CWP-996

Free-surface multiple attenuation using convolutional neural networks

Mert S. R. Kiraz & Roel Snieder
Center for Wave Phenomena, Colorado School of Mines, Golden CO 80401, USA
kiraz@mines.edu

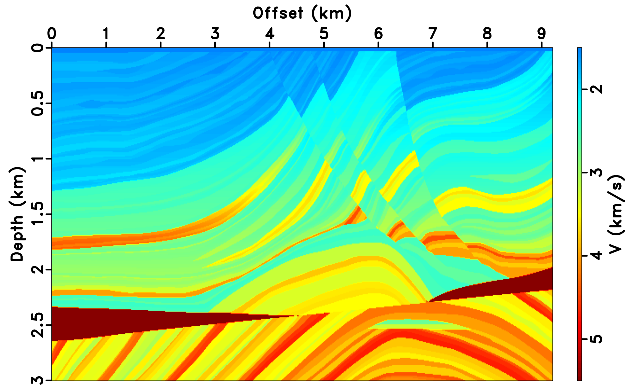
ABSTRACT
In marine seismic data, the most significant set of multiples are those which reflect at least once from the free-surface. The elimination of this type of multiples has been extensively studied and numerous approaches have been widely used to tackle free-surface multiples. However, free-surface multiple elimination is an expensive processing step and it often results in removing the primary events from the seismic data along with the free-surface multiples when primary and multiple events overlap. We present an algorithm to attenuate free-surface multiples using convolutional neural networks (CNNs), and show that data from the CNN-based free-surface multiple elimination results match the numerically modeled data without free-surface multiples. We train a network using subsets of the Marmousi and Pluto velocity models, and make predictions using subsets of the Sigsbee velocity model. We demonstrate the robustness of CNNs for free-surface multiple elimination using two numerical examples and show that CNNs are able to attenuate surface-related multiples even in the presence of overlapping primary and multiple events without removing the primary reflections.

Key words: convolutional neural networks, surface-related multiple elimination, multiples

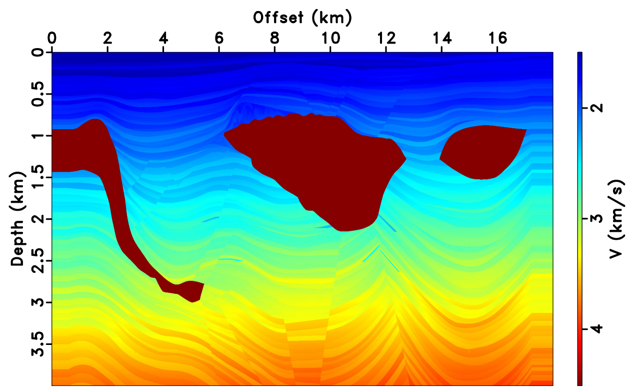
1 INTRODUCTION

Although once the presence of multiples in seismic data was a doubtful question and it was not considered as a serious limitation (Dix, 1948; Ellsworth, 1948), there is no doubt now that the removal of multiples from seismic data is one of the most challenging problems in seismic data processing. In general, there are three main categories on which the multiple attenuation algorithms rely: (1) Periodicity - so that the multiples can be attenuated if they have a constant periodic behavior. To address this type of multiple, one can use deconvolution, τ-p deconvolution, deconvolution after stack, etc. (2) Moveout - so that the multiples can be attenuated if they have a moveout that is sufficiently different than that of the primaries. To address this type of multiple, one can use stacking, f-k filtering, etc. (3) Primary & multiple reflections - so that the multiples can be attenuated based on their physical relationship between the primary and multiple reflections. Although there are several ways to address this type of multiple (e.g., inverse scattering multiple attenuation (ISMA), interbed multiple prediction (IMP)), they are out of the scope of this paper. We only focus on surface-related multiple elimination (SRME) method on this paper.

Multiple reflections are one of the most challenging problems in recorded marine seismic data, and they contaminate the primary reflections. The multiples directly related to the free-surface (or the sea surface) have at least one downward reflection at the free-surface and are the most significant multiples to address. SRME (Verschuur et al., 1992) is a data-driven technique, and it predicts multiples by exploiting the physical relationship between multiple reflections and their sub-events. The arrival time of a particular surface multiple can be predicted by convolving the traces that contain the multiples' sub-events. SRME has been widely used to predict and iteratively subtract the multiples from the seismic data, and it is one of the most effective tools in seismic processing. However, it is



(a)



(b)

Figure 1. (a) Marmousi velocity model. (b) Pluto velocity model.

computationally demanding and it does not always guarantee perfect solutions. For instance, during the subtraction step, one can remove the primary reflections from the seismic data in addition to the unwanted multiple reflections because of the overlapping primary and multiple events. Additionally, a profound limitation of SRME is that it requires a dense source and receiver distribution which makes it computationally demanding and which makes missing data a complication that must be dealt with. Although 2-D versions of SRME can be computationally efficient, they fall short in the presence of complex subsurface requiring 3-D spatial distribution of source and receiver distribution (Dragoset et al., 2010) (e.g., when multiples with strong cross-line raypath components present in the data). This condition becomes even more difficult to fulfill in 3-D surveys.

There have been several attempts to address free-surface multiple attenuation. Riley and Claerbout (1976) present free-surface multiple attenuation theory applied to 1-D data, and they also present an algorithm for surface-generated diffracted multiple elimination. Verschuur et al. (1992) show that free-surface multiples can be eliminated by predicting the multiples and subtracting them from the seismic data. Jakubowicz (1998) shows that the interbed multiples can be removed using a technique based on SRME. There have also been field data applications of SRME by Dragoset and MacKay (1993), Verschuur et al. (1995), Dragoset and Željko Jeričević (1998), and Kelamis and Verschuur (2000). Although there are many more research articles and studies about SRME and its applications, we do not discuss them in this paper.

We present and discuss a CNN-based approach to attenuate free-surface multiples. We train a network using 2-D subsets of Marmousi velocity model (Versteeg, 1994), and Pluto velocity model (Stoughton et al., 2001), and make predictions using two subsets of Sigsbee velocity model (Paffenholz et al., 2002) with salt bodies in the subsurface. As the training input data, we use 2-D data modeled with free-surface multiples using the finite-difference forward modeling, and also use 2-D data modeled without free-surface multiples using the finite-difference forward modeling as the training output data. Although we use 2-D subsets of the velocity models for the training, we use 1-D convolutional filters to find the optimal weights for the network between the input and output datasets. After training the network

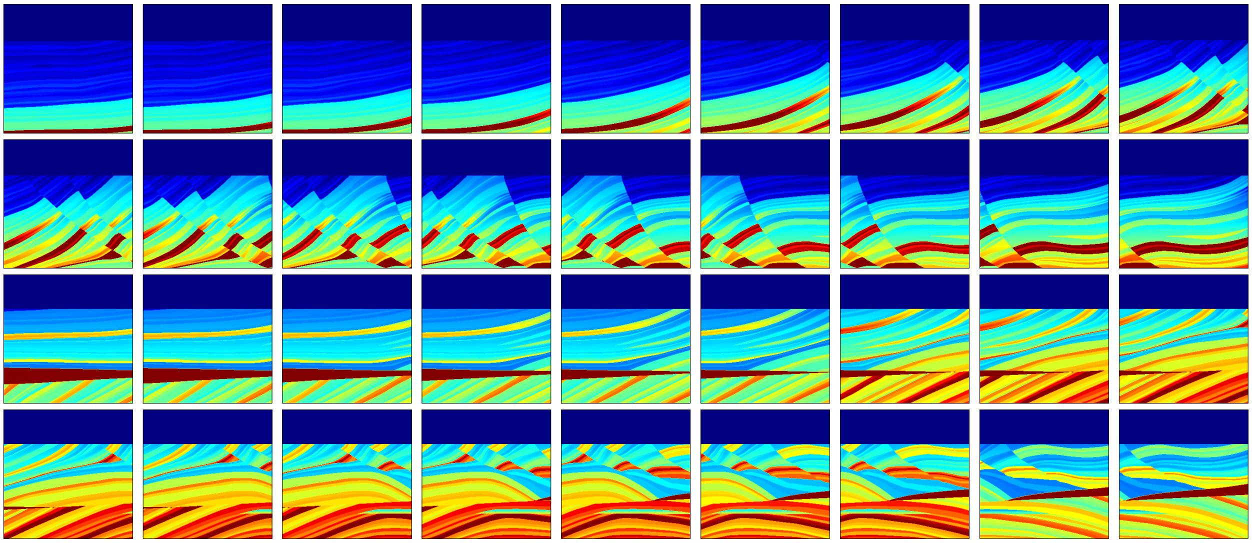


Figure 2. Subsets of Marmousi velocity model used for training.

using the input and output data, we make predictions using two new datasets which were not involved in the training process. We show that CNNs are able to attenuate free-surface multiples and when the free-surface multiples overlap with the primary events, the CNN-based solution is able to remove the unwanted multiple reflections from the seismic data while preserving the primary reflections.

2 CNN ARCHITECTURE AND TRAINING

Artificial intelligence enables computers and machines to mimic human behavior. Machine learning (ML) is a subset of artificial intelligence and it uses mathematical and statistical methods to let machines and computers improve with experience. Deep learning (DL) is a subset of ML and it is inspired by the structure of our brains and the interconnections between the neurons. DL uses multiple layers of computational units and every layer learns its own representation of the input data (Ekman, 2021).

Convolutional neural networks (CNNs) are important building blocks in DL methods inspired by the brain's visual cortex and they have been widely incorporated in image recognition studies since the 1980's (Géron, 2019). CNNs are different from the other DL algorithms because the neurons in the CNN architecture are connected to the region of pixels (receptive fields) in the input image as opposed to every single pixel. CNNs find the optimal weights to minimize an objective function between the input and output datasets. Although the convolution is a linear operator, the non-linearity in CNNs is introduced through the activation functions. Training CNNs is commonly done by backpropagation where loss function and gradient descent optimization algorithm play essential roles (Géron, 2019). A model's performance is measured by a loss function through forward propagation on a training dataset, and learnable parameters (e.g., weights), are updated according to the loss value through an optimization algorithm.

Different DL/CNN methods and their subsets have been widely used in exploration seismic studies to address different problems. Siahkoohi et al. (2018) utilize CNNs to remove the free-surface multiples and numerical dispersion; Das et al. (2019) use CNNs to obtain an elastic subsurface model using recorded normal-incidence seismic data; Wu et al. (2019) use CNNs for 3-D seismic fault segmentation; Almuteri and Sava (2021) use CNNs to address to ghost removal from seismic data. Recently, CNNs have been used to tackle free-surface multiples in various ways each with their own pros and cons (Siahkoohi et al., 2018, 2019; Ovcharenko et al., 2021; Zhang et al., 2021; Liu-Rong et al., 2021).

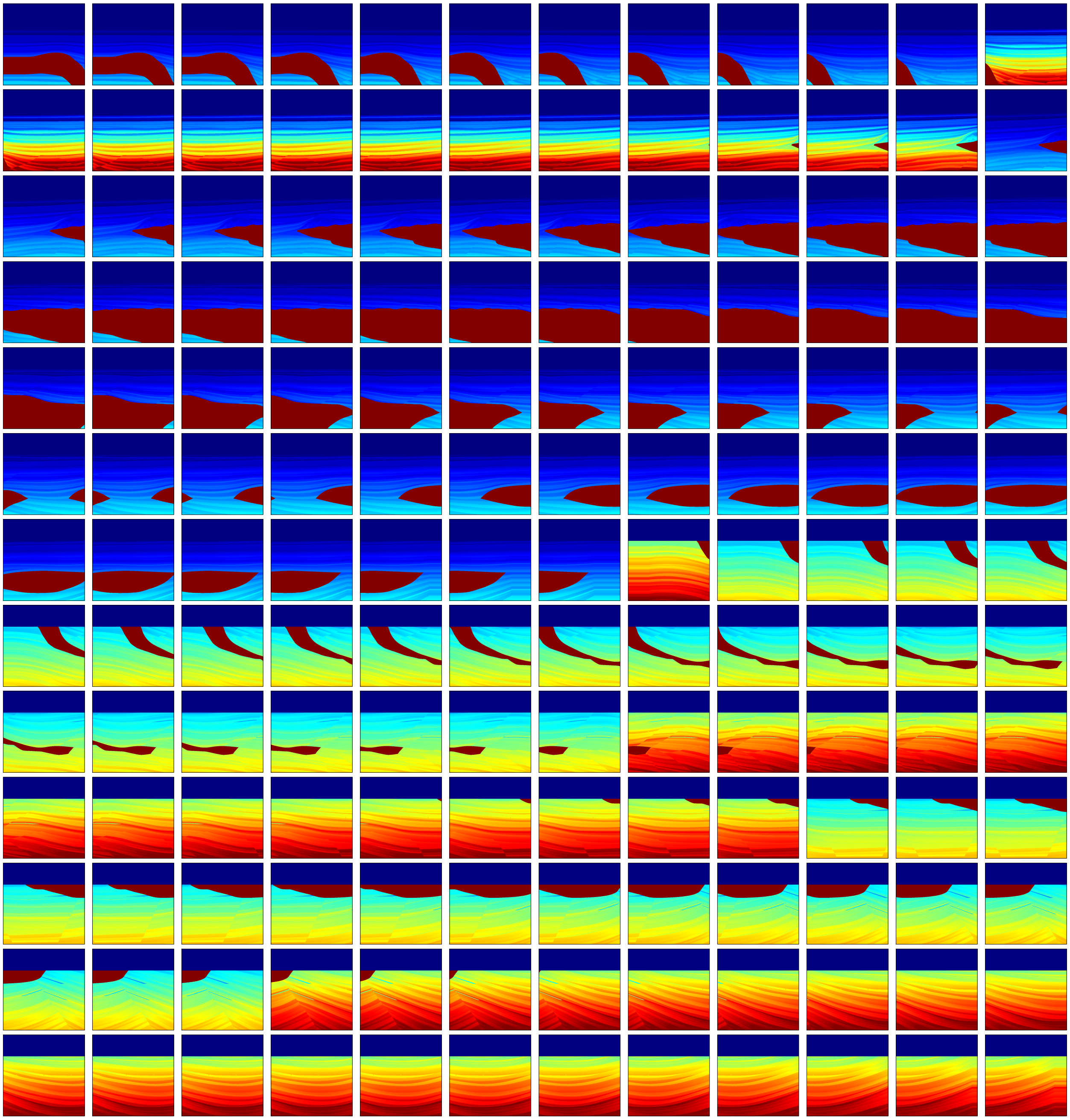
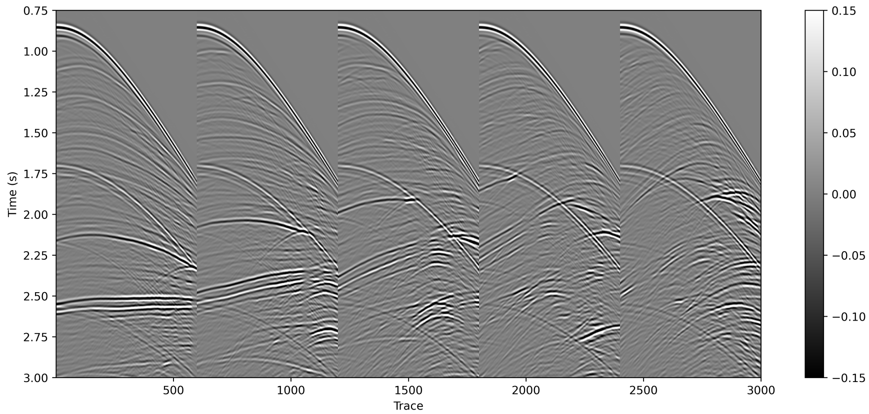


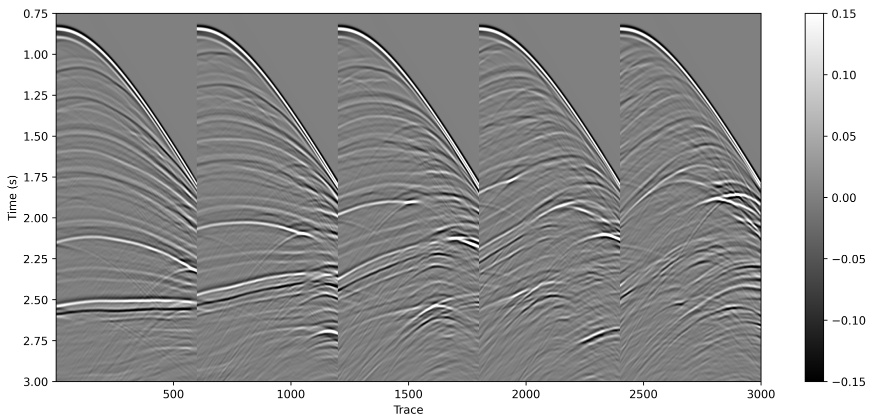
Figure 3. Subsets of Pluto velocity model used for training.

2.1 Network Architecture

For the prediction of the free-surface multiple attenuated data, the CNN architecture we use consists of an input layer, 9 hidden 1-D convolutional layers, and an output layer. The first and second, fourth and fifth, and seventh and eighth convolutional layers have 32 filters; the third, sixth, and ninth convolutional layers have 8 filters. The length of the 1-D convolutional window (kernel size) for the first 3 hidden layers is 1125 samples; for the following 5 hidden layers is 101; and for the last hidden layer is 51. Overall, the network consists of 1,773,697 total parameters all of which are trainable. For the input layer and the hidden convolutional layers, we use the rectified linear unit (ReLU) activation function (Ekman, 2021) which computes the function $f(x) = max(0, x)$, and the output layer does not



(a)



(b)

Figure 4. (a) 5 of 192 shot gathers of the input training data. (b) 5 of 192 shot gathers of the output training data.

have an activation function. We use a stochastic gradient descent optimization algorithm with momentum along with the mean-squared loss function as the objective function (Géron, 2019). To prevent our model from overfitting, and for an accurate evaluation, we split the input data into training dataset, testing dataset, and validation dataset using 60%, 20%, and 20% ratios, respectively. After defining the network parameters (or hyperparameters), we use 100 epochs (iterations) to train the network. Each epoch takes approximately 49 seconds and the whole training process takes 81 minutes to complete. For the CNN prediction results, prediction of a single trace using the trained network takes 2.4 ms, and for a shot gather of 600 traces, the prediction takes approximately 1.5 seconds.

2.2 Network Training

We create 192 input shot gathers where each gather consists of 600 traces (total of 115,200 input traces) using 2-D finite-difference forward modeling through the models with free-surface multiples, and 192 output shot gathers where each gather consists of 600 traces (total of 115,200 output traces) using 2-D finite-difference forward modeling through the models without free-surface multiples. Throughout the forward modeling for each model, 600 receivers are placed along 2.4 km horizontal extent with the receiver sampling of 4 m, and are located at 8 m depth. We use one source per model, and the coordinates of the source are given as $x_s = 25$ m and $z_s = 7$ m.

For the training model selection, we use the Marmousi and Pluto velocity models. Figure 1(a) shows the Marmousi velocity model and Figure 1(b) shows the Pluto velocity model. We extract 36 sub-models from the Marmousi model and 156 sub-models from the Pluto model. We also add 640 m of water layer to the top of each model with 1500 m/s velocity. Each one of 192 models used for training has 2.4 km horizontal extent, and 2.4 km depth which includes

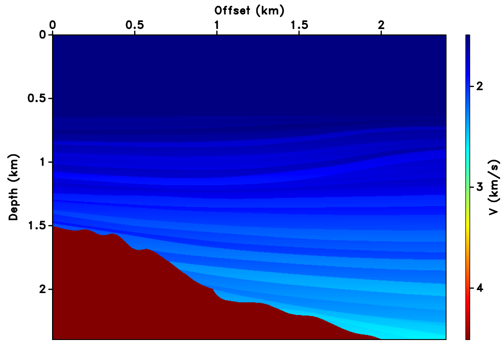


Figure 5. A sub-model of Sigsbee velocity model used for the prediction.

the water column, 600 receivers, and a single shot location. Figure 2 shows 36 sub-models of the Marmousi velocity model used for the training, and Figure 3 shows 156 sub-models of the Pluto velocity model used for the training. For each velocity model, we use a density model that has the same geological properties with the velocity model but also with a reflection coefficient of 0.33 at the water bottom to mimic the seafloor conditions (Kim, 2004). Figure 4 shows 5 of 192 shot gathers of the input and output data displayed next to each other used for training after removing the direct arrival. We remove the direct arrival from the recorded data during the training and prediction steps. Figure 4(a) shows 5 of 192 shot gathers of the data modeled with free-surface multiples which are used as the training input, and Figure 4(b) shows 5 of 192 shot gathers of the data modeled without free-surface multiples which are used as the training output. Figure 4(a) shows the first appearance of free-surface multiples at each shot gather starting around 1.7 s in the first trace of a shot gather, and ending around 2.3 s in time. Although not as strong, the shot gathers also have other types of multiples (e.g., peg-leg, internal).

3 NUMERICAL EXAMPLE AND CNN PREDICTION

For the numerical simulation, we use two different sub-models of the Sigsbee velocity model which are not used in the training process and is unknown to the network to demonstrate our CNN-based solution for free-surface multiple attenuation.

For the first example, Figure 5 shows a subset of the Sigsbee velocity model which consists of relatively flat subsurface structures along with a part of a salt body located at the deeper parts of the model. Similar to the training models, 600 receivers are placed along 2.4 km horizontal extent with the receiver sampling of 4 m, and are located at 8 m depth. The coordinates of the source are given as $x_s = 25$ m and $z_s = 7$ m. The water column has 640 m depth with a velocity of 1500 m/s. The sea bottom for this example also has a reflection coefficient of 0.33 through the density contrast at the seafloor.

Figure 6(a) shows the data modeled with free-surface multiples using the velocity model shown in Figure 5, and Figure 6(b) shows the waveforms modeled without free-surface multiples using the velocity model shown in Figure 5. Figure 6(a) shows the input data, and starting around 1.7 s in time at 0 km in offset, and ending around 2.3 s in time at 2.4 km in offset, the first multiple caused by the free-surface is located. Additionally, starting around 2.55 s in time at 0 km in offset, and ending around 2.75 s in time at 1.5 in km offset, another multiple caused by the free-surface is present. Although not as strong as these two events, there are other multiples arriving after the first multiple which are caused by the free-surface. Figure 6(b) shows the expected output, the ground truth. The free-surface multiples described in Figure 6(a) are not present in this figure because of the excluded free-surface effects. For a successful CNN network training, we expect the CNN predictions to be similar to the ground truth (Figure 6(b)). After training the network, Figure 6(c) shows our CNN-based free-surface multiple elimination results.

Figure 7 shows several trace comparisons of the input data (solid blue line), output data (or ground truth) (solid red line), and the CNN predictions (solid black line) shown in Figure 6 for the zero-, mid-, and far-offsets. Figure 7(a) shows the trace used as an input to the network which is the zero-offset trace extracted from Figure 6(a) (solid blue line), the ground truth that is the expected result from the CNN prediction which is the zero-offset trace extracted from Figure 6(b) (solid red line), and the predicted zero-offset trace using our CNN-based free-surface

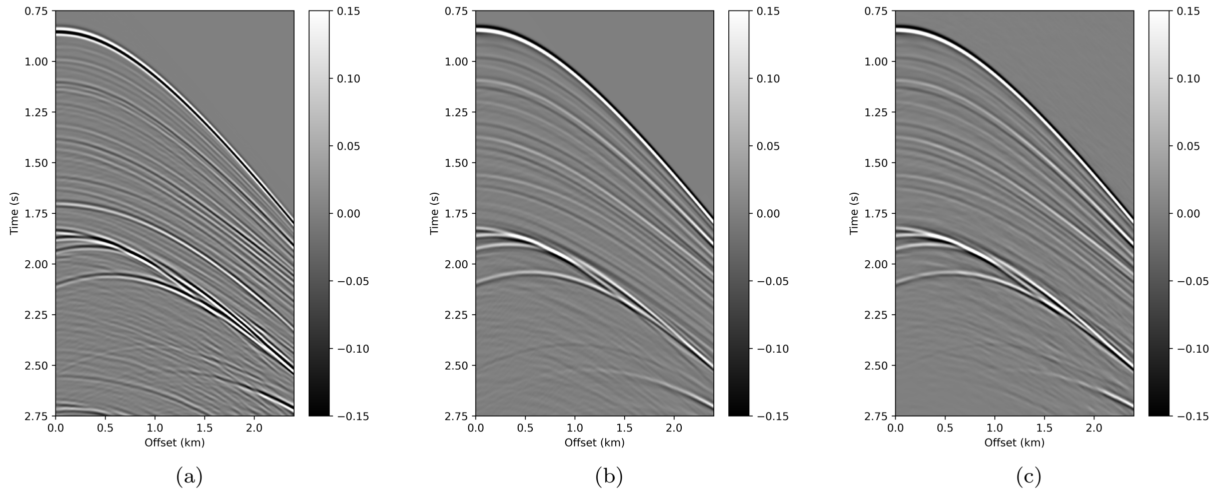


Figure 6. (a) Numerically modeled shot gather with the free-surface multiples (input data) obtained using the velocity model given in Figure 5. (b) Numerically modeled shot gather without the free-surface multiples (output data/ground truth) obtained using the velocity model given in Figure 5. (c) Predicted shot gather without the free-surface multiples using CNNs.

multiple elimination operator extracted from Figure 6(c) (solid black line). In Figure 7(a), the two strongest multiples caused by the free-surface in the input trace (solid blue line) at times around 1.7 s and 2.7 s have been successfully eliminated in the CNN prediction trace (solid black line) which matches the ground truth trace (solid red line). Figure 7(b) shows the mid-offset (offset 1.2 km) trace used as an input to the network from Figure 6(a) (solid blue line), the ground truth that is the expected result from the CNN prediction which is the mid-offset (offset 1.2 km) trace extracted from Figure 6(b) (solid red line), and the predicted mid-offset (offset 1.2 km) trace using our CNN-based free-surface multiple elimination operator extracted from Figure 6(c) (solid black line). In Figure 7(b), the strongest multiple caused by the free-surface in the input trace (solid blue line) at time around 1.85 s has been successfully eliminated in the CNN prediction trace (solid black line) which matches the ground truth trace (solid red line). Lastly, Figure 7(c) shows the far-offset (offset 2.3 km) trace used as an input to the network from Figure 6(a) (solid blue line), the ground truth that is the expected result from the CNN prediction which is the far-offset (offset 2.3 km) trace extracted from Figure 6(b) (solid red line), and the predicted far-offset (offset 2.3 km) trace using our CNN-based free-surface multiple elimination operator extracted from Figure 6(c) (solid black line). In Figure 7(c), the strongest multiple caused by the free-surface in the input trace (solid blue line) at time around 2.35 s has been successfully eliminated in the CNN prediction trace (solid black line) which matches the ground truth trace (solid red line). Figure 7 shows that the CNN-based free-surface multiple attenuation is able to attenuate the free-surface multiples despite their different arrival times due to the offset change. The CNN-based free-surface multiple attenuation is able to preserve the primary events after removing the unwanted multiples from the data. Additionally, the phase change in the input trace is successfully accounted for through the CNN prediction.

For the second example, Figure 8 shows another subset of the Sigsbee velocity model which consists of relatively flat subsurface structures with a fault near the far-offsets, and a salt body located at the deeper parts of the model. Similar to the training models, 600 receivers are placed along 2.4 km horizontal extent with the receiver sampling of 4 m, and are located at 8 m depth. The coordinates of the source are given as $x_s = 25$ m and $z_s = 7$ m. The water column has 640 m depth with a velocity of 1500 m/s. The sea bottom for this example also has a reflection coefficient of 0.33 through the density model.

Figure 9(a) shows the data modeled with free-surface multiples using the velocity model shown in Figure 8, and Figure 9(b) shows the data modeled without free-surface multiples using the velocity model shown in Figure 8. Figure 9(a) shows the input data, and starting around 1.7 s in time at 0 km in offset, and ending around 2.3 s in time at 2.4 km in offset, we see the first multiple caused by the free-surface. The challenge in this example is that the first event caused by the free-surface overlaps with a primary event around 1.5 km in offset and 1.85 s in time in Figure 9(a). Conventional SRME techniques require an adaptive subtraction step which in the presence of overlapping primary and multiple events might result in removal of the primary energy. Although not as strong as the first event caused by

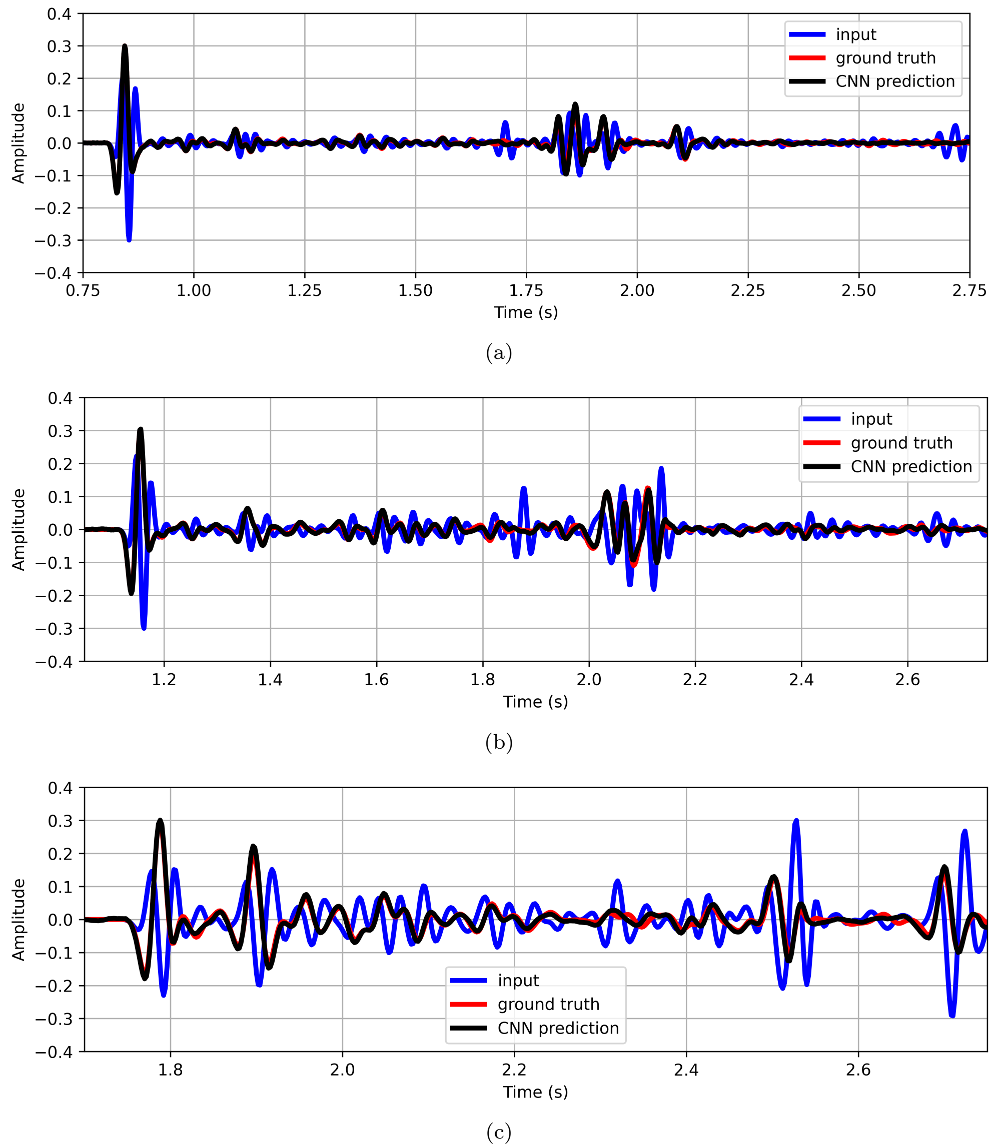


Figure 7. Trace comparison of the input (solid blue line), ground truth (solid red line), and CNN prediction (solid black line) data from the zero-, mid-, and far-offsets, respectively, extracted from the data shown in Figure 6. (a) Comparison of the zero-offset traces. (b) Comparison of the mid-offset (offset 1.2 km) traces. (c) Comparison of the far-offset (offset 2.3 km) traces.

the free-surface, there are other multiples arriving after the first multiple which are caused by the free-surface. Figure 9(b) shows the expected output, the ground truth. The events described in Figure 9(a) caused by the free-surface are not present in this figure because of the excluded free-surface effects. For a successful CNN network training, we expect the CNN predictions to be similar to the ground truth (Figure 9(b)). After training the network, Figure 9(c) shows our CNN-based free-surface multiple elimination results.

Figure 10 shows several trace comparisons of the input data (solid blue line), output data (or ground truth) (solid red line), and the CNN predictions (solid black line) shown in Figure 9 for the zero-, mid-, and far-offsets. Figure 10(a) shows the trace used as an input to the network which is the zero-offset trace extracted from Figure 9(a) (solid blue line), the ground truth that is the expected result from the CNN prediction which is the zero-offset trace extracted from Figure 9(b) (solid red line), and the predicted zero-offset trace using our CNN-based free-surface multiple elimination operator extracted from Figure 9(c) (solid black line). In Figure 10(a), the strongest multiple caused by the free-surface in the input trace (solid blue line) at time around 1.7 s has been successfully eliminated in

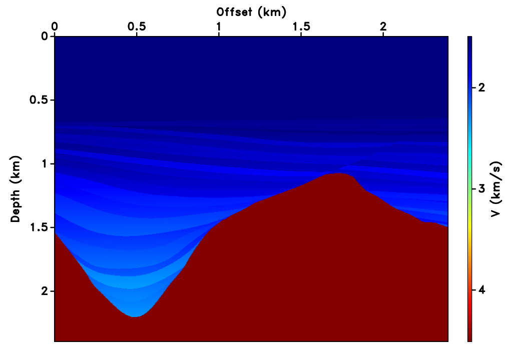
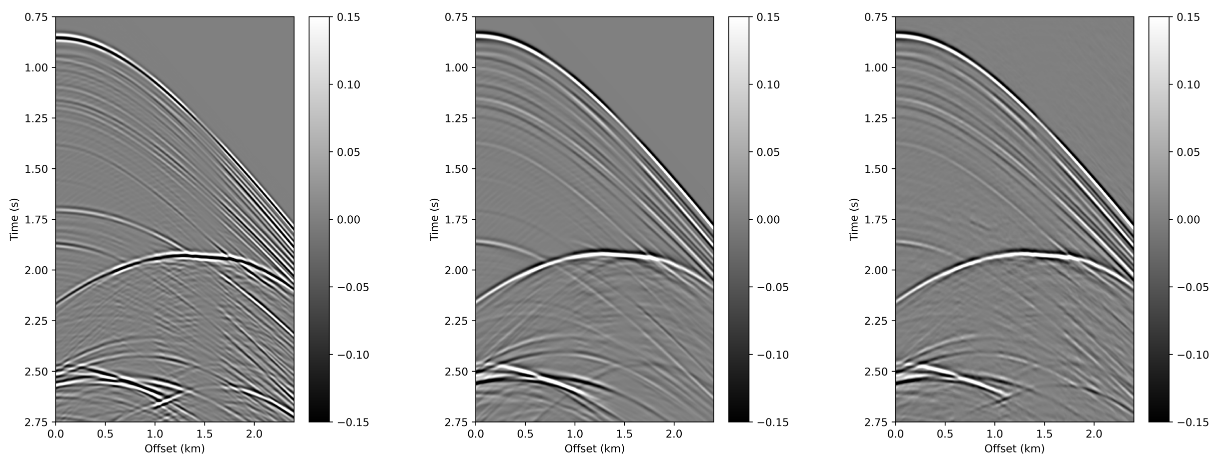


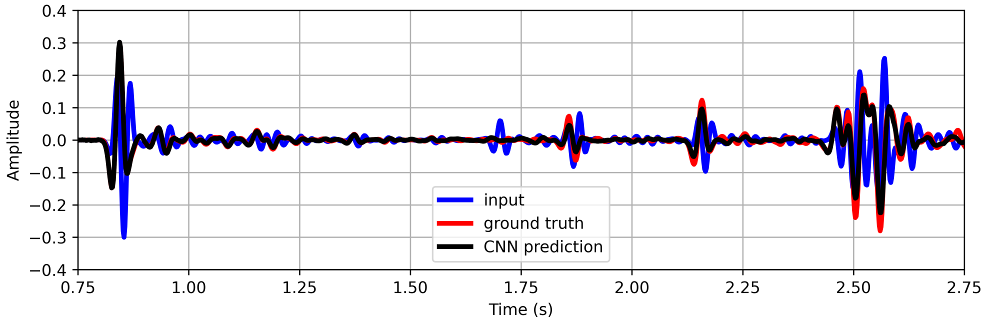
Figure 8. A sub-model of Sigsbee velocity model used for the prediction.



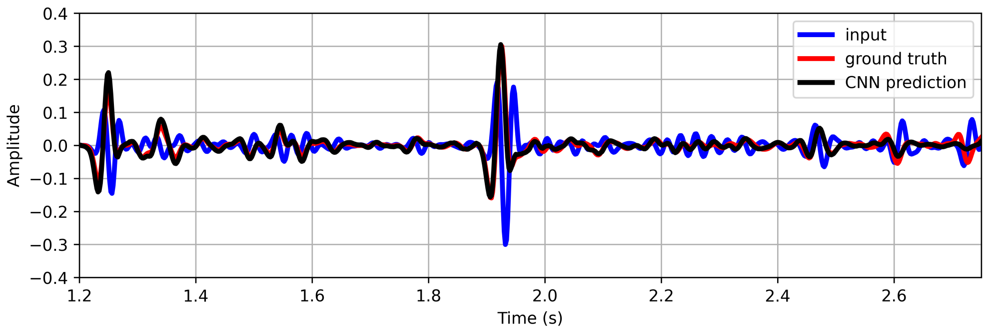
(a) (b) (c)

Figure 9. (a) Numerically modeled shot gather with the free-surface multiples (input data) obtained using the velocity model given in Figure 8. (b) Numerically modeled shot gather without the free-surface multiples (output data/ground truth) obtained using the velocity model given in Figure 8. (c) Predicted shot gather without the free-surface multiples using CNNs.

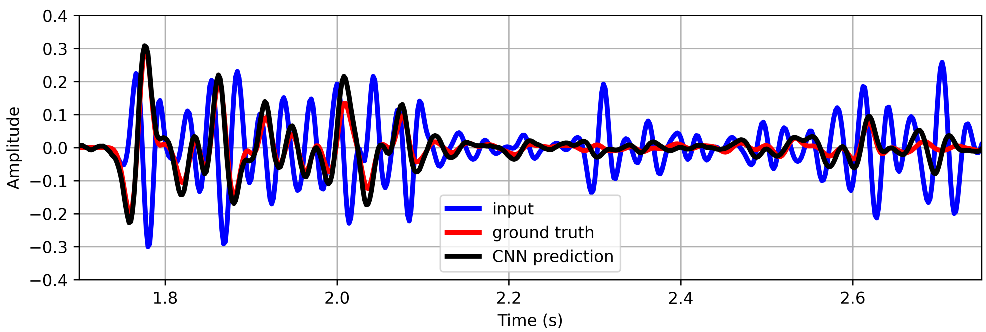
the CNN prediction trace (solid black line) which matches the ground truth trace (solid red line). Figure 10(b) shows the mid-offset (offset 1.5 km) trace used as an input to the network from Figure 9(a) (solid blue line), the ground truth that is the expected result from the CNN prediction which is the mid-offset (offset 1.5 km) trace extracted from Figure 9(b) (solid red line), and the predicted mid-offset (offset 1.5 km) trace using our CNN-based free-surface multiple elimination operator extracted from Figure 9(c) (solid black line). However, in Figure 10(b), the strongest multiple caused by the free-surface in the input trace (solid blue line) at time around 1.9 s overlaps with a primary event. Although the free-surface reflection is embedded in the high-amplitude energy around 1.9 s in Figure 10(b), the CNN prediction is able to remove the free-surface multiple without removing any of the primary events by preserving the main reflections. The CNN prediction trace (solid black line) closely matches the ground truth trace (solid red line) in Figure 10(b). Lastly, Figure 10(c) shows the far-offset (offset 2.3 km) trace used as an input to the network from Figure 9(a) (solid blue line), the ground truth that is the expected result from the CNN prediction which is the far-offset (offset 2.3 km) trace extracted from Figure 9(b) (solid red line), and the predicted far-offset (offset 2.3 km) trace using our CNN-based free-surface multiple elimination operator extracted from Figure 9(c) (solid black line). In Figure 10(c), the strongest multiple caused by the free-surface in the input trace (solid blue line) at time around 2.35 s has been successfully eliminated in the CNN prediction trace (solid black line) which matches the ground truth trace (solid red line). Figure 10 shows that the CNN-based free-surface multiple attenuation is able to predict and remove the free-surface multiples even in the presence of overlapping primary and multiple reflections. The CNN-based free-surface multiple attenuation is able to preserve the primary events after removing the unwanted



(a)



(b)



(c)

Figure 10. Trace comparison of the input (solid blue line), ground truth (solid red line), and CNN prediction (solid black line) data from the zero-, mid-, and far-offsets, respectively, extracted from the data shown in Figure 9. (a) Comparison of the zero-offset traces. (b) Comparison of the mid-offset (offset 1.5 km) traces. (c) Comparison of the far-offset (offset 2.3 km) traces.

multiples with different arrival times versus the offset from the data. Additionally, the phase change in the input trace is successfully accounted for through the CNN prediction.

The two examples presented in this paper show that the CNN-based free-surface multiple attenuation is able to remove the multiples caused by the free-surface. The change in the arrival time of the free-surface multiples versus offset is correctly predicted by the CNN and removed from the shot gathers without removing the primary events even when the primary and multiple reflections overlap. Although the data obtained by the CNN prediction mostly match the numerically modeled data, for the late times of the shot gathers (i.e., for times around 2.75 s), there are some mismatches between the CNN prediction and the ground truth. This is due to the length of the 1-D convolutional window (kernel size) running out of space towards the end of each trace in a shot gather. This causes the late arrivals of the CNN predictions to have more mismatches than the rest of the data. Although the same happens for the very early times of the shot gathers (i.e., for times around 0.75 s), there are no events recorded by the early times of the shot gathers. The use of 1-D convolutional filters makes our CNN-based free-surface multiple attenuation computationally efficient (e.g., training takes 81 minutes and prediction of a trace takes 2.4 ms). Additionally, trace-by-trace attenuation of free-surface multiples does not require a dense spatial distribution of sources and receivers. This method is suitable for irregular and sparse source/receiver acquisition geometry. In our CNN-based free-surface multiple attenuation algorithm, the prediction of the free-surface multiple is an independent process from the availability of the near-offset traces which is a traditional requirement for the conventional SRME algorithms.

4 CONCLUSIONS

We present a free-surface multiple attenuation approach in 2-D using 1-D convolutional networks. Our results show that free-surface multiples for a subsurface model with variable velocity and density profiles with salt structures can be attenuated once a network is trained. We show that when the primary and multiple reflections overlap, our CNN-based free-surface multiple attenuation method successfully removes the multiples while preserving the primary reflections. Unlike the conventional SRME methods, our algorithm does not require subtraction of the multiples from the full dataset for free-surface multiple elimination, and eliminates the risk of removing the primary energy from recorded seismic data by providing a fast and effective alternative to the existing free-surface multiple elimination methods. Our algorithm does not require dense spatial distribution of source and receivers, and it is also independent from the availability of the near-offset traces.

5 ACKNOWLEDGMENTS

We thank Jonathan Sheiman for the fruitful discussions and his expertise that helped shape this paper immensely. We also thank Khalid Almuteri from the Center for Wave Phenomena for his help with the CNN setup. This work is supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at the Colorado School of Mines. The finite-difference forward modeling examples in this paper were generated using the Madagascar software package (Fomel et al., 2013).

REFERENCES

Almuteri, K., and P. Sava, 2021, A convolutional neural network approach for ghost removal: First International Meeting for Applied Geoscience & Energy Expanded Abstracts, 2550–2554.

Das, V., A. Pollack, U. Wollner, and T. Mukerji, 2019, Convolutional neural network for seismic impedance inversion: GEOPHYSICS, **84**, R869–R880.

Dix, C. H., 1948, The existence of multiple reflections: GEOPHYSICS, **13**, 49–50.

Dragoset, B., and S. MacKay, 1993, Surface multiple attenuation and sub-salt imaging: Exploration Geophysics, **24**, 463–472.

Dragoset, B., E. Verschuur, I. Moore, and R. Bisley, 2010, A perspective on 3D surface-related multiple elimination: GEOPHYSICS, **75**, 75A245–75A261.

Dragoset, W. H., and Željko Jeričević, 1998, Some remarks on surface multiple attenuation: GEOPHYSICS, **63**, 772–789.

Ekman, M., 2021, Learning deep learning: Theory and practice of neural networks, computer vision, NLP, and transformers using TensorFlow: Addison-Wesley Professional.

Ellsworth, T. P., 1948, Multiple reflections: GEOPHYSICS, **13**, 1–18.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**, e8.

Géron, A., 2019, Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow, 2nd edition: O'Reilly Media, Inc.

Jakubowicz, H., 1998, Wave equation prediction and removal of interbed multiples: SEG Technical Program Expanded Abstracts, 1527–1530.

Kelamis, P. G., and D. J. Verschuur, 2000, Surface-related multiple elimination on land seismic data—strategies via case studies: GEOPHYSICS, **65**, 719–734.

Kim, G.Y., R. M. B. D., 2004, Sediment types determination using acoustic techniques in the Northeastern Gulf of Mexico: Geosciences Journal, **8**, 95–103.

Liu-Rong, T., J. Jin-Sheng, R. Hao-Ran, Y. Yue-Ming, and W. Bang-Yu, 2021, The separation of seismic surface-related multiples based on CAE-SAGAN: First International Meeting for Applied Geoscience & Energy Expanded Abstracts, 2914–2918.

Ovcharenko, O., A. Baumstein, and E. Neumann, 2021, Surface-related multiple elimination through orthogonal encoding in the latent space of convolutional autoencoder: First International Meeting for Applied Geoscience & Energy Expanded Abstracts, 1355–1359.

Paffenholz, J., J. Stefani, B. McLain, and K. Bishop, 2002, Sigsbee_2a synthetic subsalt dataset - image quality as function of migration algorithm and velocity model error: European Association of Geoscientists & Engineers, 2214–4609.

Riley, D., and J. Claerbout, 1976, 2-D multiple reflections: GEOPHYSICS, **41**, 1942.

Siahkoohi, A., M. Louboutin, R. Kumar, and F. J. Herrmann, 2018, Deep-convolutional neural networks in prestack seismic: Two exploratory examples: SEG Technical Program Expanded Abstracts, 2196–2200.

Siahkoohi, A., D. J. Verschuur, and F. J. Herrmann, 2019, Surface-related multiple elimination with deep learning: SEG Technical Program Expanded Abstracts, 4629–4634.

Stoughton, D., J. Stefani, and S. Michell, 2001, 2D elastic model for wavefield investigations of subsalt objectives, deep water Gulf of Mexico: SEG Technical Program Expanded Abstracts, 1269–1272.

Verschuur, D. J., A. J. Berkhout, and P. G. Kelmis, 1995, Estimation of multiple scattering by iterative inversion, part II: Examples on marine and land data: SEG Technical Program Expanded Abstracts, 1470–1473.

Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: GEOPHYSICS, **57**, 1166–1177.

Versteeg, R., 1994, The Marmousi experience: Velocity model determination on a synthetic complex data set: The Leading Edge, **13**, 927–936.

Wu, X., L. Liang, Y. Shi, and S. Fomel, 2019, Faultseg3d: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation: GEOPHYSICS, **84**, IM35–IM45.

Zhang, D., M. de Leeuw, and E. Verschuur, 2021, Deep learning-based seismic surface-related multiple adaptive subtraction with synthetic primary labels: First International Meeting for Applied Geoscience & Energy Expanded Abstracts, 2844–2848.