MIT EARTH RESOURCES LABORATORY
ANNUAL FOUNDING MEMBERS MEETING 2018

Earth Resources Laboratory

# Low frequency extrapolation with deep learning


**Hongyu Sun**

GRADUATE STUDENT [ EARTH, ATMOSPHERIC AND PLANETARY SCIENCES ]

*In collaboration with Prof.* ***Laurent Demanet***

# Motivation:
# Full waveform inversion (FWI)

# Brief introduction to FWI

- Forward modeling

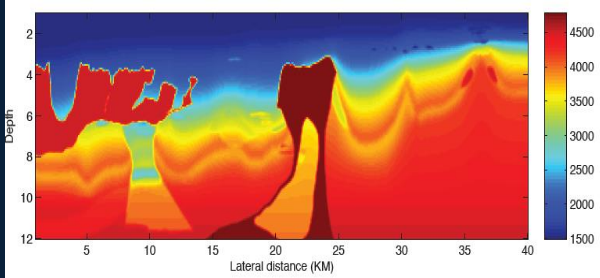
$$\nabla^2 u(\mathbf{x},\omega) + \frac{\omega^2}{v_p^{\,2}(\mathbf{x})} u(\mathbf{x},\omega) = -s(\mathbf{x},\omega)$$

- Inversion objective function

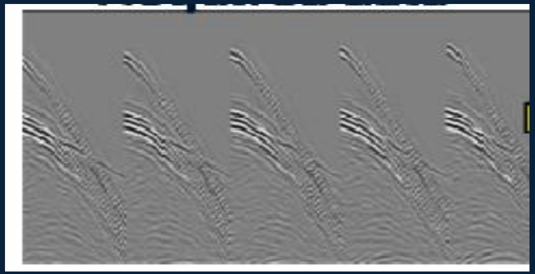
$$C(\mathbf{m}) = \left\|\Delta\mathbf{d}\right\|_p = \left\|\mathbf{d}_{cal}(\mathbf{m}) - \mathbf{d}_{obs}\right\|_p$$

- Optimization

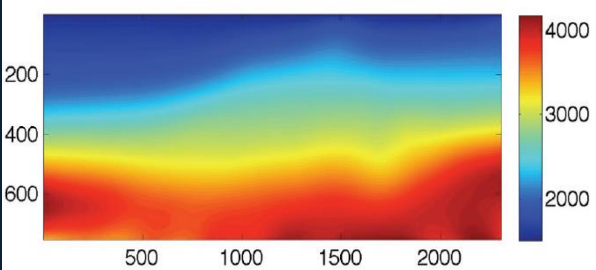
$$\Delta\mathbf{m} = -\alpha\left(diag(\mathbf{H}_a) + \lambda\mathbf{I}\right)^{-1}\nabla C(\mathbf{m})$$



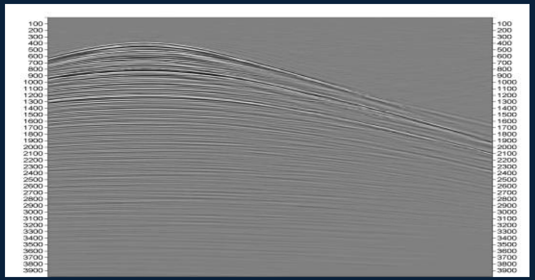

**true velocity model (unknown)**



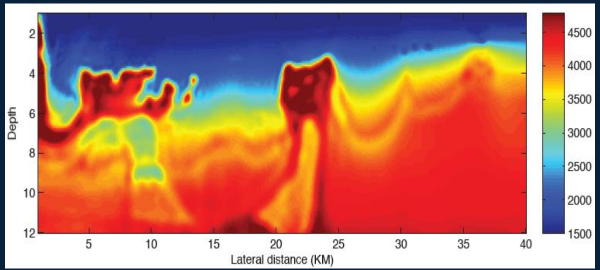
**observed data**




**initial velocity model**



**calculated data**




**inversion result**

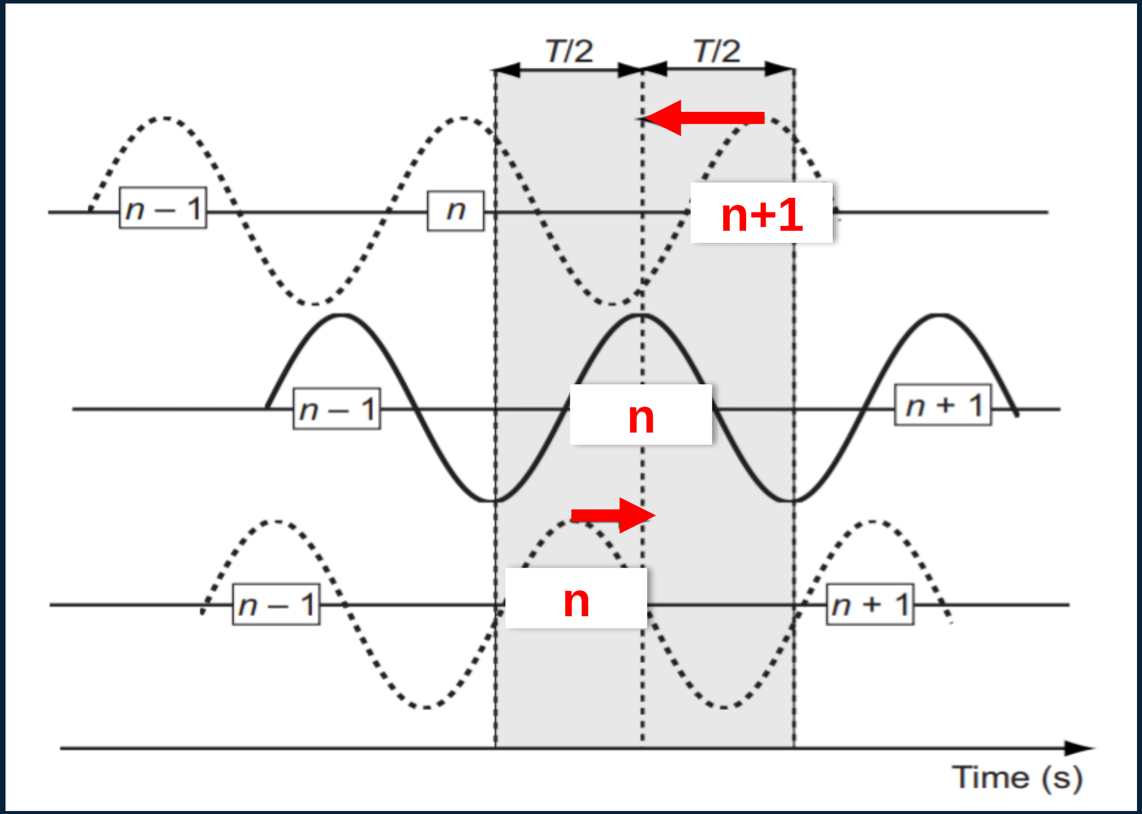
# Problem: Cycle skipping

# Cycle skipping in FWI

(a) Calculated seismogram: from **a 'bad' initial model**

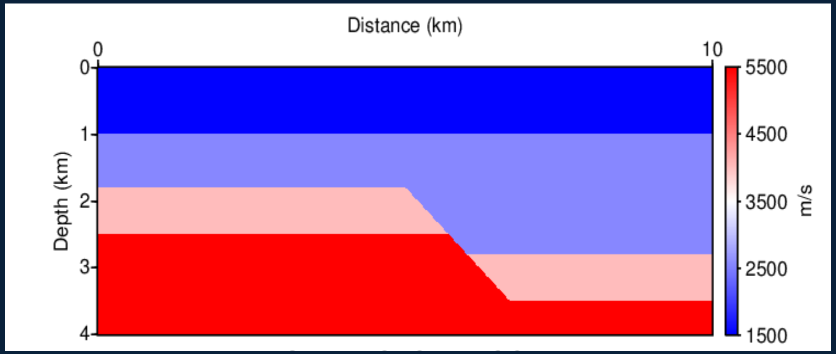
(b) Observed seismogram

(c) Calculated seismogram: from **a 'good' initial model**

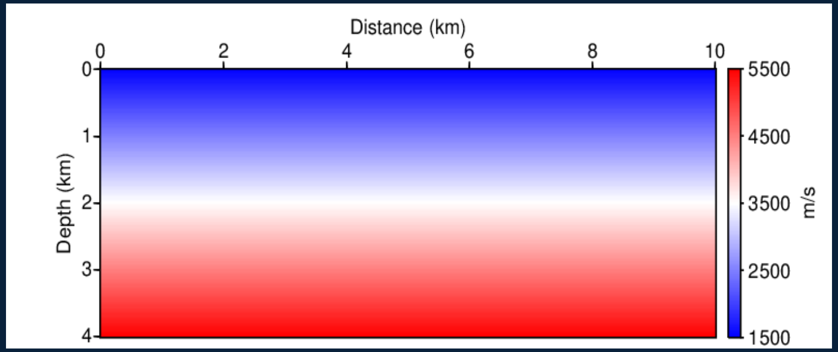



schematic of cycle-skipping artifacts in full waveform inversion (Virieux and Operato, 2009)

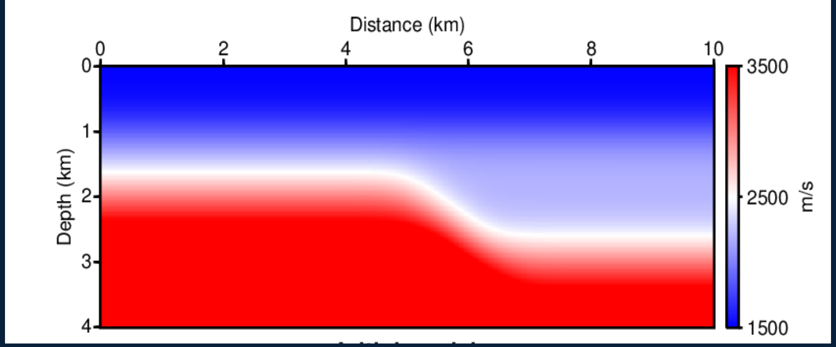
# Cycle skipping



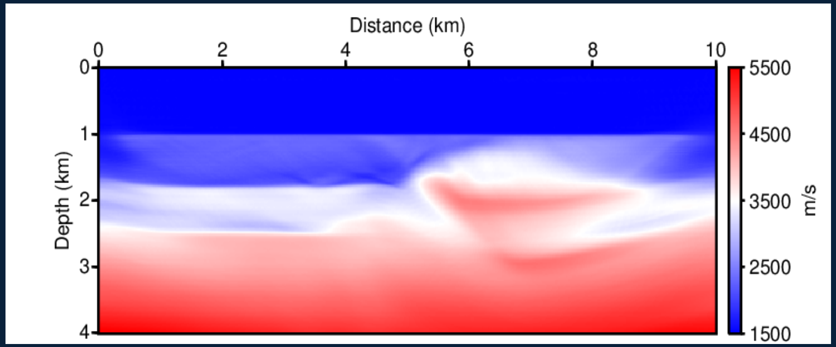

true velocity model



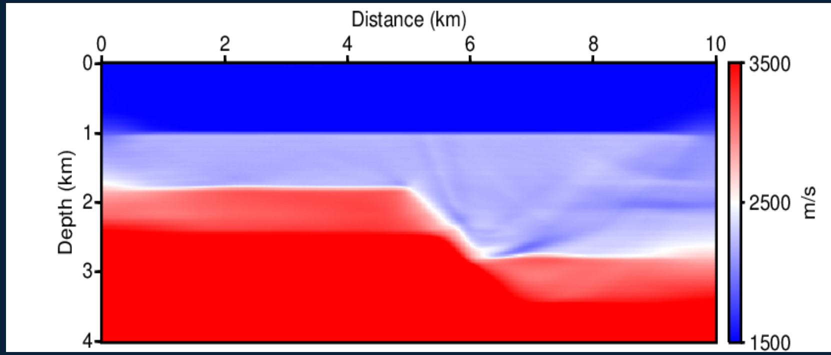

'bad' initial velocity model




'good' initial velocity model




inversion result with the 'bad' initial velocity model

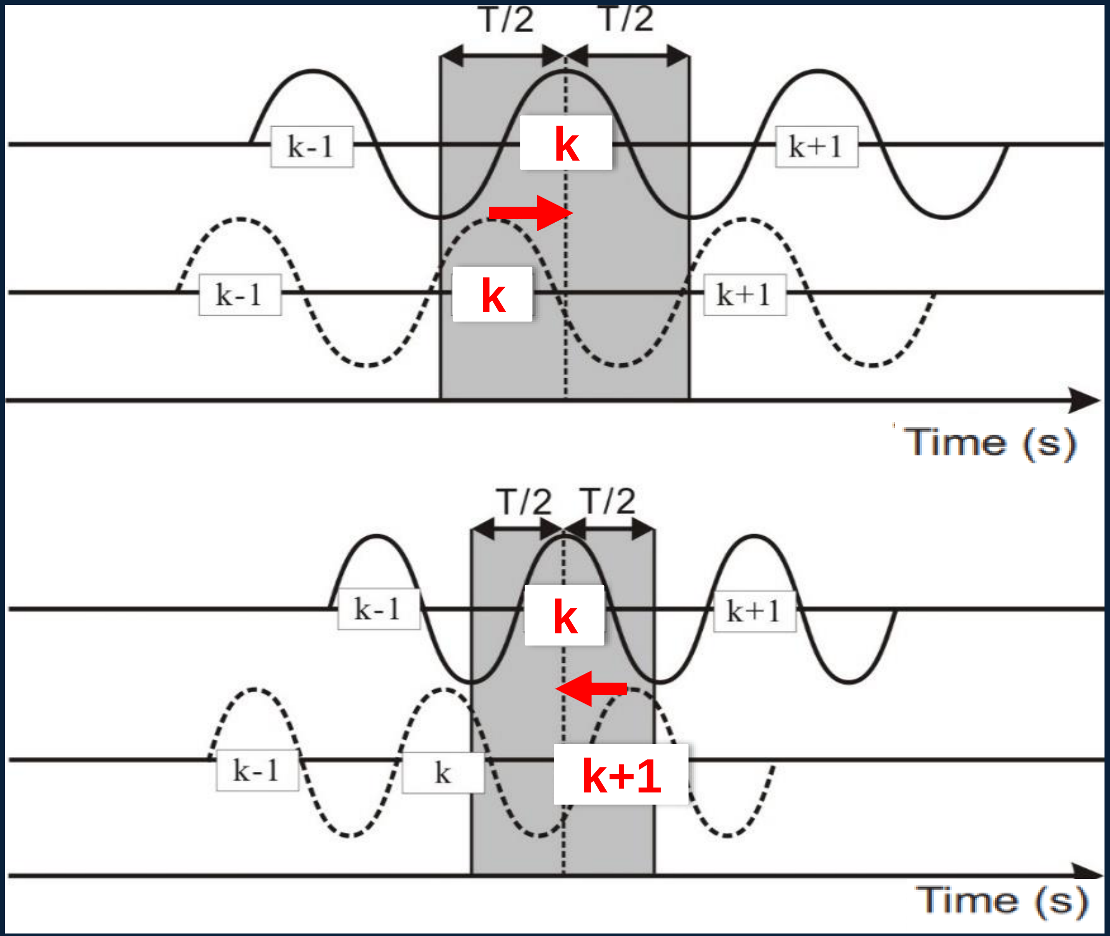



inversion result with the 'good' initial velocity model

# Cycle skipping

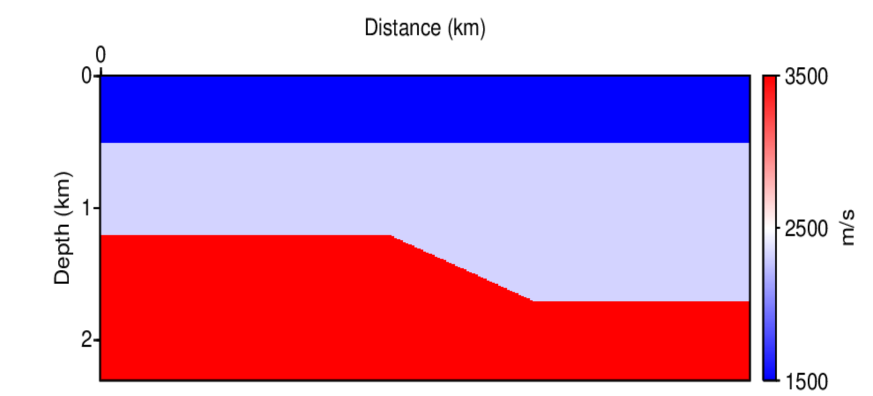
(a) **Low frequencies**
Observed seismogram
Calculated seismogram

(b) **High frequencies**
Observed seismogram
Calculated seismogram

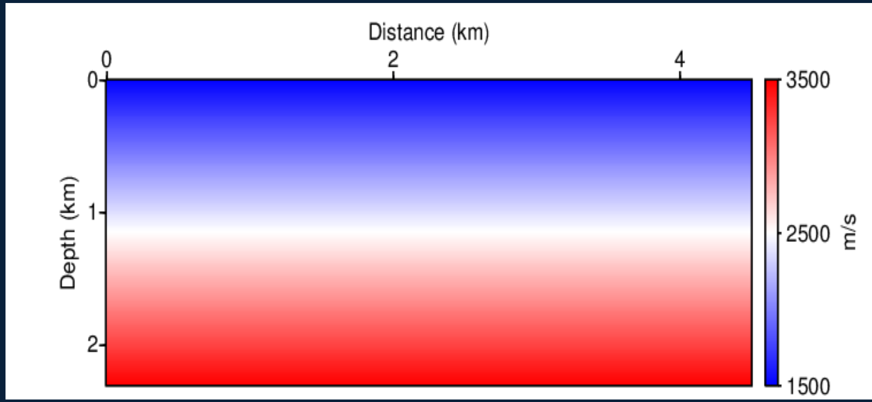



schematic of cycle-skipping artifacts in full waveform inversion (Han, 2013)

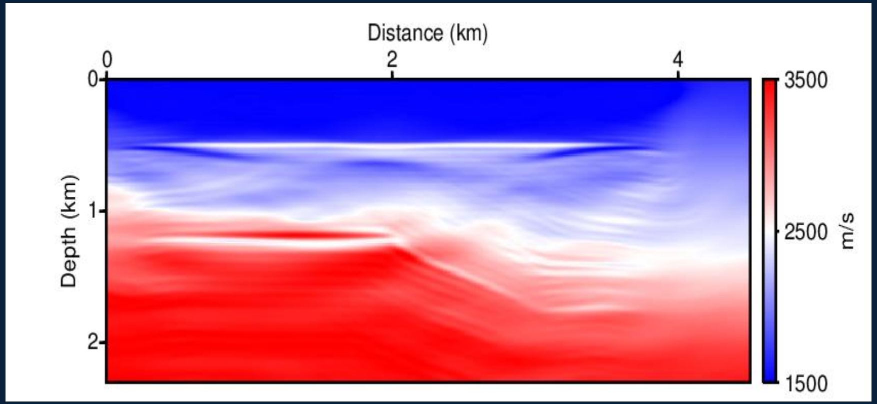
# Cycle skipping



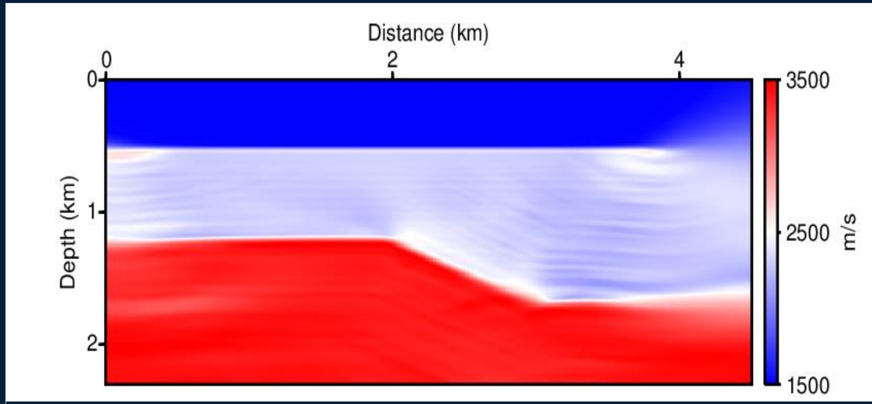

true velocity model




the ‘bad’ initial velocity model



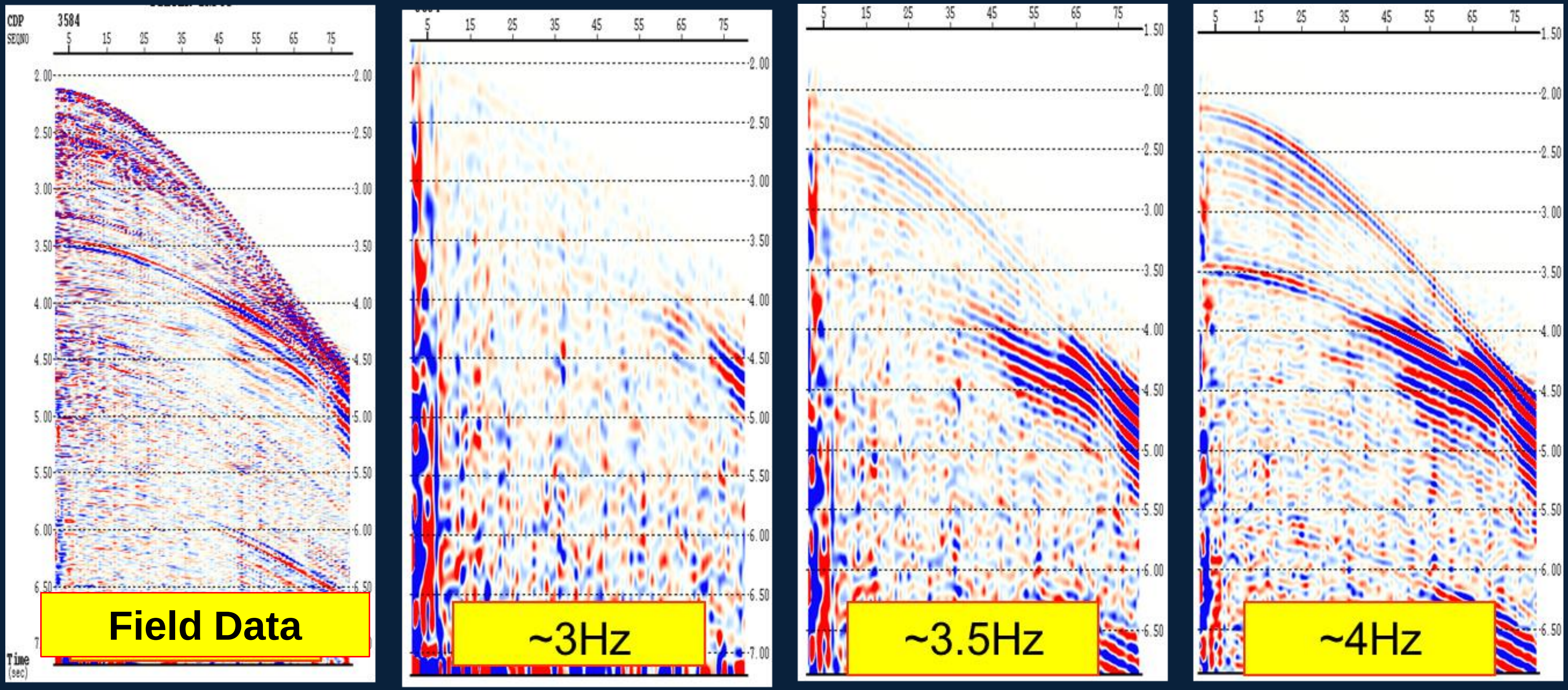

inversion result started from 5Hz




inversion result started from 1Hz

# Cycle skipping

**Low frequencies: hard to acquire in field acquisition**



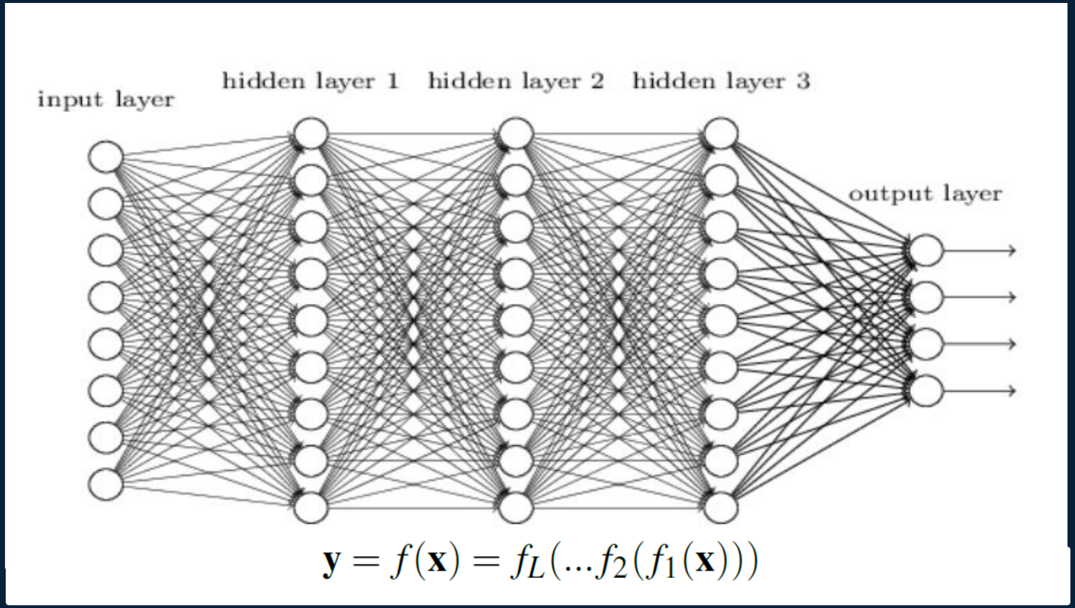

(Han, 2013)

# Solution: Deep learning

# Deep learning

**Deep neural networks**

- Data-driven method: a list of layers that transform the input data into an output prediction
- The universal approximation theorem (Hornik et al., 1989)



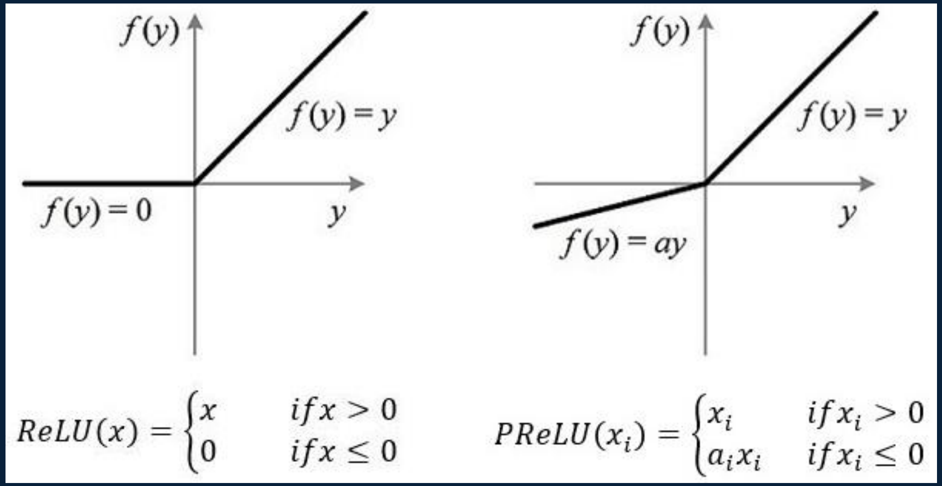

**A supervised learning method: two steps**

- **Training**: learning the coefficients of the neural networks with known low frequencies (output) and high frequencies (input)
- **Test**: feeding high frequencies into the pretrained neural networks and predicting their unknown low frequency data

# Deep learning

Convolutional Neural Networks (CNNs)

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




- Optimizer：Adam (Kingma and Ba, 2014)

  with a mini-batch of 20 samples

**Input: high frequency recordings**

Sum up five combined units

**Convolution layer**
**Batch normalization layer**
**PReLU layer**

**Fully connected layer**

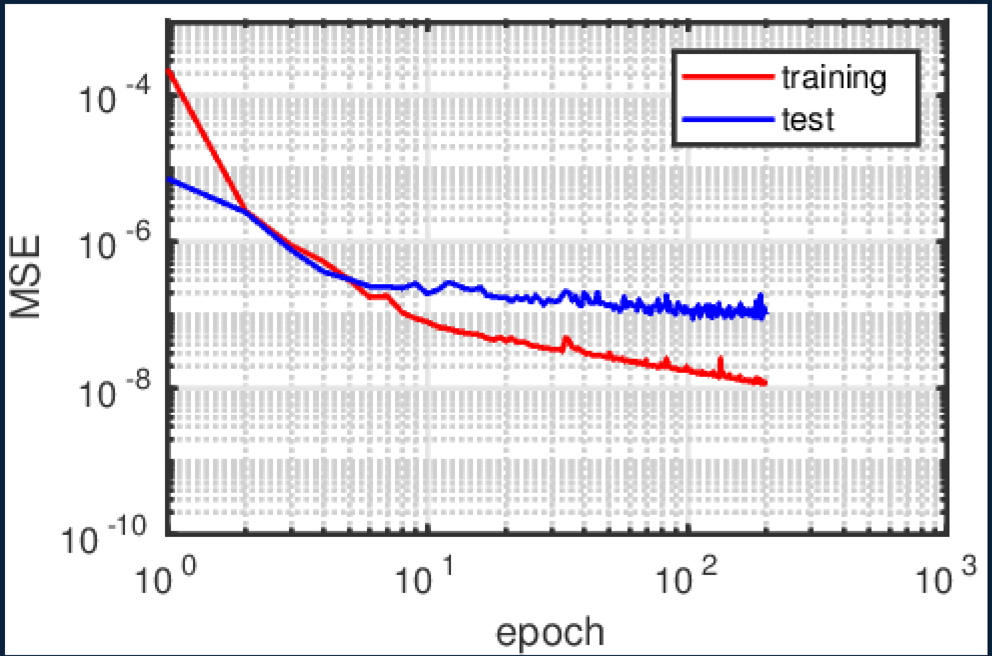
**Output: low frequency recordings**

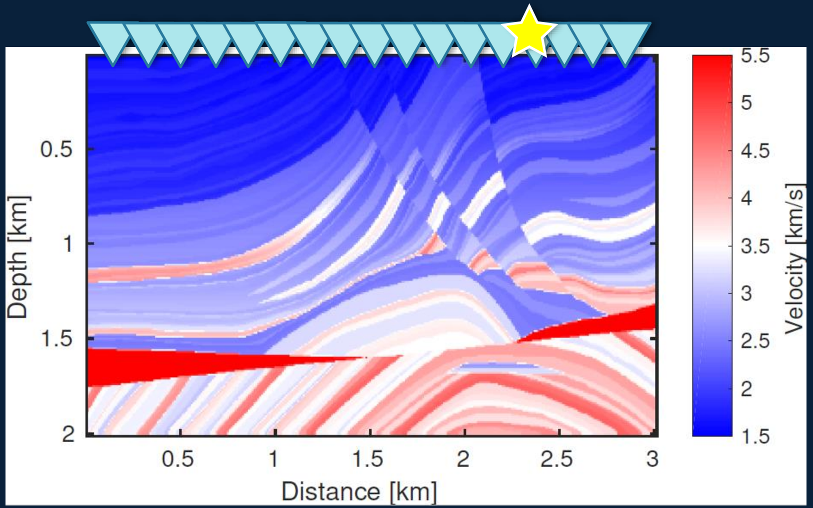
The architecture of CNNs for LF extrapolation

# Numerical experiment

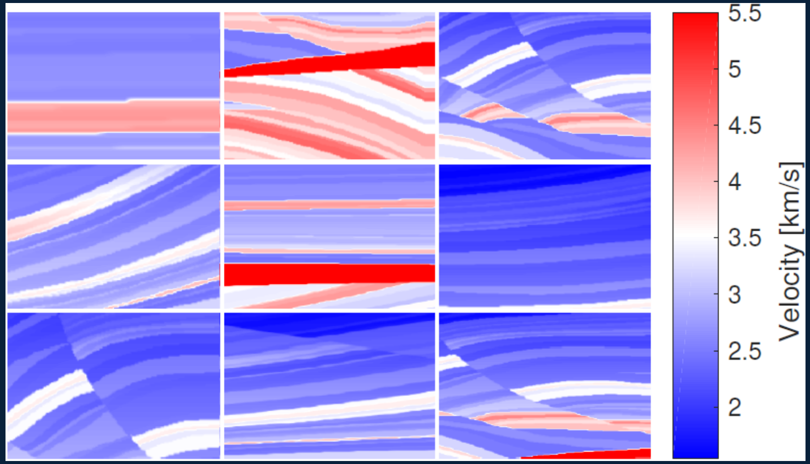
- **Training dataset：** 81,000 samples
- **Test dataset：** 9000 samples
- **Trainable parameters：** 3,290,946

**Be careful of overfitting!**



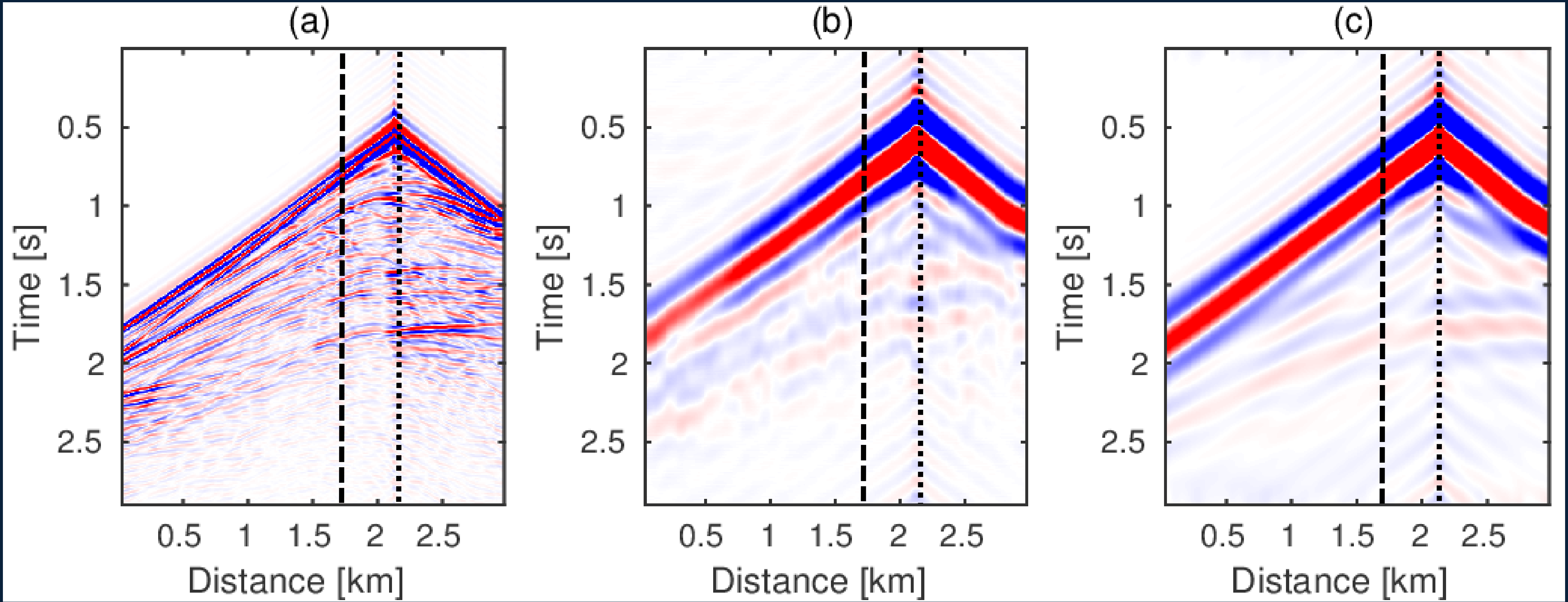

The full-size model to synthesize the test data




The nine submodels to synthesize the training data

# Numerical experiment

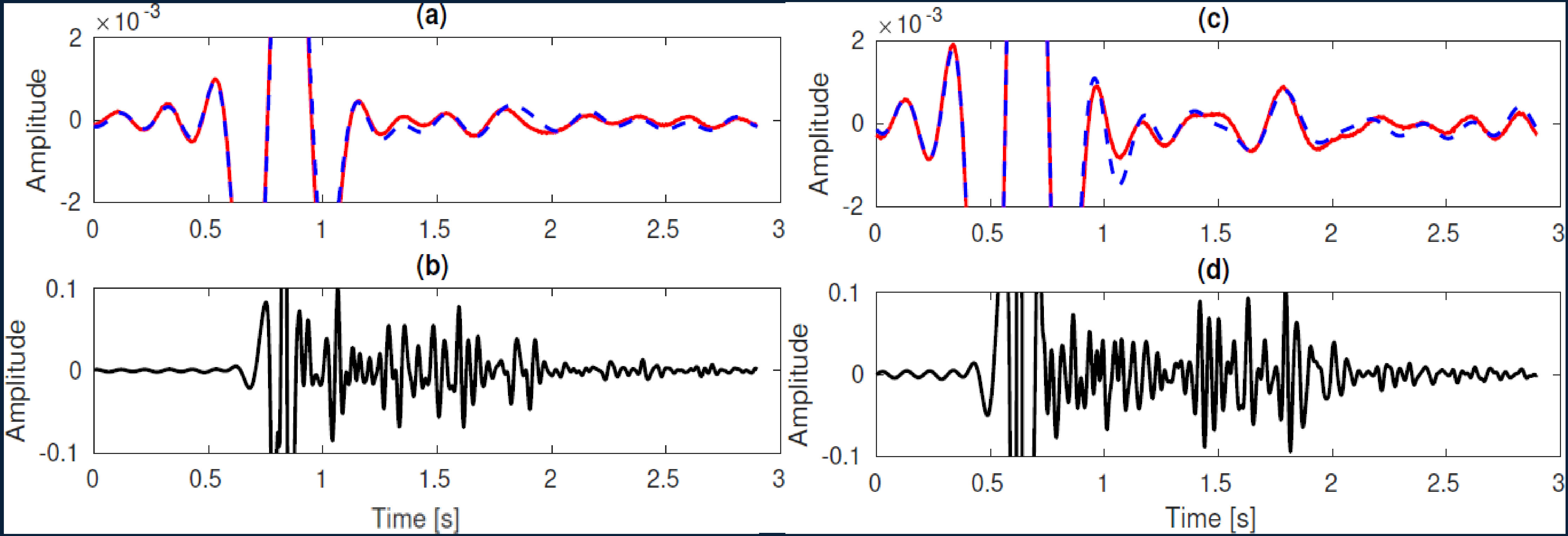
Input high frequencies — Predicted low frequencies — True low frequencies



# Numerical experiment

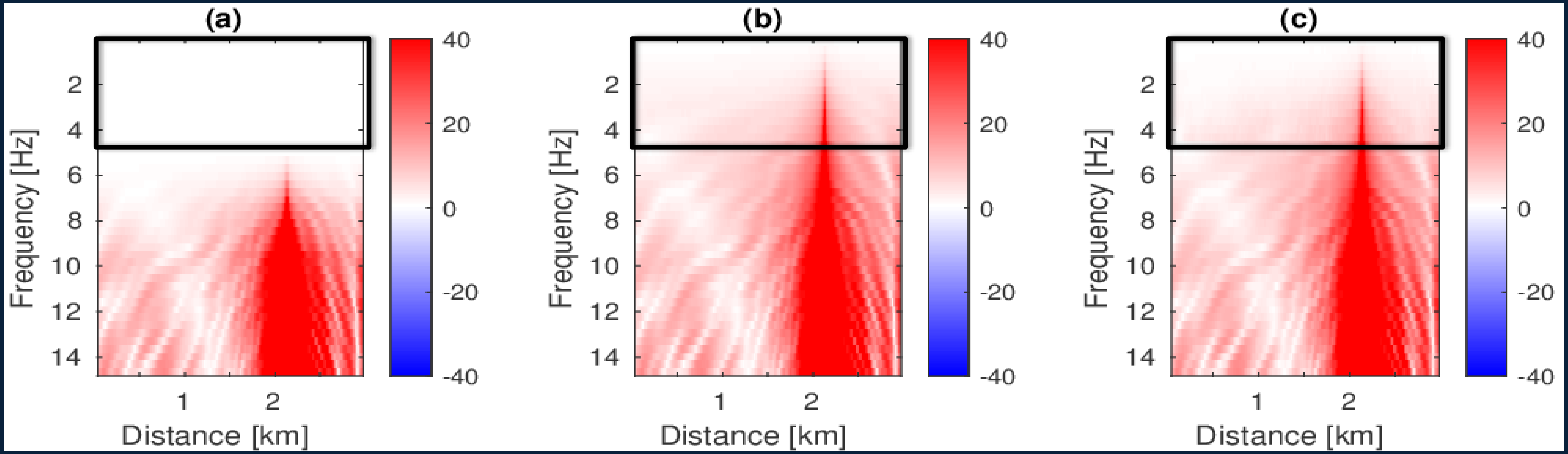
**Comparison of one trace at the horizontal distance:**

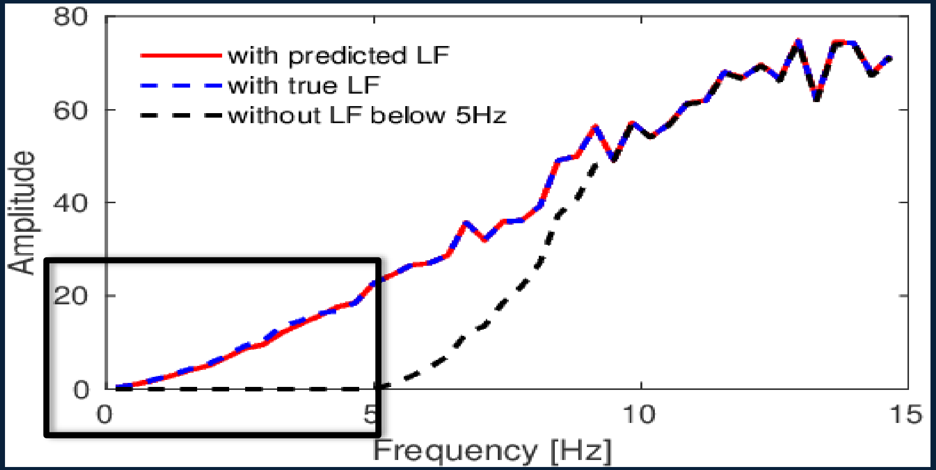
(a) (b) x = 1.73km (c) (d) x = 2.25km

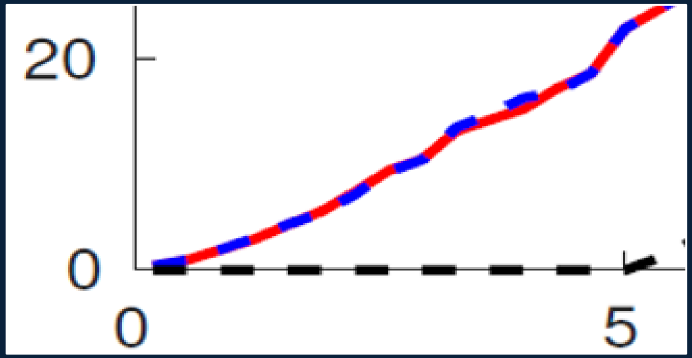



**Output:** Predicted (red line) and true (blue dash line) low frequency recording (0.3−5Hz)
**Input:** Band-limited (Black line) recording (5−50Hz)

# Numerical experiment



# Conclusion

### METHOD

- an architecture of DNNs to extrapolate the low frequencies of band-limited seismic recordings without any preprocessing and post-processing procedures;

### IMPACT

- solution to cycle-skipping problem in FWI;

### LIMITATION

- the lack of generalizability guarantees;
- demanding training time;

Thanks to MIT ERL and Total S A for support.

Thanks for your attention.

# Acknowledgements