Elastic full waveform inversion with extrapolated low-frequency data

Hongyu Sun and Laurent Demanet, Earth Resources Laboratory, Massachusetts Institute of Technology*

SUMMARY

Full waveform inversion (FWI) strongly depends on an accurate starting model to succeed. This is particularly true in the elastic regime: The cycle-skipping phenomenon is more severe in elastic FWI compared to acoustic FWI, due to the short S-wave wavelength. In this note, we extend our work on extrapolated FWI (EFWI) by proposing to synthesize the low frequencies of multi-component elastic seismic records, and use those "artificial" low frequencies to seed the frequency sweep of elastic FWI. By leveraging deep learning technologies, we separately train two neural networks to extrapolate the low frequencies of elastic data (vertical and horizontal components of particle velocity), respectively. Numerical example on the Marmousi2 model shows that the 2-4Hz low frequency data extrapolated from band-limited data above 4Hz provide good starting models for elastic FWI of P-wave and S-wave velocities. Additionally, we study the generalization ability of the proposed neural network over different physical models. For elastic test data, collecting the training dataset by elastic simulation shows better extrapolation accuracy than acoustic simulation, i.e., a smaller generalization gap.

INTRODUCTION

Full waveform inversion is well-known for its potential to provide quantitative earth properties of complex subsurface structures. Acoustic FWI is widely used and has been successfully applied to real seismic data. However, most seismic data have strong elastic effects. Acoustic approximation is insufficient to estimate correct reflections and introduces additional artifacts to FWI results (Plessix et al., 2013; Stopin et al., 2014). Therefore, there is an urgent need to develop a robust elastic FWI method for high-resolution earth model building.

Theoretical studies have shown the ability of elastic FWI to retrieve realistic properties of the subsurface (Tarantola, 1986; Mora, 1987). However, it has difficulty handling real data sets. Elastic FWI is very sensitive to accuracy of starting model, correct estimation of density, proper definition of multi-parameter classes, and noise level (Brossier et al., 2010). The complex wave phenomena in elastic wavefields bring new challenges to FWI.

Among many factors that affect the success of elastic FWI, the lowest starting frequency is an essential one, given an accurate starting model is unavailable. Compared to acoustic FWI, the nonlinearity of elastic FWI is more severe due to the short S-wave propagating wavelength. Therefore, elastic FWI always requires an lower starting frequency compared to acoustic FWI. Additionally, the parameter cross-talk problem exists in elastic FWI and becomes more pronounced at higher frequencies, so ultra-low frequencies are required for a successful inversion of S-wave velocity and density.

In synthetic studies of elastic FWI, Brossier et al. (2009) invert the overthrust model (Aminzadeh et al., 1997) from 1.7Hz. Brossier et al. (2010) invert the Valhall model (Sirgue et al., 2009) from 2Hz. Both inversion workflows start from Gaussian smoothing of true models. Moreover, Choi et al. (2008) invert the Marmousi2 model (Martin et al., 2006) using a velocity-gradient starting model but a very low frequency (0.16Hz). For a successful inversion of the Marmousi2 density model, Köhn et al. (2012) use 0-2Hz in the first stage of multi-scale FWI. Jeong et al. (2012) invert the same model from 0.2Hz.

Only few applications of elastic FWI to real data sets are reported (Crase et al., 1990; Sears et al., 2010; Marjanović et al., 2018). Vigh et al. (2014) use 3.5Hz as the starting frequency of elastic FWI given that the initial models are accurate enough. Raknes et al. (2015) apply 3D elastic FWI to update P-wave velocity and obtain S-wave velocity and density using empirical relationships. Borisov et al. (2020) perform elastic FWI involving surface waves in the band of 5-15Hz for a land data set.

New developments in acquisition enhance the recent successes of FWI by measuring data with lower frequencies and longer offsets (Mahrooqi et al., 2012; Brenders et al., 2018). However, only acoustic FWI is applied to the land data set with low frequencies down to 1.5 Hz (Plessix et al., 2012). In addition to the expensive acquisition cost for the low-frequency signals, direct use of the field low-frequency data requires dedicated pre-processing steps, including travel-time tomography, for an accurate enough model to initialize FWI. The final inversion results strongly rely on the starting tomography model. Hence, retrieving reliable low-frequency data remains challenging for elastic FWI to relieve its dependency on starting models.

Deep learning is an emerging technology in many aspects of exploration geophysics. In seismic inversion, several groups have experimented with directly mapping data to model using deep learning (Araya-Polo et al., 2018; Yang and Ma, 2019). Other groups use deep learning as a signal processing step to acquire reasonable data for inversion. For instance, Li et al. (2019) use deep learning to remove elastic artifacts for acoustic FWI. Sun and Demanet (2018); Hu et al. (2019); Ovcharenko et al. (2019) propose different architecture of convolutional neural networks (CNN) to extrapolate the missing low-frequency data from band-limited recordings.

In this note, we extend our workflow of extrapolated FWI with deep learning (Sun and Demanet, 2020) into elastic regime. We train two separate neural networks, one to predict the low-frequency data of the horizontal components (v_x) and one to predict the low frequencies of the vertical components (v_y). The extrapolated low frequency data are used to initialize elastic FWI from a crude starting model. Moreover, to investigate the generalization ability of neural networks over different physical models, we compare the extrapolation results of

the neural networks trained on elastic data and acoustic data to predict the elastic low-frequency data.

METHOD

A brief review of elastic FWI

The elastic FWI is implemented in the time domain to invert the P-wave velocities ($\mathbf{v}_p$) and S-wave velocities ($\mathbf{v}_s$) simultaneously. The object function E is formulated as

$$E = \frac{1}{2}\delta\mathbf{d}^T\delta\mathbf{d} = \frac{1}{2}\sum_s\sum_r\int[\mathbf{u}_{cal}-\mathbf{u}_{obs}]^2dt, \qquad (1)$$

where $\mathbf{d}$ are the residuals between observed wavefields $\mathbf{u}_{obs}$ and calculated wavefields $\mathbf{u}_{cal}$. In 2D, both $\mathbf{u}_{obs}$ and $\mathbf{u}_{cal}$ contain v_x and v_y components of elastic wavefields. The gradient $\frac{\partial E}{\partial \mathbf{m}}$ of model parameter $\mathbf{m}$ is calculated in terms of $\mathbf{v}_p$ and $\mathbf{v}_s$ using the velocity-stress formula of the elastic wave equation (Köhn et al., 2012). The starting models $\mathbf{m_0}$ are updated using L-BFGS method (Nocedal and Wright, 2006).

CNN architecture for low-frequency extrapolation

We use the architecture in Sun and Demanet (2020). The output and input are the same seismic recording in the low and high frequency band, respectively. Although only one trace is plotted in Figure 1, the proposed neural network can easily explore multiple traces of the shot gather for multi-trace extrapolation by increasing the size of kernel from 200×1 to $200\times ntr$, where ntr is the number of the input traces. In 2D, the elastic data contain horizontal and vertical components. Therefore, we separately train the same neural network twice on two different training datasets: one contains v_x and the other contains v_y. In this way, we are able to predict the low-frequency signals of both components of the elastic data with enough accuracy for elastic FWI.

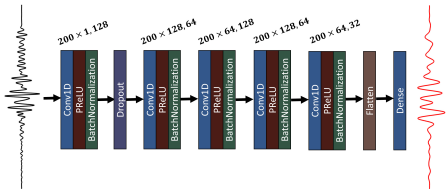

Figure 1: The architecture of CNN (Sun and Demanet, 2020).

Training and test datasets

The training and test datasets are simulated on the elastic training and test models with three parameters: $\mathbf{v}_p$, $\mathbf{v}_s$ and ρ. The Marmousi2 model is referred to as the test model (Figure 2). The training models (Figure 3) are six batches randomly extracted from the Marmousi2 model. The size of each model is 500×174 with a grid spacing of 20m, including a water layer on the top of each model with a depth of 460m. 100 shots are excited evenly from 800m to 8720m in the water layer at the same depth of 40m. A Ricker wavelet with a dominant frequency of 10Hz is used as the source signal. 400 receivers are placed under the water layer with a depth of 460m to record v_x and v_y of the elastic wavefields. The sampling rate and the recording time is 0.02s and 6s, respectively. Two training datasets are collected to separately process each component of the 2D elastic data. By trace-by-trace extrapolation setup, there are $6\times 100\times 400 = 240{,}000$ training samples in each training dataset and $1\times 100\times 400 = 40{,}000$ test samples in each test dataset. Each sample in the training and test dataset is separated into a low-frequency signal and a high-frequency signal using a smooth window in the frequency domain. Then, the time series in the high-frequency band is fed into the neural network to predict the low-frequency time series.

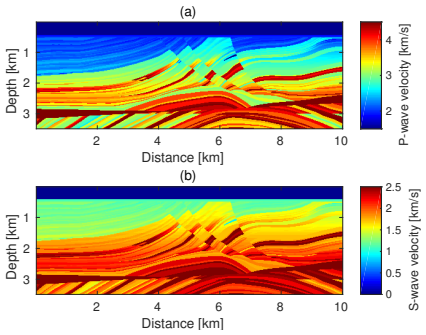

Figure 2: The Marmousi2 model (test model): (a)$\mathbf{v}_p$ and (b)$\mathbf{v}_s$.

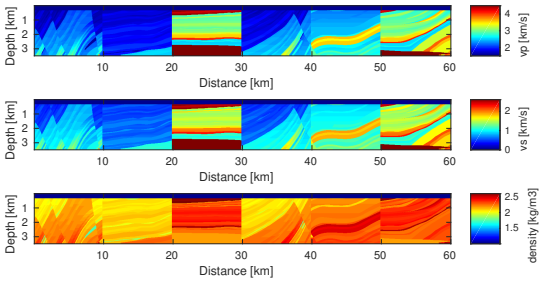

Figure 3: The training models randomly extracted from the Marmousi2 model. Each training model contains three parameters: $\mathbf{v}_p$, $\mathbf{v}_s$ and ρ.

NUMERICAL EXAMPLES

We first extrapolate the low frequency data below 5Hz on the Marmousi2 model (Figure 2) using 5-25Hz band-limited data. The neural networks are trained using ADAM method with a mini-batch of 32 samples. We refer readers to Sun and Demanet (2020) for more details about training. Figures 4(a) and 4(b) show the training process over 40 epochs. Figure 5 shows the extrapolation results of both v_x and v_y where the source is located at 7.04km. Figures 6(a) and 6(b) compare the amplitude and phase spectrum of v_y and v_x at $x = 6.18$km among the band-limited recording ($5.0-25.0$Hz), the full-band recording with true and predicted low frequencies ($0.1-5.0$Hz). Both neural networks can successfully recover the low frequencies of v_x and v_y recordings with satisfactory accuracy.

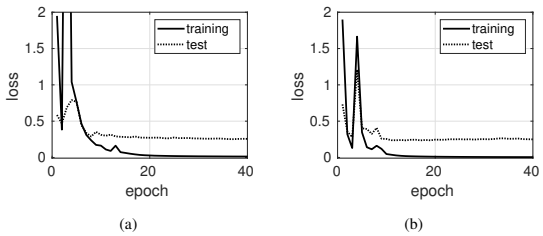

Figure 4: The learning curves of the CNN trained to extrapolate the low frequencies of (a) v_y and (b) v_x of the band-limited elastic recordings

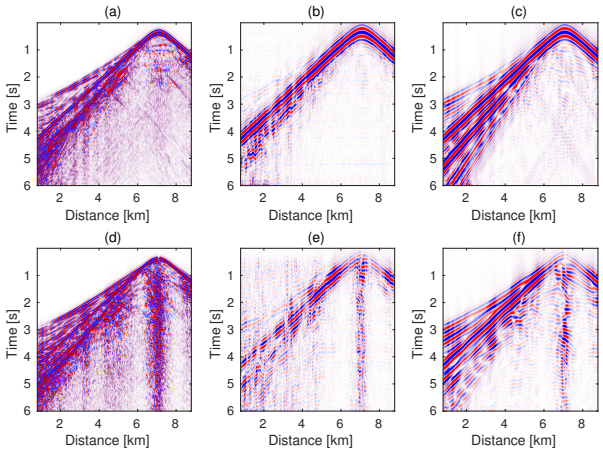

Figure 5: The extrapolation result on the Marmousi2 model: comparison among the (a) band-limited recordings ($5.0-25.0$Hz), (b) predicted and (c) true low-frequency recordings ($0.1-5.0$Hz) of v_y and (d) band-limited recordings ($5.0-25.0$Hz), (e) predicted and (f) true low-frequency recordings ($0.1-5.0$Hz) of v_x.

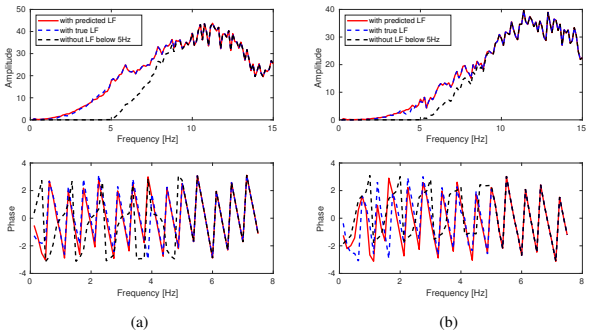

Figure 6: Extrapolation results of CNN *trained on elastic data*: comparison of the amplitude and phase spectrum of (a) v_y and (b) v_x at $x = 6.18$km among the band-limited recording ($5.0-25.0$Hz), the recording ($0.1-25.0$Hz) with true and predicted low frequencies ($0.1-5.0$Hz).

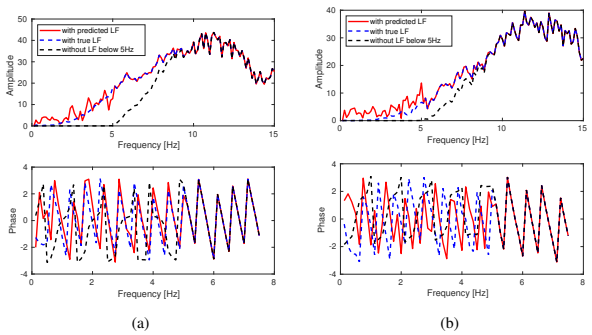

Figure 7: Extrapolation results of CNN *trained on acoustic data*: comparison of the amplitude and phase spectrum of (a) v_y and (b) v_x at $x = 6.18$km among the band-limited recording ($5.0-25.0$Hz), the recording ($0.1-25.0$Hz) with true and predicted low frequencies ($0.1-5.0$Hz).

To study the generalization ability of the proposed neural network over different physical models, we train the same neural network on acoustic training dataset and predict the low frequencies of both v_x and v_y in the same elastic test dataset. The acoustic training dataset is simulated using the acoustic wave equation on only the P-wave velocity model. Figures 7(a) and 7(b) show the amplitude and phase spectrum of v_y and v_x at $x = 6.18$km after training with the same procedure. Compared with the results in Figure 6, the extrapolation accuracy of the same trace in the test data is poorer on acoustic training dataset than elastic training dataset. This is an indicator that the neural network is difficult to generalize to different physical models.

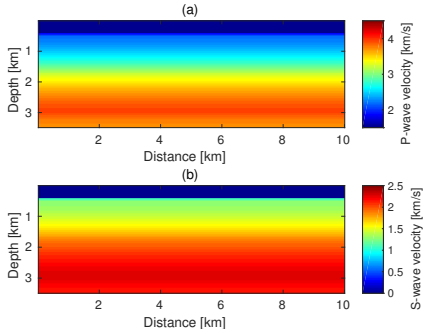

Figure 8: The initial models for elastic FWI: (a)$\mathbf{v}_p$ and (b)$\mathbf{v}_s$.

Finally, we perform EFWI using 4-20Hz band-limited data on the Marmousi2 model. Different from previous examples, a free surface boundary condition is applied to the top of the model. Starting from the crude initial models in Figure 8, Figure 9 and Figure 10 show the resulting inversion models after 30 iterations using extrapolated and true 2-4Hz low-frequency data, respectively. Then the inversion is continued using a multiscale method with the band-limited data (4-6Hz, 4-10Hz and 4-20Hz). After 80 iterations, the inversion result started from 2-4Hz extrapolated data (Figure 11) is very close to the result started from 2-4Hz true data (Figure 12). Instead, elastic FWI directly starting from the crude initial models using the band-

limited data shows large errors (Figure 13).

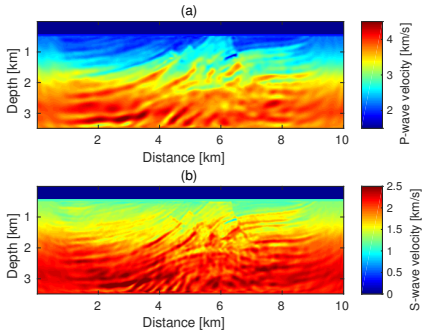

Figure 9: The inverted (a)$\mathbf{v}_p$ and (b)$\mathbf{v}_s$ models using 2-4Hz extrapolated low-frequency data.

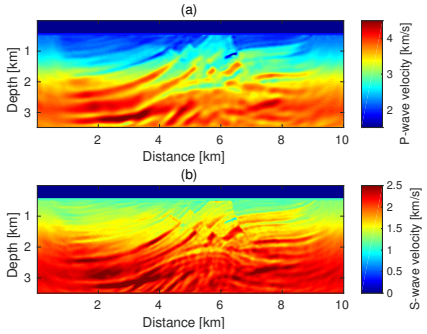

Figure 10: The inverted (a)$\mathbf{v}_p$ and (b)$\mathbf{v}_s$ models using 2-4Hz true low-frequency data.

CONCLUSION

To relieve the dependency of elastic FWI on starting models, low-frequency extrapolation of multi-component seismic recordings is implemented to computationally recover the missing low frequencies from band-limited elastic data. By training the neural network twice, once with a dataset of horizontal components and once with a dataset of vertical components, we can extrapolate the low frequencies of multi-component band-limited recordings separately. The extrapolated 0-5Hz low frequencies match well with the true low-frequency data on the Marmousi2 model. Elastic FWI using 2-4Hz extrapolated data shows similar results to the true low frequencies. The accuracy of the extrapolated low frequencies is enough to provide low-wavenumber starting models for elastic FWI on band-limited data above 4Hz.

The generalization ability of the neural network over different physical models is studied in this note. The neural network trained on purely acoustic data shows larger prediction error on elastic test dataset compared to the neural network trained on elastic data. Therefore, collecting more realistic elastic training dataset will help to process the field data with strong elastic effects.

ACKNOWLEDGMENTS

The authors thank Total SA for support. LD is also supported by AFOSR grant FA9550-17-1-0316. Tensorflow (Abadi et al., 2015) and Keras (Chollet et al., 2015) are used for deep learning. Elastic FWI in this note is implemented using the open source code DENISE (https://github.com/daniel-koehn/DENISE-Black-Edition).

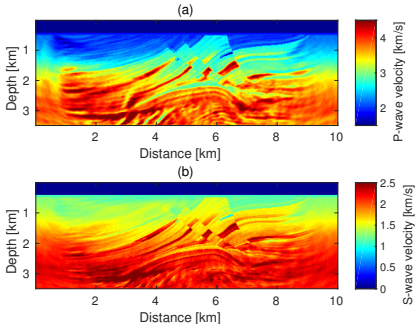

Figure 11: The inverted (a)$\mathbf{v}_p$ and (b)$\mathbf{v}_s$ models started from 2-4Hz extrapolated low-frequency data.

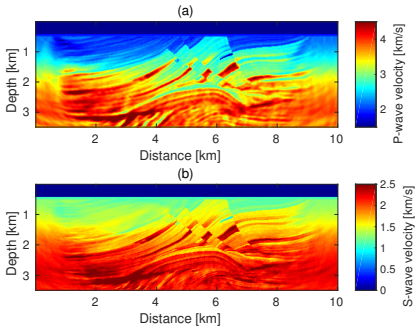

Figure 12: The inverted (a)$\mathbf{v}_p$ and (b)$\mathbf{v}_s$ models started from 2-4Hz true low-frequency data..

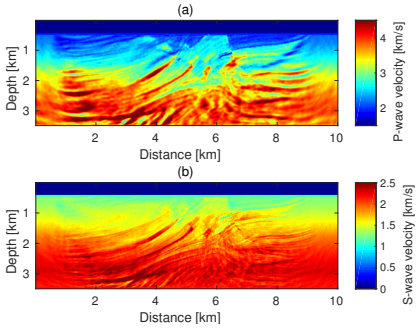

Figure 13: The inverted (a)$\mathbf{v}_p$ and (b)$\mathbf{v}_s$ using only the band-limited data.

REFERENCES

Abadi, M., A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, 2015, TensorFlow: Large-scale machine learning on heterogeneous systems. (Software available from tensorflow.org).

Aminzadeh, F., B. Jean, and T. Kunz, 1997, 3-d salt and overthrust models: Society of Exploration Geophysicists.

Araya-Polo, M., J. Jennings, A. Adler, and T. Dahlke, 2018, Deep-learning tomography: The Leading Edge, **37**, 58–66.

Borisov, D., F. Gao, P. Williamson, and J. Tromp, 2020, Application of 2d full-waveform inversion on exploration land data: Geophysics, **85**, R75–R86.

Brenders, A., J. Dellinger, C. Kanu, Q. Li, and S. Michell, 2018, The wolfspar® field trial: Results from a low-frequency seismic survey designed for fwi, *in* SEG Technical Program Expanded Abstracts 2018: Society of Exploration Geophysicists, 1083–1087.

Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2d elastic frequency-domain full-waveform inversion: Geophysics, **74**, WCC105–WCC118.

———, 2010, Which data residual norm for robust elastic frequency-domain full waveform inversion?: Geophysics, **75**, R37–R46.

Choi, Y., D.-J. Min, and C. Shin, 2008, Frequency-domain elastic full waveform inversion using the new pseudo-hessian matrix: Experience of elastic marmousi-2 synthetic data: Bulletin of the Seismological Society of America, **98**, 2402–2415.

Chollet, F., et al., 2015, Keras.

Crase, E., A. Pica, M. Noble, J. McDonald, and A. Tarantola, 1990, Robust elastic nonlinear waveform inversion: Application to real data: Geophysics, **55**, 527–538.

Hu, W., Y. Jin, X. Wu, and J. Chen, 2019, A progressive deep transfer learning approach to cycle-skipping mitigation in fwi, *in* SEG Technical Program Expanded Abstracts 2019: Society of Exploration Geophysicists, 2348–2352.

Jeong, W., H.-Y. Lee, and D.-J. Min, 2012, Full waveform inversion strategy for density in the frequency domain: Geophysical Journal International, **188**, 1221–1242.

Köhn, D., D. De Nil, A. Kurzmann, A. Przebindowska, and T. Bohlen, 2012, On the influence of model parametrization in elastic full waveform tomography: Geophysical Journal International, **191**, 325–345.

Li, D., F. Gao, and P. Williamson, 2019, A deep learning approach for acoustic fwi with elastic data, *in* SEG Technical Program Expanded Abstracts 2019: Society of Exploration Geophysicists, 2303–2307.

Mahrooqi, S., S. Rawahi, S. Yarubi, S. Abri, A. Yahyai, M. Jahdhami, K. Hunt, and J. Shorter, 2012, Land seismic low frequencies: acquisition, processing and full wave inversion of 1.5-86 hz, *in* SEG Technical Program Expanded Abstracts 2012: Society of Exploration Geophysicists, 1–5.

Marjanović, M., R.-É. Plessix, A. Stopin, and S. C. Singh, 2018, Elastic versus acoustic 3-d full waveform inversion at the east pacific rise 9 50 n: Geophysical Journal International, **216**, 1497–1506.

Martin, G. S., R. Wiley, and K. J. Marfurt, 2006, Marmousi2: An elastic upgrade for marmousi: The leading edge, **25**, 156–166.

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228.

Nocedal, J., and S. Wright, 2006, Numerical optimization: Springer Science & Business Media.

Ovcharenko, O., V. Kazei, M. Kalita, D. Peter, and T. Alkhalifah, 2019, Deep learning for low-frequency extrapolation from multioffset seismic data: Geophysics, **84**, R989–R1001.

Plessix, R., P. Milcik, H. Rynja, A. Stopin, K. Matson, and S. Abri, 2013, Multiparameter full-waveform inversion: Marine and land examples: The Leading Edge, **32**, 1030–1038.

Plessix, R.-É., G. Baeten, J. W. de Maag, F. ten Kroode, and Z. Rujie, 2012, Full waveform inversion and distance separated simultaneous sweeping: a study with a land seismic data set: Geophysical Prospecting, **60**, 733–747.

Raknes, E. B., B. Arntsen, and W. Weibull, 2015, Three-dimensional elastic full waveform inversion using seismic data from the sleipner area: Geophysical Journal International, **202**, 1877–1894.

Sears, T. J., P. J. Barton, and S. C. Singh, 2010, Elastic full waveform inversion of multicomponent ocean-bottom cable seismic data: Application to alba field, uk north sea: Geophysics, **75**, R109–R119.

Sirgue, L., O. Barkved, J. Van Gestel, O. Askim, and J. Kommedal, 2009, 3d waveform inversion on valhall wide-azimuth obc: 71st EAGE Conference and Exhibition incorporating SPE EUROPEC 2009, European Association of Geoscientists & Engineers, cp–127.

Stopin, A., R.-É. Plessix, and S. Al Abri, 2014, Multiparameter waveform inversion of a large wide-azimuth low-frequency land data set in oman: Geophysics, **79**, WA69–WA77.

Sun, H., and L. Demanet, 2018, Low frequency extrapolation with deep learning, *in* SEG Technical Program Expanded Abstracts 2018: Society of Exploration Geophysicists, 2011–2015.

———, 2020, Extrapolated full waveform inversion with deep learning: Geophysics, **85**, 1–71.

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903.

Vigh, D., K. Jiao, D. Watts, and D. Sun, 2014, Elastic full-waveform inversion application using multicomponent measurements of seismic data collection: Geophysics, **79**, R63–R77.

Yang, F., and J. Ma, 2019, Deep-learning inversion: A next-generation seismic velocity model building method: Geophysics, **84**, R583–R599.