A field-data case study of extrapolated full-waveform inversion

Hongyu Sun

Graduate Student, Dept. of Earth, Atmospheric, and Planetary Sciences

In collaboration with

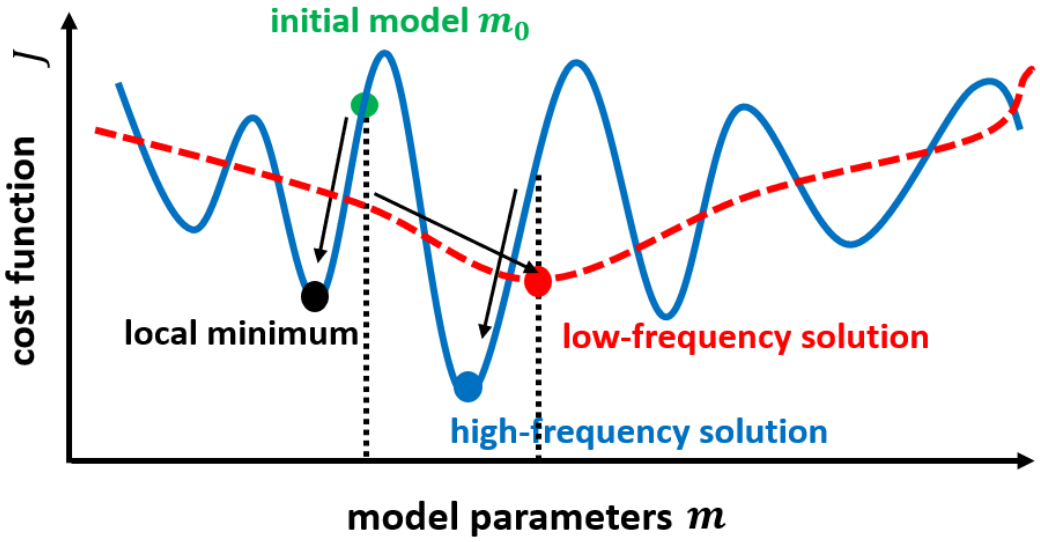
Rami Nammour, Christian Rivera, Paul Williamson, and Laurent Demanet

May 25, 2022

Thank TotalEnergies for permission to publish the work.

Extrapolated full-waveform inversion with deep learning

Cycle-skipping problem of FWI



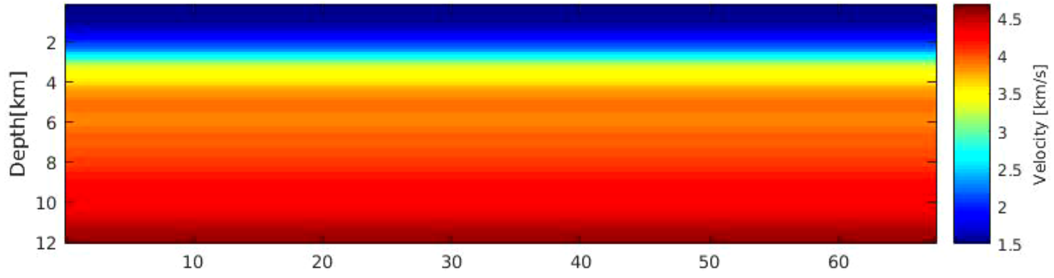
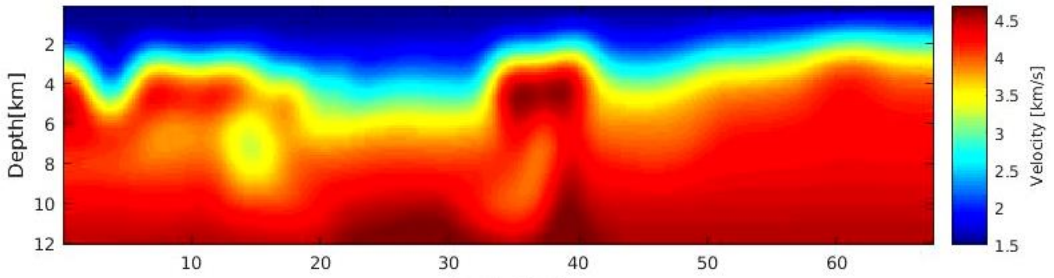
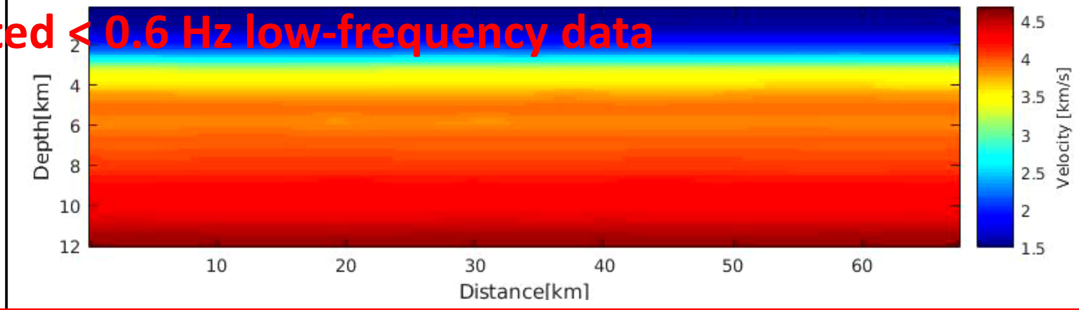
Low-frequency extrapolation

- Signal processing

Hu, 2014; Li and Demanet, 2016

- Deep learning

Sun and Demanet, 2018; Jin et al., 2018; Ovcharenko et al., 2018;

true low-freq	good initial model	full-waveform inversion (FWI) video
No	No	
No	Yes	
No	No	

with extrapolated < 0.6 Hz low-frequency data

Sun and Demanet, 2020, Geophysics

Related contributions and challenges on field data

Field-data examples:

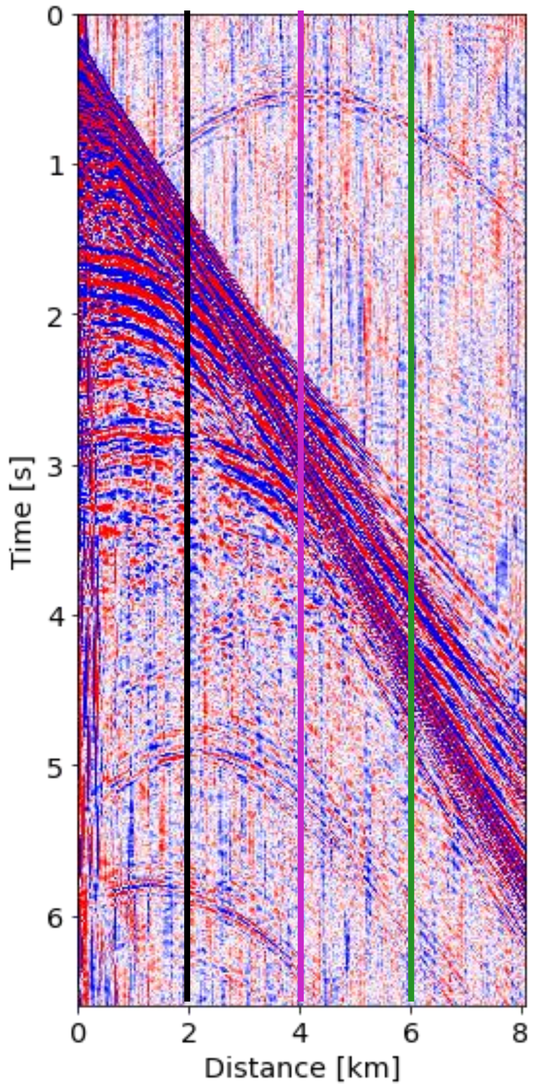
- Probably insufficient accuracy for full-waveform inversion
  - *Wang et al., 2020, SEG Technical Program Expanded Abstracts 2020*
  - *…*
- Training on real data (collected from the same region as the test data)
  - *Aharchaou and Baumstein, 2020, The Leading Edge*
  - *Zhang et al., 2021, IEEE Geoscience and Remote Sensing Letters*
- Training on synthetic data
  - *Fang et al., 2020, Geophysics*
  - *Ovcharenko, 2021, KAUST Ph.D. thesis*

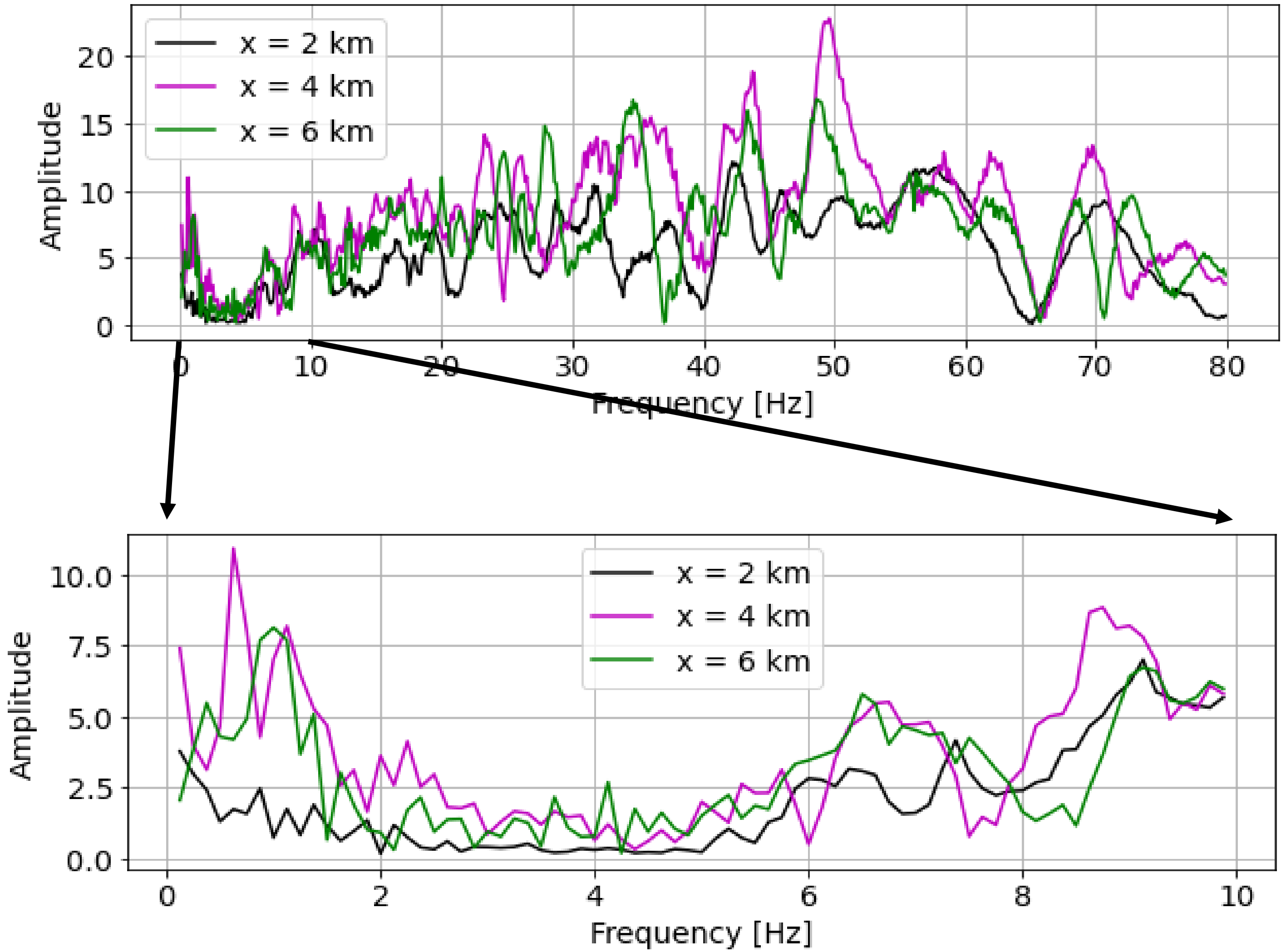
Challenges:

- Unavailability of real low-frequency data for training, in particular < 2 Hz
- Poor generalization from synthetic to real data

Our strategy: Semi-supervised learning with real data without real labels

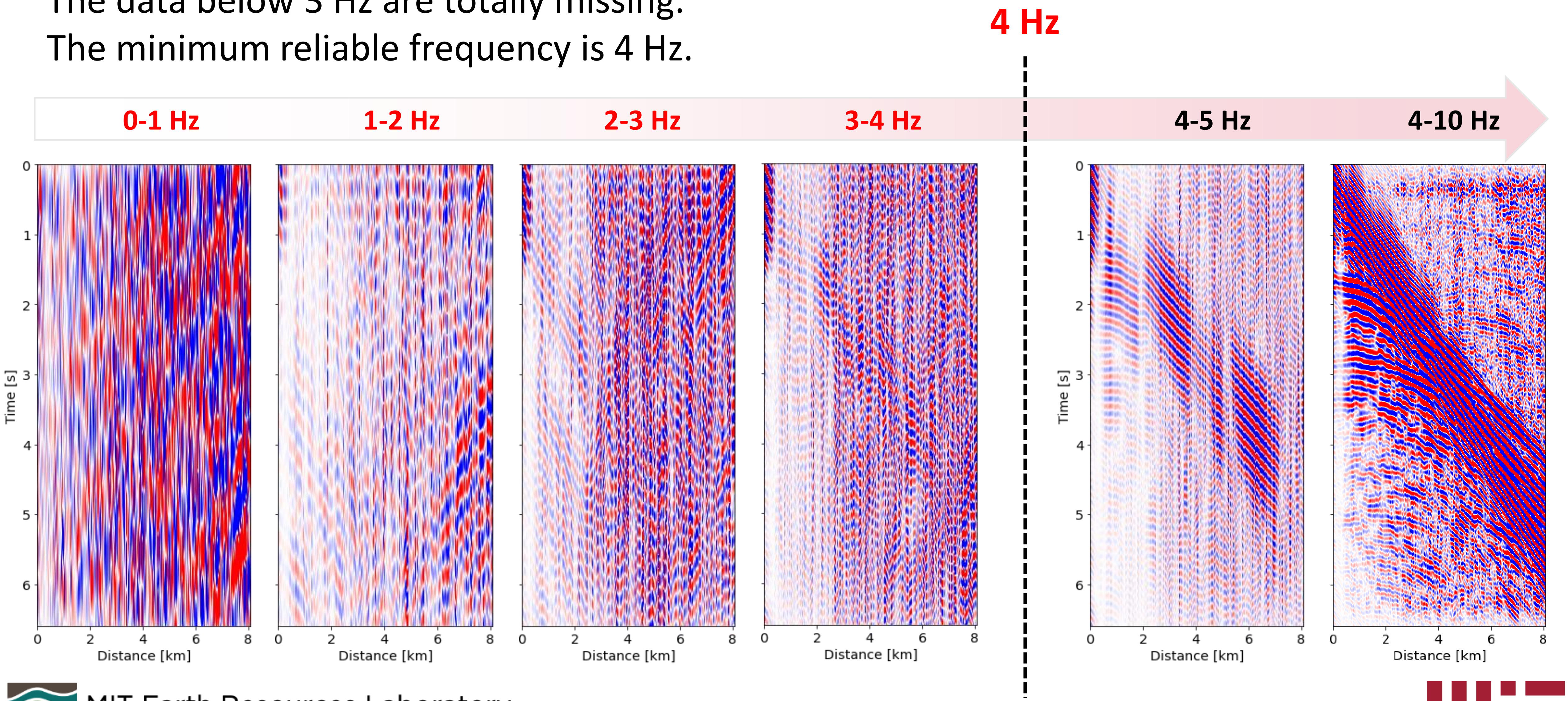
Field raw data in the time and frequency domain





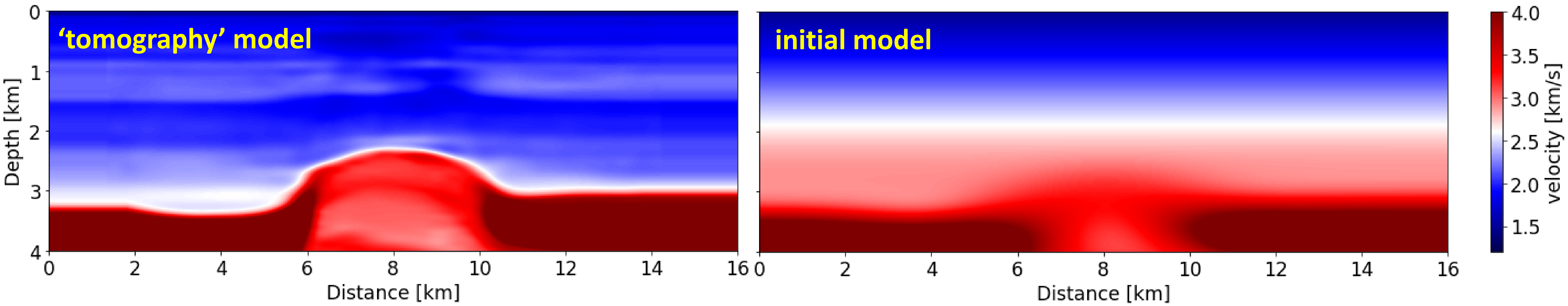
Field raw data in different frequency bands

The data below 3 Hz are totally missing.
The minimum reliable frequency is 4 Hz.

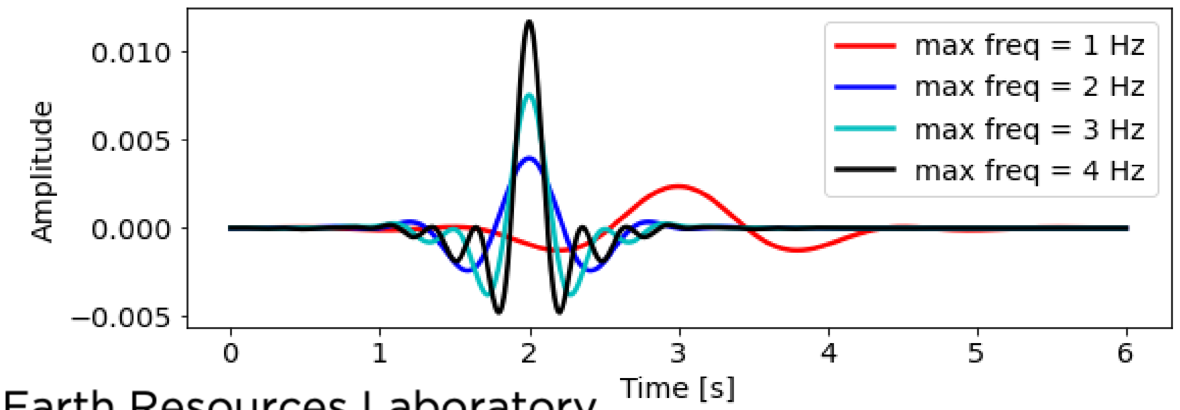


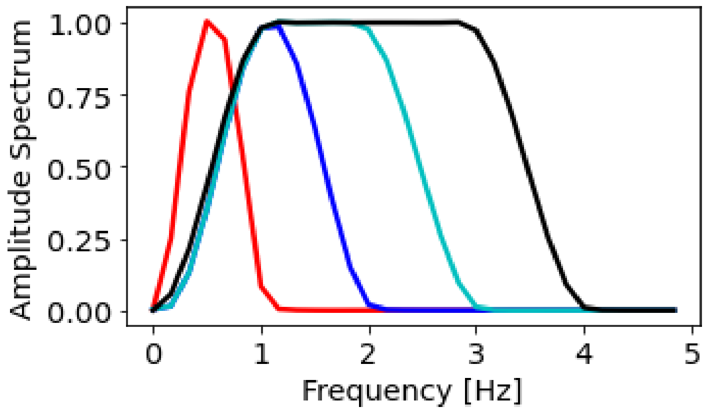
Determination of starting frequency

- **Starting from the given initial model, how low should the starting frequency be to avoid the cycle-skipping problem?**



- **Simulate low-frequency data on the ‘tomography’ model**

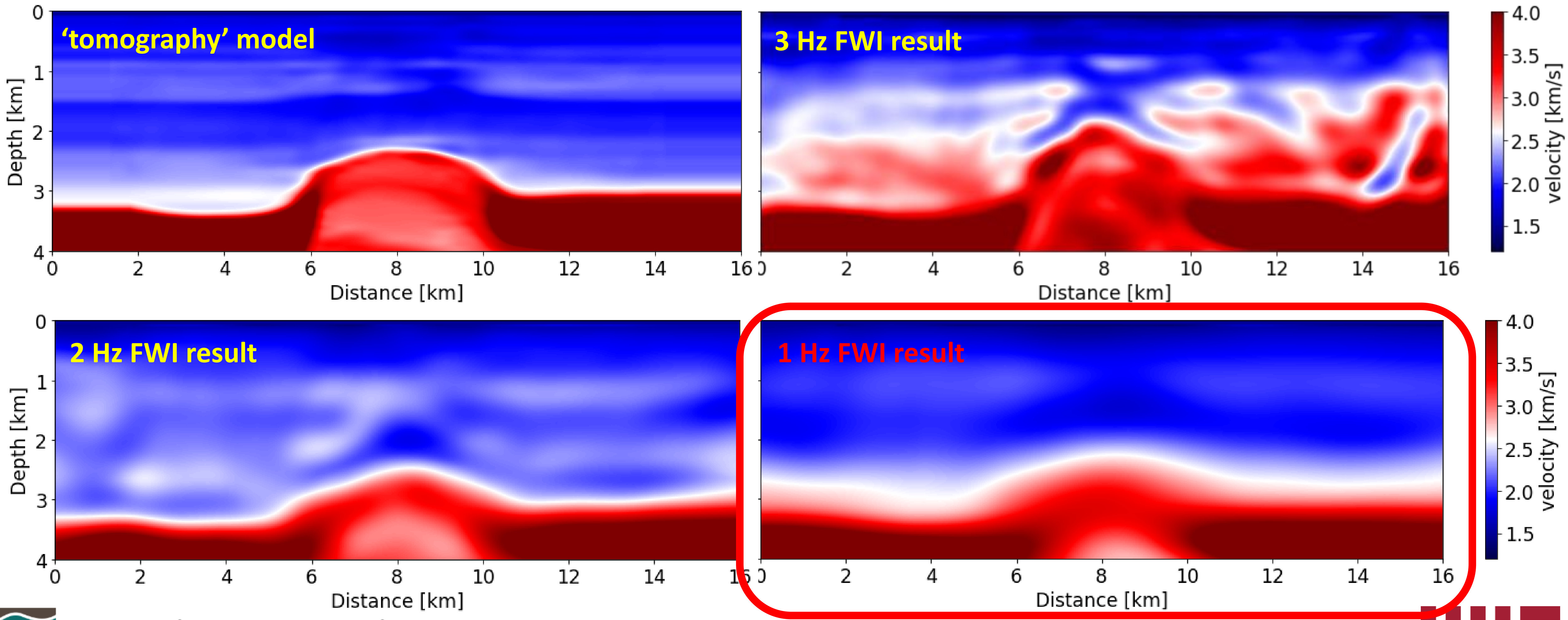




Determination of starting frequency

1 Hz

- **Starting from the given initial model, how low should the starting frequency be to avoid the cycle-skipping problem?**



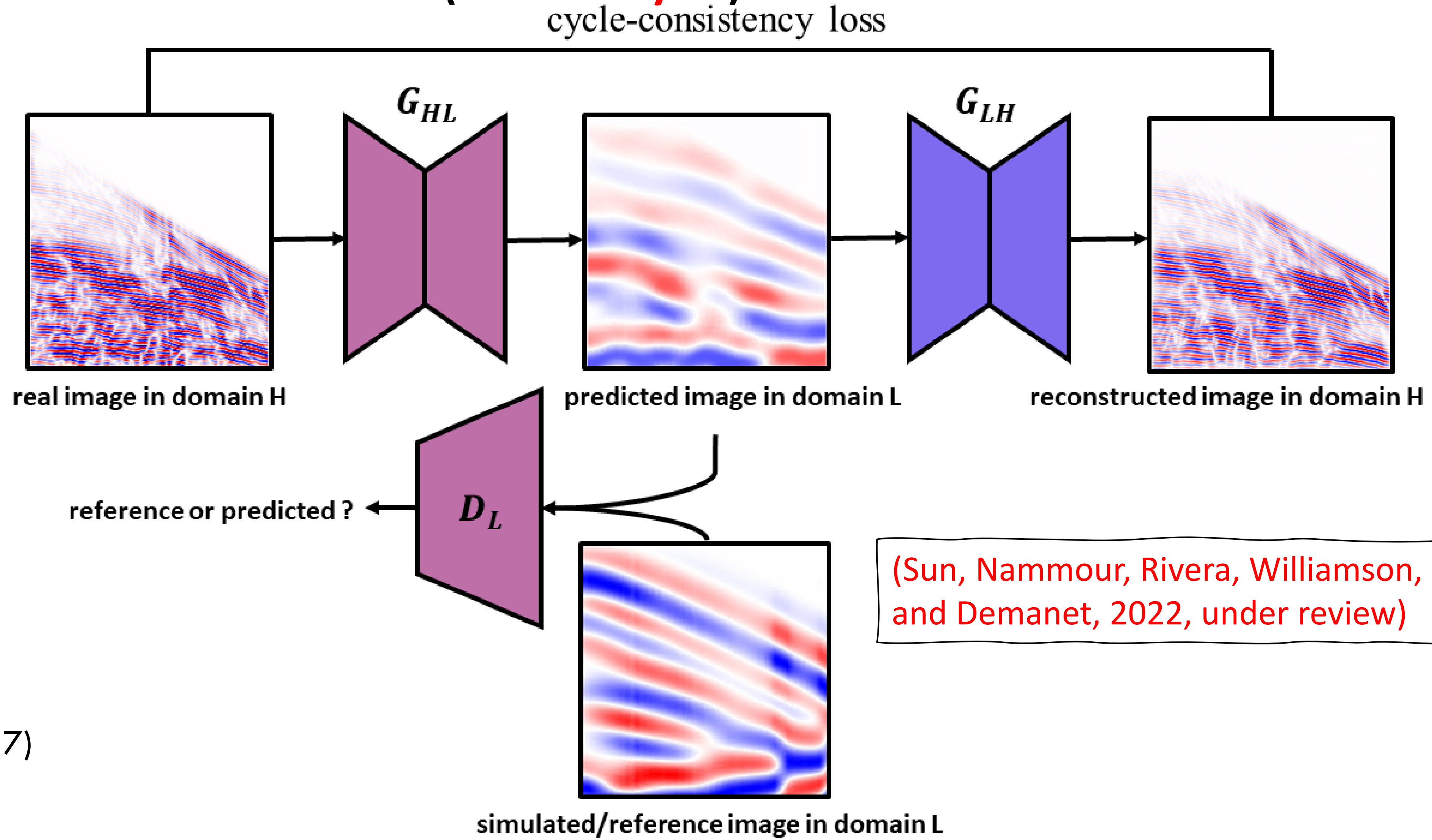
Learning with real data without real labels

Cycle-Consistent Adversarial Networks (forward cycle)

$$L(G_{HL}, D_L, H, L) =$$

$$\begin{cases} \lambda_1 \, L_{cycle}(G_{HL}, G_{LH}, H) \\ \lambda_2 \, L_{GAN}(G_{HL}, D_L, H) \\ \lambda_3 \, L_{identity}(G_{HL}, L) \end{cases}$$

$$L_{disc}(G_{HL}, D_L, H, L)$$



(Sun, Nammour, Rivera, Williamson, and Demanet, 2022, under review)

CycleGAN (Zhu et al., 2017)

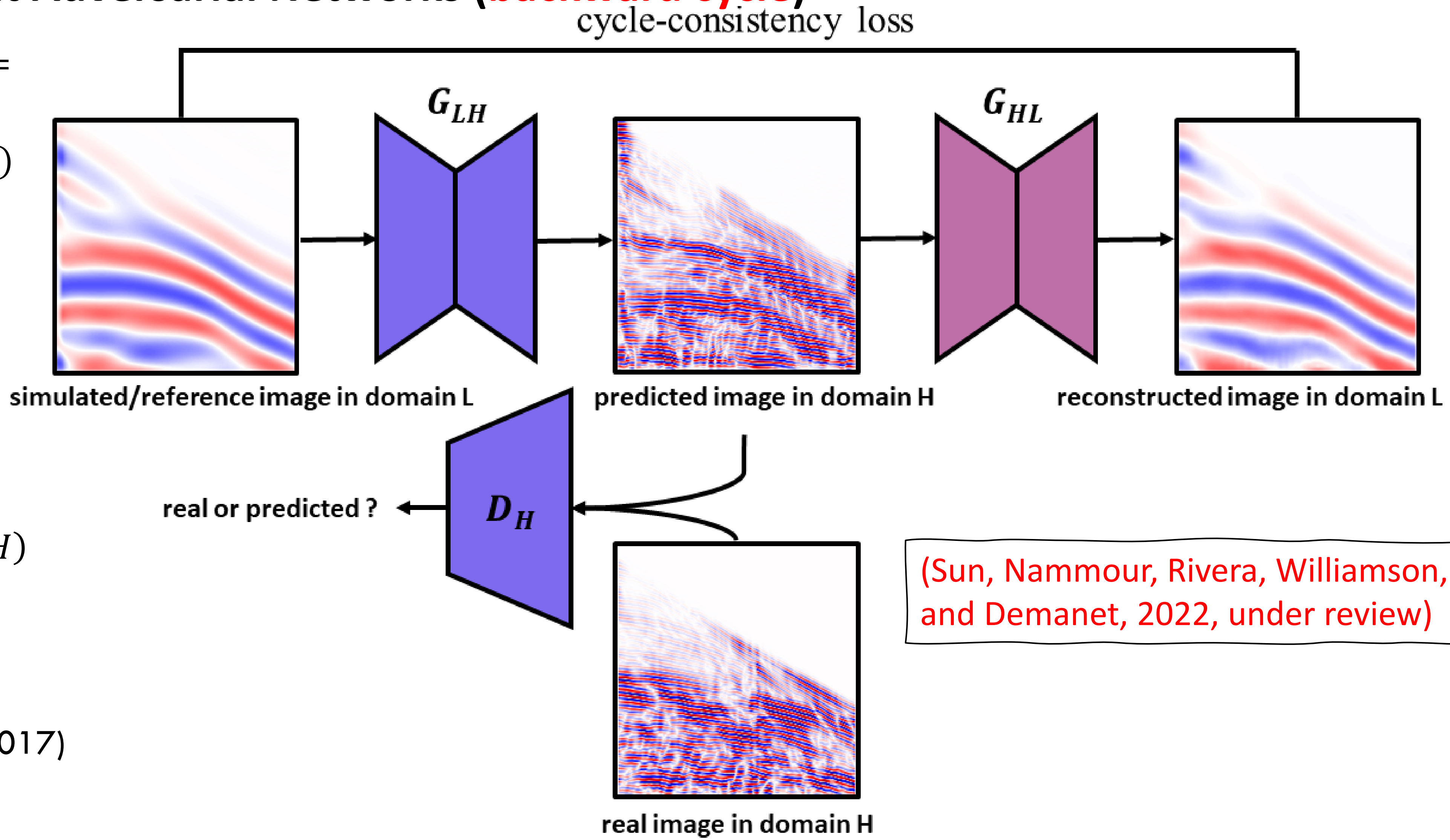
Learning with real data without real labels

- **Cycle-Consistent Adversarial Networks (backward cycle)**

$$L(G_{LH}, D_H, L, H) =$$

$$\begin{cases} \lambda_1 \; L_{cycle}(G_{LH}, G_{HL}, L) \\ \lambda_2 \; L_{GAN}(G_{LH}, D_H, L) \\ \lambda_3 L_{identity}(G_{LH}, H) \end{cases}$$

$$L_{disc}(G_{LH}, D_H, L, H)$$

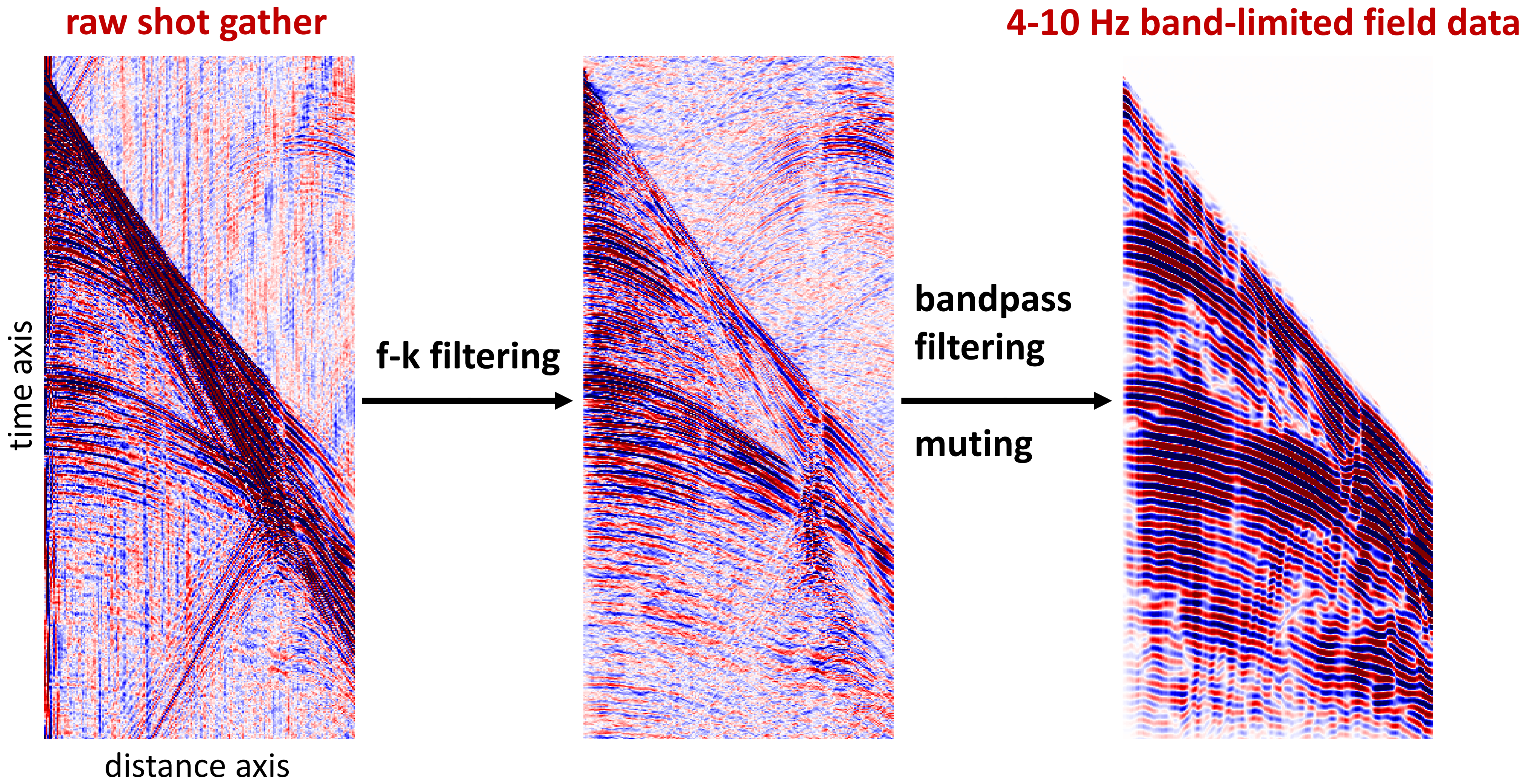


(Sun, Nammour, Rivera, Williamson, and Demanet, 2022, under review)

CycleGAN (Zhu et al., 2017)

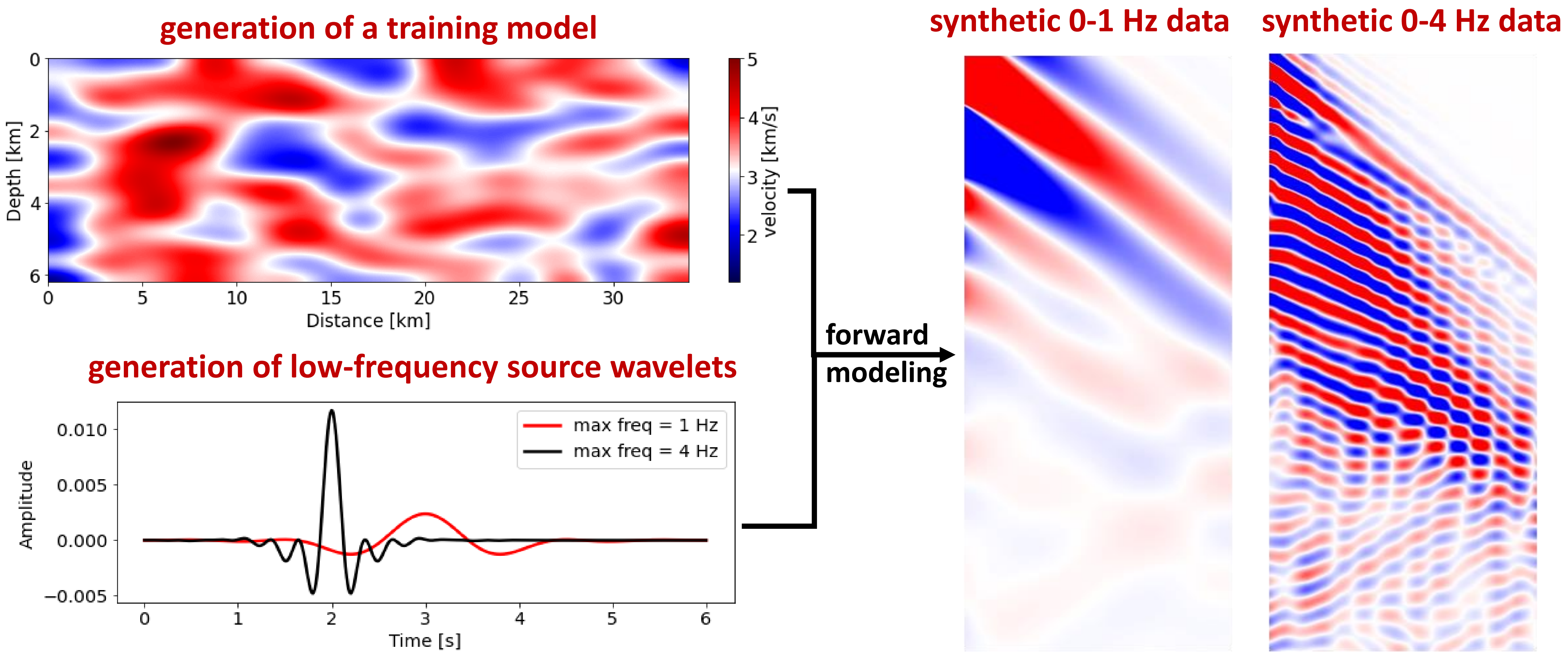
Training data preparation for the field data

- **Preprocess field data for 4-10 Hz band-limited data as image domain H**



Training data preparation for the field data

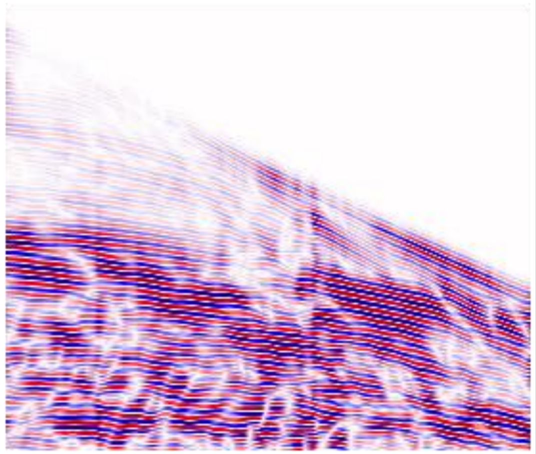
- **Simulate synthetic low-frequency data as image domain L**



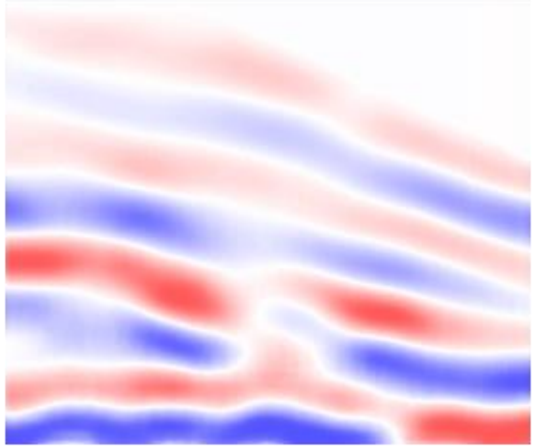
Field-data extrapolation result: from 4-10 Hz to **0-1 Hz**

- **Forward Cycle: Starting from field 4-10 Hz band-limited data**

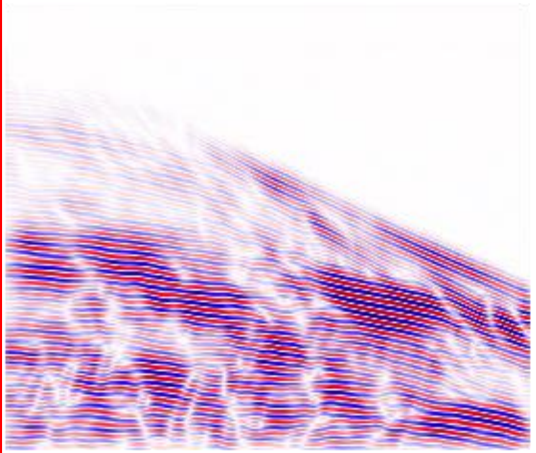
real image (field band-limited data)



predicted image (extrapolated data)

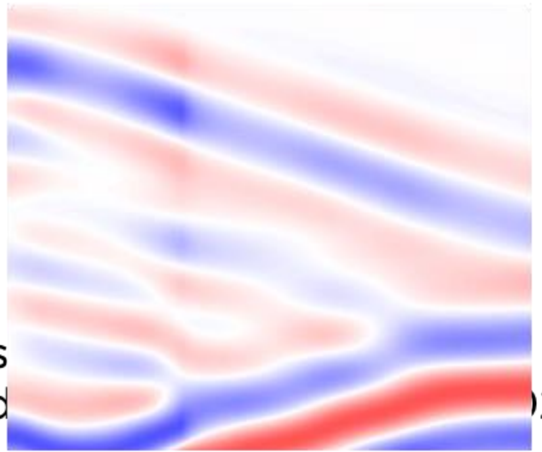


reconstructed image

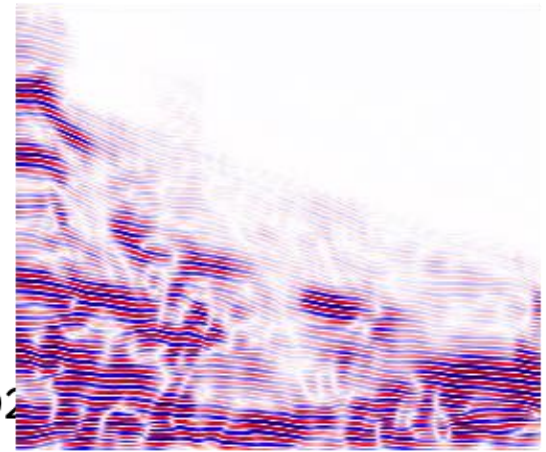


- **Backward Cycle: Starting from synthetic 0-1 Hz low-frequency data**

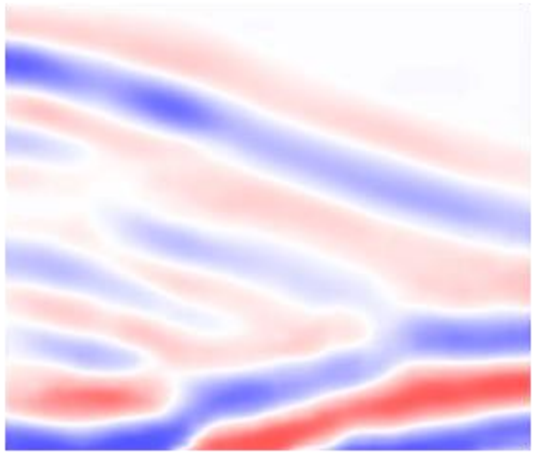
synthetic image (simulated low-frequency data)



predicted image (extrapolated data)



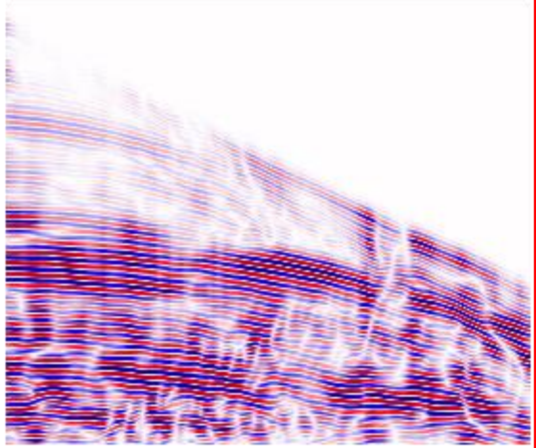
reconstructed image



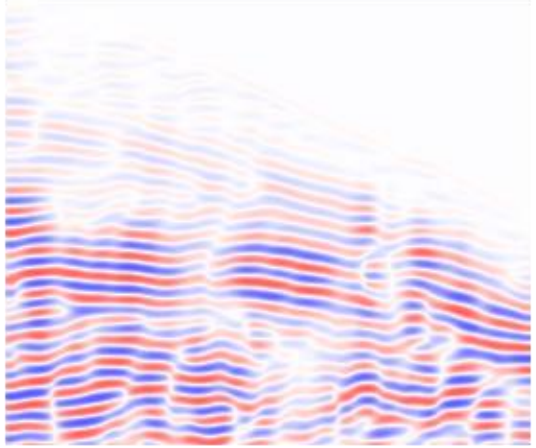
Field-data extrapolation result: from 4-10 Hz to **0-4 Hz**

- **Forward Cycle: Starting from field 4-10 Hz band-limited data**

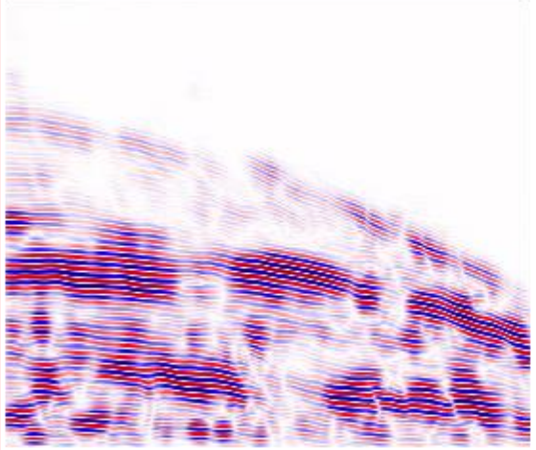
real image
(field band-limited data)



predicted image
(extrapolated data)

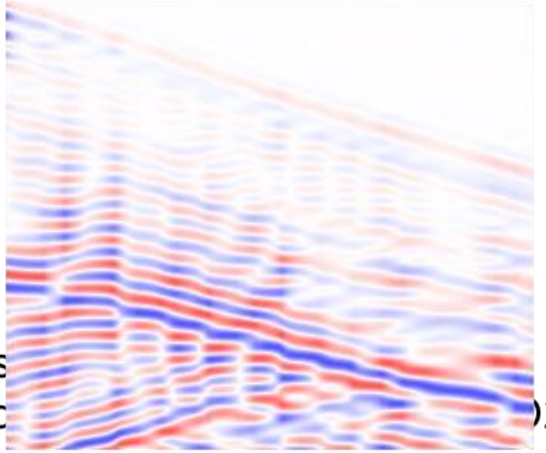


reconstructed image

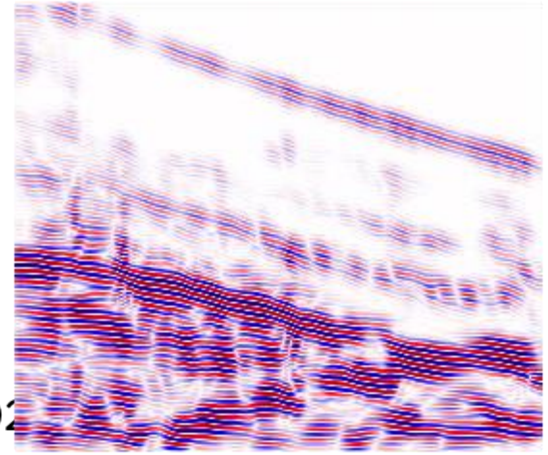


- **Backward Cycle: Starting from synthetic 0-4 Hz low-frequency data**

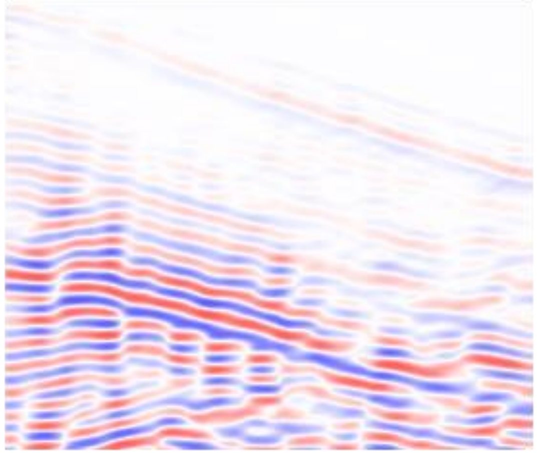
synthetic image
(simulated low-frequency data)



predicted image
(extrapolated data)

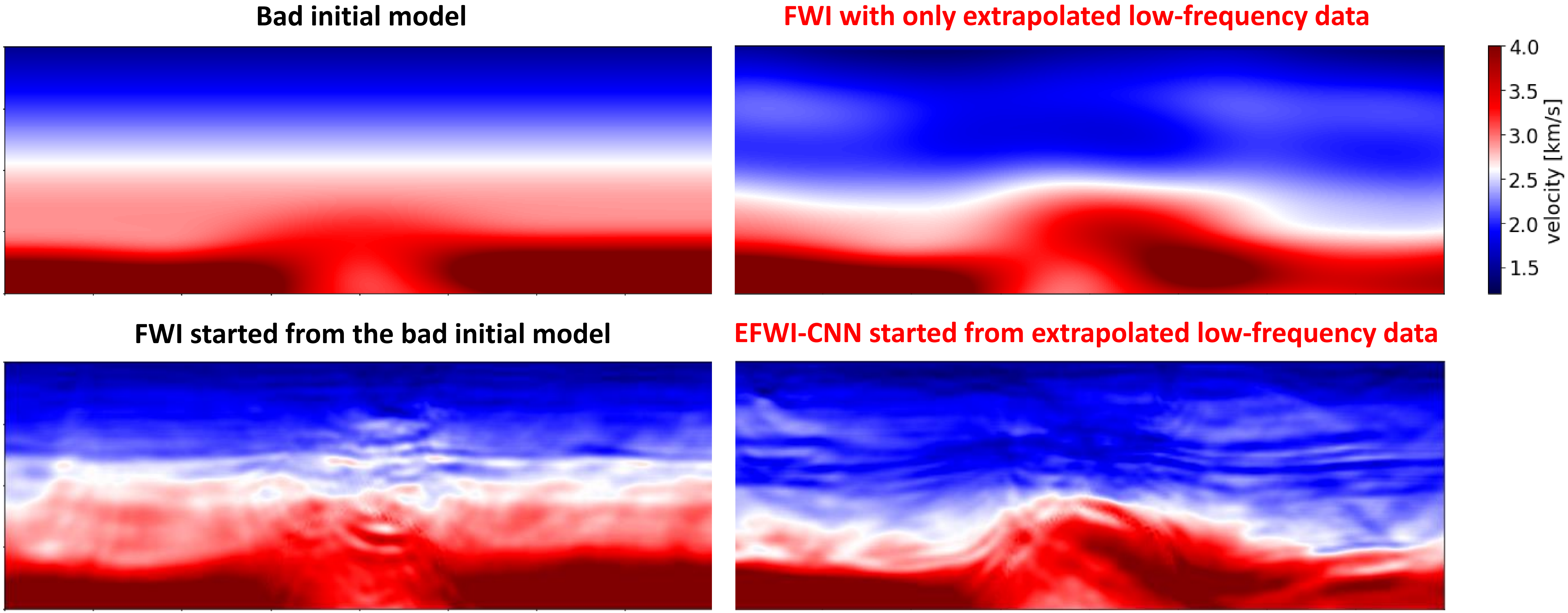


reconstructed image



Extrapolated FWI with 3-10 Hz band-limited field data

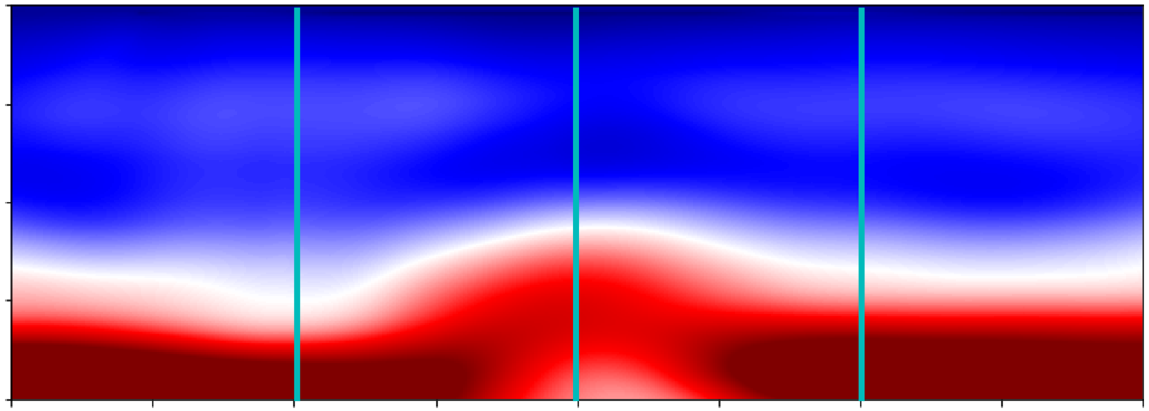
Cost function: L2 norm of normalized (trace by trace) difference in time domain



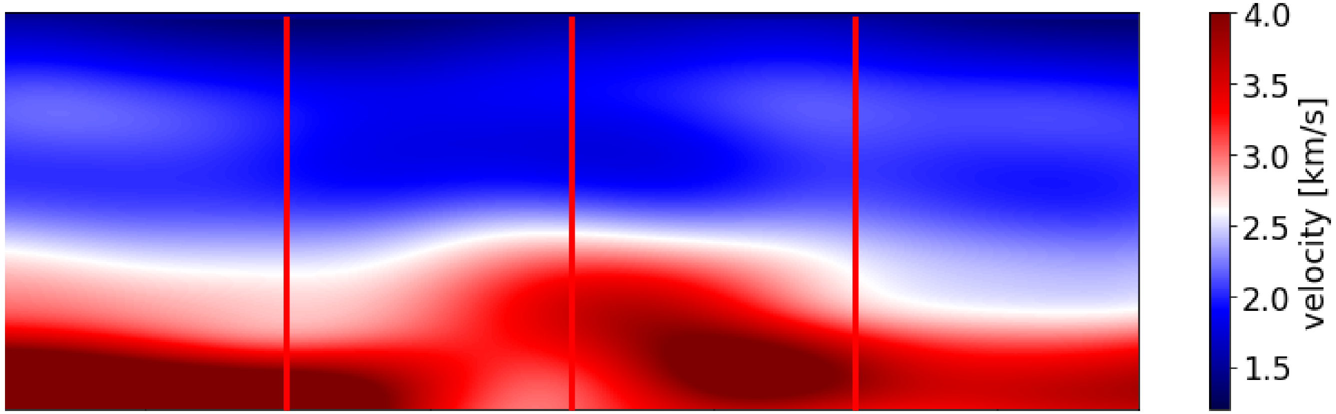
Extrapolated FWI with 3-10 Hz band-limited field data

Cost function: L2 norm of normalized (trace by trace) difference in time domain

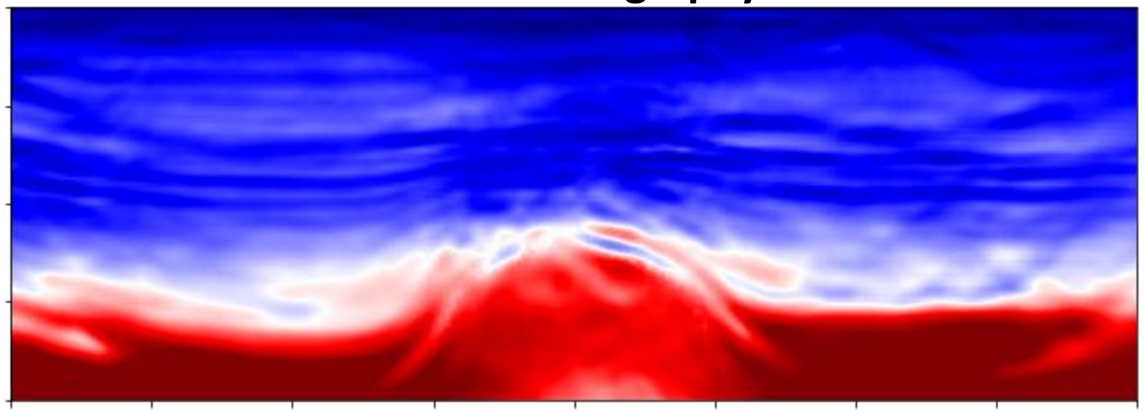
FWI with synthetic data simulated on ‘tomography’ model



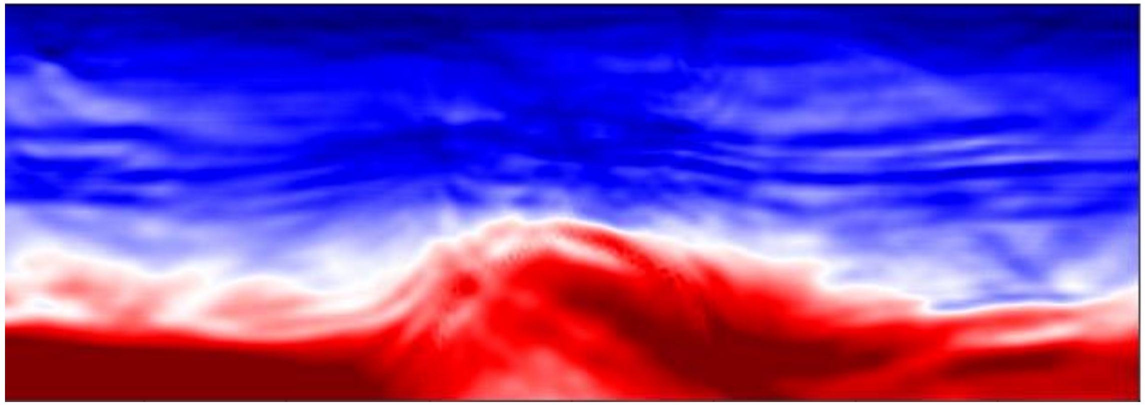
FWI with only extrapolated low-frequency data



FWI started from synthetic data simulated on ‘tomography’ model

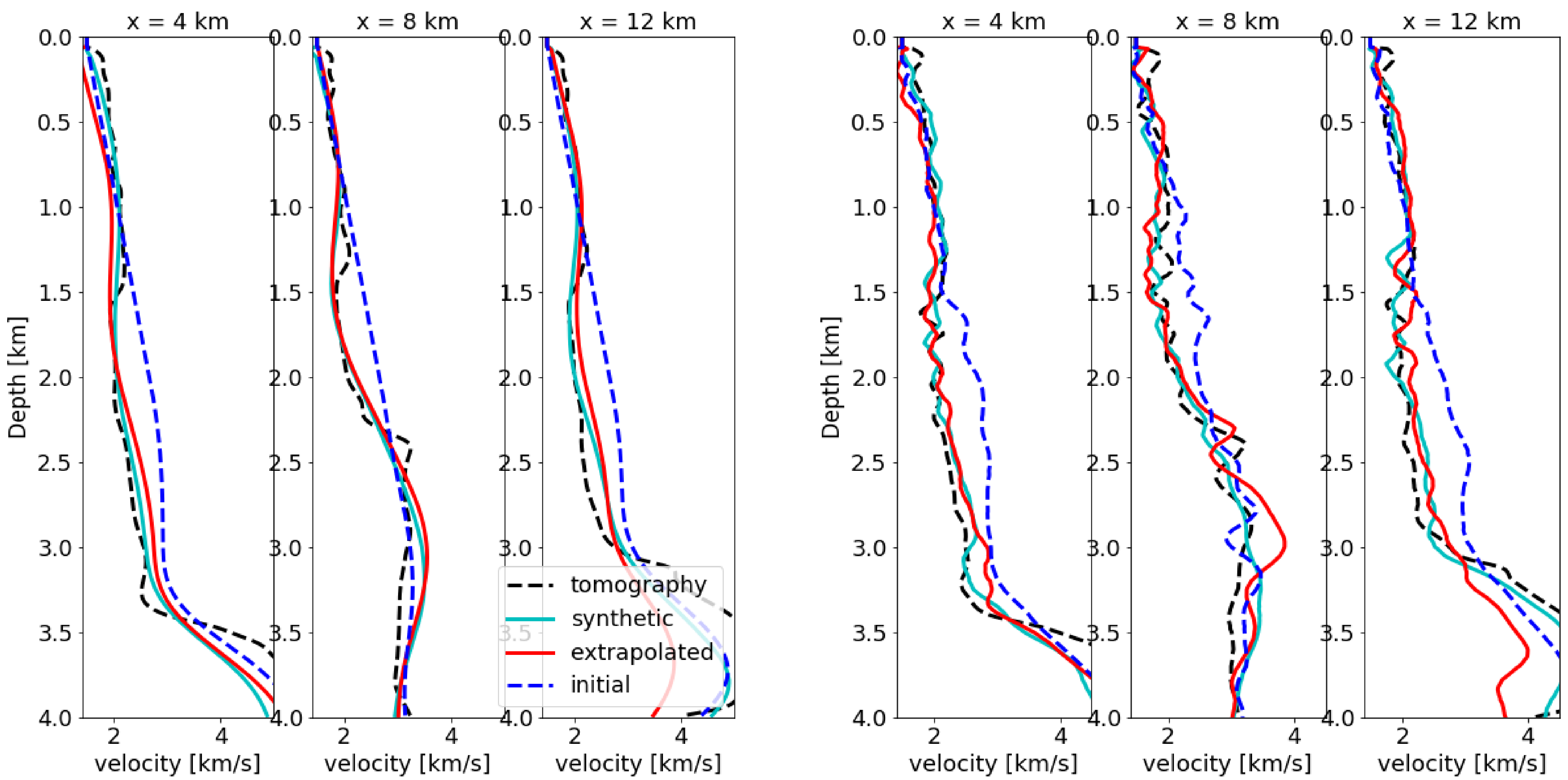


EFWI-CNN started from extrapolated low-frequency data



Quality control: Vertical section comparison

starting models | **FWI using 3-10 Hz band-limited field data**



Summary

- CycleGAN can be trained using unpaired images of field band-limited and synthetic low-frequency shots in the time domain for robust low-frequency extrapolation of the field data.

- Our field-data results validate the benefit of extrapolated low-frequency data for mitigating the cycle-skipping problem of full-waveform inversion.

- The wavelet to synthetize the low-frequency data on the training model should be used for full-waveform inversion using the extrapolated low frequencies.

Limitations and future work

- more robust FWI with the extrapolated data
- QC with RTM and image gathers
- interpretable AI?
- surface waves

Sun and Demanet, 2022, **Extrapolated surface-wave dispersion inversion**: SEG, Expanded Abstracts, in press.

Second International Meeting for Applied Geoscience & Energy (IMAGE22)

- ❑ Session Style: Oral Presentation
- ❑ Session ID: NS 2 (Near Surface)
- ❑ Session Title: Advanced Processing and Machine Learning 2
- ❑ Presentation Date and Time: August 31, 2022 from 1:45 PM to 2:10 PM

Acknowledgements

- This research is funded by TotalEnergies. We thank TotalEnergies for permission to publish this work.

- We thank Fuchun Gao, Mohamed Dolliazal, Yen Sun, and Russell Jones from TotalEnergies for additional technical support of this work.

- Thank you for your attention.