MIT EARTH RESOURCES LABORATORY
ANNUAL FOUNDING MEMBERS MEETING 2020

Elastic full waveform inversion with extrapolated low frequency data

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

Motivation: elastic full waveform inversion

Why elastic FWI

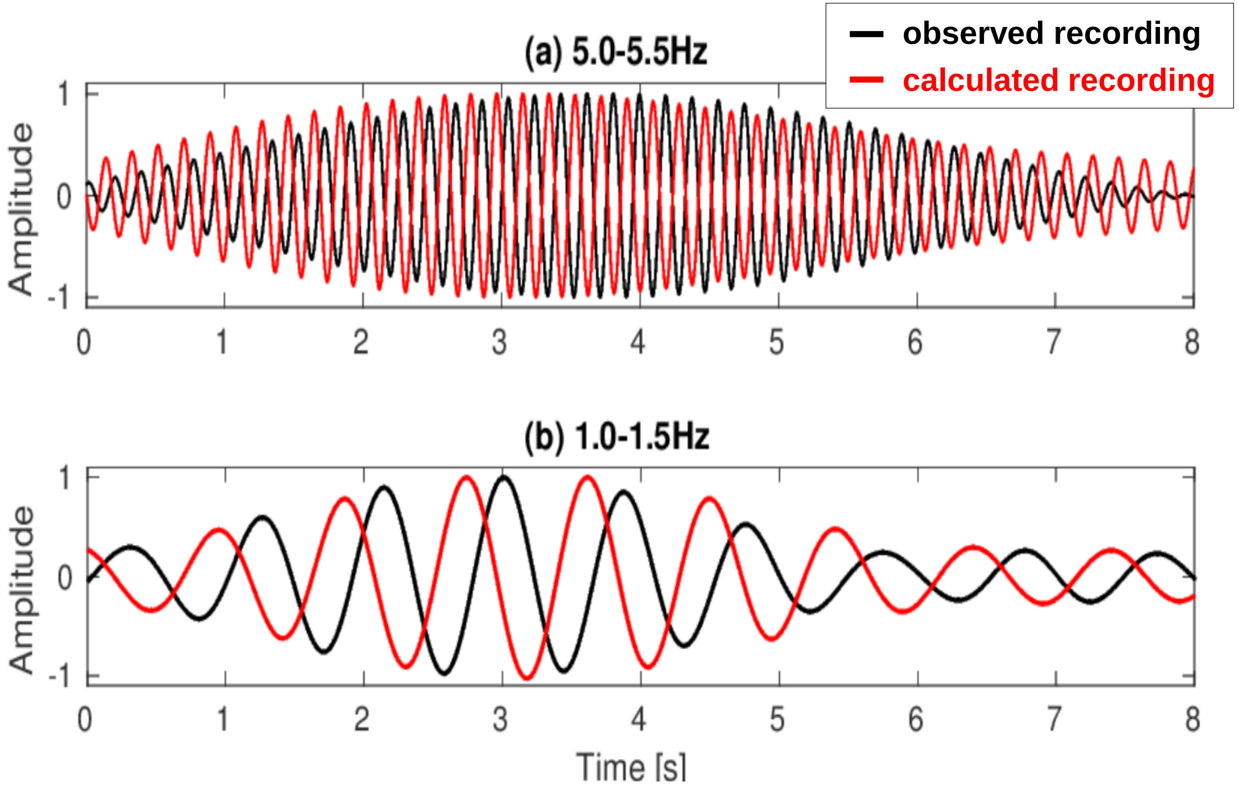
- strong elastic effects;
- reservoir characterization;
- near-surface investigations;
- …

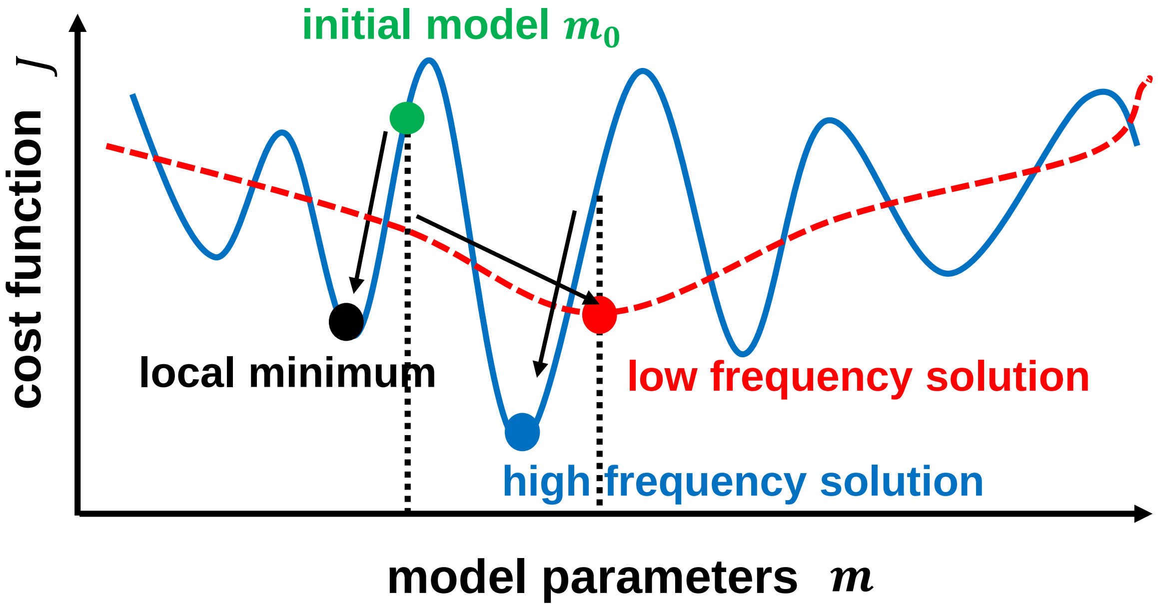
How (Tarantola, 1986; Mora, 1987; K*ö*hn et al., 2012)

$$J = \frac{1}{2}\delta \boldsymbol{d}^T \delta \boldsymbol{d} = \frac{1}{2}\sum_s \sum_r \int [\underbrace{\boldsymbol{u}_{cal}}_{v_x \quad v_y} - \underbrace{\boldsymbol{u}_{obs}}_{v_x \quad v_y}]^2 dt$$

$$\boldsymbol{m}_{k+1} = \boldsymbol{m}_k - \mu_k \left(\frac{\partial J}{\partial \boldsymbol{m}}\right)_k \longrightarrow \boldsymbol{m}\text{: } v_p,\ v_s,\ \rho$$

Motivation: cycle-skipping





Cycle-skipping is more severe in elastic FWI compared to acoustic FWI, due to the short S-wave wavelength.

Motivation: elastic full waveform inversion

Cycle-skipping is more serve in elastic FWI and requires lower starting frequency.

Synthetic data studies:

	name of benchmark model	starting model	starting frequency
Brossier et al., 2009	Overthrust	Gaussian smoothing	**1.7Hz**
Brossier et al., 2010	Valhall	Gaussian smoothing	**2.0Hz**
Choi et al., 2008	Marmousi2	velocity-gradient	**0.16Hz**
Köhn et al., 2012	Marmousi2	velocity-gradient	**0-2Hz**
Jeong et al., 2012	Marmousi2	velocity-gradient	**0.12Hz**

Field data studies: lack of low frequencies

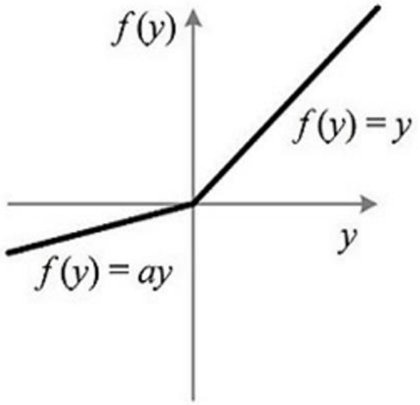
Crase et al., 1990; Sears et al., 2010; Vigh et al.,2014; Raknes et al.,2015; Marjanovi´c et al., 2018; Borisov et al., 2020, etc.

Method: 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



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

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

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

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

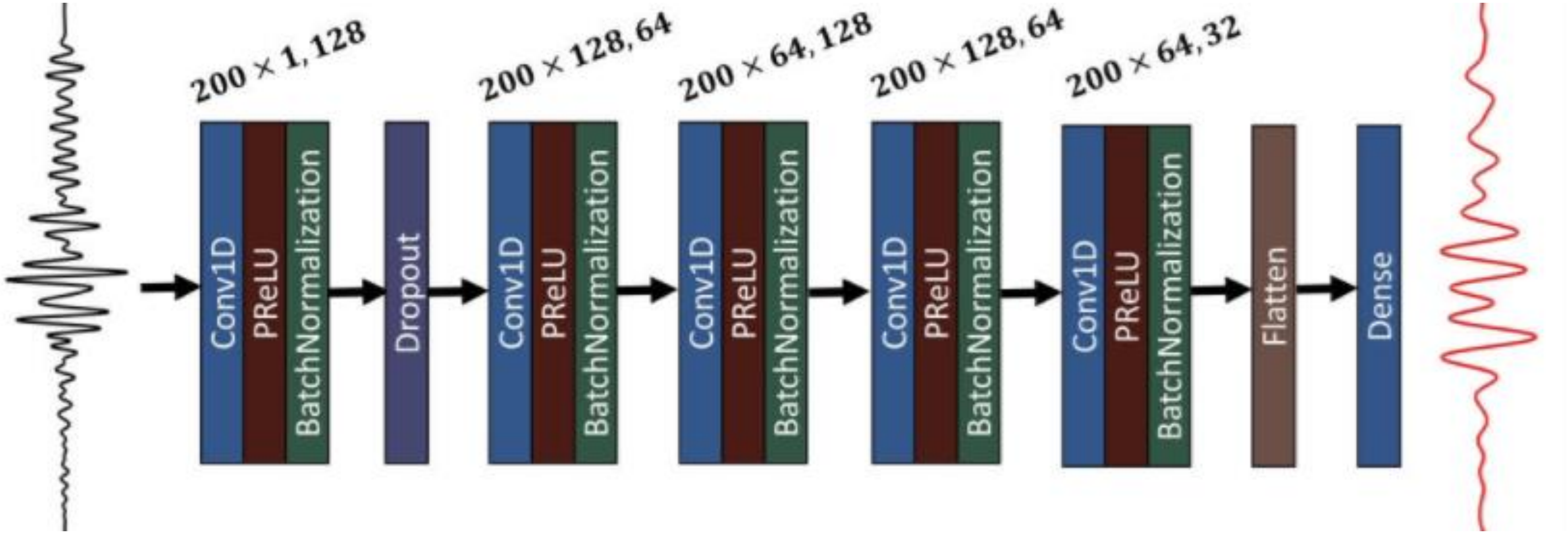
- **Test** (predict): $f(\boldsymbol{x}, \boldsymbol{w})$

(Sun and Demanet, 2018)

Method: architecture of convolutional neural networks

A large receptive field is achieved by directly using a large filter**:**

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



(Sun and Demanet, 2020)

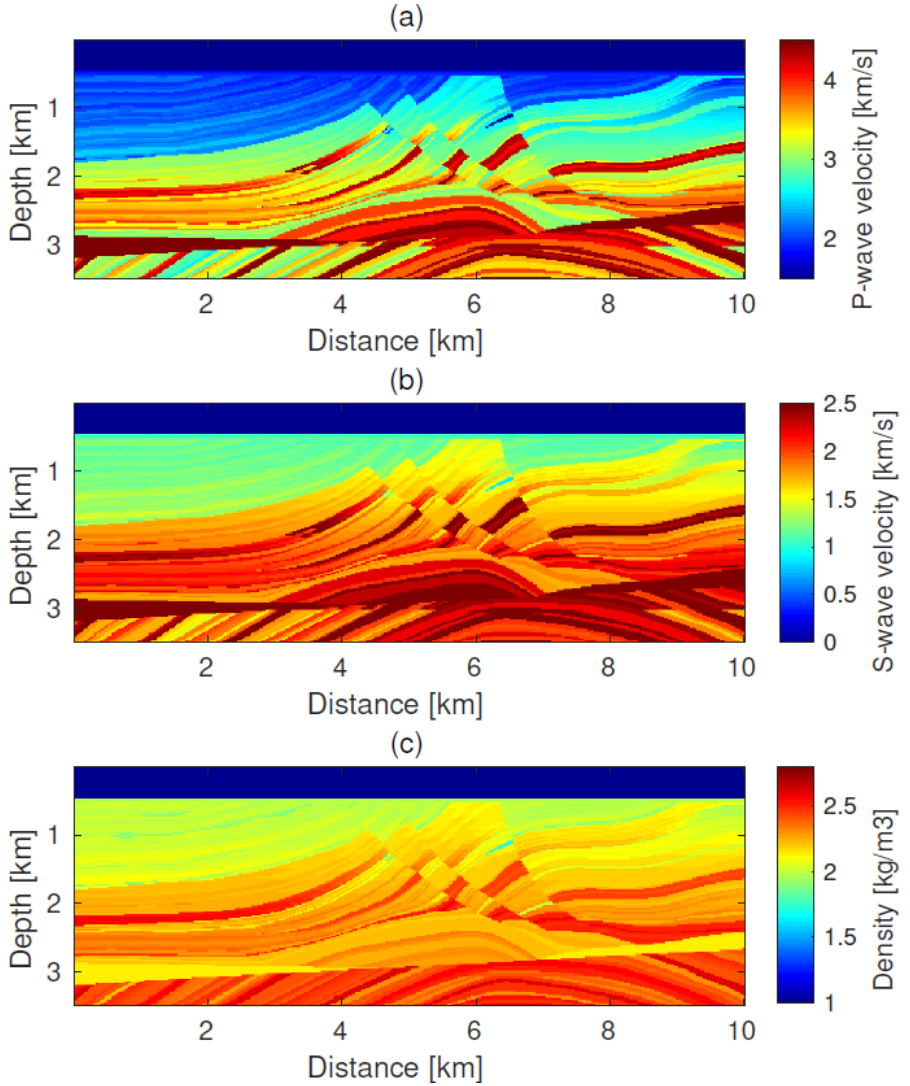
Method: training and test datasets

Test model: unknown low frequencies

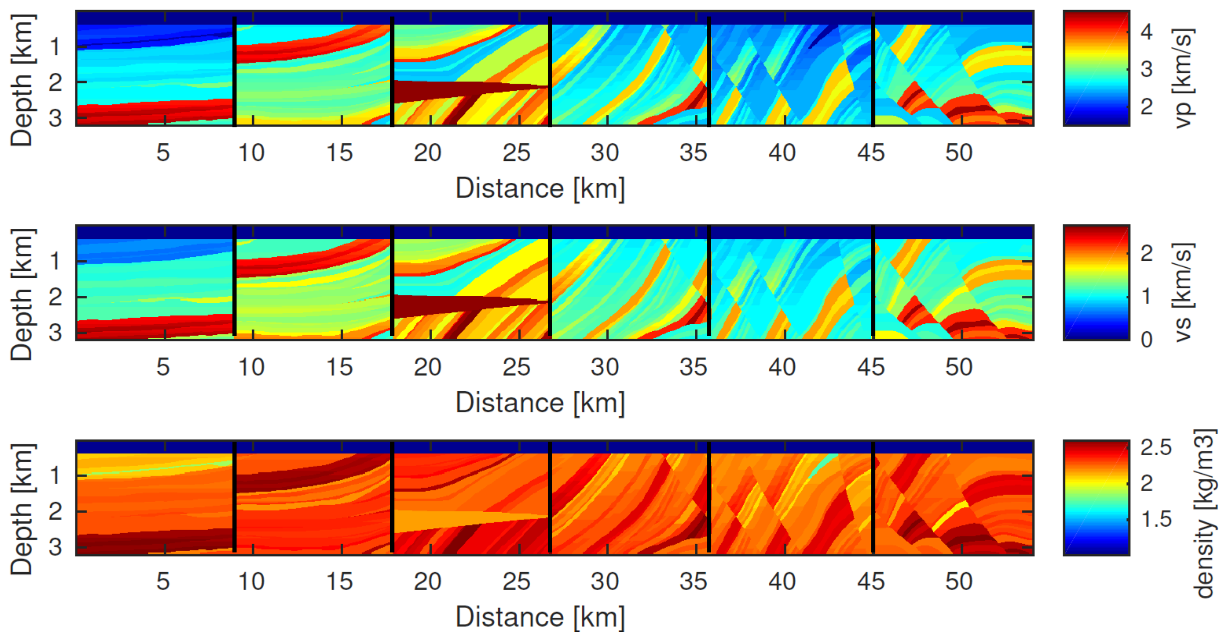
v_p

v_s

ρ



Training model: known low frequencies



- **horizontal component v_x**

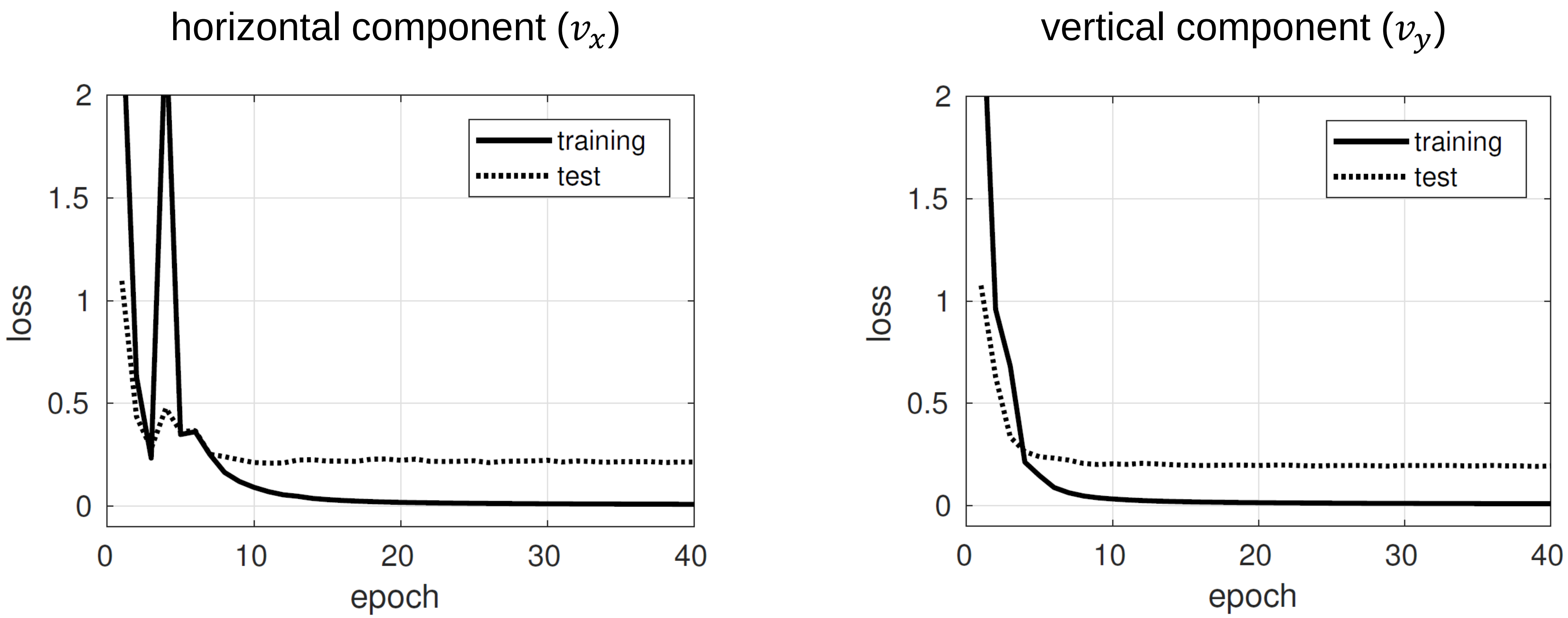
  training dataset: 6 models × 100 shots × 400 receivers
  test dataset: 1 models × 50 shots × 400 receivers

- **vertical component v_y**

  training dataset: 6 models × 100 shots × 400 receivers
  test dataset: 1 models × 50 shots × 400 receivers

Low frequency extrapolation of multicomponent data

Extrapolate 0.1 - 5Hz low frequency data from 5 - 25Hz bandlimited data using ARCH1



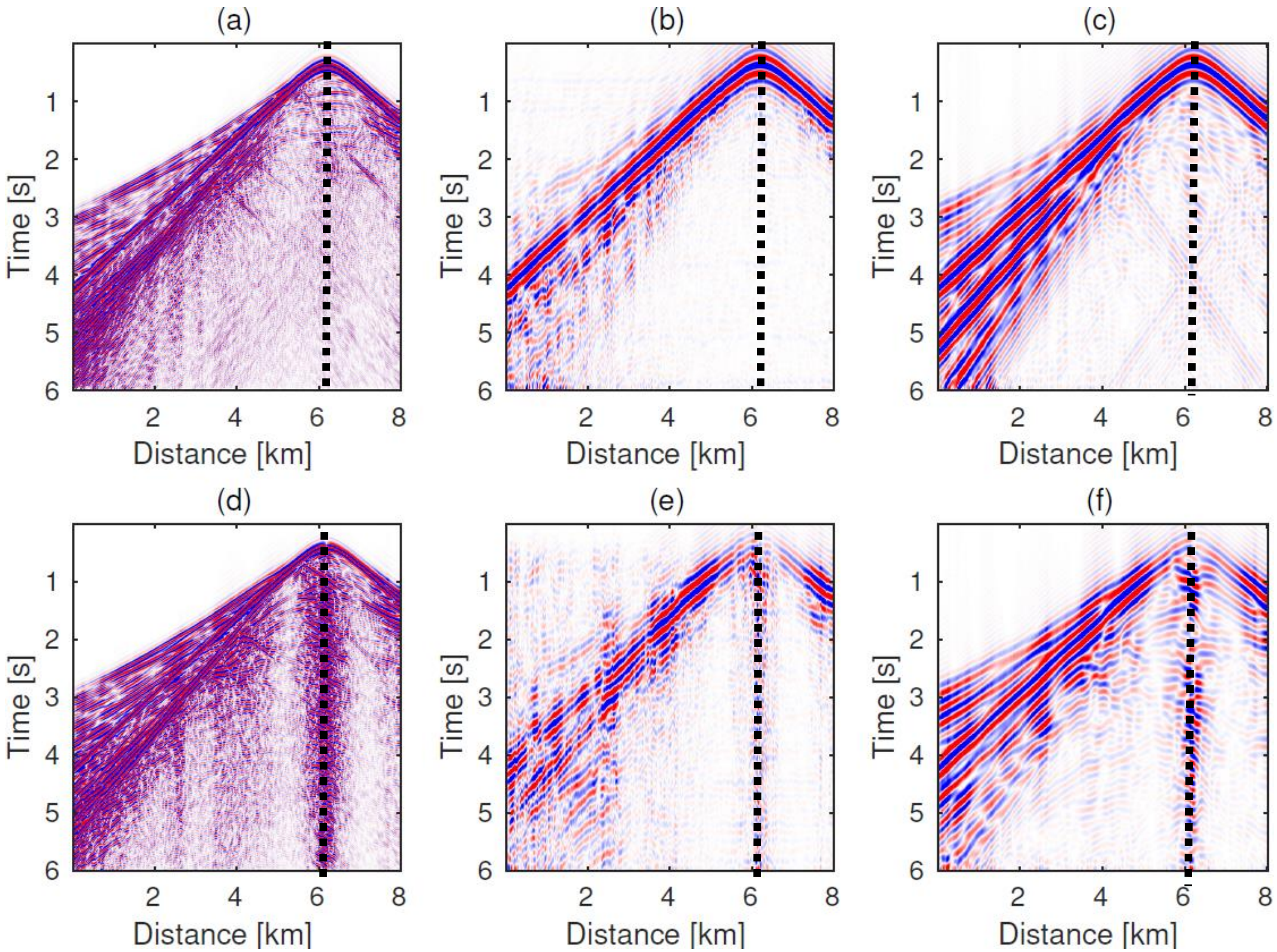
Low frequency extrapolation of multicomponent data

Extrapolation results of ARCH1 **trained on elastic data**

Input 5-25Hz | **Predicted 0.1-5Hz** | **True 0.1-5Hz**

vertical component (v_y)

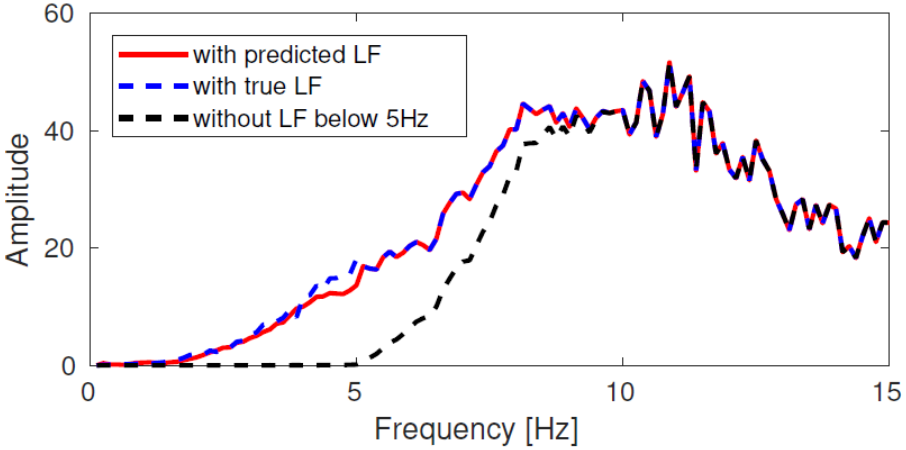
horizontal component (v_x)

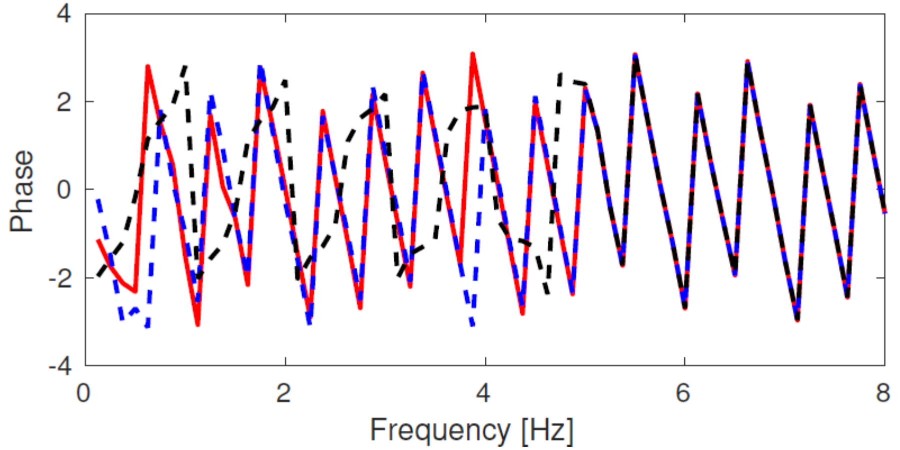


Low frequency extrapolation of multicomponent data

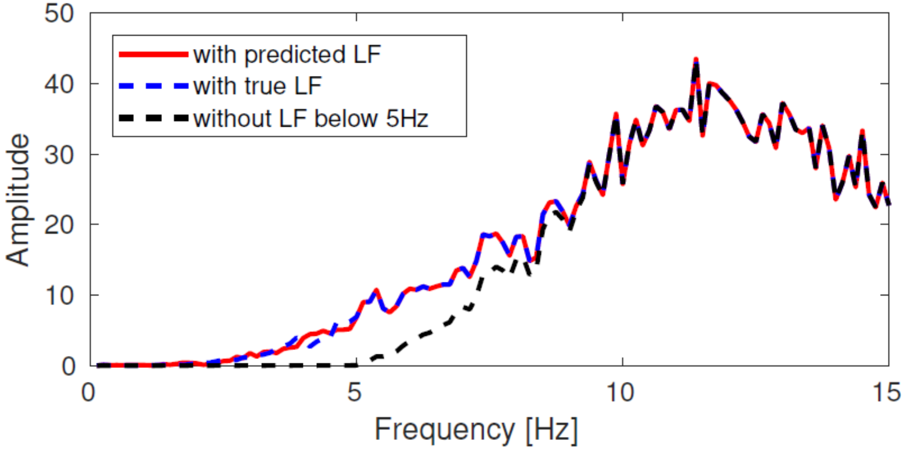
Extrapolation results of ARCH1 **trained on elastic data**

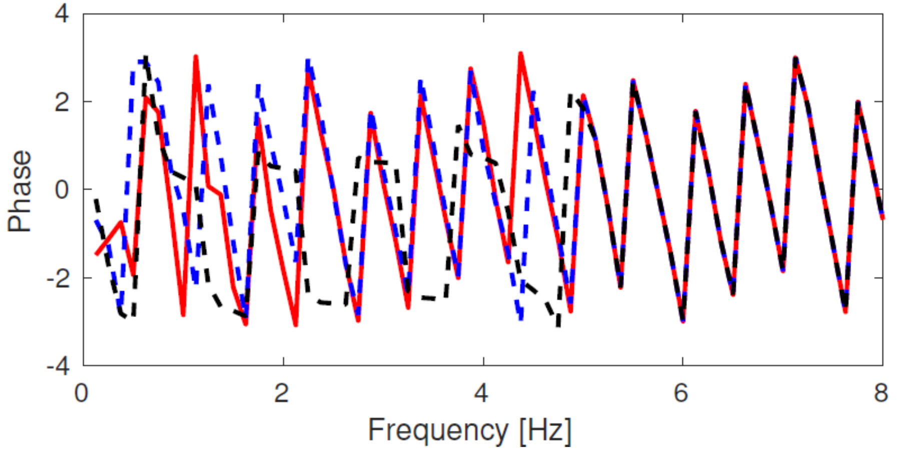
vertical component (v_y)





horizontal component (v_x)

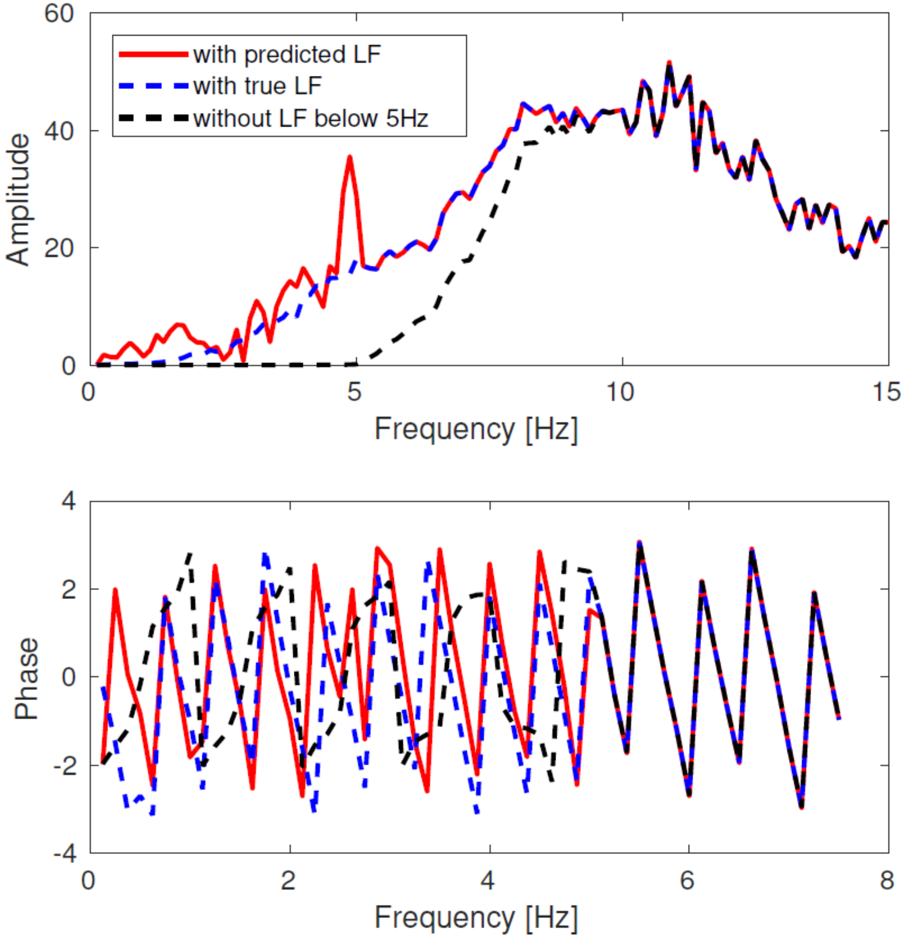




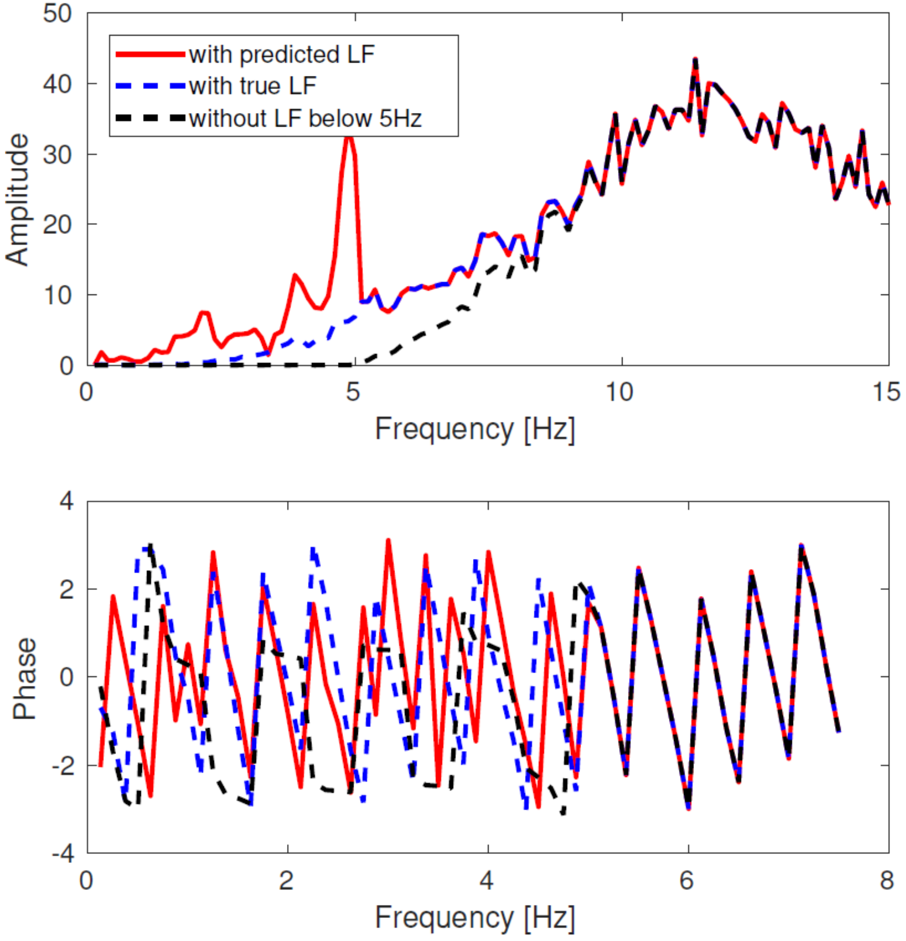
Generalization over physical models

Extrapolation results of ARCH1 **trained on acoustic data**

vertical component (v_y)



horizontal component (v_x)



Extrapolated elastic FWI

- source: Ricker wavelet with a dominant frequency of 10Hz;
- extrapolate 0.1-4Hz low frequency data from 4-25Hz bandlimited data using ARCH1 (the lower band of the bandlimited data is 4Hz);
- a free surface condition is applied to the top of models (multiples in training and test datasets);
- multi-scale FWI: 2-4Hz, 4-6Hz, 4-10Hz and 4-20Hz (Bunks et al.,1995);
- optimizer: L-BFGS with 30 iterations in each frequency band;

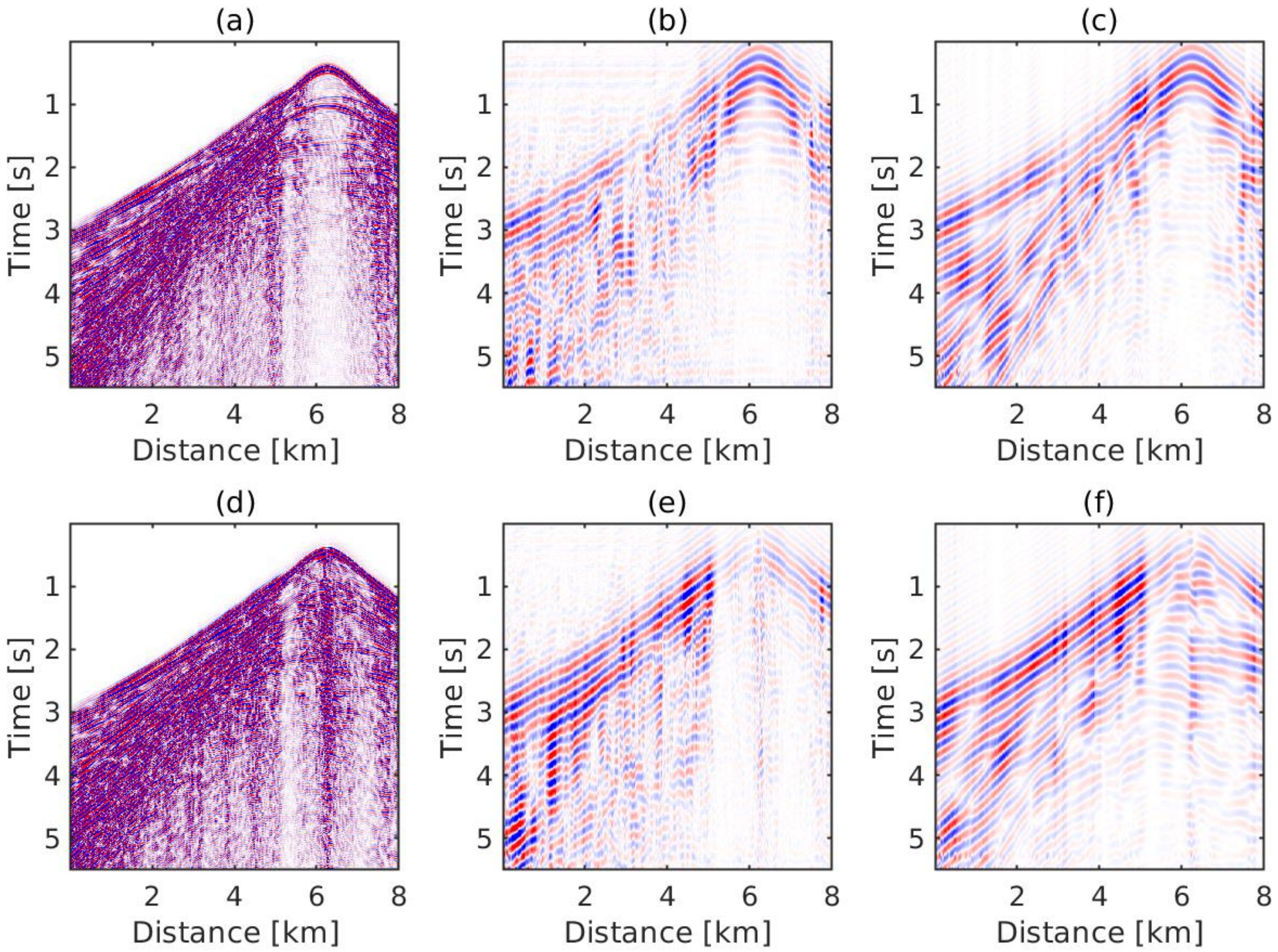
Extrapolated elastic FWI

Extrapolation results of ARCH1 **trained on elastic data with multiples**

Input **4-25Hz** | **Predicted** **0.1-4Hz** | **True** **0.1-4Hz**

vertical component
(v_y)

horizontal component
(v_x)



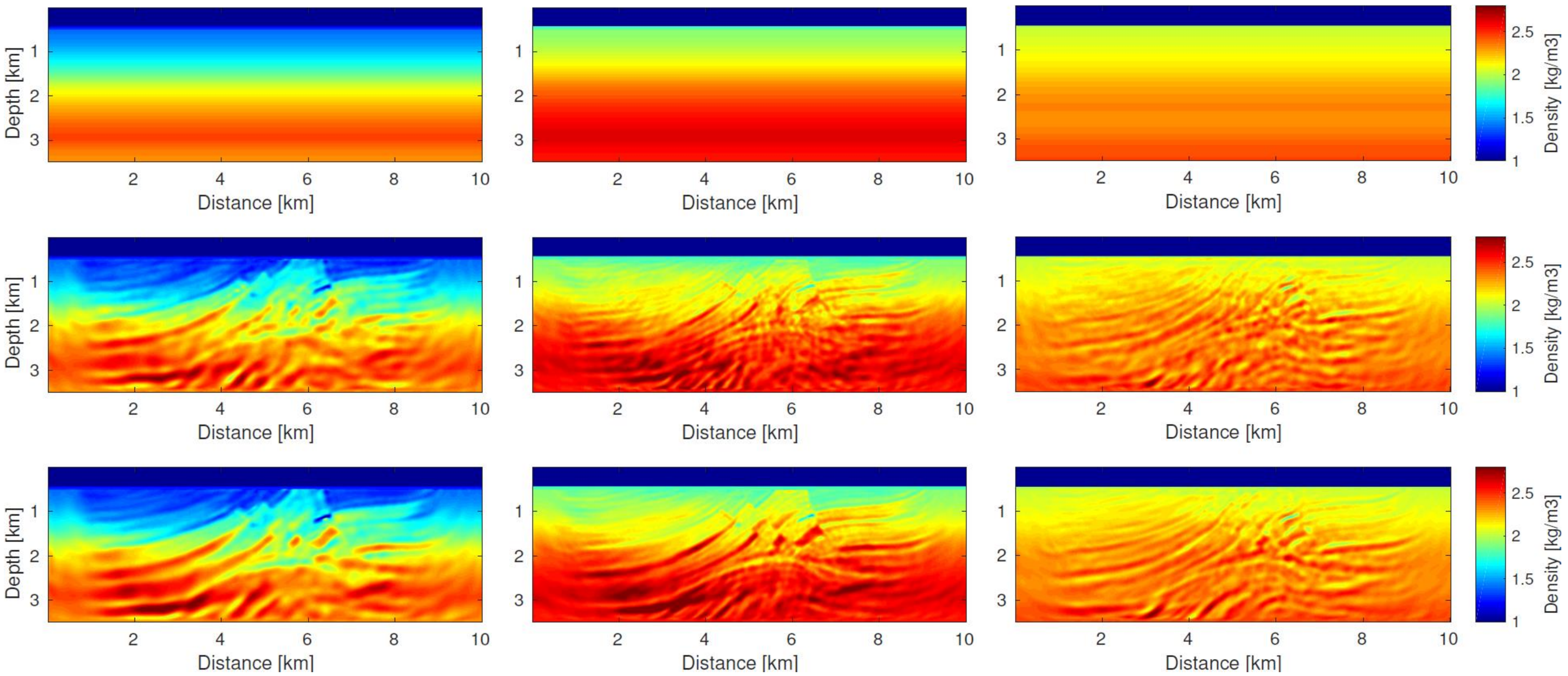
Elastic FWI using 2-4Hz low frequency data

P-wave velocity | S-wave velocity | Density

initial model

elastic FWI with **extrapolated** 2-4Hz data

elastic FWI with **true** 2-4Hz data



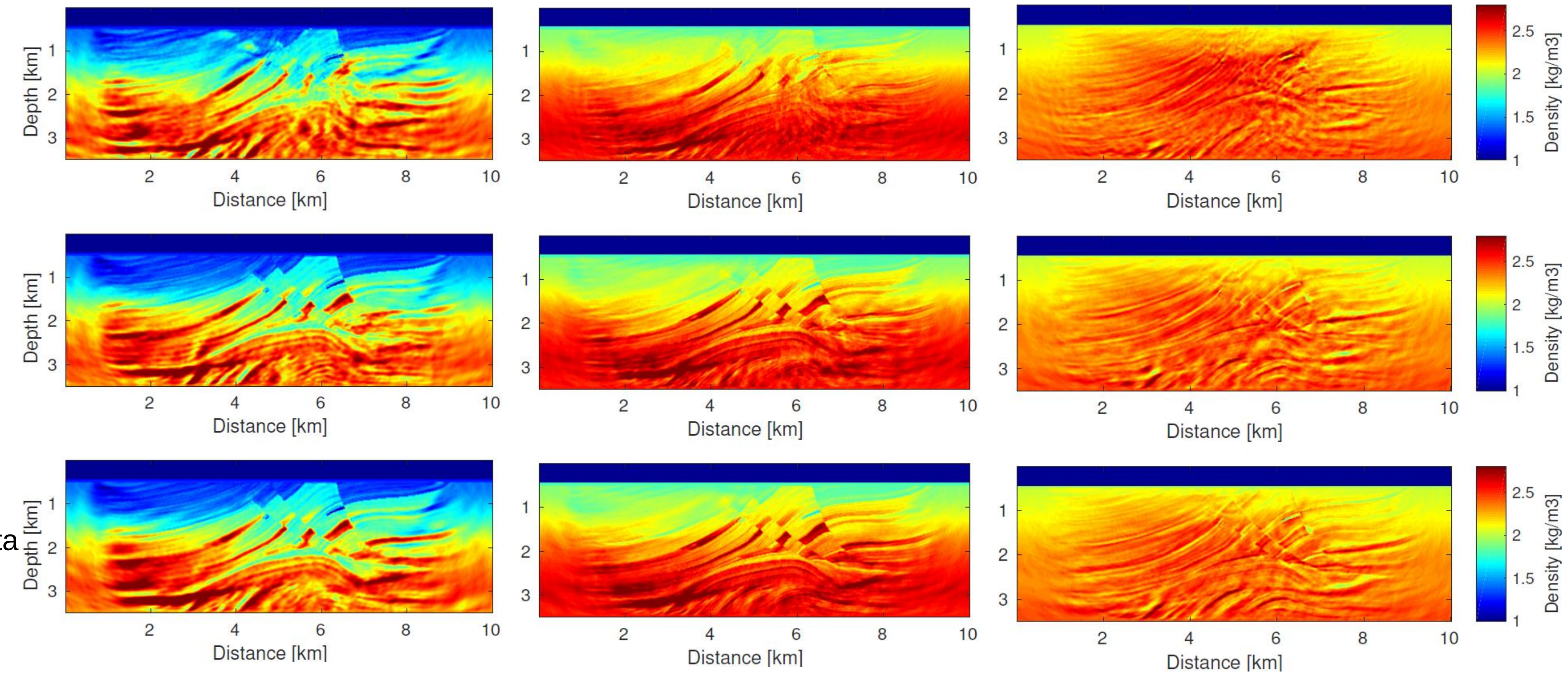
Elastic FWI continued with 4-20Hz bandlimited data

P-wave velocity | S-wave velocity | Density

started from initial model

started from **extrapolated** 2-4Hz data

started from **true** 2-4Hz data



Conclusions

- The deep learning model is designed with a large receptive field by directly using a large filter on each convolutional layers.
- 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 accuracy of the extrapolated low frequencies is enough to provide low-wavenumber starting models for elastic FWI of P-wave and S-wave velocities using band-limited data above 4Hz.
- 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.

Limitations

- Although the accuracy of extrapolated low frequency data is sufficient for elastic FWI of P-wave and S-wave velocities started from 4Hz band-limited data, challenges remain to enable the neural network to work for the data band-limited above 4Hz.
- Starting from the 2-4Hz extrapolated data, the inversion of density model still surfers from the cycle-skipping problem and lack of the low-wavenumber structures.
- In addition to the wave propagation driven by different physics, another factor that makes the generalization fail could be numerical modeling.
- The source signal is assumed to be known for extrapolated elastic FWI.
- The absence of a physical interpretation for the operations performed by the network.

Acknowledgements

- Tensorflow (Abadi et al., 2015) and Keras (Chollet et al., 2015) are used for deep learning.
- Elastic FWI is implemented using the open source code DENISE (https://github.com/daniel-koehn/DENISE-Black-Edition).
- Acoustic training datasets are simulated using Pysit (Hewett & Demanet, 2013).

- **Thanks to MIT ERL and Total S A for support.**

- **Thanks for your attention.**

Questions and Comments?