AEGC2019 Data to Discovery
Australasian Exploration Geoscience Conference 2-5 September 2019 • Perth, Western Australia
Incorporating the AIG, ASEG, PESA, and WABS

Seismic model low wavenumber extrapolation by a deep convolutional neural network

Pavel Plotnitskii
KAUST
KAUST, Thuwal,
23955-6900, Saudi Arabia
pavel.plotnitskii@kaust.edu.sa

Tariq Alkhalifah
KAUST
KAUST, Thuwal
23955-6900, Saudi Arabia.
tariq.alkhalifah@kaust.edu.sa

Oleg Ovcharenko
KAUST
KAUST, Thuwal,
23955-6900, Saudi Arabia
oleg.ovcharenko@kaust.edu.sa

Vladimir Kazei
KAUST
KAUST, Thuwal,
23955-6900, Saudi Arabia
vladimir.kazei@kaust.edu.sa

SUMMARY

Conventional seismic data are naturally mainly sensitive to the very smooth velocity variations that alter transmission traveltimes (low-model wavenumbers) and very abrupt discontinuities that cause reflections (high-model wavenumbers). Full-waveform inversion (FWI) of seismic data inherits this lack of middle model wavenumber illumination, which results into ringy artifacts in the gradients. Multiple methods have been suggested to overcome this issue. Here we tackle the problem of missing wavenumbers with a deep-learning approach. Namely, we filter out the wavenumbers that are expected to be missing from the acquisition design and then train a deep convolutional neural network to provide the missing wavenumbers trace-by-trace. We test several network configurations and several training sets derived from the Marmousi II model. The neural network shows limited capabilities in generalizing from the input data sets. We also report a tradeoff between the generalization abilities and accuracy on the training data set.
Key words: machine learning, signal processing, velocity analysis, full-waveform inversion.

INTRODUCTION

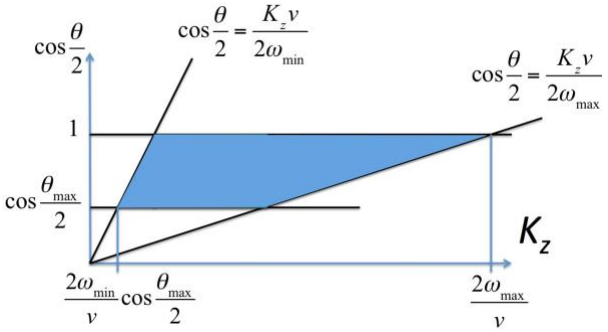

Figure 1: Illumination of vertical wavenumbers K_z under different opening angles for FWI (Alkhalifah and Kazei, 2018). Here θ is the angle between incident and reflected rays, v is the background velocity, θ_{max} is the maximum opening angle defined by aperture and the target perturbation depth, ω_{min} and ω_{max} are the minimum and maximum angular frequencies used in the FWI. On this diagram for each K_z the blue area covers the opening angles, under which K_z is illuminated. Hereinafter in all figures and equations, we will use notation θ for reflection angle, which is an angle between incident angle and vertical instead of opening angle.

Claerbout (1985) showed, that seismic reflection data are not sensitive to middle vertical wavenumbers in the model spectrum. Devaney (1984) estimated the missing part of the spectrum for the velocity models, used in diffraction tomography; Mora (1989) expanded this technique to wavenumber coverage of FWI. In (Alkhalifah and Kazei, 2018) illumination of vertical wavenumbers Kz under different opening angles for FWI was calculated (see Fig. 1) and is useful for this paper, as we consider recovery of vertical middle wavenumbers in this paper.

Seismic FWI attempts to match synthetic and observed waveform data through iterative refining of the subsurface velocity model (Virieux and Operto (2009)). Despite significant efforts and progress made over the past decades in FWI methodology (e.g. Pratt et al., 1996; Warner et al., 2013; van Leeuwen and Herrmann, 2013; Alkhalifah, 2016; Kazei et al., 2016; Gray, 2016; Alkhalifah et al., 2018; Kazei and Alkhalifah, 2018; Ovcharenko et al., 2018a; Kalita et al., 2018; Yao et al., 2019; Guo and Alkhalifah, 2017), the non-linear iterative optimization procedure is still prone to stagnation in local minima when initiated from a poor assumption and applied to a geologically complex region. Thus, building a plausible initial model becomes an important task and extrapolation of missing middle wavenumbers in the model could significantly improve the convergence. For extrapolation of middle wavenumbers different tools can be used, but we intend to discuss deep learning capabilities for that.

A number of deep-learning driven approaches proposed to reconstruct the missing seismic data frequency is growing lately. This includes two-dimensional shot-by-shot (Ovcharenko et al., 2018b; Jin et al., 2018) and one-dimensional trace-by-trace (Sun and Demanet, 2018) data bandwidth extension. With respect to accuracy one-dimensional approaches seem less effective, than two-dimensional, because of an accumulation of predicted errors reduces the coherence of data across traces. However, trace-by-trace techniques seem more stable and easier to tune, as fewer parameters are involved in a neural network training process than for the two-dimensional problem. In this paper, we discuss only one-dimensional approach, although we plan to develop two-dimensional one for the meeting.

THEORY

We consider this technique to be applied in the FWI workflow, namely improving the inverted velocity model. So to test our technique we suggest to generate synthetic velocity perturbations, which could be acquired in real FWI applications. Perturbations can be simulated by subtracting the background velocity model from the true model, having the full broadband of wavenumbers, e.g. Marmousi. As the background model can be some modification of the true model: for example its linear trend, moving average etc. In the examples below we calculate background models, taking a linear trend of the true one (see Fig. 3 (a)).

After that, according to diagram on Fig. 1, we construct a high-pass filter, that *deletes low and middle wavenumbers from the data trace-by-trace*. Namely, we delete vertical wavenumbers smaller than

$$K_z < \frac{2f_{min}\cos(\theta_{max}(z))}{v(z)} \quad (1)$$

from each log of the perturbation model.

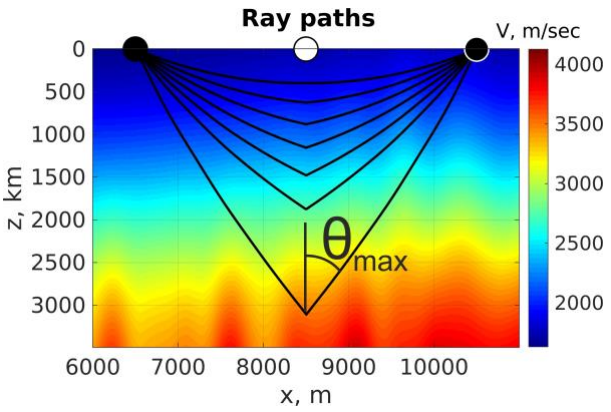

Figure 2: Acquiring of maximum reflection angles in equation 2 for maximum offset equal to 4 km using ray tracing.

The chosen cut-off value of the filter corresponds to the lower boundary of high wavenumbers (reflection part of the data). We don't consider low wavenumbers corresponding to the transmission part of the data in our filter, since it is not easy to formulate an analytic expression for them, as for the high wavenumbers. Thus we intend to extrapolate middle and low wavenumbers from high wavenumbers with the proposed deep-learning approach. We can formulate the filter in this form:

$$K_z \in \left[\frac{2f_{min}\cos(\theta_{max}(z))}{v(z)}, F_{max}\right], \quad (2)$$

where K_z is the vertical wavenumber, θ_{max} is the maximum reflection angle defined by the aperture and target perturbation depth, F_{max} is the maximum frequency available in the data, v is the background velocity.

For each depth of the log we calculate the maximum reflection angle θ_{max}, as the angle under which the ray arrives at a certain depth from the maximum offset point, noted on Fig. 2 with a black dot. As the white dot on the Fig. 2 we mark the log for which we perform ray tracing. In our approach, presented in Fig. 2, we trace rays in the smoothed with 2-dimensional Gaussian filter Marmousi. We assume that we can conduct ray tracing only for one log from smoothed Marmousi and use the acquired θ_{max} for all logs in our data set, because angles from ray tracing will not differ very much along profile in laterally homogeneous strongly smoothed Marmousi.

This is a vertical wavenumbers K_z filter, which we apply to the depth-frequency transform of each velocity log. In this paper we chose to use the so-called wavelet synchrosqueezed transform (WSST). This time-frequency (TF) decomposition is used to characterize the non-stationary relation between the depth and instantaneous frequency, which can be very useful in the processing and interpretation of seismic data.

The application of the filter on a single Marmousi perturbation log, taken from the testing data set from Fig. 4, is shown on Fig. 3 on the (b) as the half-transparent gray mask, overlaid on the WSST spectrum. We apply the mask, constructed with equation 2 to the WSST spectrum of log and cut the part of the spectrum, lying outside the mask. Then we calculate the inverse WSST transform of the filtered with mask spectrum and acquire filtered perturbation (see (c) in Fig. 3).

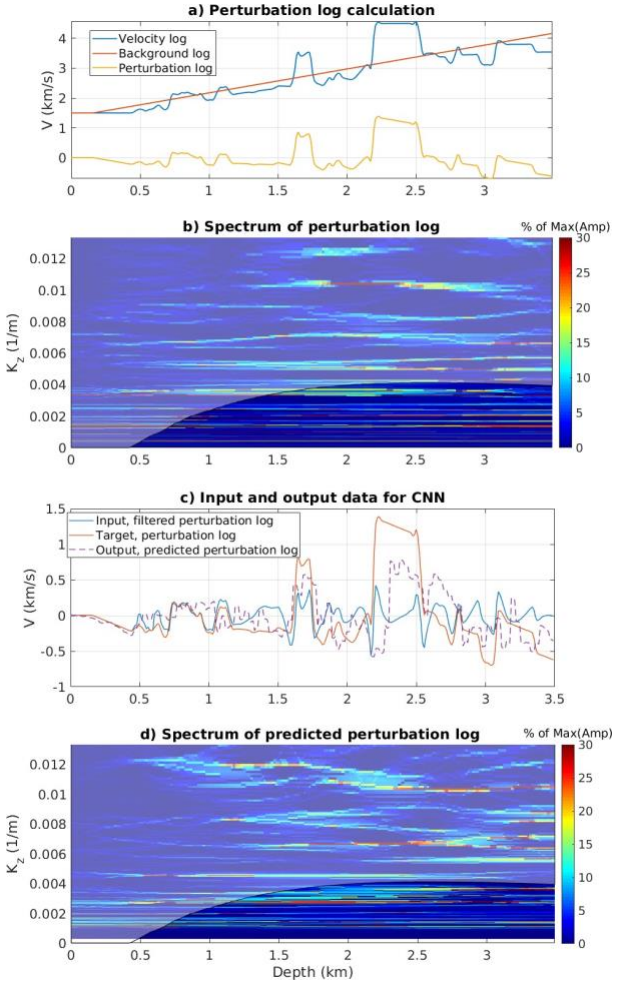

Figure 3: a) True, background and perturbation logs; b) WSST spectrum of perturbation log with an overlaid filter mask; c) Filtered, perturbation and predicted perturbation logs; d) WSST spectrum of predicted perturbation log. $f_{min} = 10$ Hz.

We can notice (on (c) in Fig. 3) , that the low-frequency part of the data have been filtered from the log. Also, we can notice that low frequencies deleted with filter are recovered on the predicted perturbation log (see Fig. 3). On (c), we can see that

CNN predicts some features of perturbation log, but with the wrong amplitude.

The idea of our approach is to train 1D convolutional neural network (CNN) to extrapolate low and middle wavenumbers from filtered perturbation logs.

CNN FRAMEWORK

Convolutional neural networks are mathematical models which are designed to mimic the workflow of visual data processing in the human brain. At the training stage a supervised CNN derives dependencies in the data set given pairs of input and target data. The trained network then might be used to infer predictions from the previously unknown data. In data-driven approaches the data are of higher importance, than a method that it is approached. Thus, proper composition and preprocessing of training data set should be addressed in the first place.

Data

Our motivation is to produce a depth log with full wavenumber coverage given an equivalent log with missing middle and low wavenumbers. Thus, input and target data for the network are depth logs taken at the same spatial location however with different wavenumber content. Namely, the target data for CNN are the filtered perturbation logs shown in Fig. 4 in (b). The input data are equivalent logs, however with filtered out middle wavenumbers; the output data are shown on (c) in Fig. 4. Both input, target and output (predicted) vectors are 300 elements long.

We generate the data set of 13601 logs from the original Marmousi II (Bourgeois et al., 1991) and BP 2004 (Billette and Brandsberg-Dahl, 2005) benchmark models and split it into training, validation and testing parts as 80/10/10. Full dimensions of inputs and target data sets are 13601×300. To ensure even contribution of features from the data set scale values in each column to fit the [-1, 1] interval.

Network Architecture

We design a generic CNN which takes a vector of 300 elements as input and produces a vector of the same size as an output. The architecture of the CNN consist of a sequence of 5 convolutional layers each followed by a max pooling layer. Convolutional layers derive spatial patterns in the data by convolving the input volume with a set of filters. There are 4, 8, 16 and 4 filters in the convolutional layers, respectively. Elements of these filters are trainable parameters which will learn patterns in the data during the iterative optimization at the training stage. Max pooling layers reduce spatial dimensionality of the data by drawing a single maximum value from the nearest neighborhood. We add a fully-connected layer of 100 neurons at the end of the network to interconnect feature representations from higher convolutional levels. We use rectified linear units as activation functions in all convolutional layers and a hyperbolic tangent in the dense layer. At the last stage, we upsample the 100 element output from the dense layer by doing another convolution with 3 filters and then reshape the resulting volume to produce the target log of 300 elements.

The proposed network is numerically implemented in Python, using Keras library (Chollet et al., 2015) with Tensorflow Abadi et al. (2016) backend.

EXAMPLES

We test our approach by inference the logs with full wavenumber content from equivalent logs with missing wavenumbers.

We want to present the example of application the CNN to the Marmousi velocity data set (see Fig. 4). Fig. 4 consists of 3 panels: a) filtered velocity perturbations; b) velocity perturbations obtained by the subtracting of the background model from original Marmousi; c) data, predicted by CNN from input data. Predicted data should be similar to the true one. You can notice, that CNN predicts quite well the part of the data, where it was trained, and the testing data set as well, because it manages to spread existing trend in the training data set layers to the right edge of the model. We have made another experiment, there the testing data set is larger than on 4 and is extended to the left (see Fig. 5). We can observe, that CNN still shows promising results.

DISCUSSION & CONCLUSIONS

Neural network manages to map filtered perturbations to the real perturbations quite well for the samples that were used directly in training. There is some minor degradation on the validation data set, as it is quite similar to some samples in the training set. The training data set is rather small and simple and overfitting is obvious even with a rather simple neural network architecture and few parameters. Reducing the number of parameters in the network helps increase the generalization abilities of the network at the expense of overall (training and testing) prediction degradation.

The one-dimensional trace-by-trace technique demonstrated potential in the recovery of the missing middle and low wavenumbers. The continuity of the non-horizontal structures is not taken into account though. We mainly attribute it to the use of limited training data set and 1D convolutional network. These to some extent contradicts the successful application of 1D extrapolation to seismic data reported by Sun and Demanet (2018). Therefore, as the next step we consider in the development of two-dimensional approach for middle wavenumbers extrapolation.

ACKNOWLEDGEMENTS

The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, Saudi Arabia.

REFERENCES

Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., 2016, Tensorflow: A system for large-scale machine learning: 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), 265–283.

Alkhalifah, T., 2016, Full-model wavenumber inversion: An emphasis on the appropriate wavenumber continuation: GEOPHYSICS, 81, R89–R98.

Alkhalifah, T., and V. Kazei, 2018, Waveform inversion for orthorhombic anisotropy with P waves: feasibility and resolution: Geophysical Journal International, 213, 963–982.

Alkhalifah, T., B. B. Sun, and Z. Wu, 2018, Full model wavenumber inversion: Identifying sources of information for

the elusive middle model wavenumbers: GEOPHYSICS, 83, R597–R610.

Billette, F., and S. Brandsberg-Dahl, 2005, The 2004 bp velocity benchmark: Presented at the 67th EAGE Conference & Exhibition. Bourgeois, A., M. Bourget, P. Lailly, M. Poulet, P. Ricarte, and R. Versteeg, 1991, Marmousi, model and data: The Marmousi Experience, 5–16.

Chollet, F., et al., 2015, Keras: https://keras.io.

Claerbout, J. F., 1985, Imaging the Earth's Interior: Blackwell scientific publications.

Devaney, A. J., 1984, Geophysical Diffraction Tomography: IEEE Transactions on Geoscience and Remote Sensing, GE-22, 3–13. Gray, S. H., 2016, *in* 4. Seismic imaging: S1–1–S1–16.

Guo, Q., and T. Alkhalifah, 2017, Elastic reflection-based waveform inversion with a nonlinear approach: Geophysics, 82, R309– R321.

Jin, Y., W. Hu, X. Wu, and J. Chen, 2018, *in* Learn low-wavenumber information in FWI via deep inception-based convolutional networks: 2091–2095.

Kalita, M., V. Kazei, Y. Choi, and T. Alkhalifah, 2018, *in* Regularized full-waveform inversion for salt bodies: 1043–1047.

Kazei, V., and T. Alkhalifah, 2018, Time-lapse waveform inversion regularized by spectral constraints and sobolev space norm, *in* SEG Technical Program Expanded Abstracts 2018: Society of Exploration Geophysicists, 5487–5491.

Kazei, V., E. Tessmer, and T. Alkhalifah, 2016, Scattering angle-based filtering via extension in velocity, *in* SEG Technical Program Expanded Abstracts 2016: Society of Exploration Geophysicists, 1157–1162.

Mora, P., 1989, Inversion = migration + tomography: Geophysics, 54, 1575–1586.

Ovcharenko, O., V. Kazei, D. Peter, and T. Alkhalifah, 2018a, Variance-based salt body reconstruction for improved full-waveform inversion: Geophysics.

Ovcharenko, O., V. Kazei, D. Peter, X. Zhang, and T. Alkhalifah, 2018b, Low-Frequency Data Extrapolation Using a Feed-Forward ANN: Presented at the 80th EAGE Conference and Exhibition 2018.

Pratt, R. G., Z.-M. Song, P. R. Williamson, and M. R. Warner, 1996, Two-dimensional velocity models from wide-angle seismic data by waveform inversion: Geophysical Journal International, 124, 323–340.

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.

van Leeuwen, T., and F. J. Herrmann, 2013, Mitigating local minima in full-waveform inversion by expanding the search space: Geophysical Journal International, 195, 661–667.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, 74, WCC1–WCC26.

Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Stekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand, 2013, Anisotropic 3D full-waveform inversion: Geophysics, 78, R59–R80.

Yao, G., N. V. da Silva, V. Kazei, D. Wu, and C. Yang, 2019, Extraction of the tomography mode with non-stationary smoothing for full-waveform inversion: Geophysics, 84, 1–47.

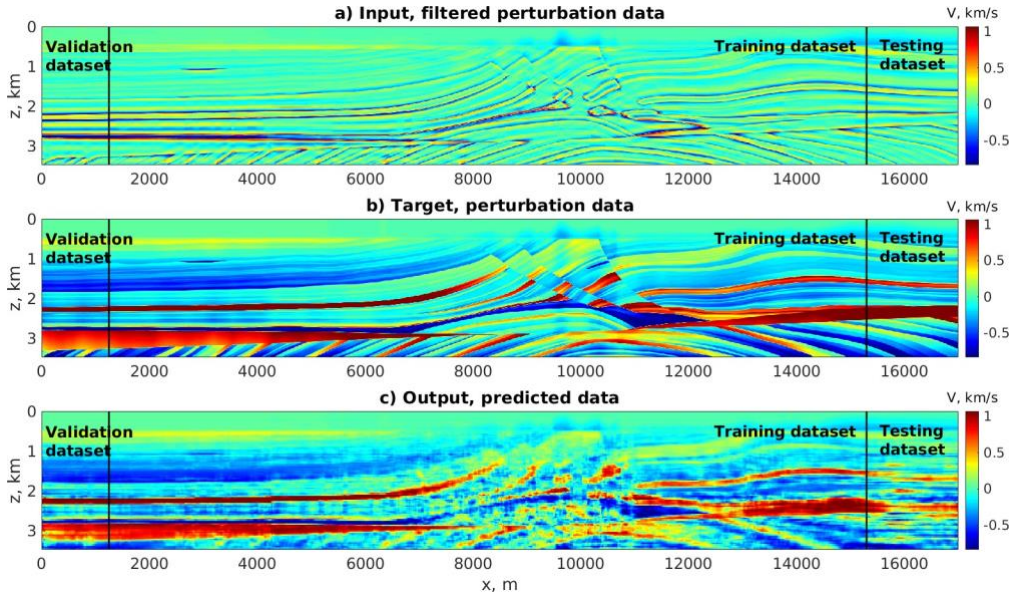

Figure 4: a) Input, filtered perturbation data, b) Target, perturbation data, c) Output, predicted data. Number of logs in validation, training and testing datasets is 1000, 11240 and 1362 logs, correspondingly.

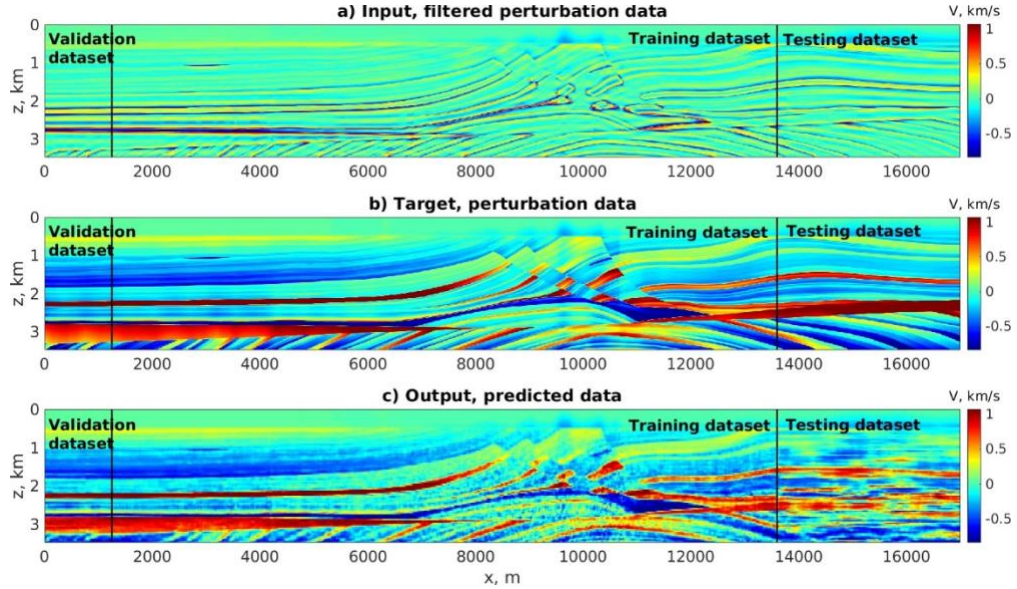

Figure 5: a) Input, filtered perturbation data, b) Target, perturbation data, c) Output, predicted data. Number of logs in validation, training and testing datasets is 1000, 9878 and 2724 logs, correspondingly.