MIT EARTH RESOURCES LABORATORY
ANNUAL FOUNDING MEMBERS MEETING 2019

Extrapolated Full Waveform Inversion with Deep Learning

Hongyu Sun and Laurent Demanet
[EARTH, ATMOSPHERIC AND PLANETARY SCIENCES]

Motivation: full waveform inversion

- **Forward modeling**

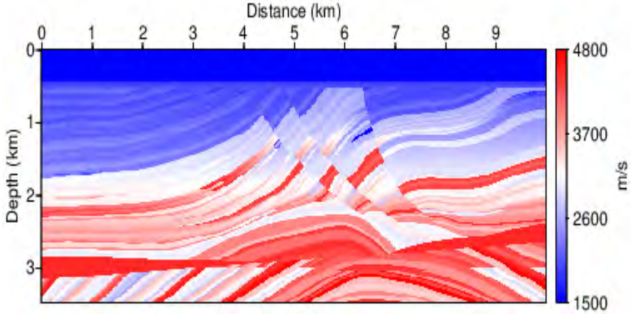
$$\boldsymbol{m}(\boldsymbol{x})\frac{\partial^2 u(\boldsymbol{x},t)}{\partial t^2} - \Delta u(\boldsymbol{x},t) = f(\mathbf{x},\mathrm{t})$$

- **Inversion objective function**

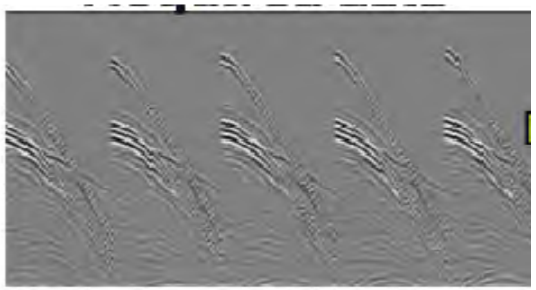
$$J(\boldsymbol{m}) = \|d_{cal}(\boldsymbol{m}) - d_{obs}\|_2$$

- **Optimization**

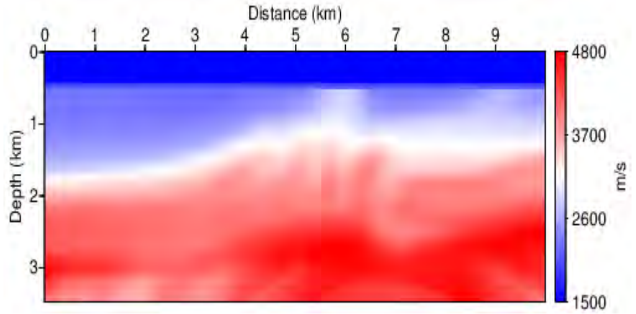
$$\boldsymbol{m}_{k+1} = \boldsymbol{m}_k - \boldsymbol{H}^{-1}\,\nabla J(\boldsymbol{m}_k)$$



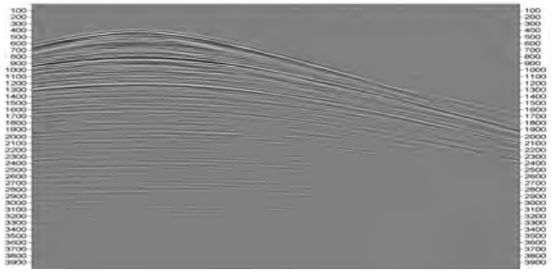
true velocity model (unknown $\boldsymbol{m}$)



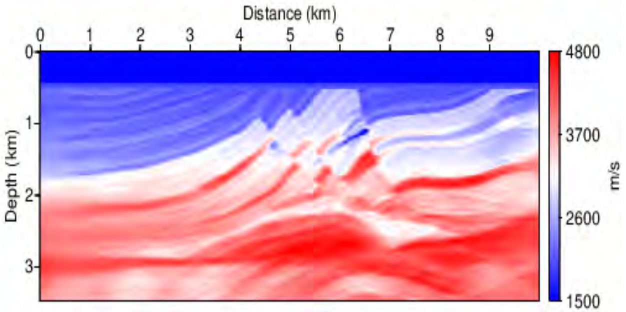
observed data $\boldsymbol{d_{obs}}$



initial velocity model $\boldsymbol{m_0}$

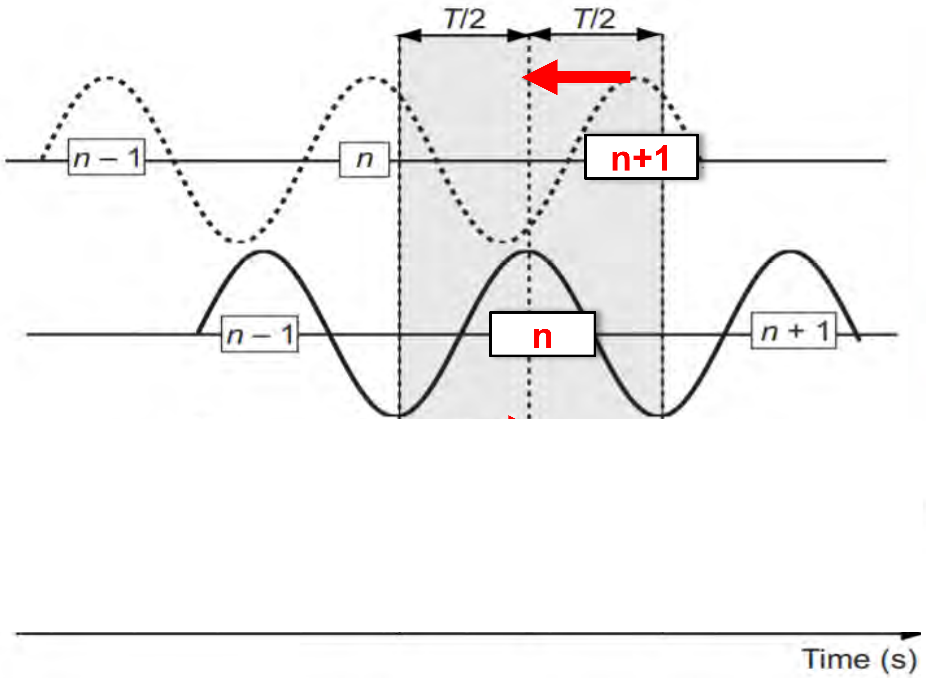


calculated data $\boldsymbol{d_{cal}(m)}$



inversion result

Motivation: Cycle-skipping

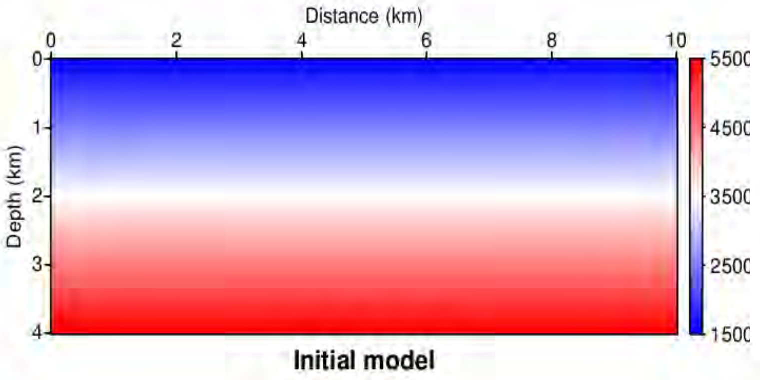


(Virieux and Operato, 2009)

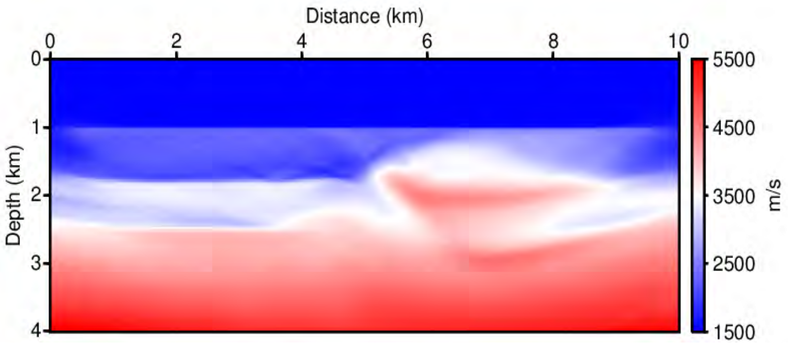
—— **observed recording**

·········· **calculated recording**

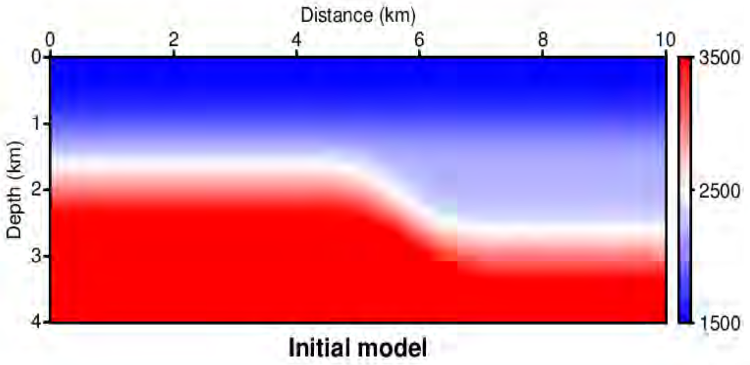
the ‘bad’ initial model



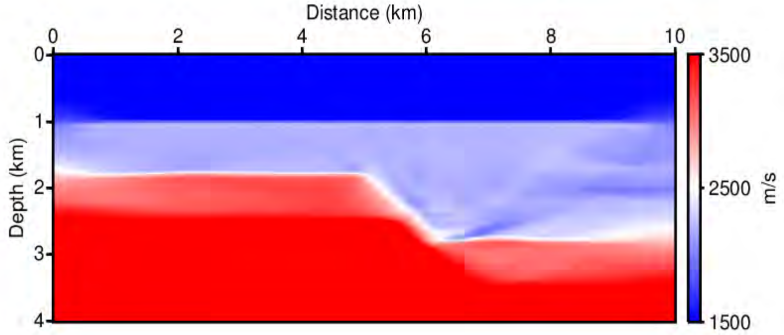
FWI result



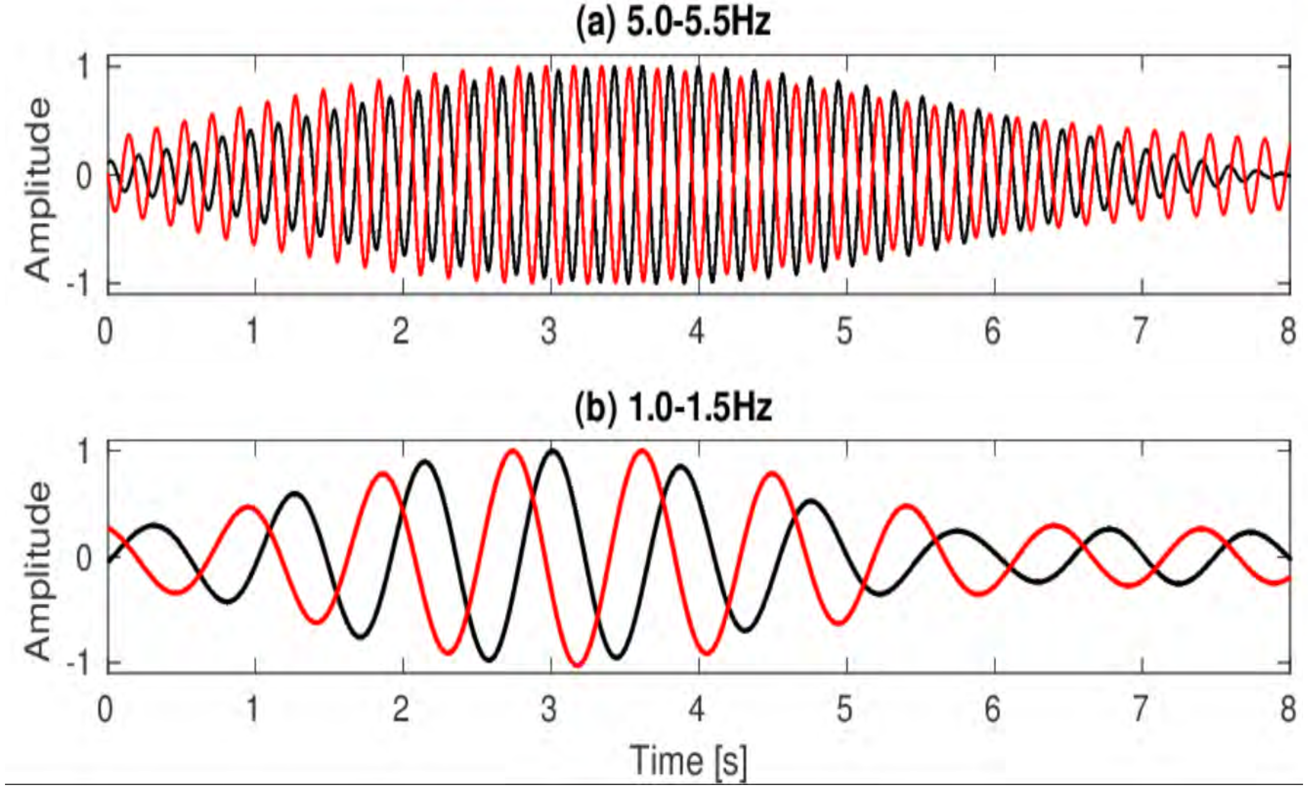
the ‘good’ initial model

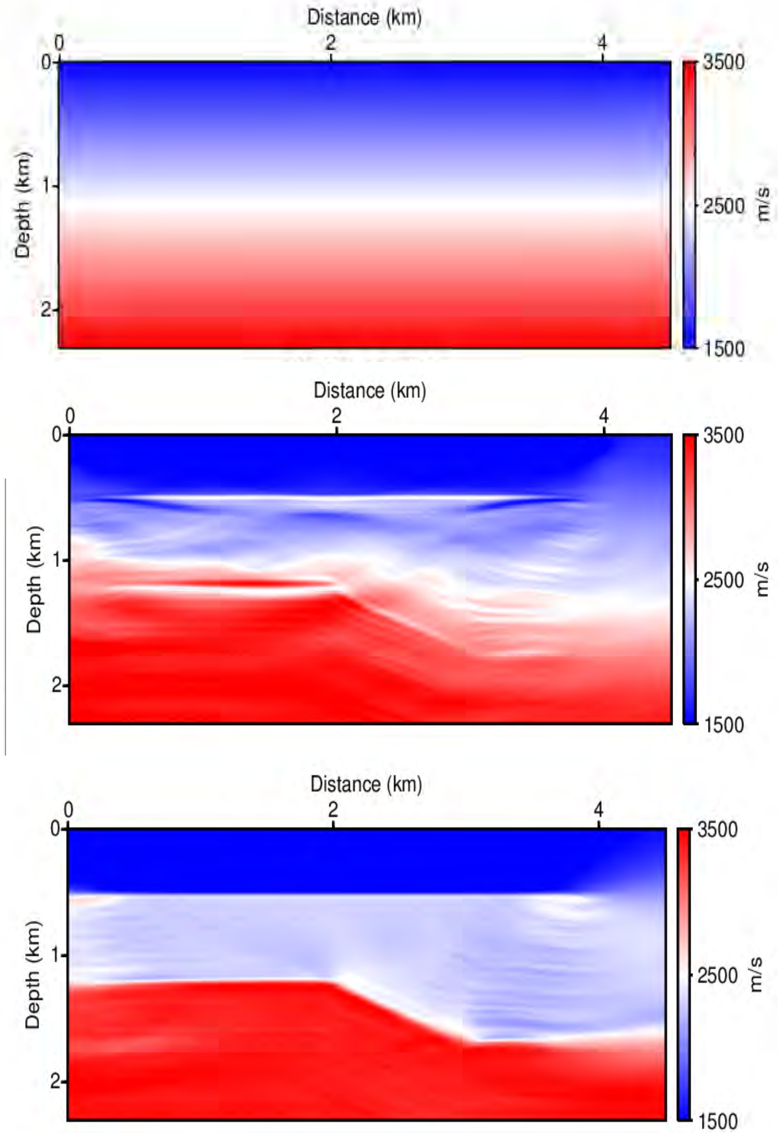


FWI result



Motivation: Cycle-skipping

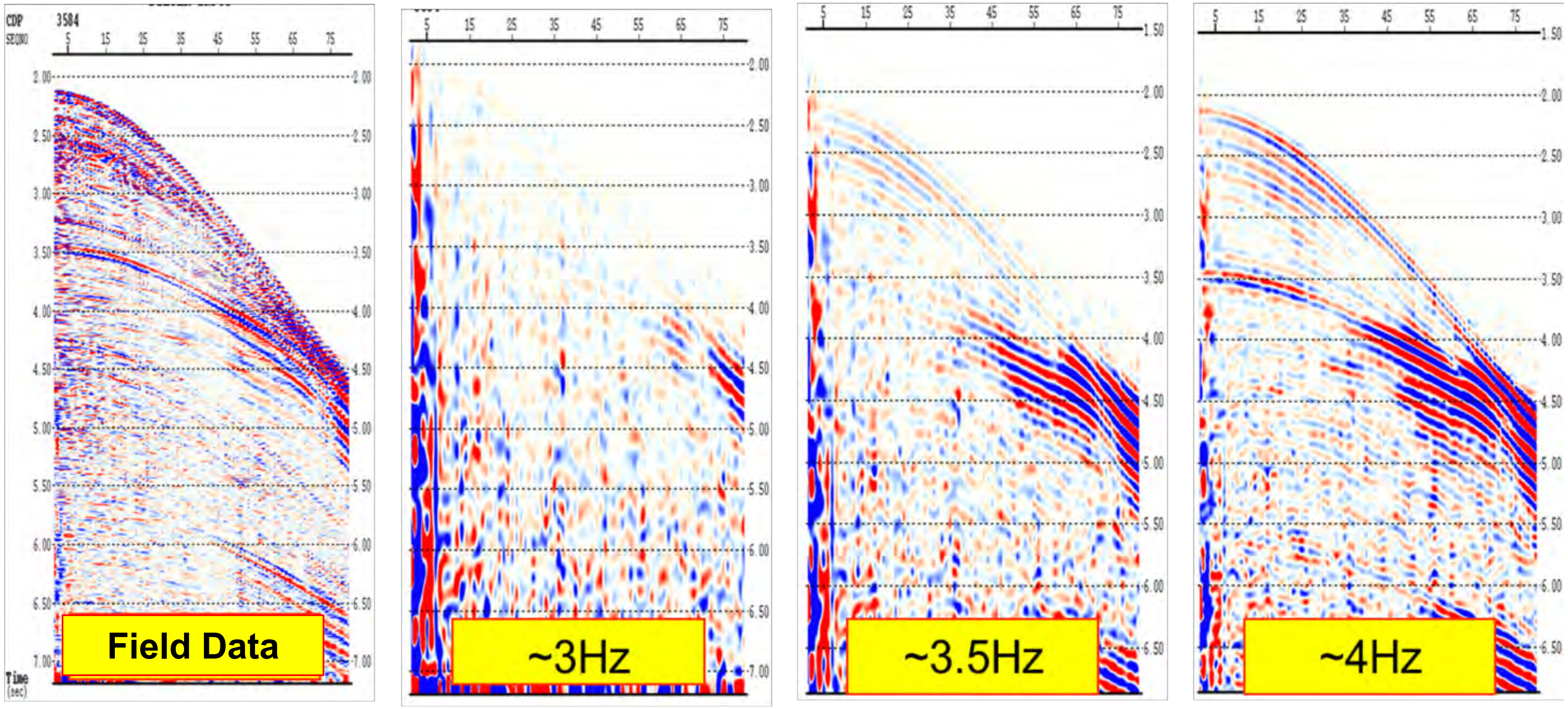




observed recording calculated recording

Motivation: Cycle-skipping

Low frequency data (<3Hz) are hard to acquire in the field



(Han, 2014)

Bandwidth Extension with Deep learning

- **Deep neural networks (DNN):**

$$\boldsymbol{y} = f(\boldsymbol{x}, \boldsymbol{w}) = f_L(\dots f_2(f_1(\boldsymbol{x})))$$

where

- $\boldsymbol{x}$: seismograms bandlimited to high frequencies
- $\boldsymbol{y}$: the same seismograms bandlimited to low frequencies
- $\boldsymbol{w}$: parameters of DNN to be learned

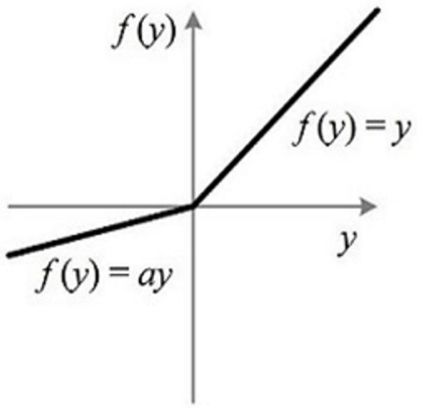
- **Training:** learning $\boldsymbol{w}$ with known $\boldsymbol{y}$

$$\boldsymbol{J}(\boldsymbol{w}) = \frac{1}{m}\sum_{i=1}^{m} L(y_i, f(x_i, w))$$

- **Test** (predict) $f(\boldsymbol{x}, \boldsymbol{w})$

Convolutional Neural Networks

- **Convolution filter：** 128 – 64 – 128 – 64 –1
- **Activation function：** PReLU (He et al., 2015)

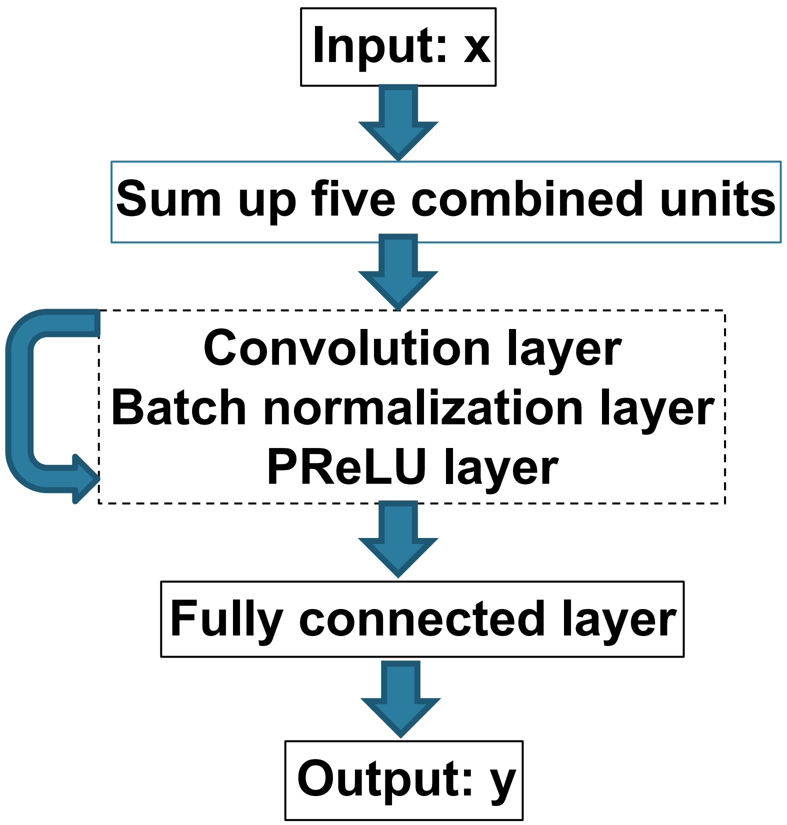


$$PReLU(x_i) = \begin{cases} x_i & if\, x_i > 0 \\ a_i x_i & if\, x_i \leq 0 \end{cases}$$

- **Optimizer：** Adam (Kingma and Ba, 2014)

  with a mini-batch of 20 samples

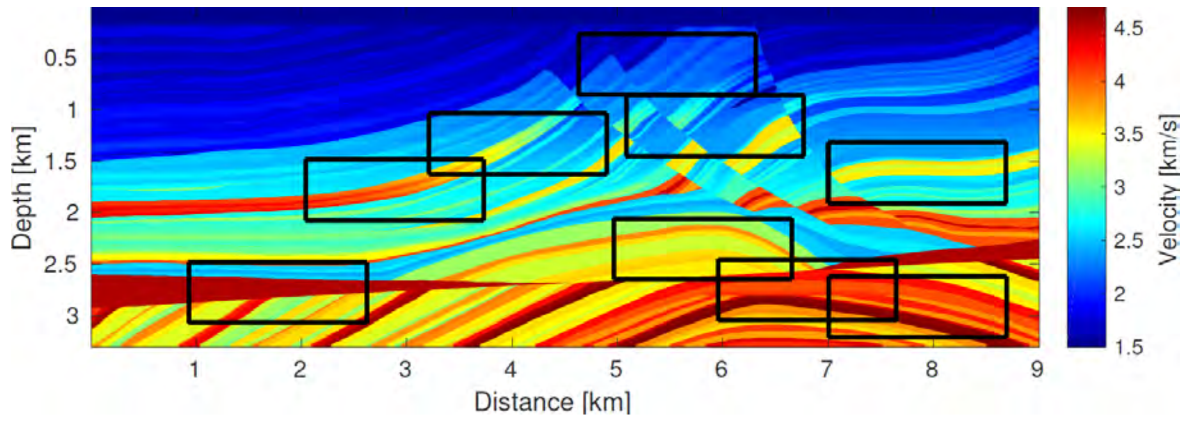
Architecture



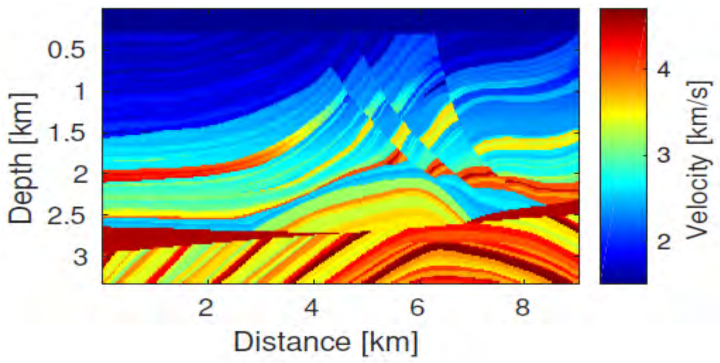
Numerical experiments

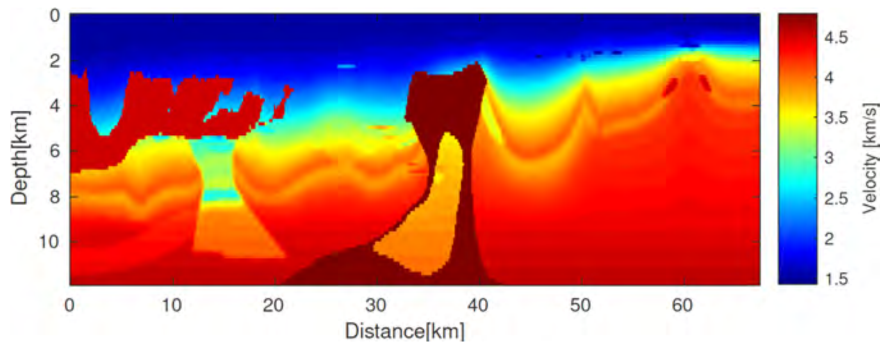
How to collect the training data?

Training model: known low frequencies

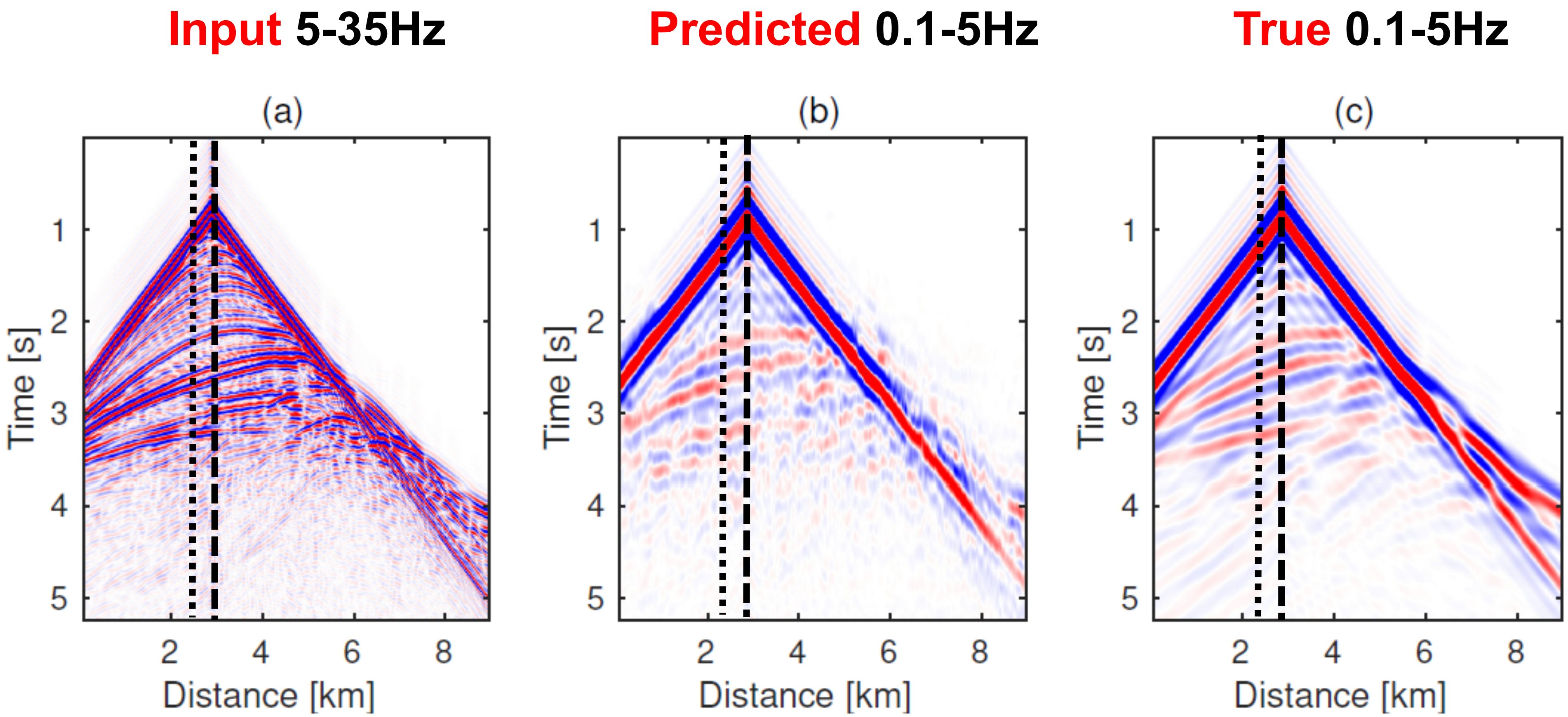


Test model: unknown low frequencies

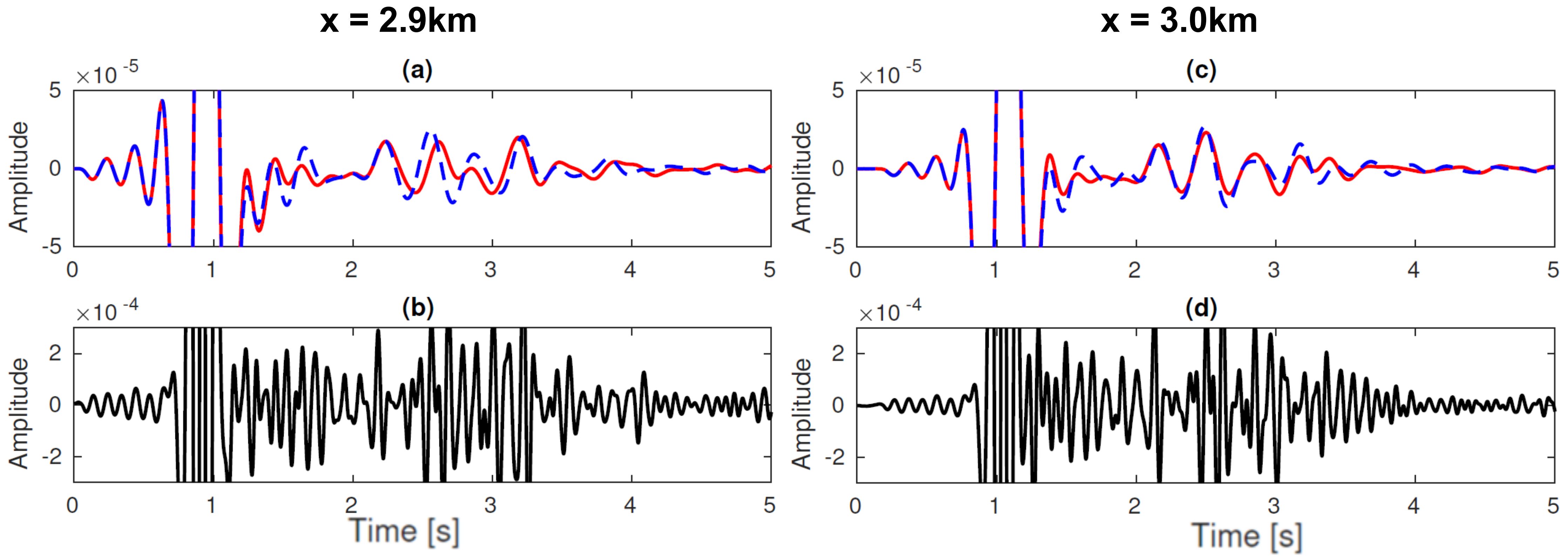




Test error on Marmousi2



Test error on Marmousi2



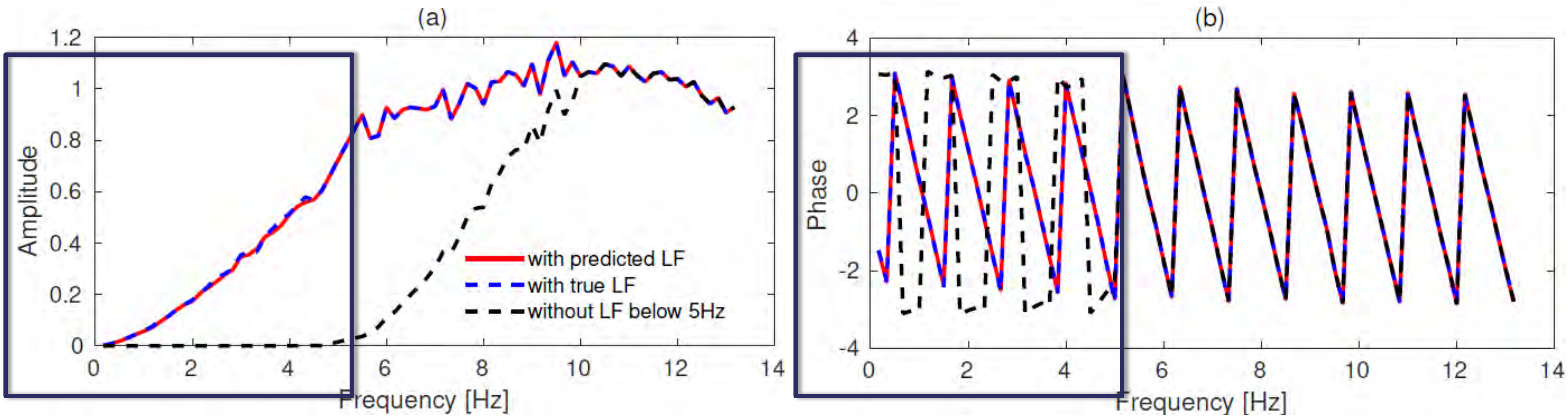
input of CNN, recording **bandlimited** in **5-35Hz**

output of CNN, **predicted** low frequency recording in **0.1-5Hz**

true low frequency recording in **0.1-5Hz**

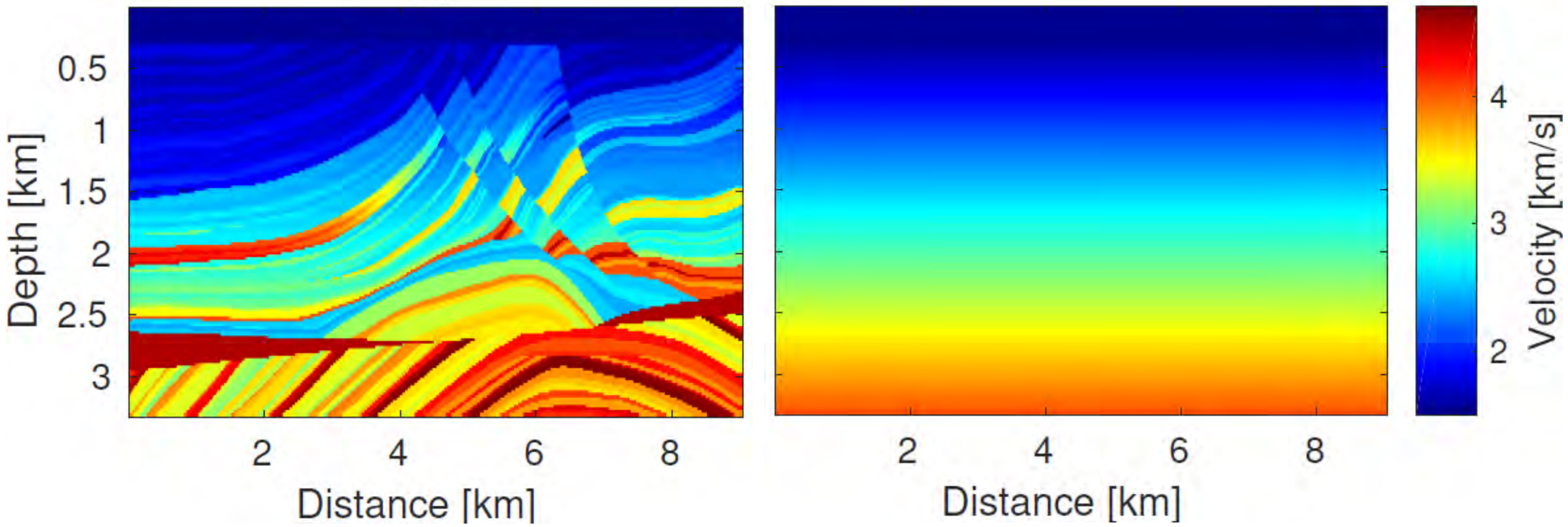
Test error on Marmousi2

Comparison of the amplitude and phase spectrum at the horizontal distance x = 2.9km



Extrapolated FWI with CNN

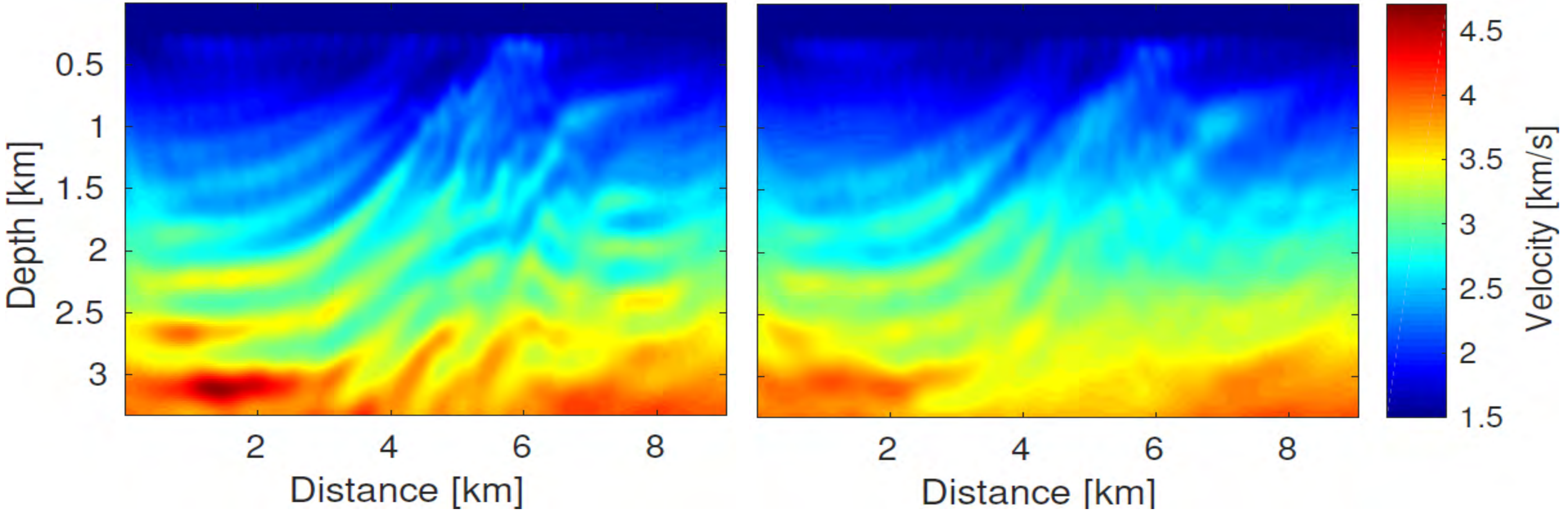
☐ Marmousi2 P-wave velocity model



- maximum offset: 1km
- acoustic modeling
- optimizer: L-BFGS
- surface acquisition with 30 sources

Extrapolated FWI with CNN

FWI using true 0.1-5Hz data | **EFWI using predicted 0.1-5Hz data**

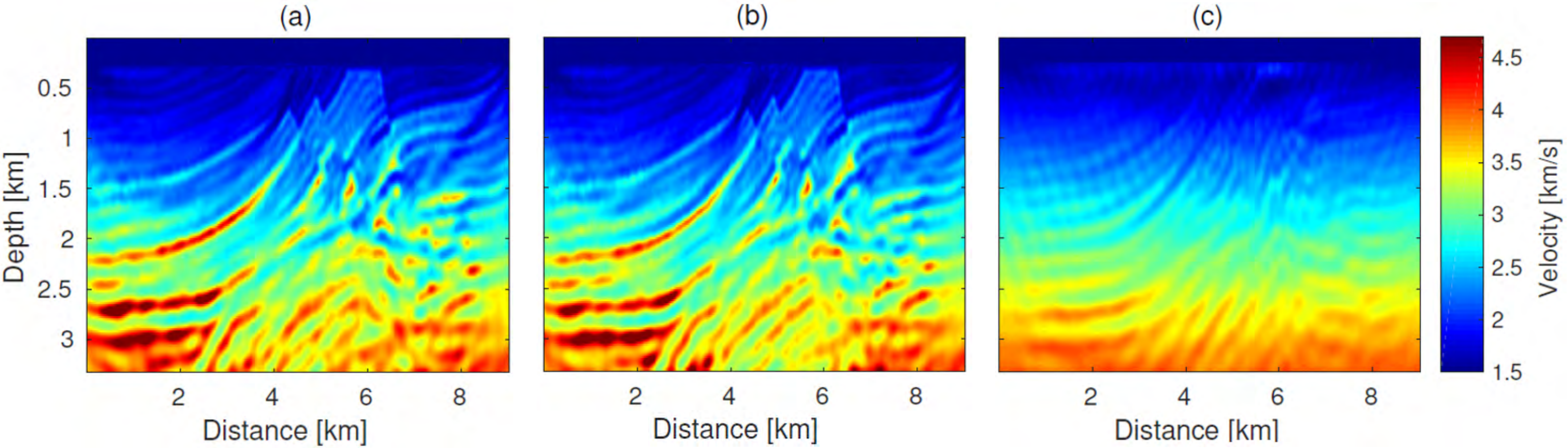


Extrapolated FWI with CNN

FWI started from 0.1-5Hz true data

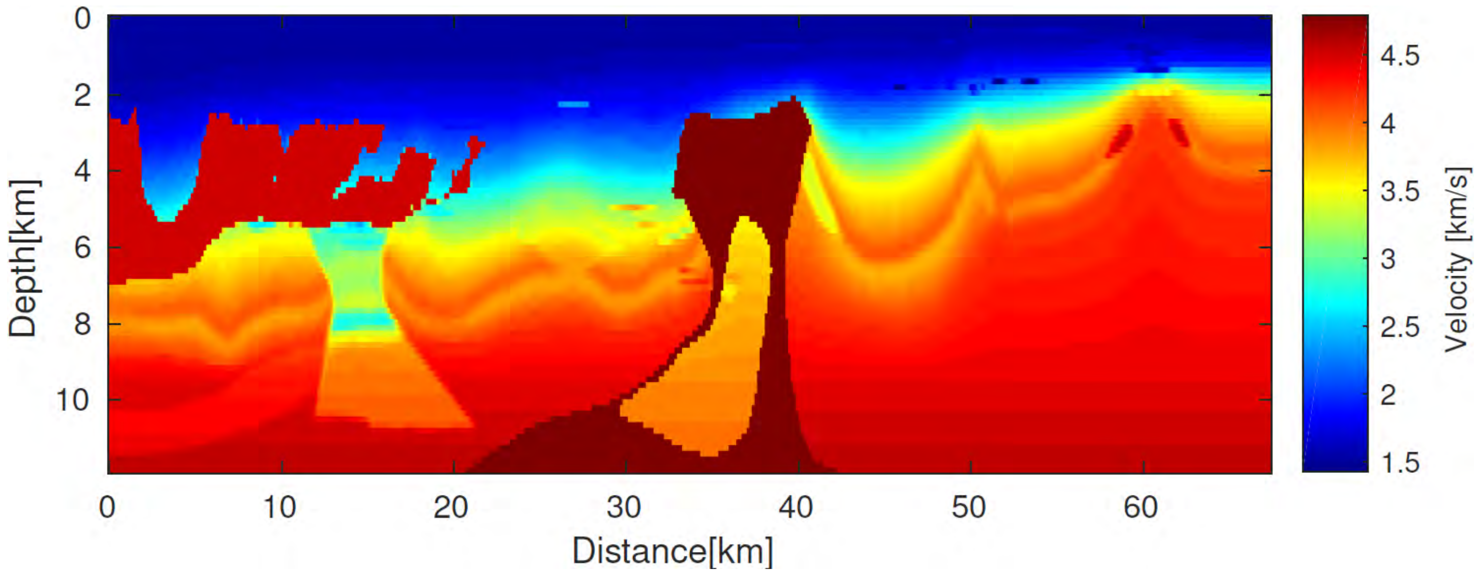
EFWI started from 0.1-5Hz predicted data

FWI started from 5Hz bandlimited data



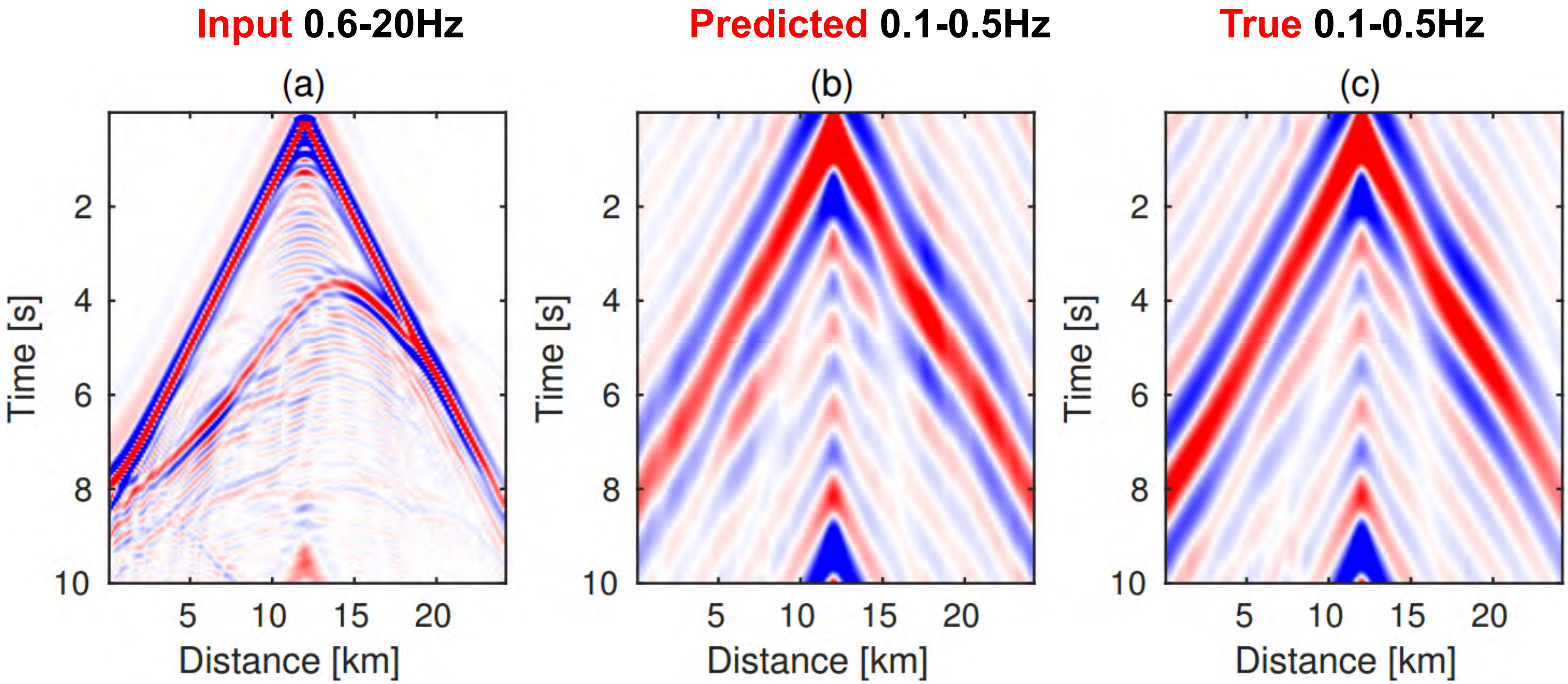
Extrapolated FWI with CNN

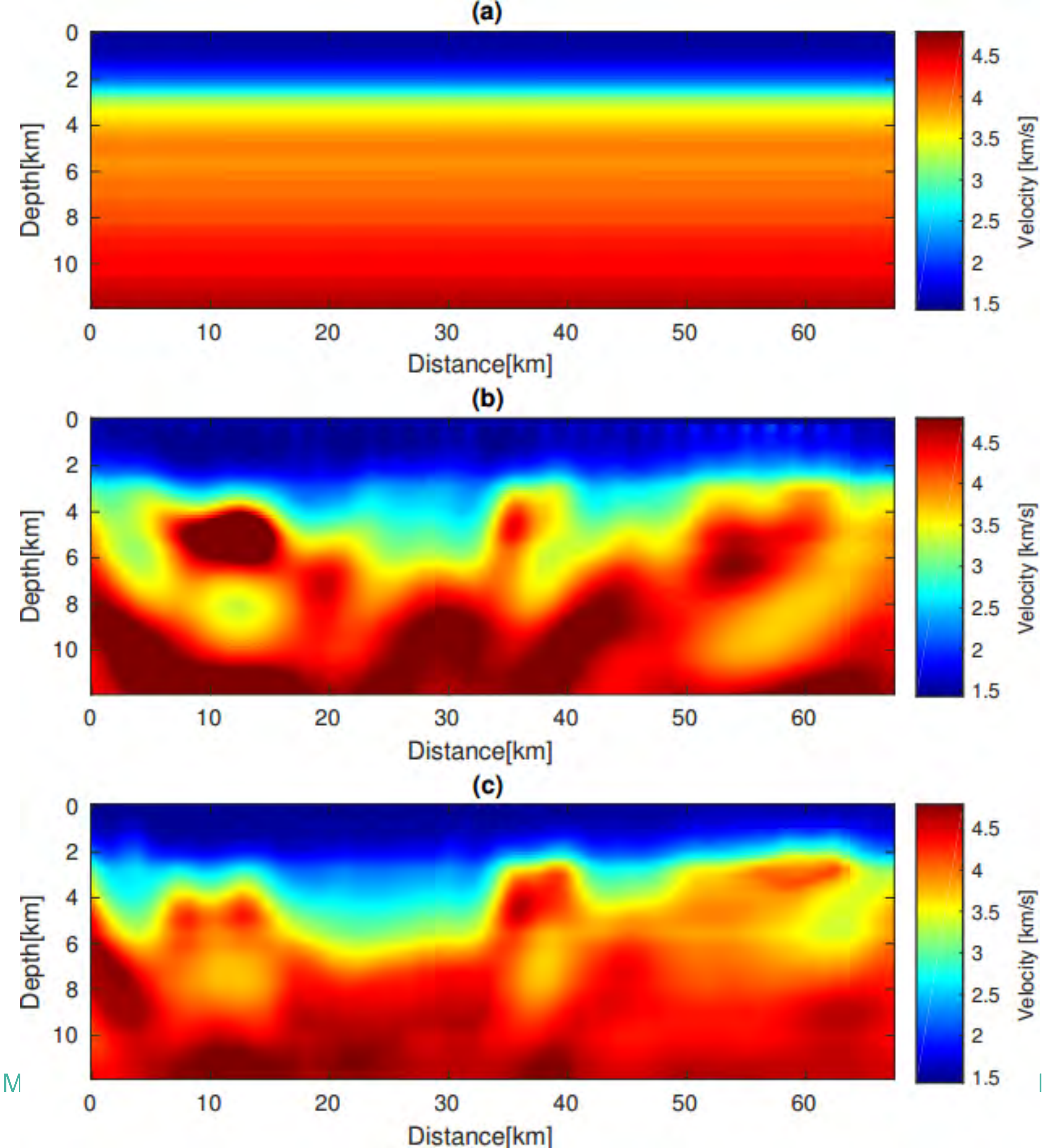
BP 2004 Benchmark model



- optimizer: L-BFGS
- maximum offset: 12km
- surface acquisition with 30 sources (10Hz Ricker wavelet)
- **predict 0.1-0.5Hz low frequency data using 0.6-20Hz bandlimited data**
- **training data are collected from submodels of Marmousi2**

Extrapolated low frequency data



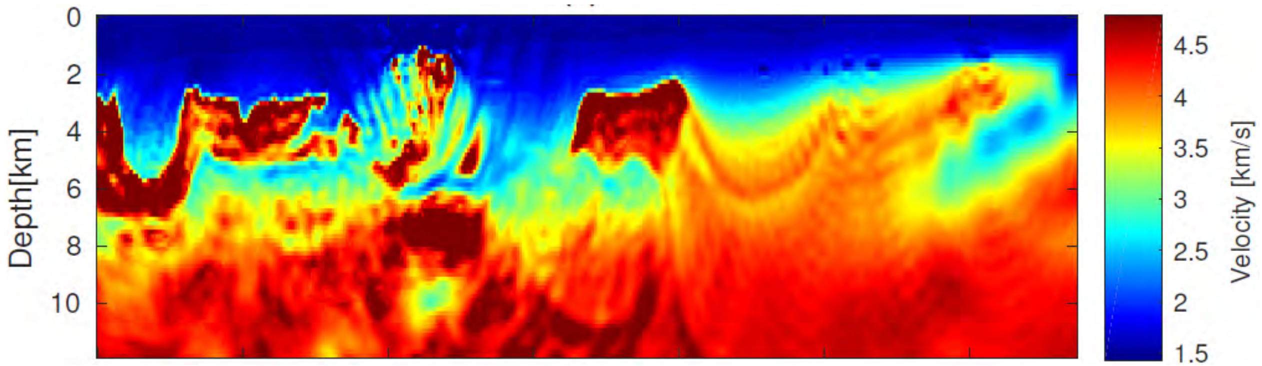


initial model

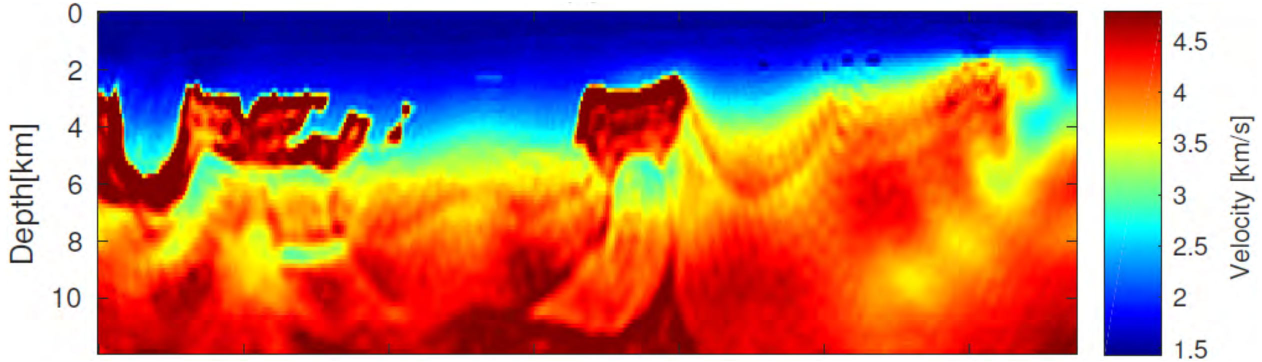
EFWI using 0.3Hz predicted data

FWI using 0.3Hz true data

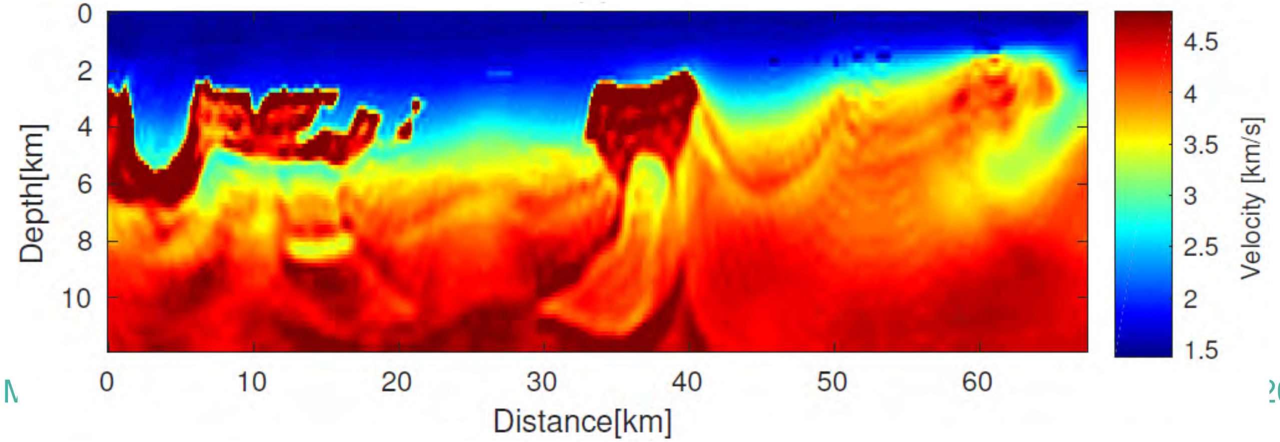
FWI started from 0.6Hz bandlimited data



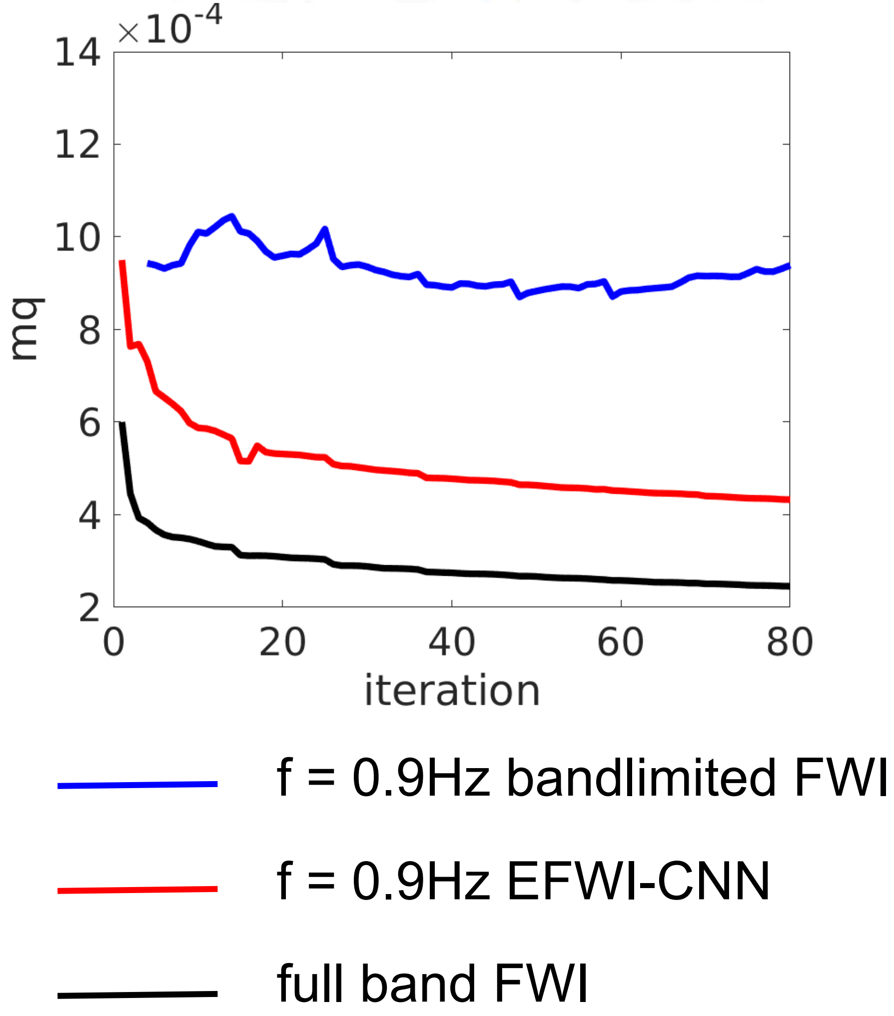
EFWI started from 0.3Hz predicted data



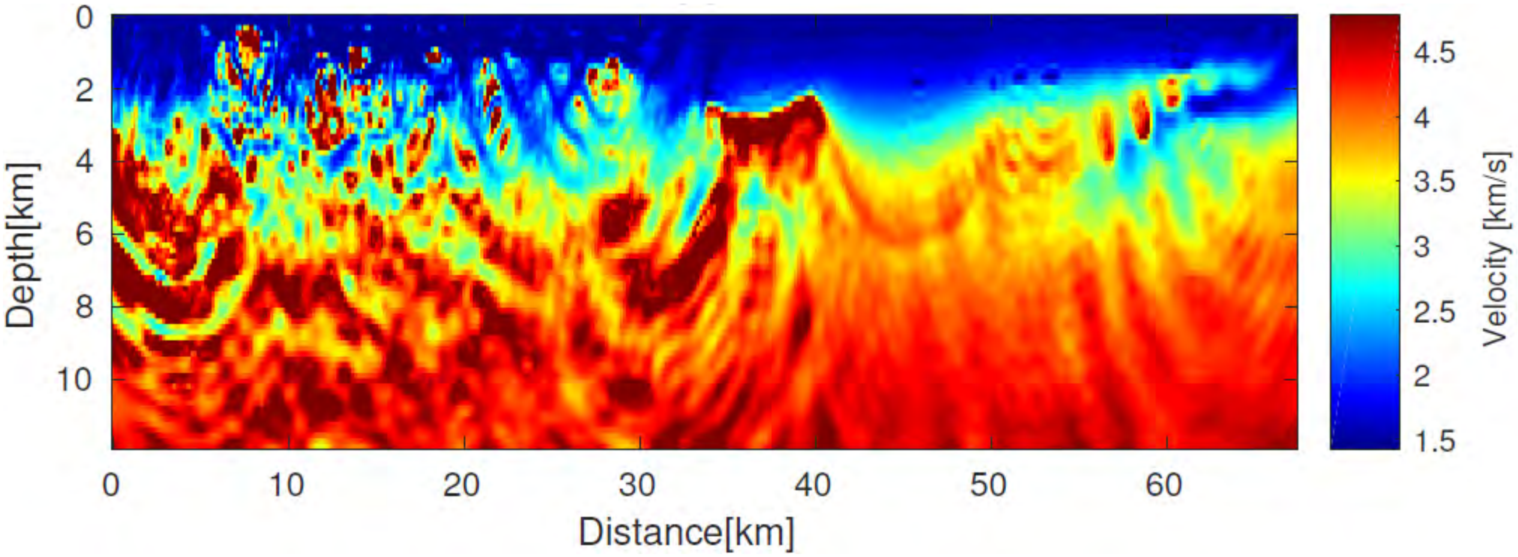
FWI started from 0.3Hz true data



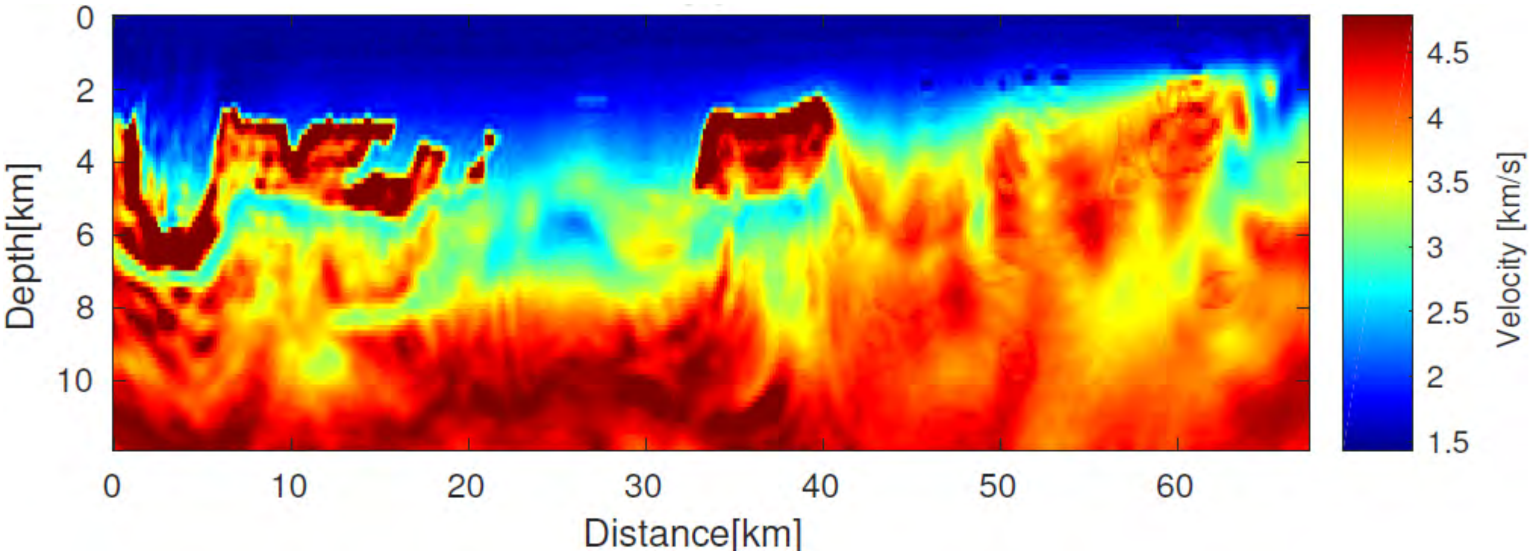
$$mq = \frac{1}{N}\left\|\frac{\mathbf{m}_{inv} - \mathbf{m}_{true}}{\mathbf{m}_{true}}\right\|_2$$



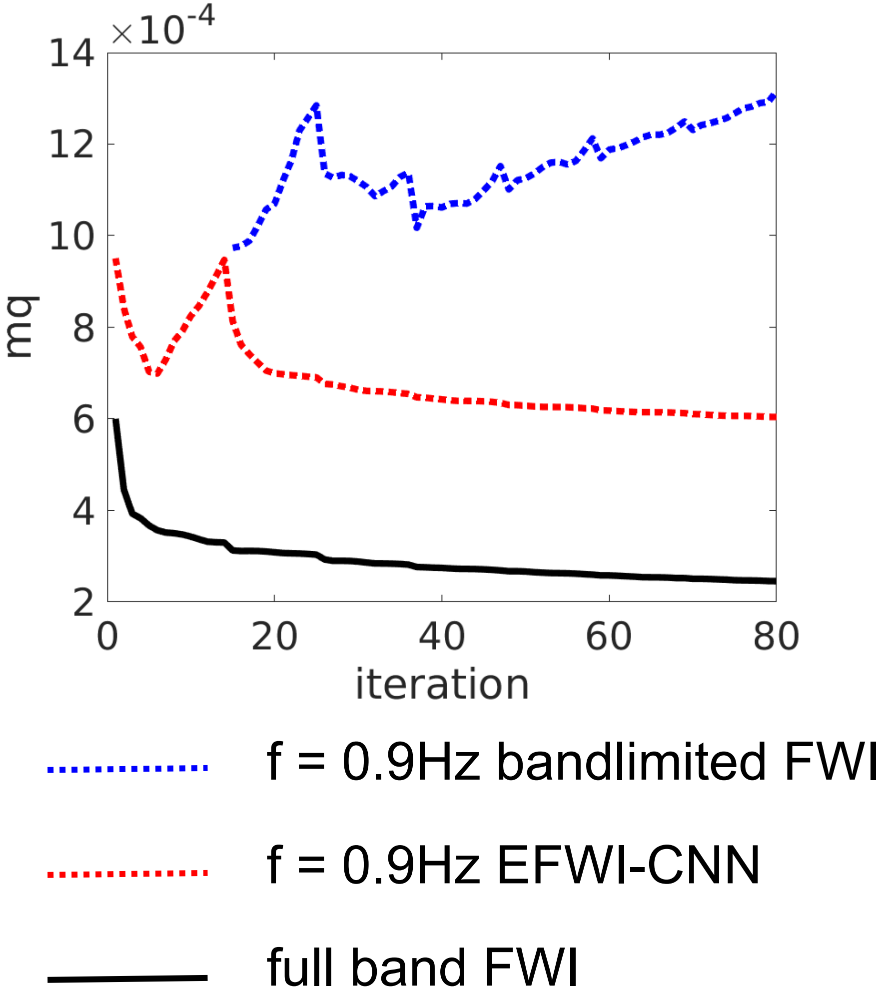
FWI started from 0.9Hz bandlimited data



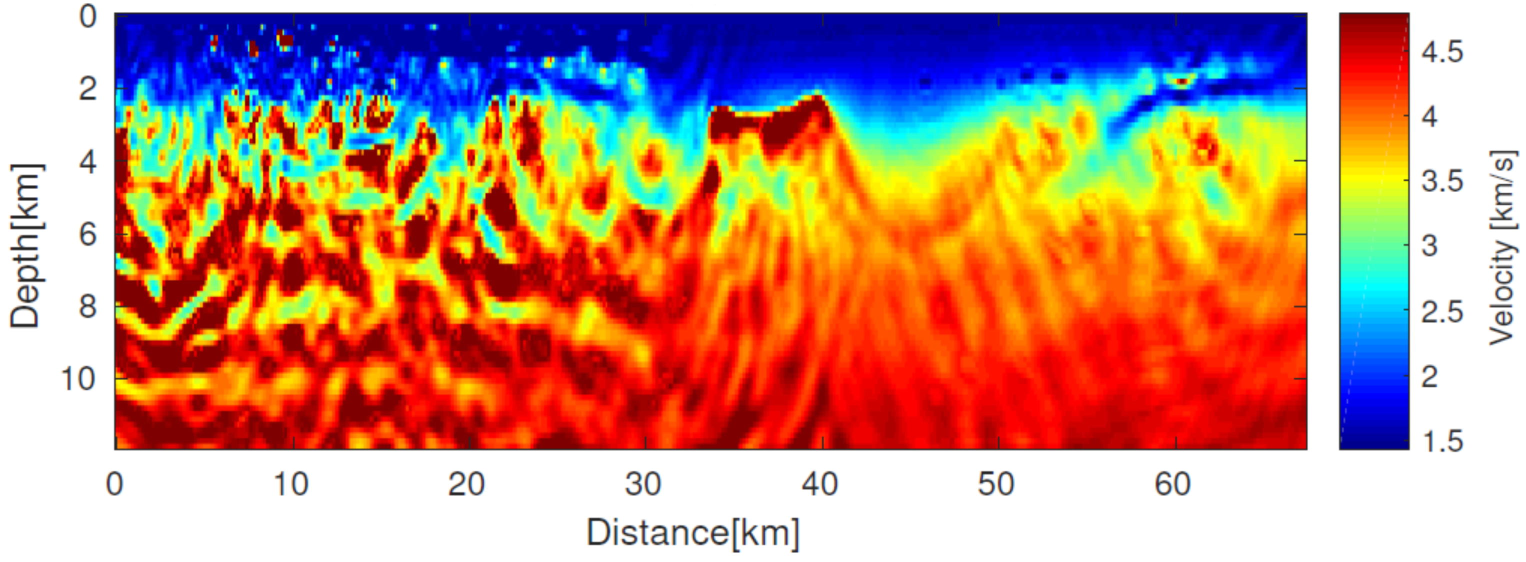
EFWI started from 0.3Hz and 0.6Hz predicted data



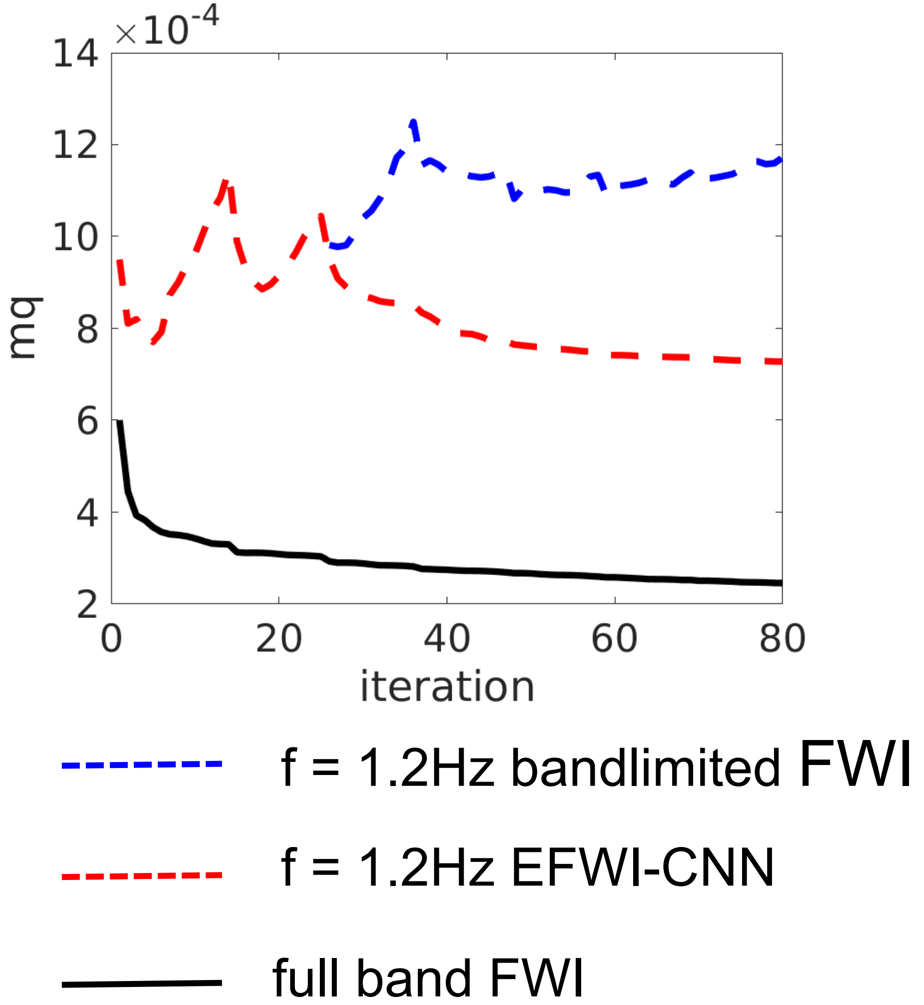
$$mq = \frac{1}{N}\left\|\frac{\mathbf{m}_{inv} - \mathbf{m}_{true}}{\mathbf{m}_{true}}\right\|_2$$



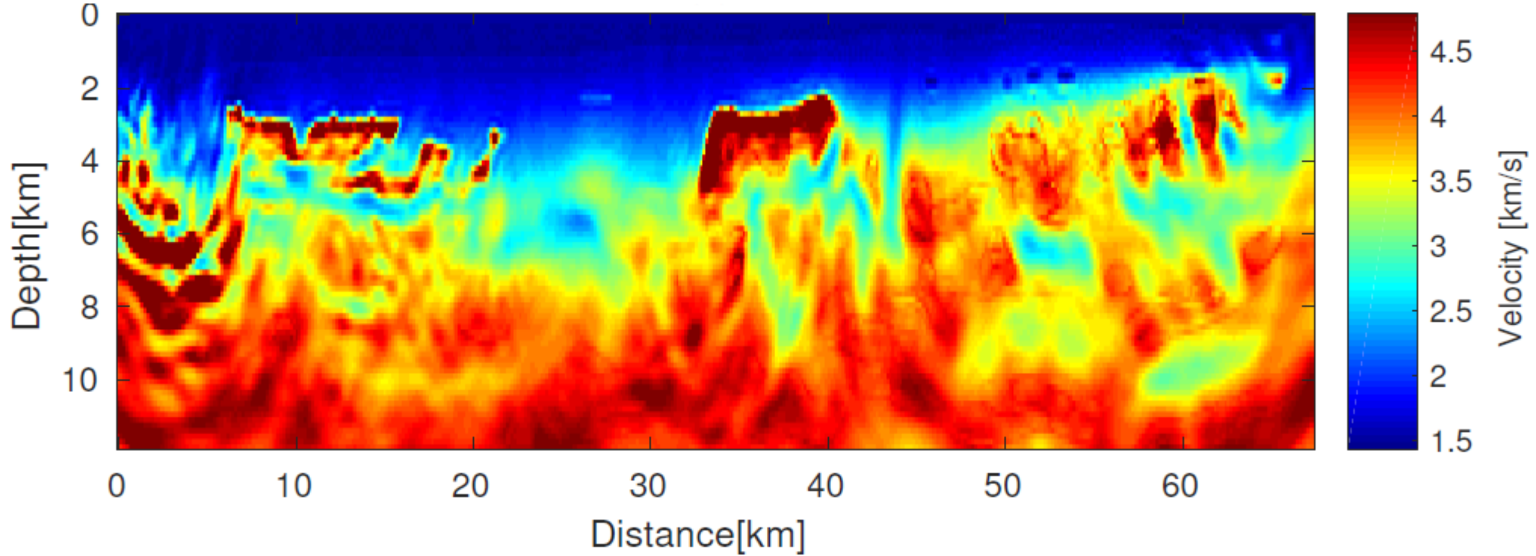
FWI started from 1.2Hz bandlimited data



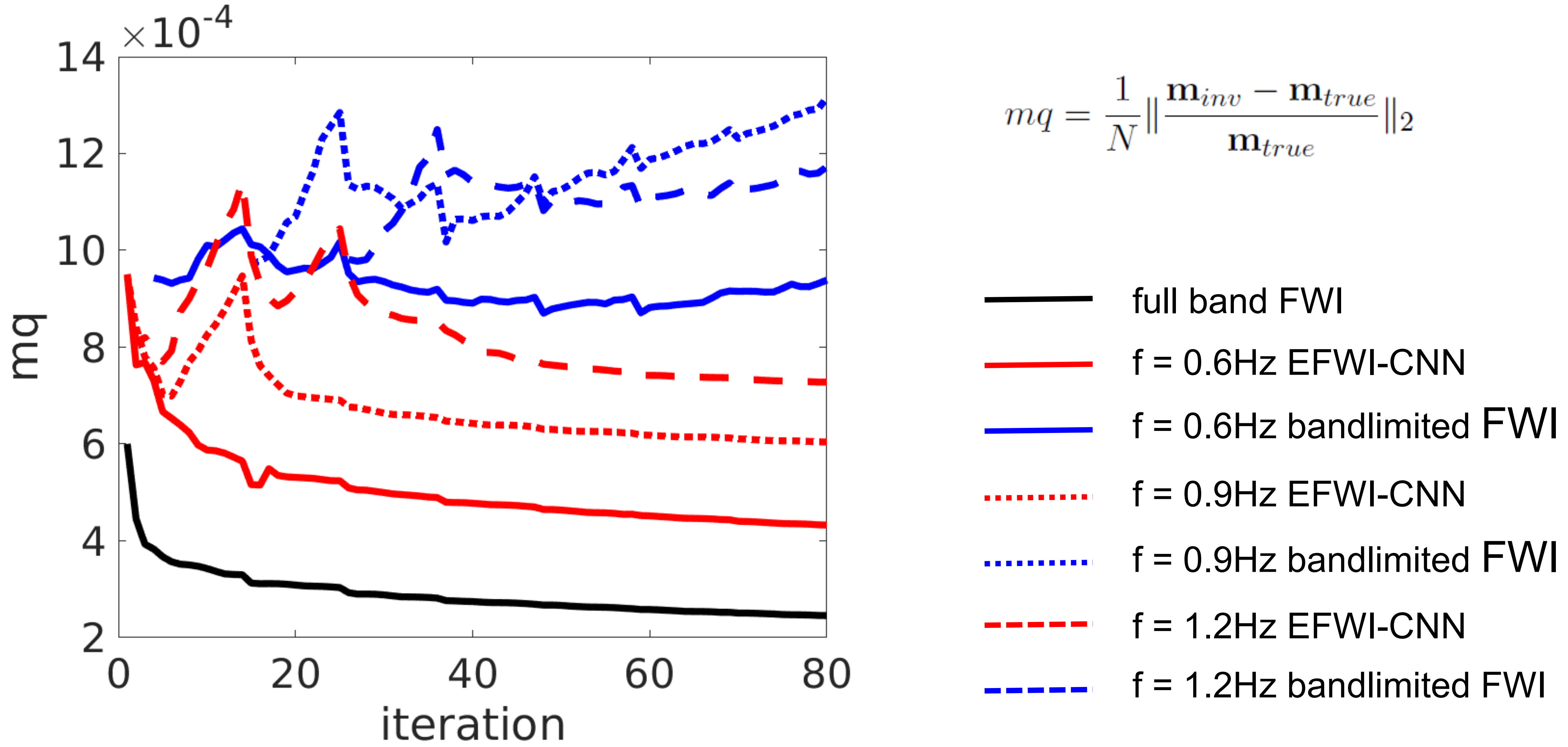
$$mq = \frac{1}{N} \| \frac{\mathbf{m}_{inv} - \mathbf{m}_{true}}{\mathbf{m}_{true}} \|_2$$



EFWI started from 0.3Hz, 0.6Hz and 0.9Hz predicted data



Quality of EFWI-CNN



$$mq = \frac{1}{N}\left\|\frac{\mathbf{m}_{inv} - \mathbf{m}_{true}}{\mathbf{m}_{true}}\right\|_2$$

Conclusions

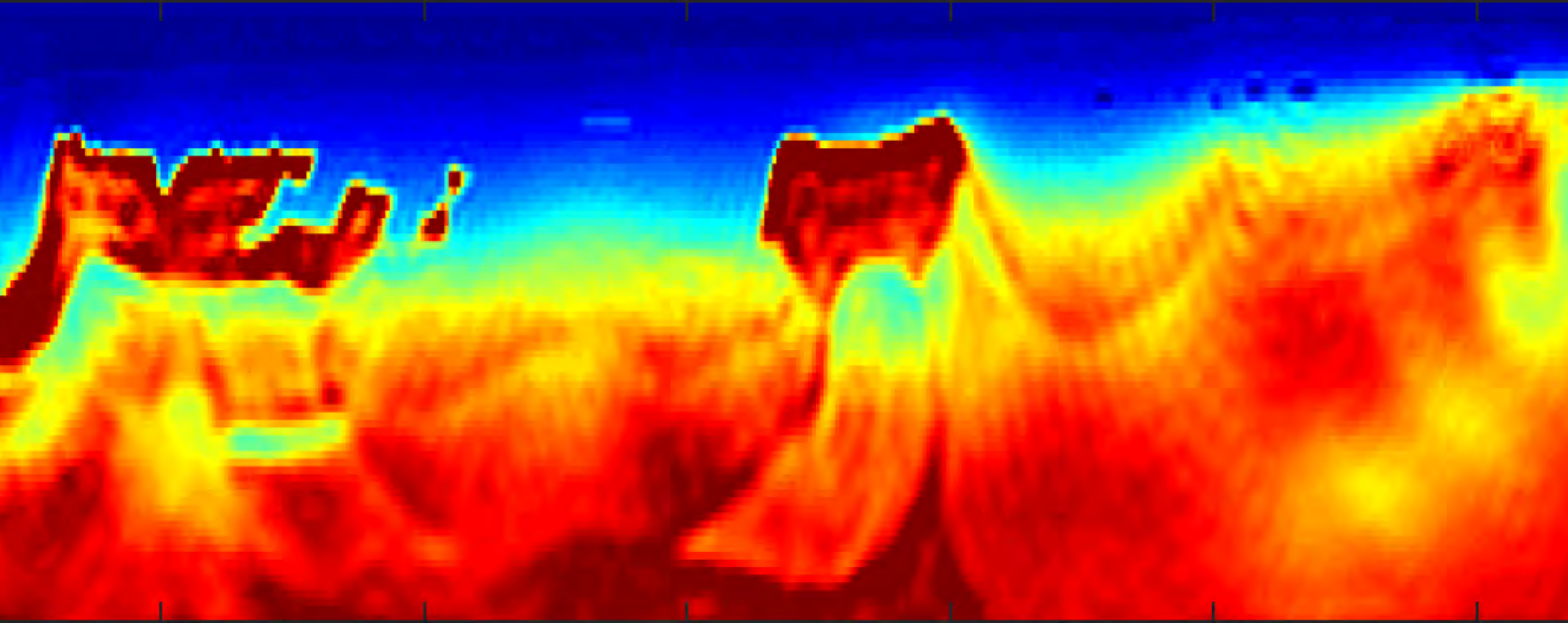
- CNNs have the ability to recover the low frequencies of unknown subsurface structure that are completely missing at the training stage.
- The extrapolated low frequency data can be reliable to seed FWI and mitigate cycle-skipping.
- The choice of the architectural parameters of the deep learning model is non-unique.

- the absence of a physical interpretation for the operations performed by the network

Acknowledgements

Thanks to MIT ERL and Total S A for support.

Thanks for your attention.

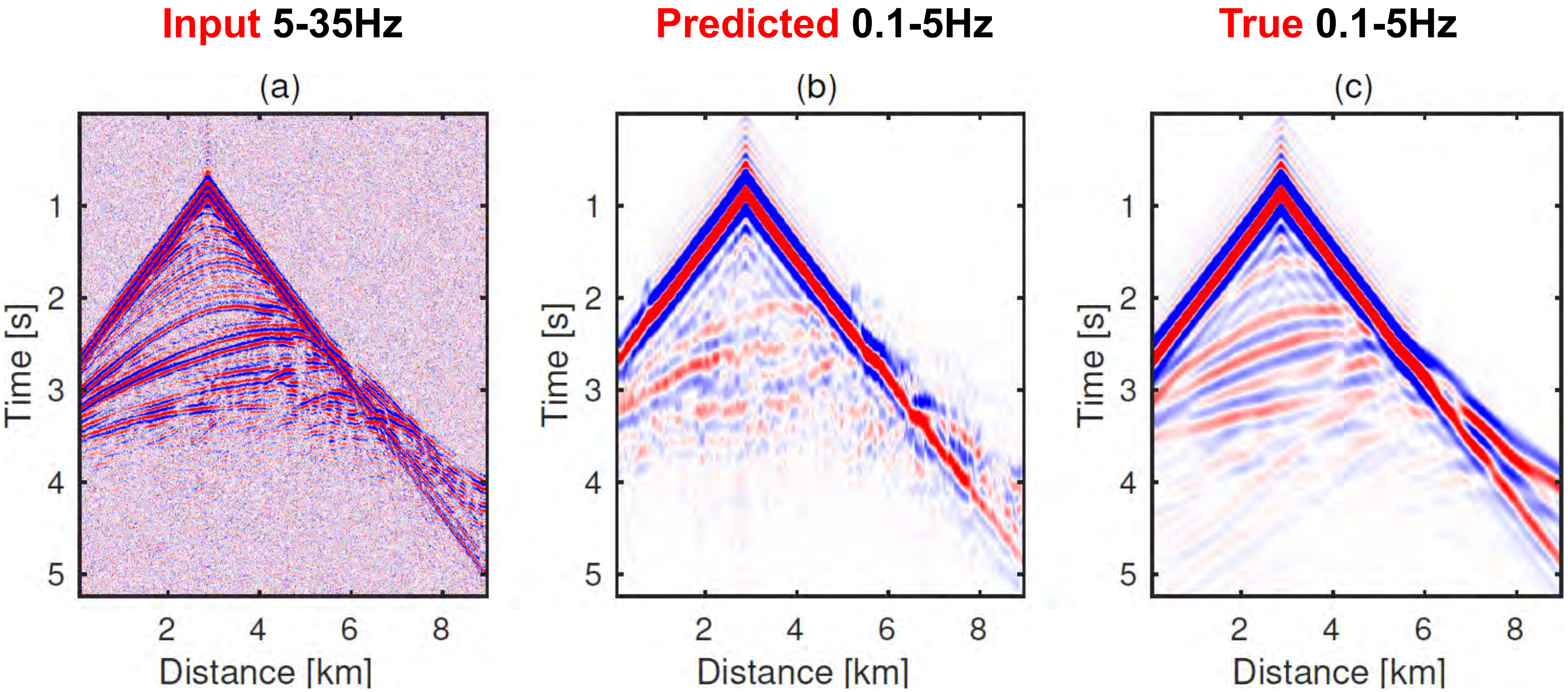


Extrapolated Full Waveform Inversion with Deep Learning

Hongyu Sun and Laurent Demanet

Uncertainty Analysis

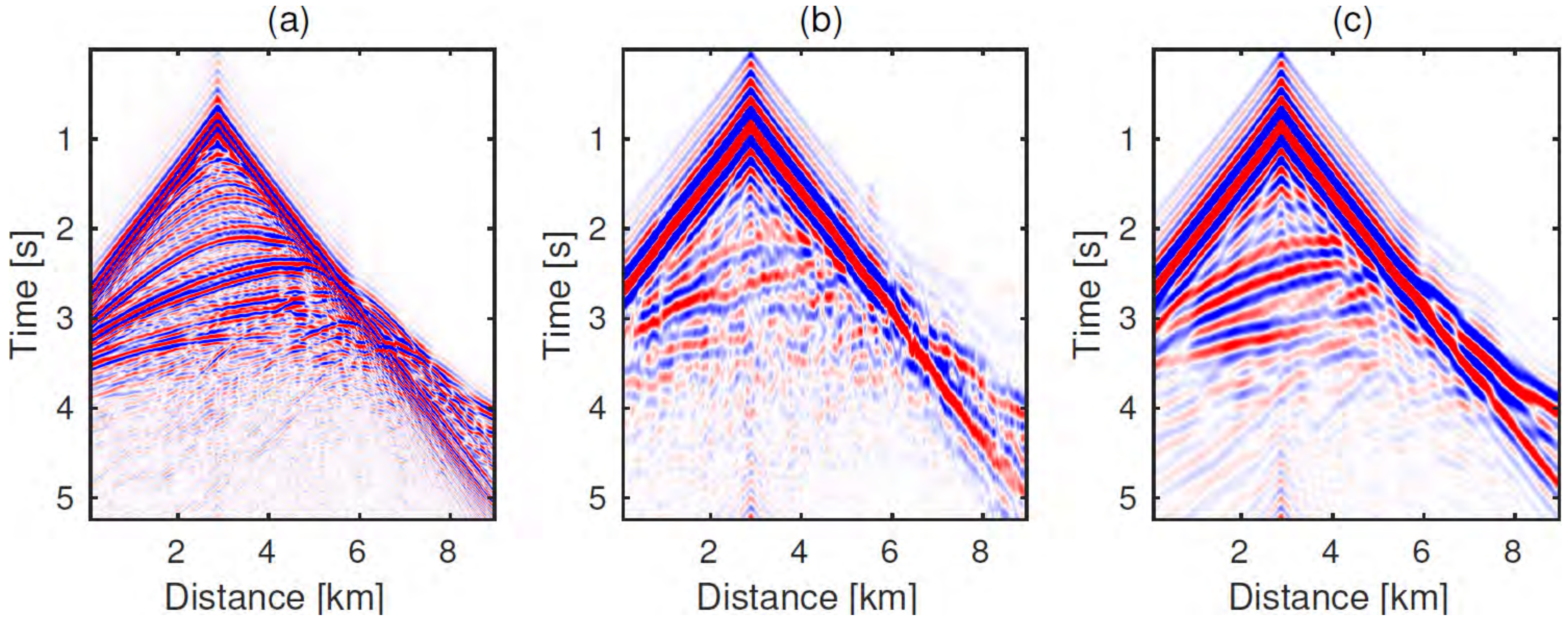
□ Robustness with noise



Uncertainty Analysis

□ Different forward modeling solver

Input 5-35Hz **Predicted** 0.1-5Hz **True** 0.1-5Hz



Uncertainty Analysis

□ Unknown source wavelet