Full waveform inversion of VSP accelerometer data from the CAMI field site

Scott Keating, Matt Eaid and Kris Innanen

ABSTRACT

The 2018 CaMI VSP survey provided a data-set suitable for creating a baseline subsurface model for later monitoring comparisons. Here, we implement elastic full waveform inversion to obtain an estimate of subsurface properties. Several challenges exist in applying full waveform inversion on the field data, including in particular a complex, difficult to characterize near-surface and discrepancies between geometrical spreading in 2D simulations and field data. We develop approaches for coping with these challenges, and obtain an inversion result that accurately reproduces the measured data. The inversion results in a subsurface model that is plausible, but difficult to verify.

INTRODUCTION

In 2018, a vertical seismic profile (VSP) survey was carried out at the Containment and Monitoring Institute (CaMI) Field Research Station near Brooks, Alberta. One of the goals of this survey was to obtain a baseline seismic data set to compare against later monitoring data, gathered during the course of carbon dioxide sequestration. Remote monitoring of sequestered CO_2 plays a very important role in CO_2 sequestration applications because it allows for confidence that the sequestered CO_2 does not migrate back to the atmosphere, or alternately, allows for rapid detection of leaks. Seismic surveys are a promising approach for CO_2 sequestration monitoring because they are capable of measuring large areas with good spatial resolution and can be related to physical properties that are highly sensitive to the presence of CO_2 (particularly acoustic impedance).

In this report, we focus on using the technology of full-waveform inversion (FWI) to estimate the subsurface properties from the measured data. This approach should allow for a relatively well-resolved model of the study area to be constructed, with the potential to combine information from several available data sets, including multiple components of the accelerometer data as well as distributed acoustic sensing measurements. While using FWI offers advantages, it also introduces complications, such as a significant sensitivity to data features introduced by the complex and

This report is focused on the application of an effective source FWI approach to a subset of the accelerometer measurements obtained in the 2018 CaMI VSP data set. Two companions papers, Eaid et al. (2021b) and Eaid et al. (2021a) detail the preliminary processing performed on the data prior to its use in inversion, and the full waveform inversion of both DAS and hybrid DAS-accelerometer data subsets from this survey.

THEORY

In this report, we perform two-dimensional, frequency-domain FWI on a set of VSP accelerometer data. Because the near-surface introduces many complications in land seismic data, we consider here an effective sources approach for VSP FWI, which attempts

to remove the near surface from the inversion problem. In Keating et al. (2021) this technique is developed and applied to a numerical data set; here, we will apply this approach to the 2018 CaMI monitoring data. The optimization problem considered in this approach is given by

$$\min_{m^*,f^*} \frac{1}{2}||R^*u^* - d^*||_2^2 + \phi_R \text{ subject to } S^*(m^*)u^* = f^*, \tag{1}$$

where z^* is a the depth of the effective source, m^* is the subsurface model, f^* represents the effective source amplitudes, R^* is a matrix applying receiver sampling, u^* is the seismic wavefield, d^* is the measured data, ϕ_R is a regularization term, and S^* is a Helmholtz matrix applying finite-difference elastic wave propagation. Each of the * variables is only allowed to take values at depths below z^*, which is effectively the top of our reduced model. In this formulation, instead of inverting for the (complex and difficult to characterize) near surface, we instead invert for f^*: the effective source term that generates the wavefield that would have been obtained by propagation through the near surface. In effect, this replaces the conventional optimization problem of minimization over model parameters with a known source with a problem of minimization over both model parameters below the near-surface and effective sources. The theory for this type of FWI problem has been described in Keating et al. (2018).

While initializing the subsurface model estimate, m^*, is a common problem with many well-known solutions (here we use a smoothed well-log), initialization of the effective source term, f^*, is more challenging. As detailed in Keating et al. (2021), an estimate of the wavefield at depth z^* may be available from forward modeling through the initial model, but this may not relate simply to an estimate of f^*. While several potential initializations were considered, we have not identified a more effective approach than to initialize f^* with zero amplitude at all effective source locations.

Because no effective initial estimate for the source term is available, the effective source updates at the early iterations play a much larger role in fitting the measured data than subsurface model updates do. Furthermore, before a good effective source estimate is obtained, misplacement of source energy may cause the model to update in ways that locally improve data fit, but do not improve the accuracy of the model and may hamper the inversion at later iterations. For these reasons, we consider a sequential inversion here, where at the first few iterations only the effective source term is updated, while at later iterations both the source and the model are simultaneously updated. This helps to ensure that model artifacts which improve data fit for poor source representations but not for accurate effective sources are not accidentally introduced into the inversion model.

In the companion report Keating et al. (2021), we introduced a regularization term for the sources, penalizing discrepancy between the energy distribution predicted based on the initial model and the energy distribution of the effective source. We adopt a similar regularization term here, defined as

$$\phi_{R_1} = \alpha \sum_s ||E_s(x) - ({f^*}^2_{x_s}(x) + {f^*}^2_{z_s}(x))||^2, \tag{2}$$

where $E_s(x)$ is the expected source energy at offset x, ${f^*}_{x_s}(x)$ is the x-component of the source term at offset x, ${f^*}^2_{z_s}(x))$ is the z-component of the source term at offset x, the sum is

over sources, and α is a weighting term. This regularization term allows for our knowledge about the geometry of the survey and source types to be directly considered in the inversion. Without this term, the inversion may converge to a set of effective source distributions which are not plausible, regardless of the near-surface. In applying this regularization term, we assume that the near surface does not significantly redistribute the energy distribution of the source in the x-direction (though it allows for conversions between P- and S- wave modes). In this report, we assume that the energy distribution corresponds to an explosive source. While a z-oriented point force would typically be a better approximation of a vibroseis shot, we have found that the energy distribution patterns for the two are similar, and that the explosive source energy distribution is much less sensitive to the assumed S-wave velocity of the near surface.

Due to computational cost considerations, we consider a two-dimensional implementation of full-waveform inversion here. Due to the neglect of three-dimensional effects in the field data (in particular the underestimation of geometrical spreading), we introduce an amplitude correction process. To prevent under-emphasizing measurements from deeper in the model, where amplitudes are typically smaller, we apply an overall scale to the measured data for each source-receiver pair such that the sum of all amplitudes for each pair is equal to one. This should help to prevent under-emphasis of deeper measurements, and also undo three-dimensional geometrical spreading. In order to compare these re-scaled data to our simulated data in the FWI procedure, we need to also scale the modeled data in a similar way. This scaling is introduced in two stages. At the start of the inversion, when an accurate estimate of the effective sources is unavailable, we design the source-receiver pair scaling terms based on a modeled shot using the true acquisition geometry in the initial model estimate (this necessitates including the near-surface part of the model for this stage). Later in the inversion, when a better estimate of the effective source is available, we re-calculate the scaling term based on data modeled from the effective sources. This scaling is straightforward to include in the inversion, requiring only a reweighting of the (now source-dependent) receiver sampling matrix R^*.

While we consider elastic inversion here, analysis of the well log for this site suggests that P-wave velocity, v_P, S-wave velocity, v_S, and density, ρ are strongly correlated in this area (Eaid et al., 2021b). For this reason, we consider here a single-parameter inversion, with this parameter chosen to simultaneously define each of the elastic properties we consider. This approach is described in detail in Eaid et al. (2021b).

Previous study of the CaMI field site has suggested that the geology in this area primarily consists of flat layers (e.g. Mossop and Shetsen, 1994). To incorporate this prior information into our inversion, we introduce a second regularization term, given by

$$\phi_{R_2} = \beta||m^* - Cm^*||^2, \tag{3}$$

where C is a matrix applying convolution with a Gaussian smoother in the x-direction, and β is a weighting term. This term promotes smoothness in the x-direction by penalizing the difference between m^* and its equivalent after smoothing in the x-direction. The full regularization term then includes both a source-penalizing and a layer-promoting term:

$$\phi_R = \phi_{R_1} + \phi_{R_2}. \tag{4}$$

Frequency band	1	2	3	4	5	6	7
	10	10	10	10	10	10	10
	10.5	11	11	11.5	11.5	12	12
	11	11.5	12	12.5	13	13.5	14.5
	11.5	12.5	13	14	14.5	15.5	16.5
	12	13	14.5	15.5	16.5	17.5	18.5
	12.5	14	15.5	17	18	19.5	20.5
	13	14.5	16.5	18	19.5	21	23
	13.5	15.5	17.5	19.5	21	23	25

Table 1. Frequencies considered at each band of the inversion. Frequencies are given in Hz.

FIELD DATA INVERSION

The inversion results we present here use the theoretical framework and inversion strategy described in Keating et al. (2021) and the data and preliminary processing workflow described in Eaid et al. (2021b). In this report we focus on the accelerometer data from one of the shot lines at the field site, including all shot points to the east of the well. These data include offsets from 5 m west (the closest shot to 0 offset) to 480 m east. An inversion of the distributed acoustic sensing (DAS) data from this shot line, as well as a hybridized inversion of DAS and accelerometer data are presented in Eaid et al. (2021a).

The inversion approach we consider here is an elastic, two-dimensional, frequency-domain FWI. In this report, we focus on frequencies between 10 Hz and 25 Hz. While lower frequencies would be desirable, 10 Hz was found to be the lowest frequency at which the ratio of seismic signal to noise was acceptable. The dominant frequency of this data-set is substantially higher than the highest frequency we considered here, which was limited by the computational costs of moving to the smaller finite-difference grid spacing necessary for higher frequencies. We consider measurements from both the vertical and an in-plane horizontal directions. We neglect measurements from the out-of plane horizontal direction due to the limited anisotropy expected in the study area (Hall et al., 2018) and our assumption of limited out-of-plane reflections.

We consider a model of 510 m in the x-direction, and 322.5 m depth. This model is split into grid cells 2.5 m in each direction for finite-difference modeling and inversion. For the inversion, we consider eight frequency bands, each consisting of seven frequencies, starting with the lowest frequencies and ending with a band spanning from low to high frequencies. The frequencies used in each band are summarized in Table 1. To parameterize the model, we use an approach of the type described in Keating et al. (2018) to define small, spatial Gaussian functions in two-dimensions as our basis functions. We define these basis functions with a standard deviation of 7.5 m in the z direction and 17.5 m in the x-direction to further promote layering. For our implementation of the effective sources inversion strategy, we consider an effective source depth of 40 m. The initial model we use here was based on a smoothed version of the well logs available at the field site, and is shown in Figure 1.

We consider three stages to our inversion approach. Because we initialize the effective sources at zero amplitude, source updates play a much more significant role than model

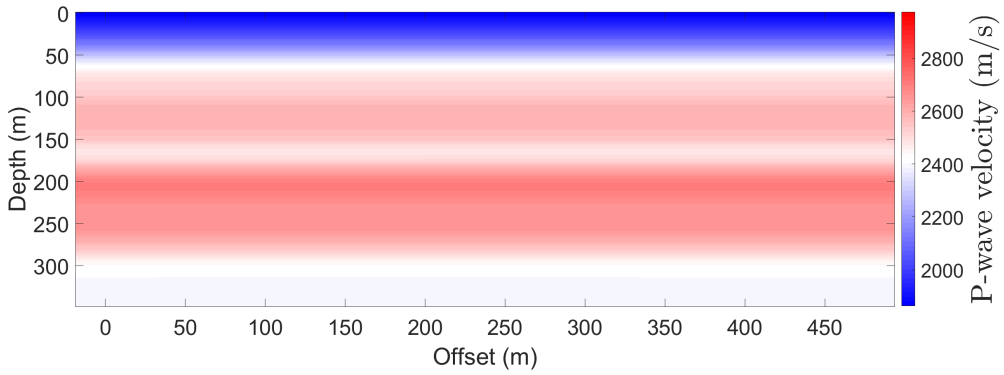

FIG. 1. v_P values of the initial model used.

updates in the first few iterations, and no confident estimate of modeled amplitudes is available to determine the amplitude scaling factors. For these reasons, we begin the inversion with five iterations of L-BFGS optimization at all frequencies, updating sources only and not including any amplitude scaling. After these initial iterations, the modeled data can be used to generate an appropriate estimate of the amplitude scaling term for each source-receiver pair. After applying this amplitude scaling, we perform 15 additional iterations of L-BFGS optimization at all frequencies, updating sources only. After this second stage of source updates, we recalculate the amplitude scaling terms and proceed to simultaneous effective source and model inversion. The simultaneous inversion is performed over seven frequency bands, with 10 iterations of L-BFGS optimization used at each band.

The real part of the measured data and modeled data, after both the source-only updates and after the whole inversion process, are shown for the shot at 450 m offset in Figure 2. As this comparison demonstrates, the data-fit is relatively good after inversion: across all shots, the data-fit objective function term after the inversion is 8.9% of the initial data-fit term (where the effective sources had zero amplitude). This figure also illustrates that much of the data fit is provided by the effective source only: the data fit term for this shot is 15.2% of the initial after updating only the sources, so much of the data residual can be eliminated just by updating the effective sources (or, equivalently by changing the modeling of wave propagation through the near-surface).

The inversion result is shown for v_P in Figure 3. This result has several positive features: the model updates are largely layer-like, with limited heterogeneity in the x-direction, and the average model profile is similar to the known well profile, represented by the initial model. The inversion also seems to recover a large contrast at about the expected depth of the reservoir of interest (approximately 300 m). Negative features of this result include some apparent artifacts near the well itself, where the impact of model updates on the data is large. This may suggest that a larger layer-promoting regularization term is needed for this problem.

A more in-depth treatment of full-waveform inversion on the 2018 CaMI data set is available in Eaid et al. (2021a), which includes hybrid inversions of DAS and accelerometer data for the entire shot line considered here.

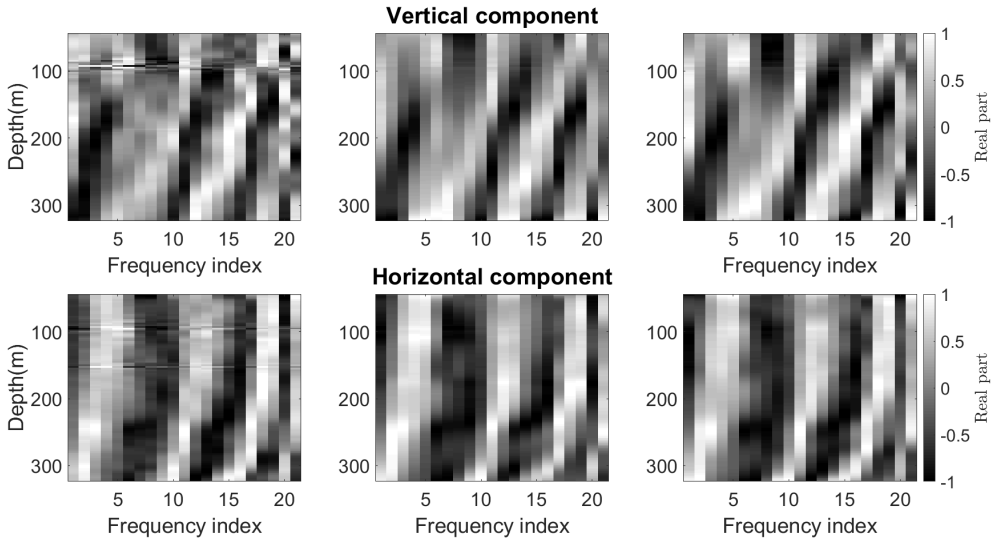

FIG. 2. Real part of frequency domain data for shot at 450 m offset. Top row: vertical component of measurements. Bottom row: horizontal component of measurements. Left column: Measured data. Middle column: Modeled data after effective source estimation (no elastic property updates). Right column: Modeled data after simultaneous inversion for effective source and elastic properties. Frequency index on the x-axis corresponds to the numbering of the 21 unique frequencies considered in Table 1, arranged from low (10 Hz) to high (25 Hz).

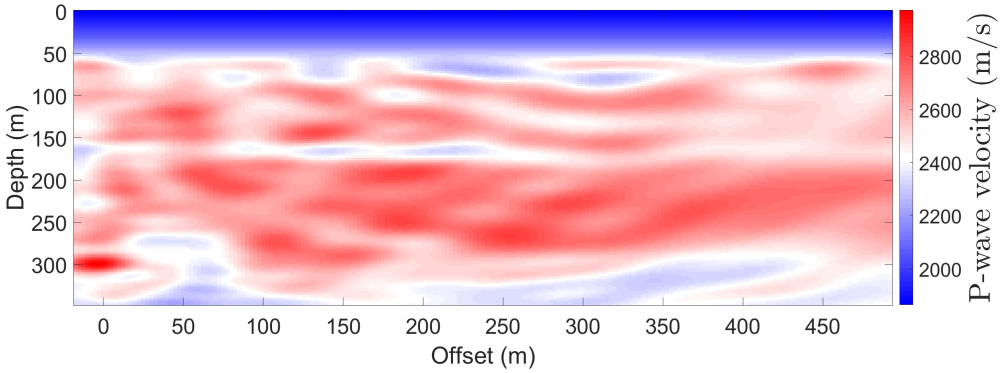

FIG. 3. v_P values for the inversion result. Compare to Figure 1.

DISCUSSION

Assessing the quality of the inversion results obtained here is difficult due to the lack of a ground-truth data source to compare against. For this data-set, we have used the available well-logs both to generate an initial model for the inversion, and as a comparison to assess the quality of the results. This situation is not ideal, especially as the inversion is intended to recover information away from the well, where this type of validation is not appropriate. For this reason, we plan to use the null-space shuttling approach described by Keating and Innanen (2020) to assess the confidence of this inversion in different model features.

The lack of information available on appropriate initial estimates for the effective source term led us to initialize this term with zero amplitude here, which, in turn, necessitated computationally expensive source-only iterations before the model-updating simultaneous iterations. In the companion report Keating et al. (2021), two effective source FWI strategies were presented, and numerical examples were performed using the wavefield (rather than source) based formulation. This formulation allowed for the modeled wavefield from the initial model to be used as a initial wavefield estimate. If this could be done effectively in the field data case, or if an equivalent initialization for the source formulation could be determined, it could significantly improve the speed and perhaps accuracy of this approach.

Currently, we estimate the weights for the regularization terms through a trial-and-error process, which may be untenable for larger problems. One objective for future work on this FWI strategy is to develop an approach for estimating these regularization terms in a consistent, computationally inexpensive way. Strength of regularization

CONCLUSIONS

In this report, we perform full waveform inversion on a subset of the data collected in the 2018 CaMI VSP survey. We implement one of the FWI strategies described in Keating et al. (2021) to minimize the complications associated with the near-surface in this problem. We recover a model in reasonable agreement with well-logs, and consistent with our prior expectation of a flatly layered medium. An application of the same approach to a larger subset of the available data from this survey (including both accelerometer and DAS data) is presented in Eaid et al. (2021b).

ACKNOWLEDGEMENTS

The authors thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19. Matt Eaid was partially supported by a scholarship from the SEG Foundation. Scott Keating was also supported by the Canada First Research Excellence Fund, through the Global Research Initiative at the University of Calgary. The data were acquired through a collaboration with the Containment and Monitoring Institute, Carbon Management Canada.

REFERENCES

Eaid, M., Keating, S., and Innanen, K. A., 2021a, Full waveform inversion of das field data from the 2018 cami vsp survey: CREWES Annual Report, **33**.

Eaid, M., Keating, S., and Innanen, K. A., 2021b, Processing of the 2018 cami vsp survey for full waveform inversion: CREWES Annual Report, **33**.

Hall, K., Bertram, K., Bertram, M., Innanen, K., and Lawton, D., 2018, CREWES 2018 multi-azimuth walk-away VSP field experiment: CREWES Research Reports, **30**, No. 16, 1–14.

Keating, S., Eaid, M., and Innanen, K. A., 2021, Effective sources: removing the near surface from the vsp fwi problem: CREWES Annual Report, **33**.

Keating, S., and Innanen, K. A., 2020, Parameter cross-talk and leakage between spatially-separatedunknowns in viscoelastic full waveform inversion, under revision: Geophysics.

Keating, S., Li, J., and Innanen, K. A., 2018, Viscoelastic fwi: solving for q_p, q_s, v_p, v_s, and density: CREWES Annual Report, **30**.

Mossop, G., and Shetsen, I., 1994, Geological atlas of the Western Canada Sedimentary Basin, `https://ags.aer.ca/reports/atlas-western-canada-sedimentary-basin`, accessed: 2021-08-30.