Implicit elastic full waveform inversion: application to the Snowflake dataset

Tianze Zhang, Xiaohui Cai, Daniel Trad, Kris Innanen

ABSTRACT

Carbon capture and storage (CCS) has risen as a critical research domain, aiming to reduce carbon dioxide emissions by confining the gas deep within subsurface reservoirs. Seismic data are indispensable in monitoring the sequestered carbon dioxide to ensure its confinement to intended zones and prevent encroachment into areas of potential risk. In 2018, the Consortium for Research in Elastic Wave Exploration Seismology executed a 3D walkaway-walkaround VSP survey, integrating both three-component accelerometers and DAS fibres. This investigation employs implicit full waveform inversion to determine the baseline model based on 2018's accelerometer data. This implicit elastic full waveform inversion harnesses neural networks to produce elastic models. The neural network's weights are optimized to generate refined elastic models that minimize data misfit, obviating the need for precise initial models. A comparison of inversion outcomes with well-log data is encouraging, and the alignment between synthetic and observed data further underscores its promise.

INTRODUCTION

Global warming poses one of the gravest threats to our planet. The overarching agreement among scientists is that human-made greenhouse gas emissions are the main culprits behind this warming trend. As these emissions rise, they float up to the atmosphere, capturing heat from the sun and upsetting the planet's climatic balance. If we neglect this issue, the continuous increase in these emissions might amplify solar heat retention, which could amplify extreme and regular weather disturbances globally. Among these emissions, CO_2, a byproduct of burning, is notably a dominant greenhouse contributor. Carbon sequestration and containment (CSC) is a promising field of study focused on curbing the harmful effects of CO2 in the atmosphere. This method aims to divert CO2 emissions from the atmosphere by utilizing advanced capture mechanisms during industrial processes. After securing the CO2, it's typically relayed through pipelines to a chosen location, where it's stored deep underground for an extended period. A paramount risk with CSC is the possibility of CO2 escaping these underground storage areas, gathering close to the ground level, and making its way back into the atmosphere. This escape would not only diminish CSC's efficacy but might also trigger an abrupt and devastating surge of CO2, presenting grave health hazards for communities nearby.

Advancing technologies to observe injected CO_2 is a primary objective in numerous CCS initiatives. The goal of such technologies is to trace the plume's trajectory over its lifespan, confirm its containment within the designated reservoir, and detect potential leaks before they escalate into significant issues. Seismic monitoring is a pivotal tool in ensuring CO2's secure confinement in deep underground reservoirs. When applying seismic monitoring to CCS endeavours, a foundational seismic assessment is carried out before introducing the fluid. This is followed by periodic monitoring at different fluid injection phases.

Assuming the seismic procedures for foundational and ongoing surveys are consistent, and there's little change in near-surface conditions over seasons, discrepancies in seismic data from one survey to another can be attributed to the CO_2 presence. These discrepancies in seismic readings help pinpoint the CO_2's position, spread, and characteristics, granting seismic monitoring the ability to produce detailed imagery of the stored fluid.

The Containment and Monitoring Institute (CaMI), operating under Carbon Management Canada (CMC) is dedicated to pioneering CO2 monitoring technologies at its Field Research Station in Brooks, Alberta. The purpose of the FRS revolves around crafting techniques to monitor the expansion of a CO2 plume housed in the water-saturated Basal Belly River sandstone layer of the Upper Cretaceous era, resting 300 meters below the surface. Aiding this mission, a foundational vertical seismic profile (VSP) survey was conducted in 2018, deploying accelerometers along with co-located distributed acoustic sensing (DAS) fibres in a surveillance well roughly 20 meters southwest from the main injection site.

2018 VSP SURVEY OF THE CONTAINMENT AN MONITORING INSTITUTE

The primary objective of the Containment and Monitoring Institute's Field Research Station, situated near Brooks, Alberta, revolves around pioneering methods and tech innovations to monitor CO_2 stored in the Basal Belly River Sandstone from the late Cretaceous period, situated 300 meters beneath the surface see Fiure 1. Of the diverse technologies suggested to achieve this aim, seismic monitoring stands out as a pivotal instrument anticipated for rendering detailed images of the CO_2 plume. In alignment with this seismic oversight, various acquisition instruments have been stationed at the FRS. This research hub encompasses three wells: the principal CO_2 injection well and two ancillary observational wells. The second observation well, informally dubbed the 'geophysics well,' continuously accommodates a linear DAS fibre in addition to a coiled fibre at a 30$\circ$ lead angle. In an affiliated study, data gleaned from the accelerometers and the linear DAS fibre underwent processing, setting the stage for FWI. Conversely, this document zooms in on the inversions originating from the DAS fibre data, individually and when integrated with accelerometer insights. Figure 2 showcases the VSP survey's source layout, accentuating the key source line in blue.

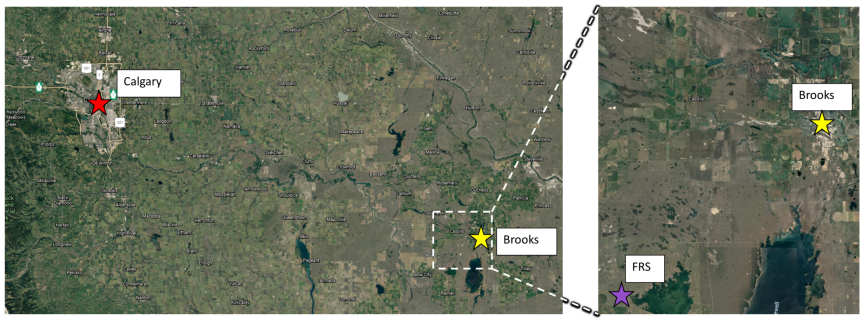

FIG. 1. Location of the CaMI FRS in Newell Country, to Calgary, Alberta, and Brooks, Alberta.

IMPLICIT ELASTIC FULL WAVEFORM INVERSION

In this study, we will employ the implicit full waveform inversion (IFWI) technique to conduct a full waveform inversion based on the CaMI.FRS 2018 dataset. The specific

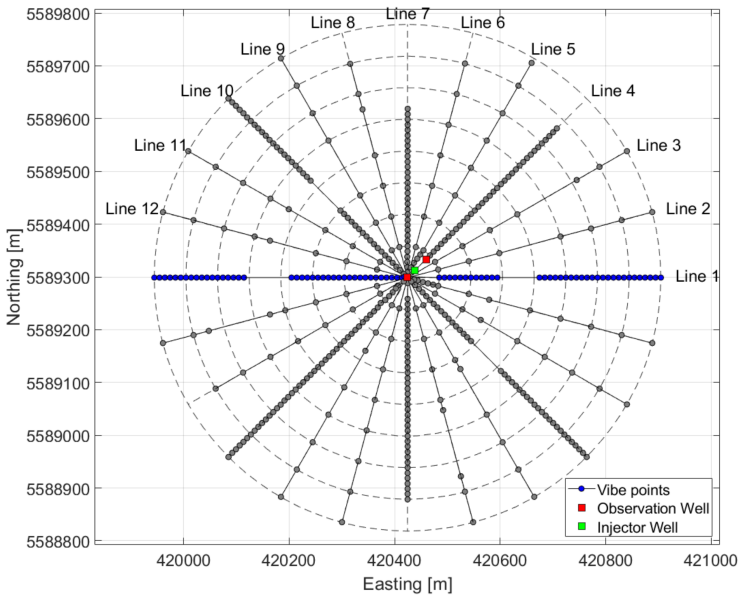

FIG. 2. Shot geometry of the CaMI-FRS 2018 3D walkaway-walkaround VSP. The blue circles represent shot point locations on source line 1, the red squares the locations of the two observation wells (Observation well 2 is at the center of the shot points), and the green squares are the location of the injector well. The dotted lines are 60-meter concentric circles centred on observation well 2.

IFWI approach adopted here is derived from the works of Sun et al. (2023) and Zhang et al. (2023). At its core, this method harnesses neural networks to produce the elastic models. This is succeeded by the application of a recurrent neural network-centric modeling procedure to yield synthetic data. Subsequently, an objective function gauges the divergence between this synthetic data and the actual observed data. The neural network's parameters are then fine-tuned, guiding it to generate elastic models that can further optimize and minimize the objective function.

The standout advantage of the IFWI method is its independence from precise initial models when computing FWI. Nevertheless, it does mandate the well-log data of the area under investigation. This data offers a comprehensive overview, particularly the mean and standard deviation of the desired elastic parameter. Such information ensures that the resulting elastic models yield values within an acceptable and plausible range. Also, the uncertainty quantification(UQ) of inversion results is relatively cheap. In this MLP-based EIFWI, we can use the dropout method, which randomly mutes the weights in the neural network during the forward passing of the neural network. In the BNN, the weights are regarded as the Gaussian distributed neural network, and each forward passing of the neural network will randomly realized the weights using the corresponding Gaussian distributed PDFs, all of which can provide computationally efficient methods for sampling models around the inversion results. A more detailed inversion process can be seen in Zhang et al. (2023).

DATA PRE-PROCESSING

The data explored in this chapter were procured using an Inova Geophysical UniVib. The device executed a linear sweep ranging from 1-150 Hz over 16 seconds, complemented

by 0.2-second half-cosine tapers and a listening period of three seconds. For every shot point, a pair of sweeps were conducted, combined in-field. Refer to Figures 3(a), and (b) for a depiction of the sweep, the theoretical Klauder wavelet—computed via the sweep's auto-correlation, respectively.

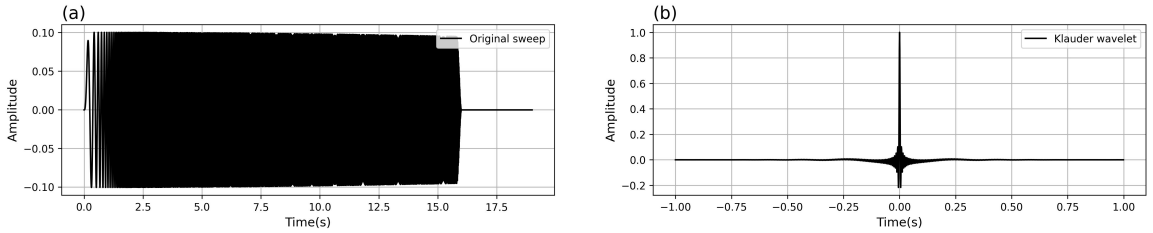

FIG. 3. The sweep and the Klauder wavelet. (a) The original sweep. (b) The Klauder wavelet is obtained from the correlation of the original sweep.

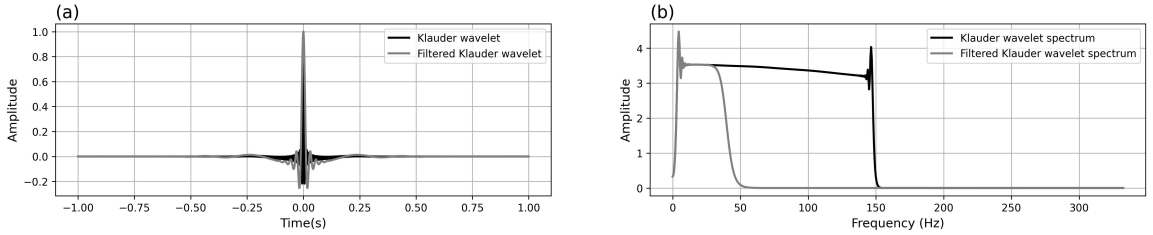

FIG. 4. Original Klauder wavelet and filtered Klauder wavelet and the corresponding spectrum. (a) The black line is the Klauder wavelet is obtained from the correlation of the original sweep. The gray line is the filtered Klauder wavelet. (b) The black line is the spectrum of the Klauder wavelet. The gray line is the spectrum of the filtered Klauder wavelet.

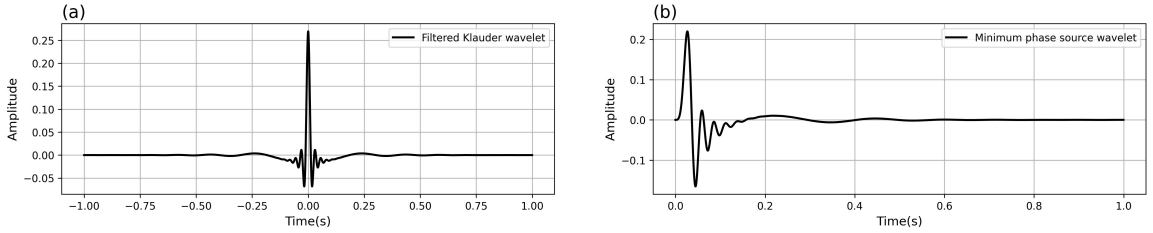

FIG. 5. (a) The zero phase Klauder wavelet. (b) The minimum phase wavelet transformed from Figure (a).

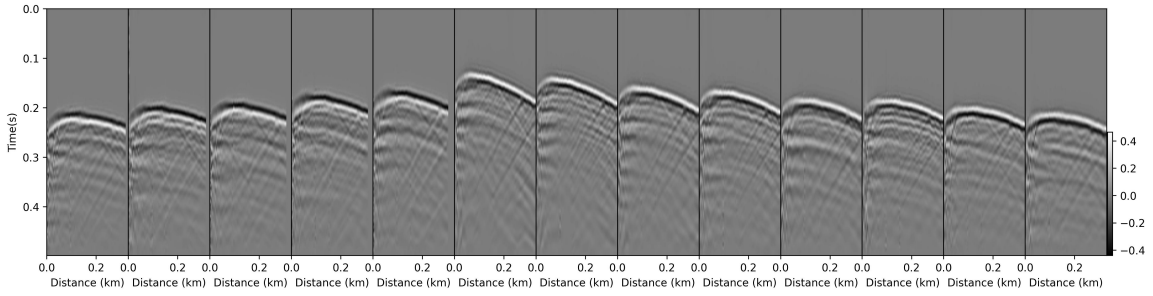

FIG. 6. The examples of the original horizontal components obtained from the field.

We further refined the theoretical Klauder wavelet by subjecting it to Butterworth filtering. Using an 8th-order filter with a cutoff frequency of 40 Hz, this process ensures

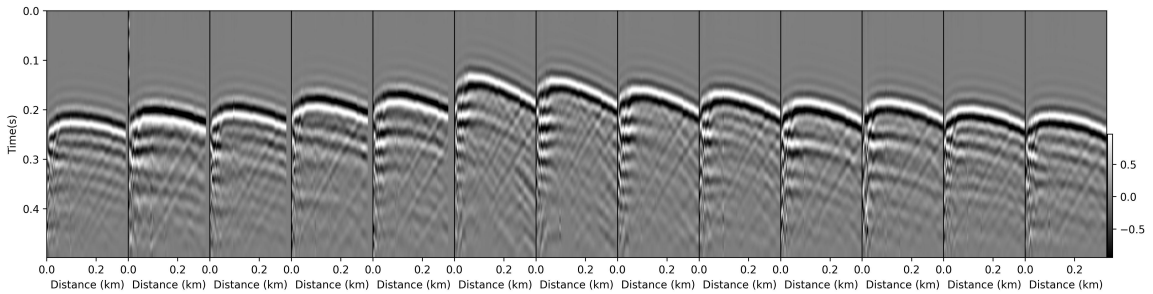

FIG. 7. The examples of the filtered horizontal components obtained used for inversion.

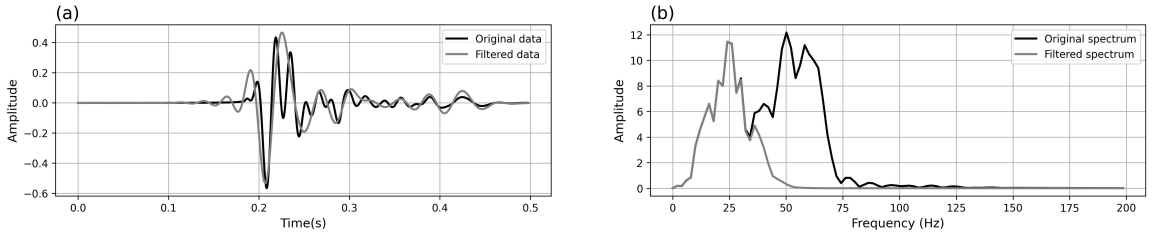

FIG. 8. A example of trace before and after the filtering process. (a) The solid black line is the original data, and the gray line is the filtered data. (b) The solid black line is the original spectrum, and the gray line is the filtered spectrum.

the finite-difference method employed in wavefield modeling experiences minimal numerical dispersion. Figures 4(a) and 4(b) depict the filtered Klauder wavelet and compare the spectrum profiles pre and post the filtering process, respectively. Leveraging the minimum phase assumption, I converted the zero-phase Klauder wavelet to its minimum phase counterpart. This transformed wavelet is presented in Figure 5(b). It's worth noting that this minimum phase wavelet, showcased in Figure 5(b), will serve as the source wavelet for inversion wavefield modeling. The observed data are also being filtered with the same Butterword filtering method, with an order of 8 and a cutoff frequency of $40Hz$. An example of the data before filtering and after filtering are plotted in Figure 6, and Figure 7, respectively. And an example of the trace, before and after the filtering process are plotted in Figure 8(a), and Figure 8(b), respectively.

MLP-BASED EIFWI FOR THE CAMI.FRS DATA

In this first test, we use the multiple layers perceptron as the neural network. The overall structure of the network is displayed in table 1. The input of the network is the coordinate information of the simulation model, with the dimension of $\mathbb{R}^{2\times n_z\times n_x}$. Then I permute the dimension of the data into $\mathbb{R}^{n_z\times n_x\times 2}$. In this table, I use the sinusoidal function as the activation function and 256 hidden layers. Thus after the first MLP hidden layer, the dimension of the data is changed into $\mathbb{R}^{n_z\times n_x\times 1024}$. This dimension of the data is kept until it reaches the last MLP layer, which changes the data into the dimension of $\mathbb{R}^{n_z\times n_x\times 3}$. The last dimension 3 indicates that I will output 3 parameter models to perform forward modeling. Then, I transform the dimension of the data back into $\mathbb{R}^{3\times n_z\times n_x}$, where the value of the first dimension is the number of the geophysical parameter models, and the second and the third dimension are the z and x coordinate grid points.

Table 1. **The structure of MLP used for training**

Layer name	Description	Weight dim.	Data dim.
Input	Coordinate in x and z direction	N/A	$\mathbb{R}^{2\times n_{\rm z}\times n_{\rm x}}$
Permute	Dimension permute	N/A	$\mathbb{R}^{n_{\rm z}\times n_x\times 2}$
Layer 1	FC layer	$\mathbf{w}_1\in\mathbb{R}^{2\times b_{\rm h}}$	$\mathbb{R}^{n_{\rm z}\times n_{\rm x}\times b_{\rm h}}$
Activation	Sinusoidal	N/A	$\mathbb{R}^{n_{\rm z}\times n_{\rm x}\times b_{\rm h}}$
Layer 2	FC layer	$\mathbf{w}_2\in\mathbb{R}^{b_{\rm h}\times b_{\rm h}}$	$\mathbb{R}^{n_{\rm z}\times n_{\rm x}\times b_{\rm h}}$
Activation	Sinusoidal	N/A	$\mathbb{R}^{n_{\rm z}\times n_{\rm x}\times b_{\rm h}}$
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
Layer N	FC layer	$\mathbf{w}_4\in\mathbb{R}^{b_{\rm h}\times b_{\rm h}}$	$\mathbb{R}^{n_{\rm z}\times n_{\rm x}\times b_{\rm h}}$
Output V_p	FC layer	$\mathbf{w}_{V_p}\in\mathbb{R}^{b_{\rm h}\times 1}$	$\mathbb{R}^{n_{\rm z}\times n_{\rm x}}$
Output V_s	FC layer	$\mathbf{w}_{V_s}\in\mathbb{R}^{b_{\rm h}\times 1}$	$\mathbb{R}^{n_{\rm z}\times n_{\rm x}}$
Output ρ	FC layer	$\mathbf{w}_\rho\in\mathbb{R}^{b_{\rm h}\times 1}$	$\mathbb{R}^{n_{\rm z}\times n_{\rm x}}$

I train the network as follows, as listed in algorithm 1. The network input is coordinate information $\mathcal{C}$, the source wavelet for each shot $\mathbf{f}$, the mean and standard deviation of the model parameter. First, I initialize the weights as determining values in the MLP, and I use the MLP network to predict the model parameters, which are V_p, V_s, and ρ. In this study, I use velocity parameterization. Second, I solve $\mathbf{F}(\mathbf{u}, N_{\rm mlp}(\mathcal{C};\mathbf{w})) = \mathbf{f}$ for $\mathbf{u}$, and the synthetic data is obtained with $\mathbf{d}_{\rm syn} = \mathbf{R}\mathbf{u}$. In this study, I solve $\mathbf{u}$ with the elastic wave equation based physics informed recurrent neural network (RNN), i.e., Zhang et al. (2020), Sun et al. (2020). Third, I calculate the residual with the ℓ_2 objective function and then calculate the gradients by using the automatic differential method. I update the weights $\mathbf{w}$ in $N_{\rm mlp}(\mathcal{C};\mathbf{w})$ with the gradient-based optimization method: $\mathbf{w} = \mathbf{w} - \eta\triangle_{\mathbf{w}}\phi_{\rm EIFWI}$, where η is the step length, and $\triangle_{\mathbf{w}}\phi_{\rm EIFWI}$ is the gradient of $\mathbf{w}$ with respect to the objective function $\phi_{\rm EIFWI}$. If biases are included in the calculation, then the bias vector will be updated similarly to the weights. This training process distinguishes it from conventional FWI, in which the elastic model parameters are solved explicitly and are based on updates of an explicitly selected initial model.

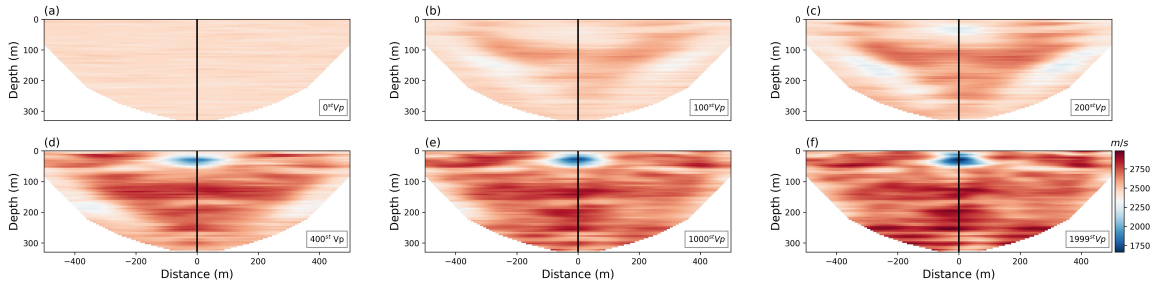

FIG. 9. MLP-based IFWI results for the V_p model. (a)-(f) are the inversion results for V_p at the 1^{st}, 100^{st}, 200^{st}, 400^{st}, 1000^{st}, 2000^{st} iteration.

I use Adam's algorithm as our optimization method (Kingma and Ba, 2014). I use the

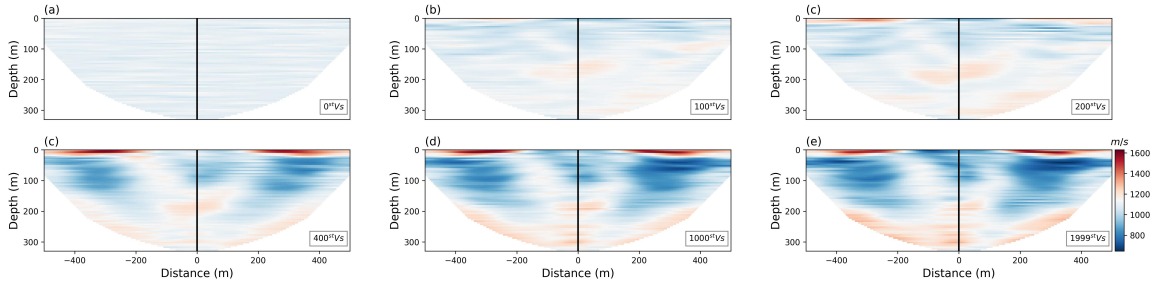

FIG. 10. MLP-based IFWI results for the V_s model. (a)-(f) are the inversion results for V_s at the 1^{st}, 100^{st}, 200^{st}, 400^{st}, 1000^{st}, 2000^{st} iteration.

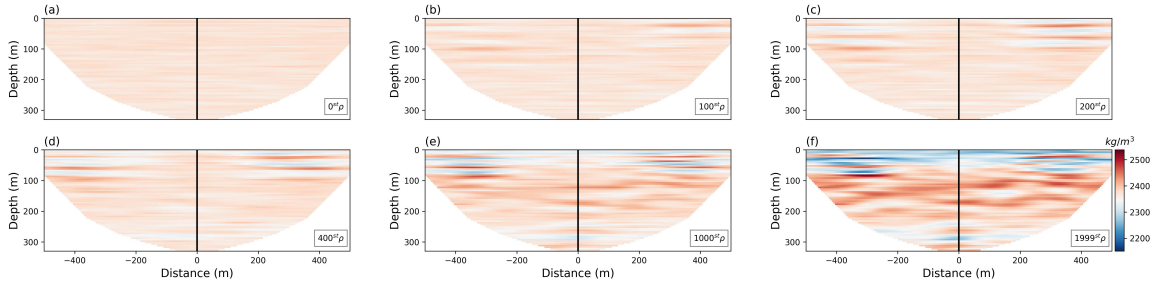

FIG. 11. MLP-based IFWI results for the ρ model. (a)-(f) are the inversion results for ρ at the 1^{st}, 100^{st}, 200^{st}, 400^{st}, 1000^{st}, 2000^{st} iteration.

learning rate 0.001, along with Adam's hyper-parameters $\beta_1 = 0.9$, and $\beta_2 = 0.990$, $\epsilon = 10^{-8}$ to update the weights in the MLP. In this test, the MLP network consists of three fully connected layers, each with 1024 weights. The activation functions I use in this test are sinusoidal activation functions. Figure 9(a), Figure 10(a), and Figure 11(a) are the V_p, V_s, and ρ model generated with the MLP network after the first iteration scaled with the mean and the standard deviation of the V_p, V_s, and ρ given by well-log. The well-log profiles for V_p, V_s, and ρ are plotted in Figure 12 I can see no meaningful geological information for the elastic models at the first iteration, and the prediction results are filled with values around the mean value. This indicates that I am starting the inversion without using accurate initial models. Figure 9(a)-(f) showcases how the prediction of the V_p changes as the iteration number increaes. Figure 9(f) is the final prediction result of V_p. The inversion results for parameters V_s and ρ are plotted in Figure 10(a)-(f) and Figure 11(a)-(f), respectively. Inversions for V_s and ρ both show similar convergence patterns to the V_p case. MLP-based EIFWI gradually updates the model from top to bottom, converging to the correct inversion results, and no explicitly accurate initial models are required.

The results of the MLP-based IFWI, when juxtaposed with the well-log, are illustrated in Figure 12. Specifically, Figure 12(a), (b), and (c) depict the vertical profile comparisons of the MLP-based EIFWI results at the well log's location. The red lines signify the initial values, while the green lines represent the IFWI inversion results after 1000 iterations. Meanwhile, the blue lines showcase the final IFWI predictive outcomes. It's imperative to note that the red line suggests that EIFWI's starting point isn't an optimal kickoff for FWI. Instead, it is more of a random model with values gravitating around the mean of the well-log. The visualization underscores that, with increasing iterations, the FWI inversion results (evident in the green and blue lines) tend to align more closely with the smoothed well-log

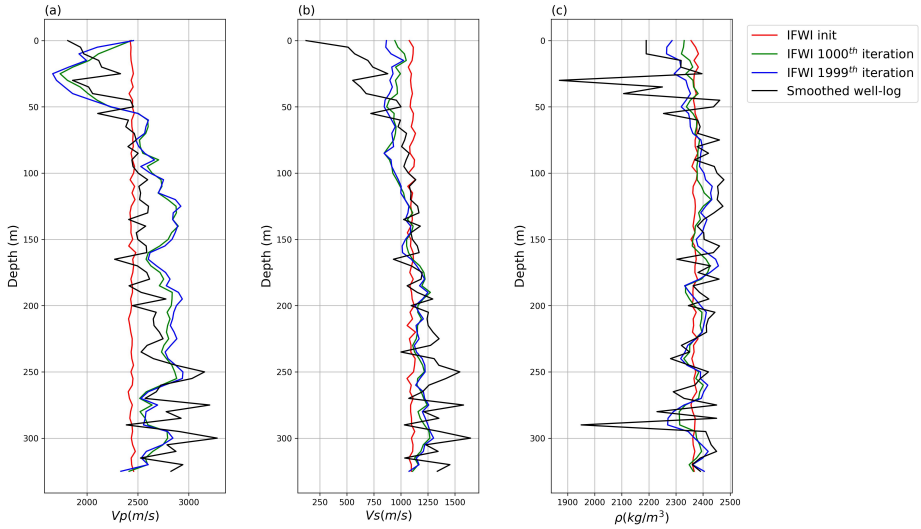

FIG. 12. MLP-based IFWI results in comparison with the well-log. Red lines are the initial value, and green lines are the IFWI inversion results at 1000 iterations. The blue lines are the final IFWI prediction results.

curve (depicted by the black line). This convergence hints at the iterative methodology's efficacy in honing the model, rendering it more congruent with the subsurface medium's intrinsic physical attributes.

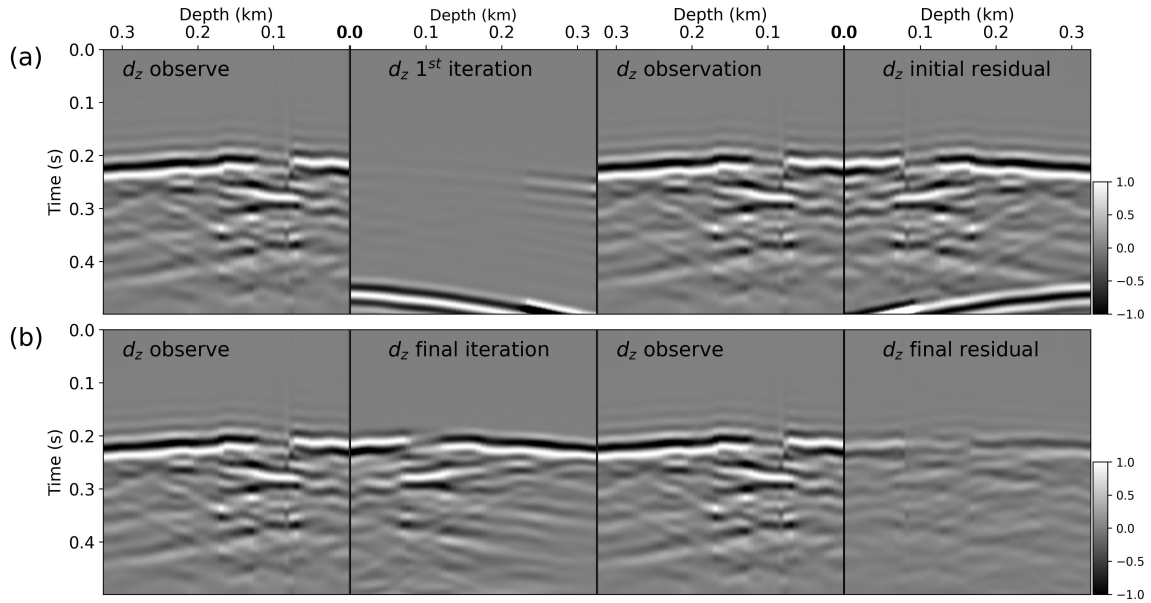

FIG. 13. MLP-based IFWI vertical component data comparison. (a) Observe accelerometer data of the vertical component, with amplitude normalization, initial IFWI vertical component, and the initial residual of the Vertical components are plotted from left to right. (b) Observe accelerometer data of the vertical components, final IFWI vertical components and the final residual of the vertical components are plotted from left to right.

To showcase the efficacy of the IFWI, I present a detailed examination of data convergence. Figure 13 encompasses two sets of panels, designated as (a) and (b), which each depict distinct phases of the waveform inversion procedure. Figure 13(a) presents the observed data (d_z observe), the results from the initial iteration (d_z 1st iteration), and the

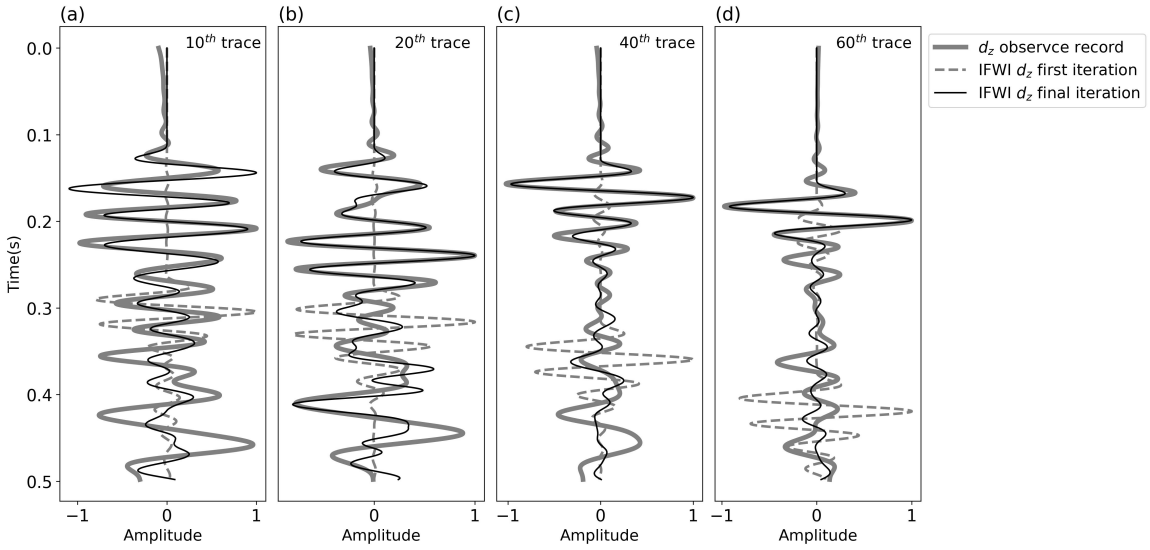

FIG. 14. MLP-based IFWI vertical component data comparison. The discrete traces selected from the horizontal accelerometer observations (solid gray lines) with amplitude normalization, the initial model synthetic shot (dashed gray lines), and inversion results synthetic (solid black lines) are plotted for comparison.

initial residuals, which quantify the deviation between the observed and calculated data (d_z initial residual). This comparative presentation facilitates an evaluation of the initial alignment between the model and the observed data. Conversely, Figure 13(b) showcases the observed data (d_z observe), the outcomes from the final iteration (d_z final iteration), and the final residuals (d_z final residual). This juxtaposition illuminates the enhancements realized through the FWI procedure, underscored by the anticipated reduction in residuals in the final iteration results. Figure 14 further elucidates the FWI results by showcasing individual seismic traces at specific intervals along the survey line. Four panels, labeled from (a) to (d), represent the 10th, 20th, 40th, and 60th traces, respectively. Each panel illustrates three seismic traces: The observed data is shown as a solid gray line. The initial iteration of FWI is depicted as a dashed line. The final iteration of FWI is displayed as a solid black line. The convergence of the dashed and solid black lines towards the solid gray line indicates the extent to which the FWI has successfully adjusted the model to fit the observed data. The closer the final iteration trace is to the observed data trace, the more accurate the resulting seismic model is considered to be. Similar observations can be seen in the comparison of the horizontal data comparisons, which are plotted in Figures 15, and Figure 16.

The post-trained MLP-based model was forwarded 500 times, with a dropout rate of 0.05, yielding 500 V_p, V_s, and ρ models. The mean and standard deviation models are depicted in Figure 17. In the top panel, Figure 17(a) through Figure 17(c) display the average inversion outcomes for V_p, V_s, and ρ respectively. The bottom row, comprising Figure 17(d) to 17(f), outlines the uncertainty quantification for these physical properties. Figure 17(d) illustrates the standard deviation in P-wave velocity (Std Vp), where darker blues indicate zones of minimized uncertainty, while lighter hues denote increased uncertainty. This distribution hints that the maximum certainty lies at superficial depths and closer to the midpoint, possibly attributed to improved data collection or sensitivity in these areas. Figure

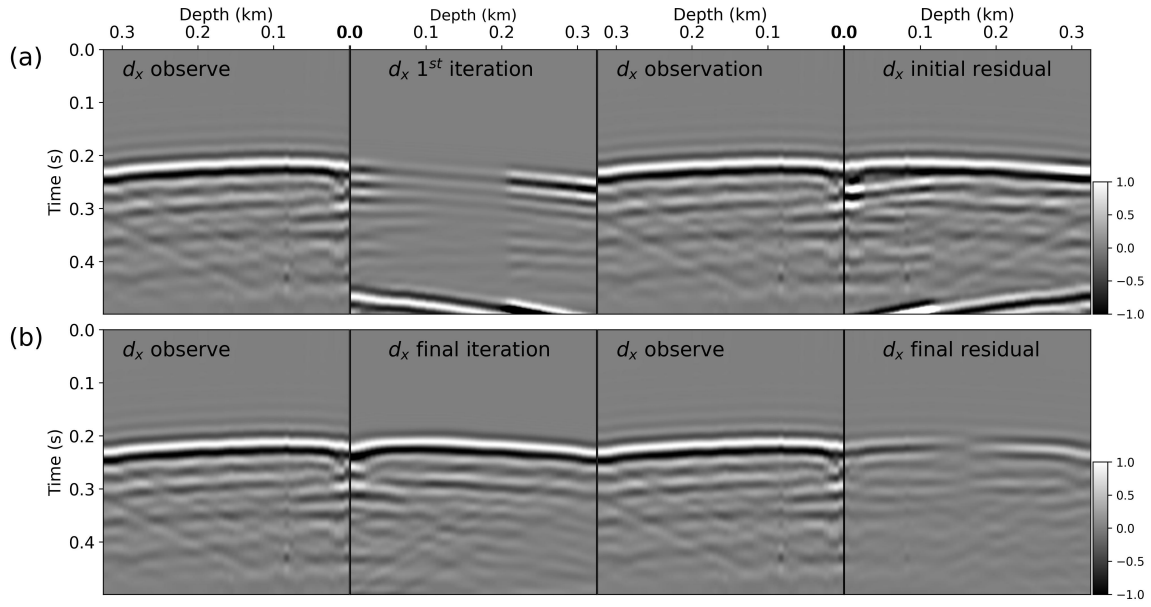

FIG. 15. MLP-based IFWI horizontal component data comparison. (a) Observe accelerometer data of the horizontal component, with amplitude normalization, initial IFWI horizontal component, and the initial residual of the horizontal components are plotted from left to right. (b) Observe accelerometer data of the horizontal components, final IFWI horizontal components and the final residual of the horizontal components are plotted from left to right.

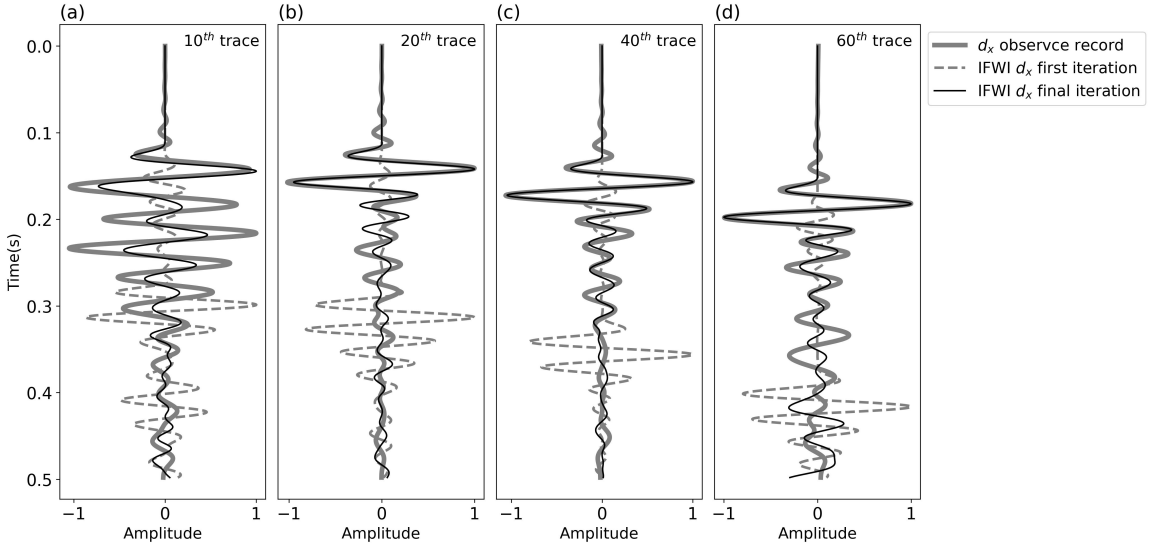

FIG. 16. MLP-based IFWI horizontal component data comparison. The discrete traces selected from the horizontal accelerometer observations (solid gray lines) with amplitude normalization, the initial model synthetic shot (dashed gray lines), and inversion results synthetic (solid black lines) are plotted for comparison.

17(e) represents the standard deviation in S-wave velocity (Std Vs), exhibiting an expanded uncertainty pattern that intensifies as one moves deeper, suggesting reduced sensitivity in these regions. Figure 17(f) portrays the uncertainty in density (Std ρ), with the overarching trend indicating ubiquitous uncertainty throughout the model, interspersed with pockets of reduced uncertainty, potentially corresponding to zones with more accurate velocity predic-

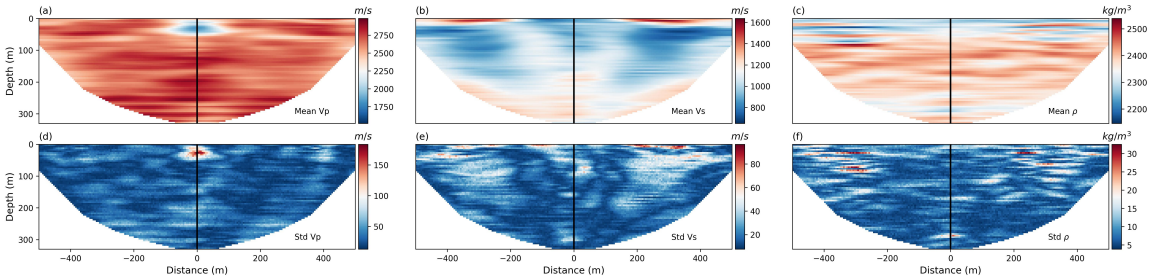

FIG. 17. MLP-based IFWI uncertainty quantification. The UQ of the IFWI results is obtained by forward passing the well-trained network 1000 times and calculating the mean and the standard deviation of the predicted models. (a), (b) and (c) are the mean model of the predicted V_p, V_s, and ρ model, respectively. (d), (e) and (f) are the standard deviation of the predicted models for V_p, V_s, and ρ, which will be regarded as the uncertainty quantification of the inversion results.

tions or heightened data constraints. This side-by-side presentation of FWI results not only offers insights into potential subterranean structures but also affords tactical recommendations for enhancing data acquisition and processing methodologies, thereby augmenting the trustworthiness and precision of inversion-based revelations. Figure 18 shows the reverse time migration results using MLP-based IFWI. The V_p, V_s, and ρ models used for reverse time migration are the mean prediction results given by the well-trained MLP. From the PP coefficient model, we can see that the main reflection, at around 150m and 300m, can be seen, demonstrating the validity of the proposed method.

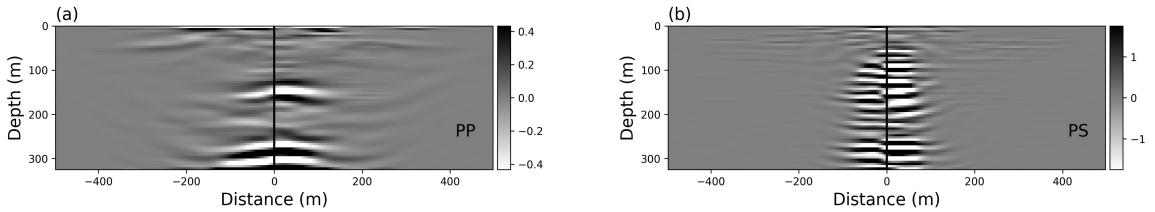

FIG. 18. Reverse time migration for the PP and PS reflectivity using MLP-based IFWI results. (a) PP reflection coefficient. (b)PS reflection coefficient models.

CONCLUSION

The Multi-Layer Perceptron (MLP) based Implicit Full Waveform Inversion (IFWI) is employed to analyze the CAMI.FRS 2018 dataset. Both the data and the source undergo a filtering process up to a frequency of 40 Hz using the Butterworth low-pass filter prior to the EIFWI computation. Accurate preliminary models are not a prerequisite for obtaining these promising inversion results. The MLP-based EIFWI showcases commendable data alignment. There is a notable reduction in the initial data residual upon implementing the MLP-based EIFWI update. The inversion's vertical profile aptly represents the primary characteristics of the well-log data, affirming the reliability of the inversion outcomes. Furthermore, the Reverse Time Migration (RTM) imaging of the inversion results distinctly indicates the reservoir's location, reinforcing the credibility of the inversion results.

ACKNOWLEDGMENTS

The sponsors of CREWES are gratefully thanked for continued support. This work was funded by CREWES industrial sponsors, NSERC (Natural Science and Engineering

Research Council of Canada) through the grant CRDPJ 543578-19.

REFERENCES

Kingma, D. P., and Ba, J., 2014, Adam: A method for stochastic optimization: arXiv preprint arXiv:1412.6980.

Sun, J., Innanen, K., Zhang, T., and Trad, D., 2023, Implicit seismic full waveform inversion with deep neural representation: Journal of Geophysical Research: Solid Earth, **128**, No. 3, e2022JB025,964.

Sun, J., Niu, Z., Innanen, K. A., Li, J., and Trad, D. O., 2020, A theory-guided deep-learning formulation and optimization of seismic waveform inversiontheory-guided dl and seismic inversion: Geophysics, **85**, No. 2, R87–R99.

Zhang, T., Innanen, K. A., Sun, J., and Trad, D. O., 2020, Numerical analysis of a deep learning formulation of multi-parameter elastic full waveform inversion, *in* SEG International Exposition and Annual Meeting, OnePetro.

Zhang, T., Sun, J., Trad, D., and Innanen, K., 2023, Multilayer perceptron and bayesian neural network based elastic implicit full waveform inversion: IEEE Transactions on Geoscience and Remote Sensing.