source position and adding positive energy at the correct source position. Therefore, we can conclude that just as for the velocity model update, lower frequencies should be used first.

Although the results give us an idea of the behaviour of this gradient, the limitations of this approach are numerous. For one, real data does not have receivers at all positions. Therefore, we would expect the gradient to have lower resolution than we have shown.

Additionally, using a point source is also unrealistic. Although theoretically nucleation of an event occurs at a point, there is a finite area (i.e. nucleation patch) that moves as a result of the earthquake. Future implementations of this approach will account for both of these limitations.

Source timing errors

Looking back at equation 5, another aspect to consider is the possibility of source timing errors. If the source origin time, t^*, is not known accurately, this may introduce error in the inversion scheme. In microseismic applications, the origin time is not known in advance, and many advanced processing techniques have been developed without the need to determine an explicit origin time. In the absence of an origin time, equation 5 would have to be modified in order add an integral over t^*. This would be computationally very expensive, and a better approach would likely be to determine an approximate origin time based on conventional microseismic processing techniques.

Cross talk

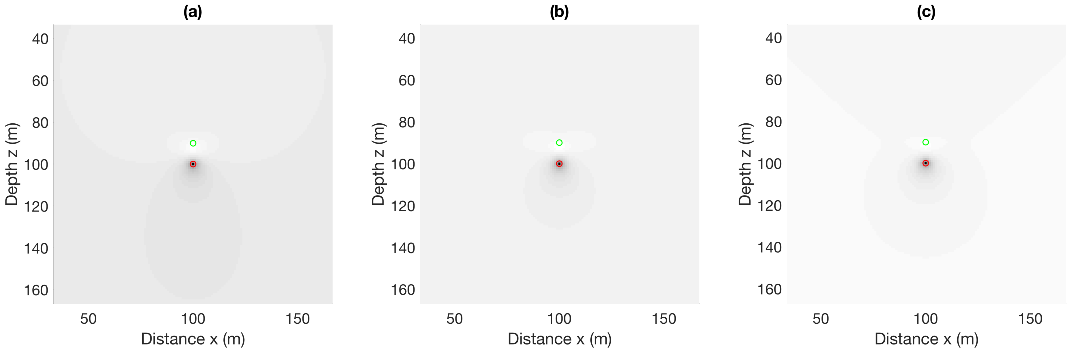

FIG. 12. Gradient as a function of change to true velocity model: (a) $v_t < v_0$, (b) $v_t = v_0$, (c) $v_t > v_0$.

As in the case of the velocity model gradient, one way to explore some of the potential cross-talk between the two gradients is to introduce error in the velocity model and study the effect on the source-term gradient. In the absence of the Hessian, this is what the update would be. Based on the previous analysis, we will refrain to using 5 Hz as the dominant frequency, and keep the difference in position between the true and starting models small. Keeping the model homogeneous, but changing the velocity of the true model, appears to change the shape and relative strength of the update, as shown in Figure 12. In each case, the true velocity has been perturbed by ± 200 m/s. We can see that having a starting velocity model that is too high, as in Figure 12a, leads to over-estimating the source region, while having a starting velocity model that is too low, as in Figure 12c, leads to underestimating the source region.

Figure 13 shows that the velocity model has been given a layer that is not present in the starting model, and the resulting gradient in that case. As expected based on the above experiments from changing the velocity model, if the starting model is too low, as it is on the right half of the model, the update overestimates the region.

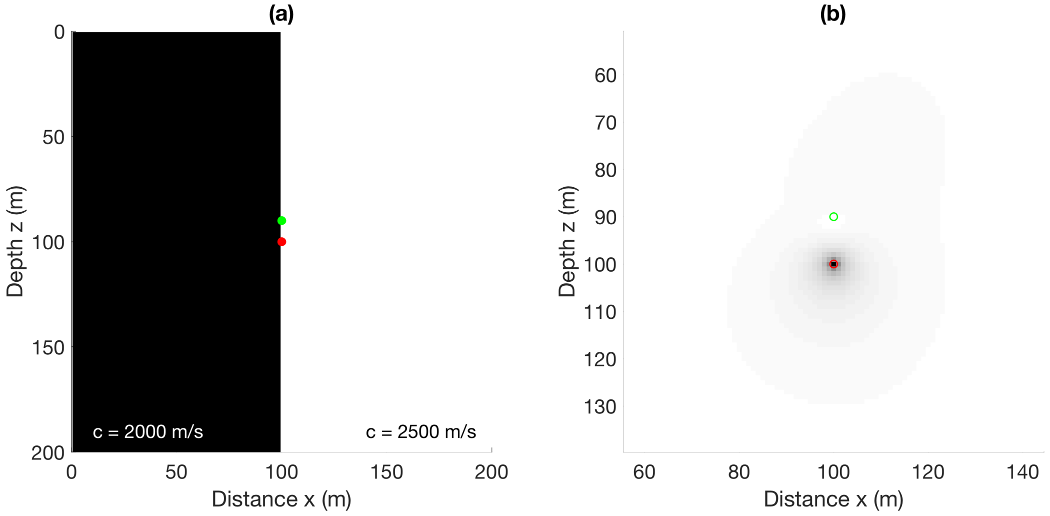

FIG. 13. Model and gradient with a velocity model discrepancy. (a) Velocity model, with true position in green and starting position in red; (b) gradient, with darker colours being negative.

MULTIPARAMETER MFWI: SIMULTANEOUS SOURCE AND VELOCITY ERRORS

The multiparameter MFWI implementation would involve simultaneous source and velocity model updates. Based on the definitions proposed previously, it would have the form of

$$\delta \mathbf{s} = -\mathbf{H}^{-1}\mathbf{g}, \tag{7}$$

where

$$\begin{bmatrix} \delta \mathbf{s}_c \\ \delta \mathbf{s}_s \end{bmatrix} = -\begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 \\ \mathbf{H}_3 & \mathbf{H}_4 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{g}_c \\ \mathbf{g}_s \end{bmatrix}. \tag{8}$$

The subscript of c refers to the velocity model update, while the subscript of s refers to the source position update.

After each iteration, both the velocity model space and source location model space would be updated. Due to the non-uniqueness of the problem, an approach without the Hessian is inconceivable. Practically speaking, conventional FWI would have to be carried out to produce a preliminary three-dimensional velocity model. Also, with the inherent problem of cycle skipping, it may be necessary to use conventional microseismic processing methods to determine a reasonable starting point for the source location. Many computationally inexpensive methods exist for doing so.

Therefore, multiparameter MFWI produces advanced processing results. Consider a starting three-dimensional velocity model and microseismic events occurring at some depth within the volume. MFWI refines the velocity model above the event along raypaths from the source to the receivers (Figure 14a). Given that there are tens of thousands of microseismic events for a given hydraulic fracturing operation, this could provide significant additional resolution.

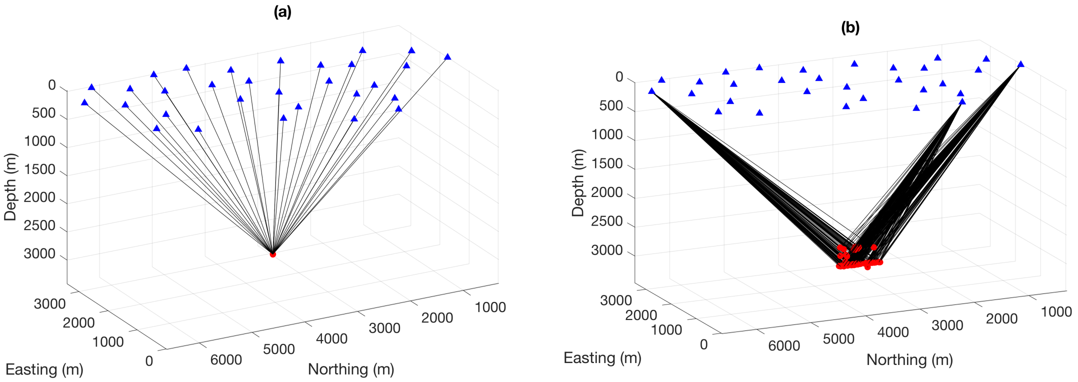

FIG. 14. Schematic illustration of ray paths for a homogeneous medium, with sources in red, receivers in blue and ray paths in black. (a) For the case of a single event; (b) for a N/S fault, with ray paths for only a few receivers shown to illustrate increased resolution near the region of the fault.

Of particular interest is the potential of this method to image faults. During fault activation, microseismic tends to illuminate the area of the fault that was activated (Figure 14b). The source position update can provide insight into the size of a nucleation patch that triggered each earthquake, and may be a high-resolution method to determine precise nucleation points. Since the faults activated during hydraulic fracturing are often ancient, inactive faults, increased understanding of the rupture processes, including the effect of fault cohesion due to fault latency, is an active area of research.

Additionally, in recent studies of induced seismicity due to fluid injection, a common theme has been that

1. Not all faults seen on seismic are activated during fluid injection.
2. Not all faults that are activated are seen on seismic.

Though the first problem cannot be addressed herein, the second may be illuminated with the velocity model update. Using a subset of microseismic events along a fault, the velocity model update will refine the region immediately around a fault. With the superposition of many such updates, the hypothesis is that the footprint of the fault, be it in density or velocity, would be found. The character of this footprint would be of great interest to the academic community. Working in the other direction, this footprint can be used to find other faults within the seismic volume, with the goal of understanding theme #2 above.

CONCLUSIONS AND FUTURE WORK

In this report, we proposed a framework for the simultaneous inversion of the velocity model and source position. Using a time-domain FWI implementation, we have demonstrated the behaviour of the velocity model gradient in various geometries and frequency content. We have also shown the character of the source position update. Finally, we explored the nature of cross-talk in an unconventional way by studying the effect of introducing error to the models, and have proposed how this framework can be extended toward simultaneous source and velocity model inversion.

Since this research is still in its infancy, there remains a broad scope for the future work. Of immediate interest is the exact nature of the cross-talk between the two parameters. Although some preliminary insights have been gained from the analyses herein, the Hessian is the key to understanding this problem more deeply. The next step is to create a functional FWI code that includes the Hessian in a Gauss-Newton framework. This will be done in Python by building on existing codes.

Though the work to date has been in the acoustic environment, in order to include a moment tensor, the next logical step is to move into an elastic environment. This will involve using more advanced wavefield modelling engines and, if necessary, modifying existing codes to include a moment tensor. Additionally, the parametrization of this FWI problem including a moment tensor will take some thought. There are also certain allowances to be made given the limited nature of real data (geometry and frequency content).

Once a framework has been established and tested on synthetic data, the goal is to create a laboratory dataset in the Crewes physical laboratory that mimics the microseismic set-up. Since such a set-up has yet to be attempted to date, this will involve novel experimental and instrumental design. Finally, the ultimate aim with this project is to be able to successfully apply it to real microseismic data in order to gain unique insight into the spatial distribution of microseismic events during high-magnitude fault activation.

ACKNOWLEDGEMENTS

We thanks the sponsors of CREWES for their continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13. Nadine Igonin is also funded through the NSERC PGS-D.

REFERENCES

Akram, J., and Eaton, D. W., 2017, 1D layered velocity models and microseismic event locations: synthetic examples for a case with a single linear receiver array: Journal of Geophysics and Engineering, **14**, No. 5, 1215.

Almuteri, K., and Innanen, K., 2016, Tutorial: acoustic full waveform inversion in time domain: CREWES Research Report, **28**, 6.1–6.17.

Hayles, K., Horine, R., Checkles, S., and Blangy, J., 2011, Comparison of microseismic results from the Bakken Formation processed by three different companies: Integration with surface seismic and pumping data: 81st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1468–1472.

Igonin, N., and Innanen, K., 2017, Toward full waveform inversion for source location and velocity model: gradient and Hessian: CREWES Research Report, **29**.

Margrave, G., Yedlin, M., and Innanen, K., 2011, Full waveform inversion and the inverse hessian: CREWES Research Report, **23**, 77.1–77.13.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, No. 8, 1259–1266.

Virieux, J., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: GEOPHYSICS, **74**, No. 6, WCC1–WCC26.