Full-waveform inversion based on gradient sampling algorithm

Jizhong Yang[1], Yunyue Elita Li[1],
Yanwen Wei[2], Haohuan Fu[2], and Yuzhu Liu[3]

[1] National University of Singapore
[2] Tsinghua University
[3] Tongji University

Outline

- Introduction
- Methodology
- Numerical Examples
  - 2004 BP model
- Conclusions

Outline

- **Introduction**
- **Methodology**
- **Numerical Examples**
  - **2004 BP model**
- **Conclusions**

Overview of FWI

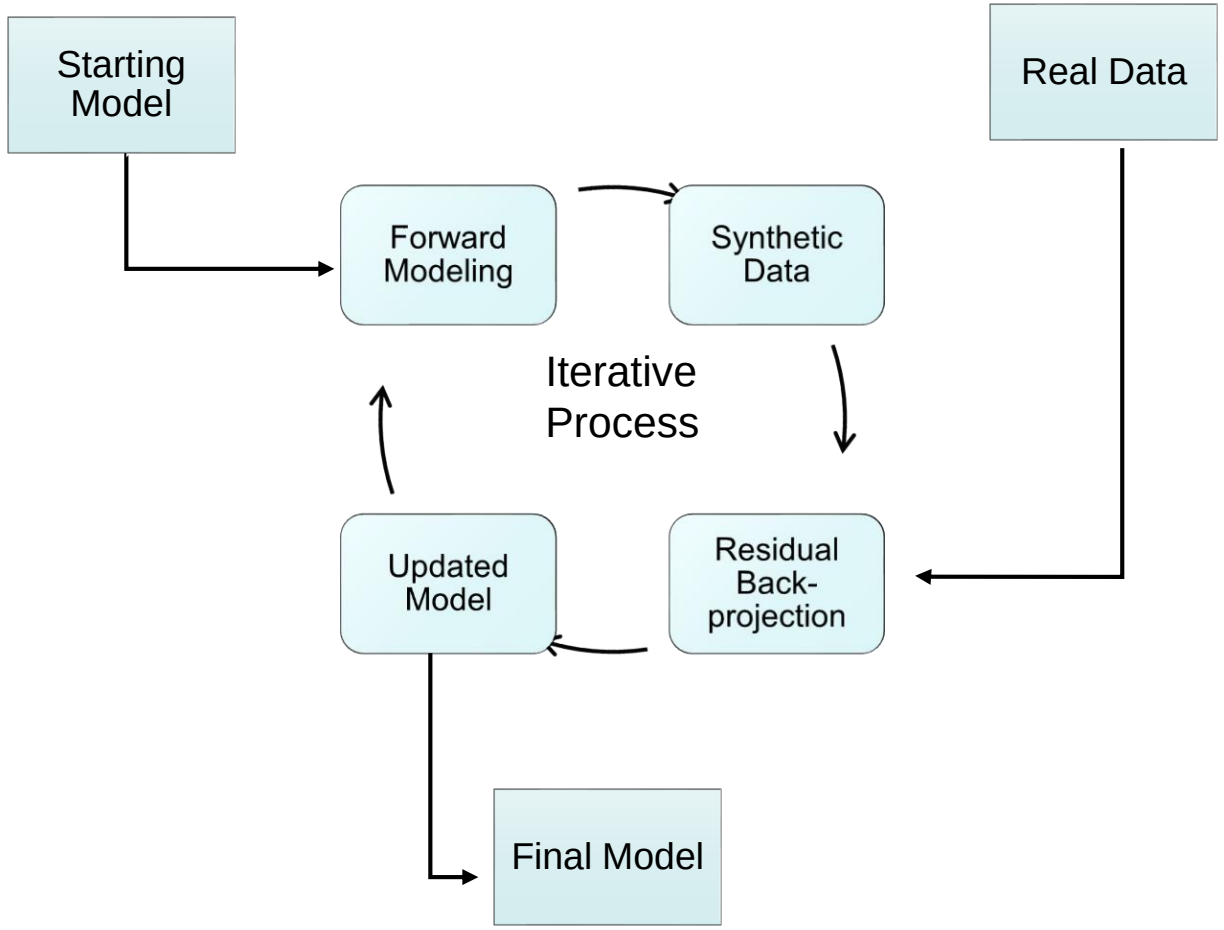
$$J(m) = \frac{1}{2}\left\| u(m) - d \right\|_2^2$$

d: recorded seismic data
u(m): modeled seismic data
m: seismic velocity

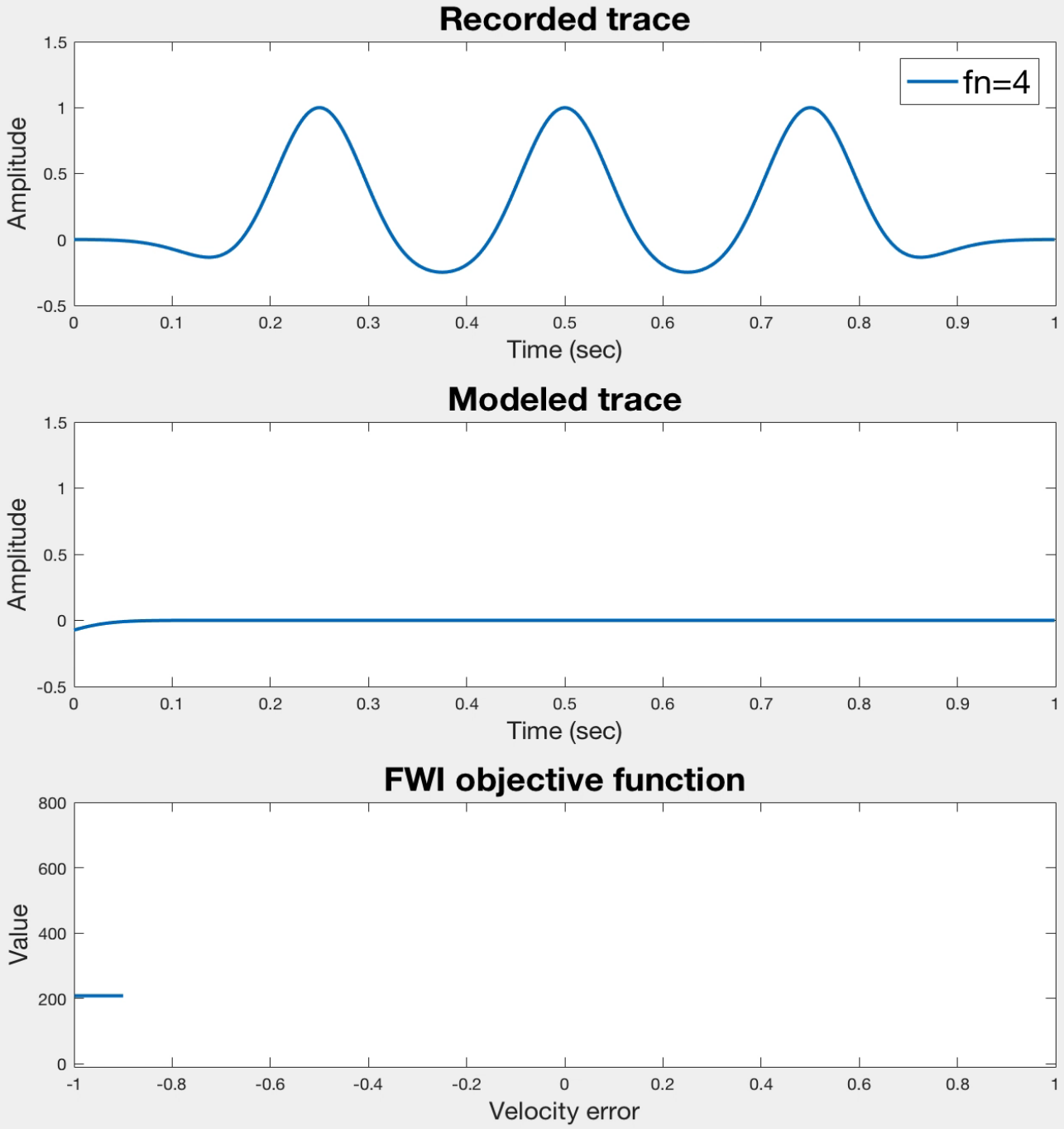
Surface acquisition geometry

Time-domain implementation

Acoustic regime

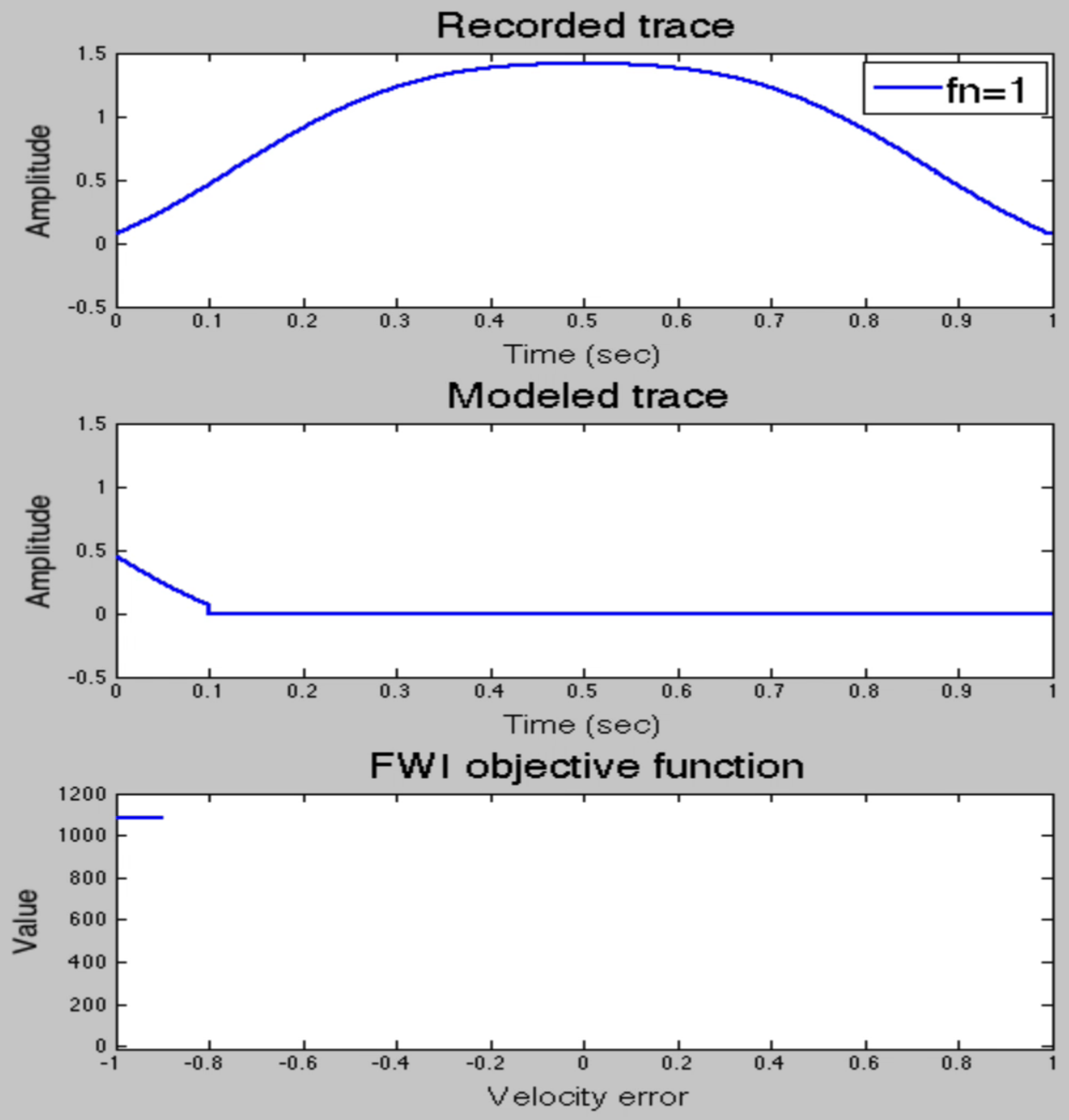


Workflow of FWI



FWI Challenges

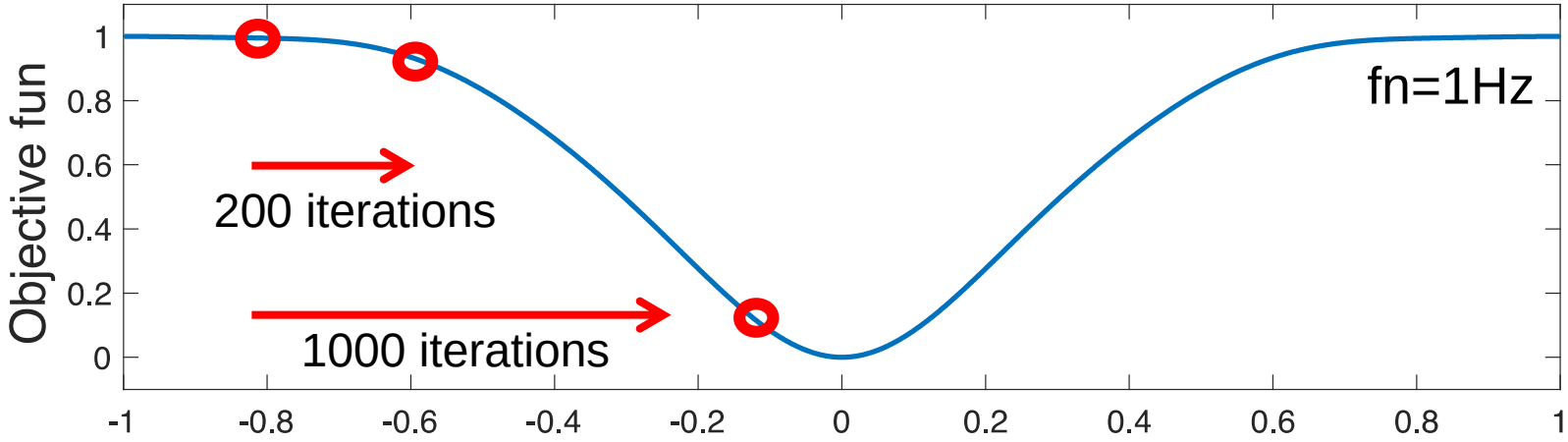
- Nonlinear inversion
  - Iterative methods
- Nonconvex
  - Local minimum



FWI Challenges

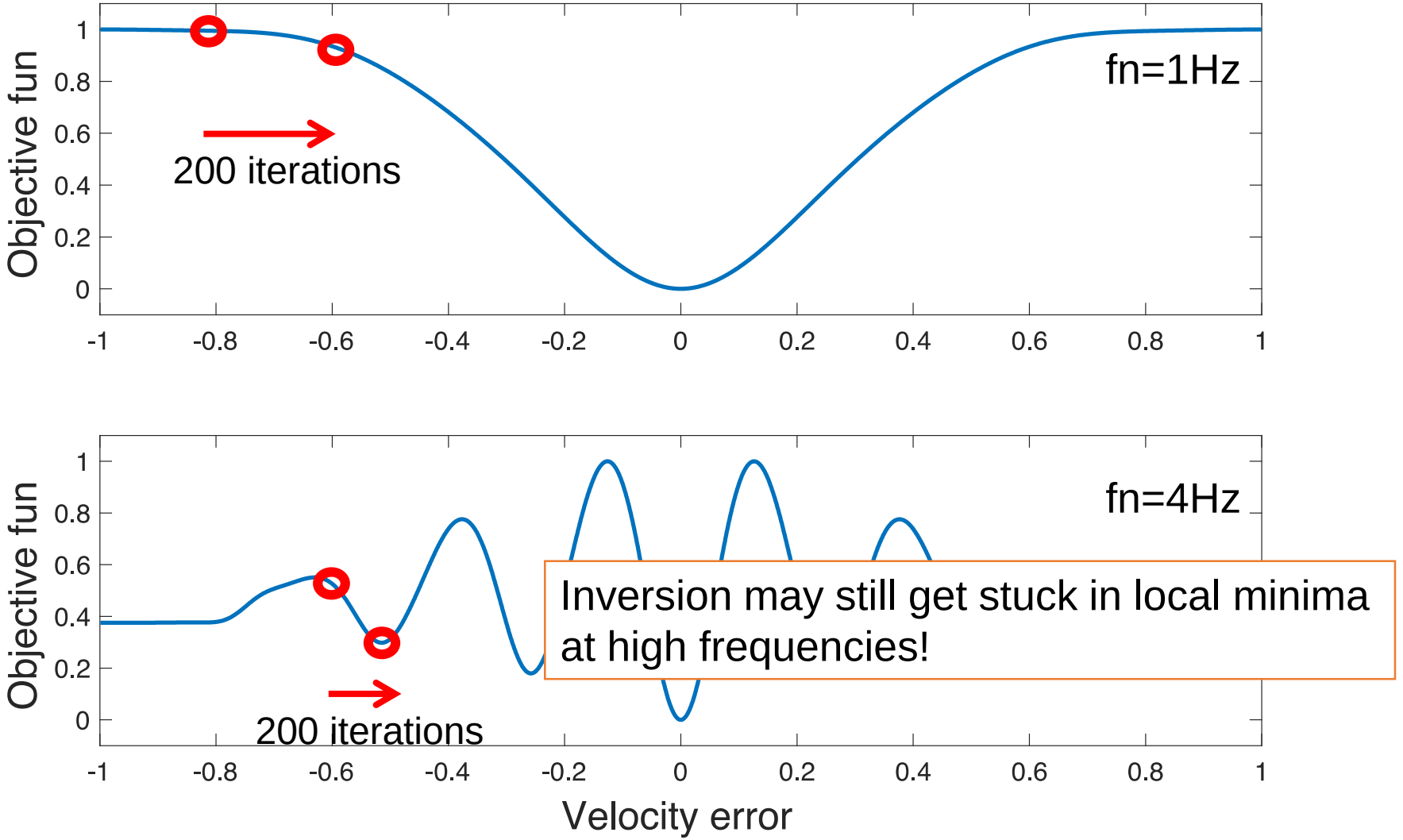
- Low frequency data
  - Missing from previous acquisitions
  - Expensive to acquire
- Very slow convergence at large velocity error

FWI challenges: slow convergence



- Small gradient at the lowest frequency
- Takes too many iterations to update

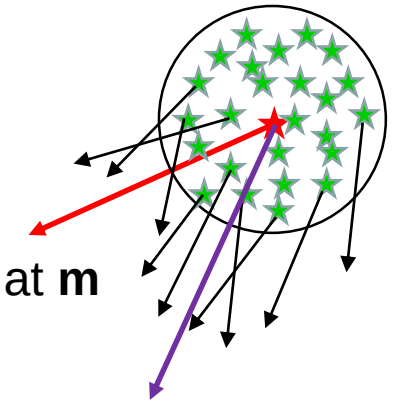
FWI challenges: slow convergence



Inversion may still get stuck in local minima at high frequencies!

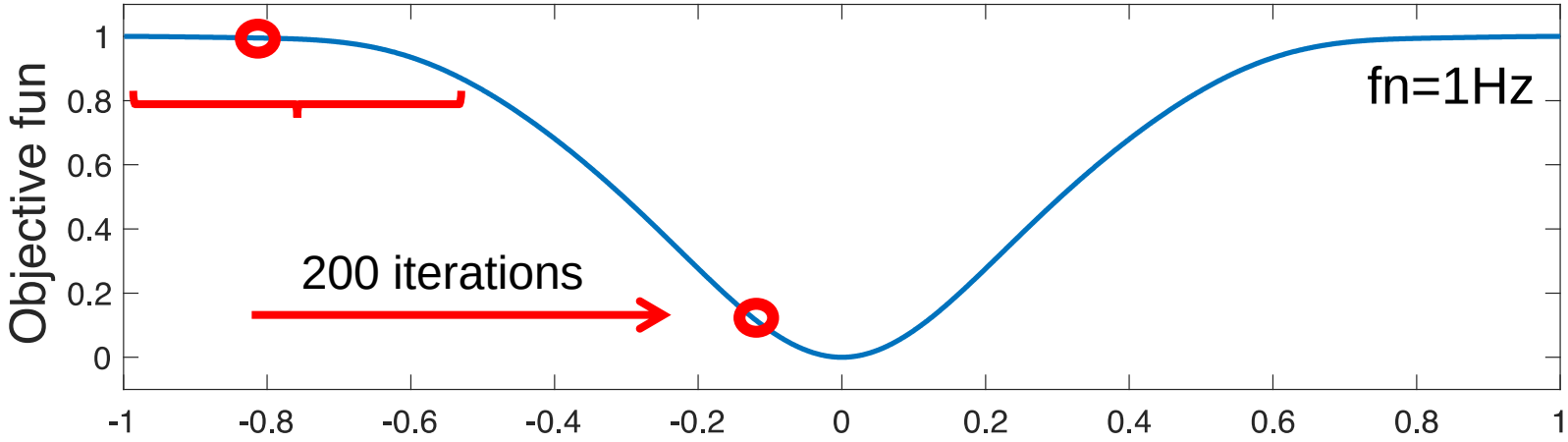
Gradient sampling algorithm

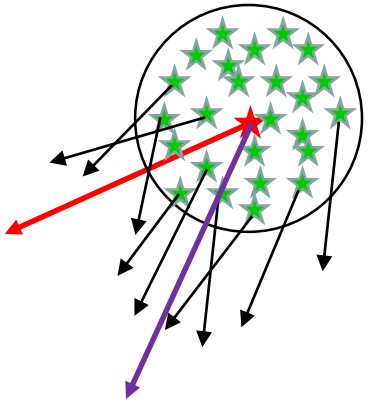
- Sampling N+1 model vectors in a region close to the current model
- Calculating gradients for each sample
- Computing descent directions as a weighted sum over all sampled gradients
- Updating the model using the gradient sampling direction



(Burke et al., 2005; Curtis and Que, 2013)

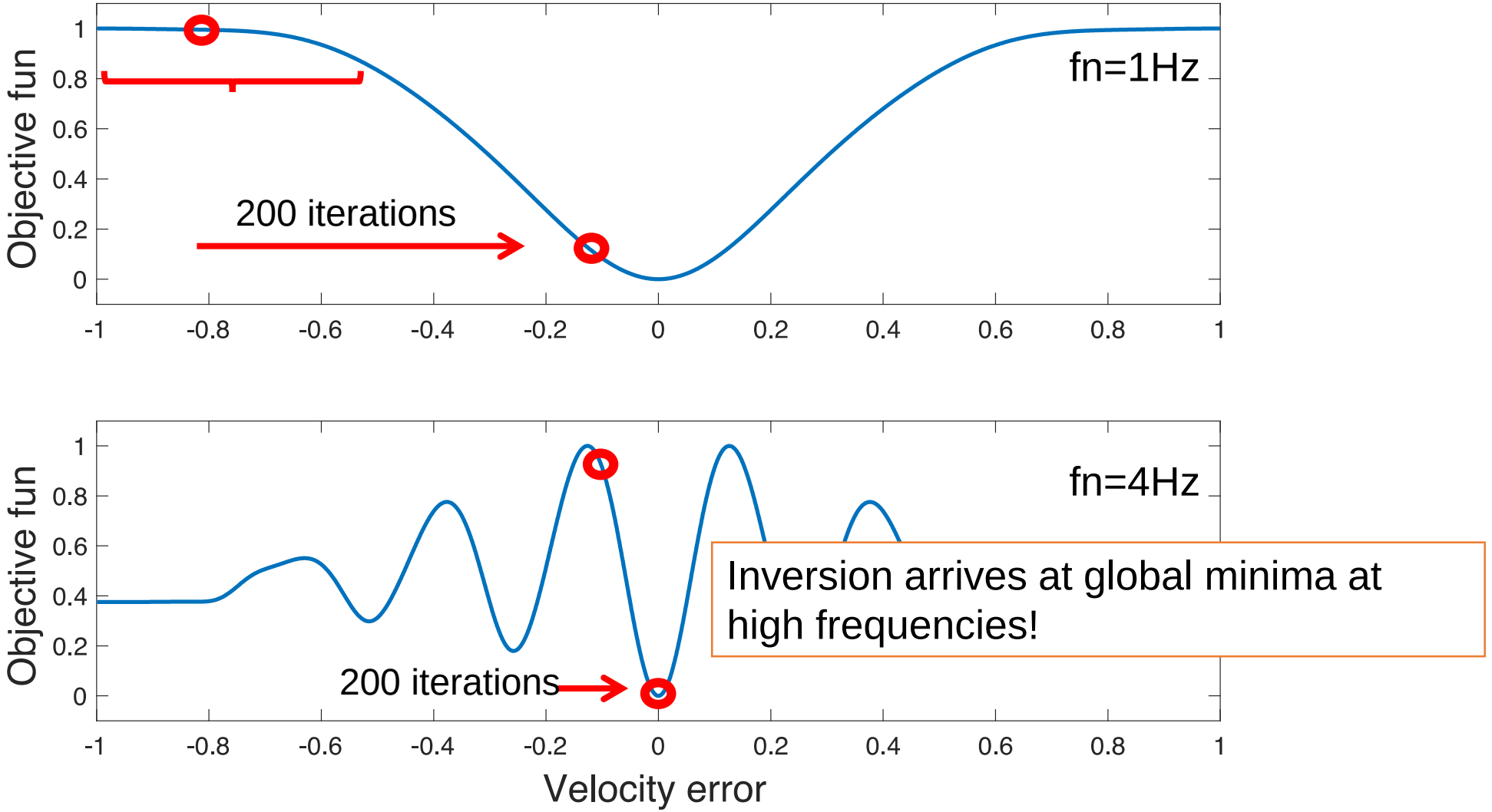
FWI with GSA: global optimal solution





- Sample the vicinity of current model
- Use the weighted gradient to update model
- Better gradients at low frequency and faster convergence

FWI with GSA: global optimal solution



Inversion arrives at global minima at high frequencies!

Outline

- Introduction
- Methodology
- Numerical Examples
  - 2004 BP model
- Conclusions

Methodology

	Gradient	Computational Cost
Conventional FWI	$g(\boldsymbol{x};\boldsymbol{m}) = 2\boldsymbol{m}(\boldsymbol{x})\int U(\boldsymbol{x},t;\boldsymbol{m})V(\boldsymbol{x},t;\boldsymbol{m})dt$	**2Ns**
GSA-FWI	$g(\boldsymbol{x};\boldsymbol{m}_i) = 2\boldsymbol{m}_i(\boldsymbol{x})\int U(\boldsymbol{x},t;\boldsymbol{m}_i)V(\boldsymbol{x},t;\boldsymbol{m}_i)dt$ $g = \sum_{i=0}^{N}\alpha_i g_{mi} \qquad \alpha_i > 0, \sum_{i=0}^{N}\alpha_i = 1$	**2Ns*(N+1)**

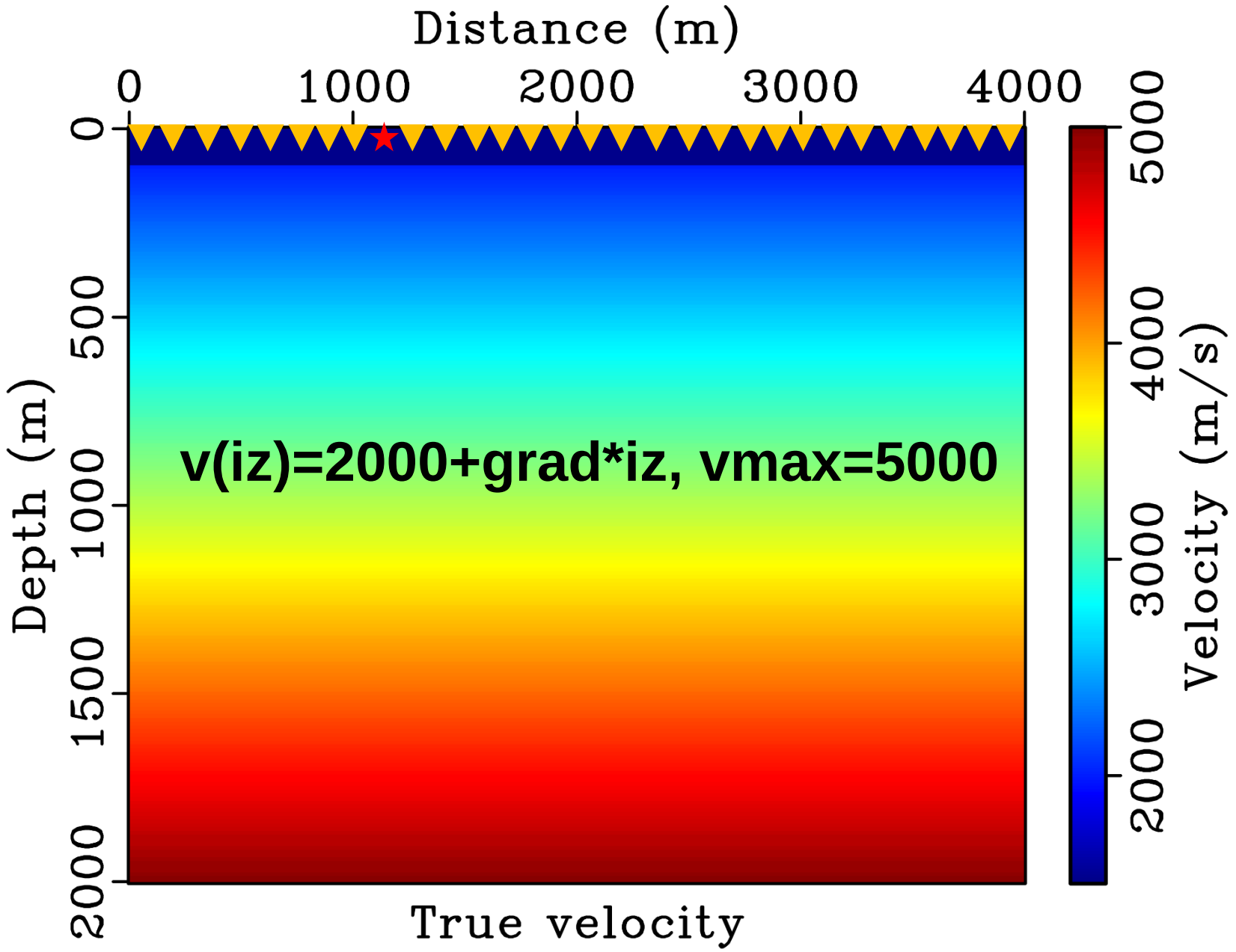
$U(\boldsymbol{x},t;\boldsymbol{m})$: Forward propagated wavefield

$V(\boldsymbol{x},t;\boldsymbol{m})$: Backward propagated wavefield

α : Weighting coefficient

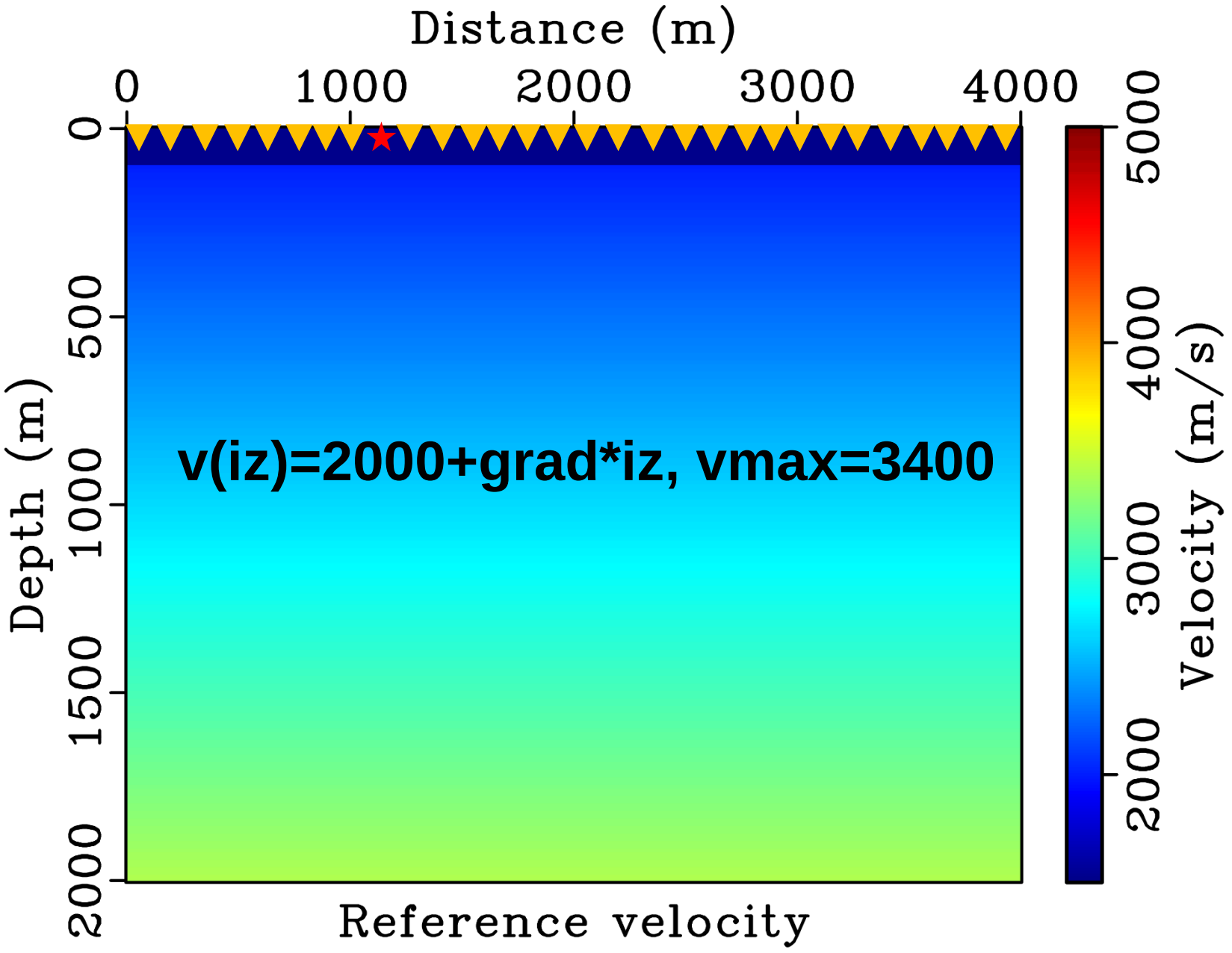
Ns is the number of sources, N+1 is the number of sampled vectors in the vicinity of the current model

How can we approximate the sampled gradient in an efficient way?



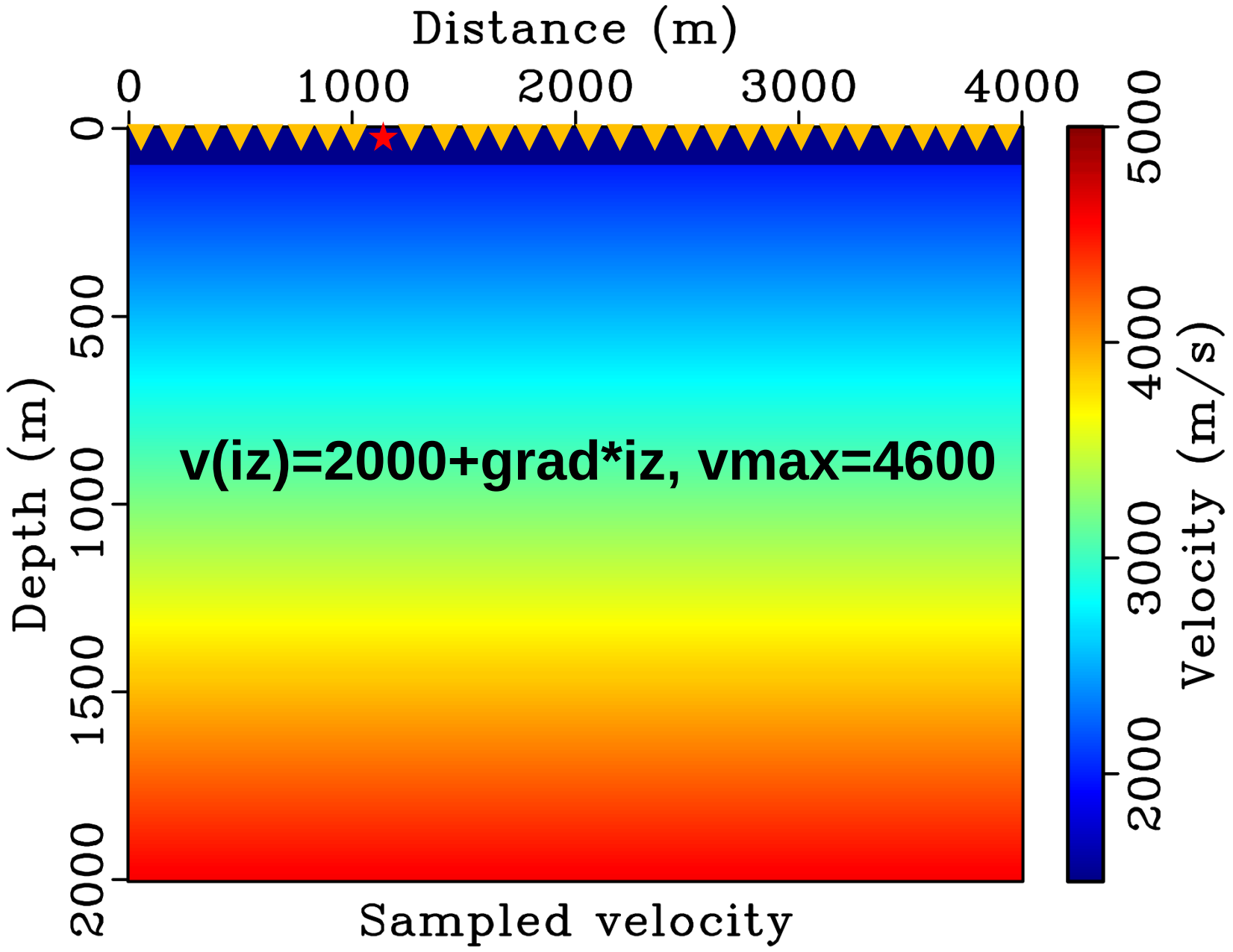
★ source

▼ receiver



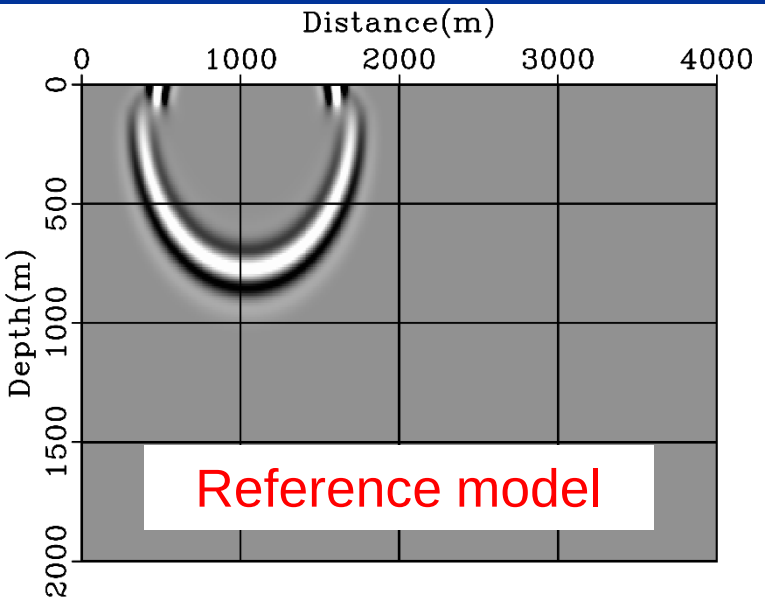
★ source

▼ receiver

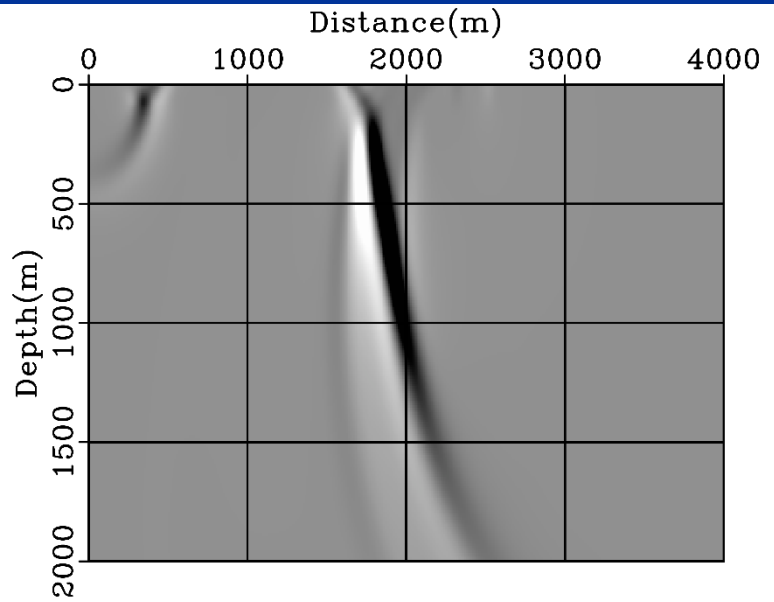


★ source

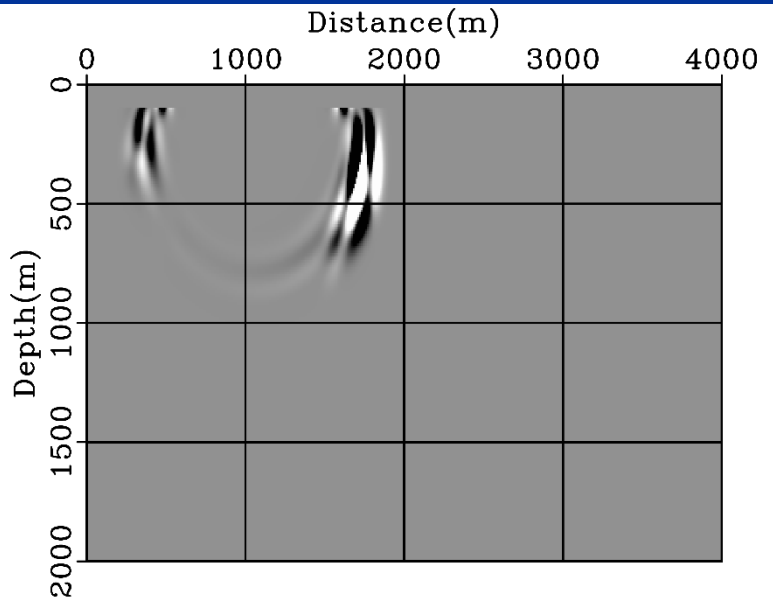
▼ receiver



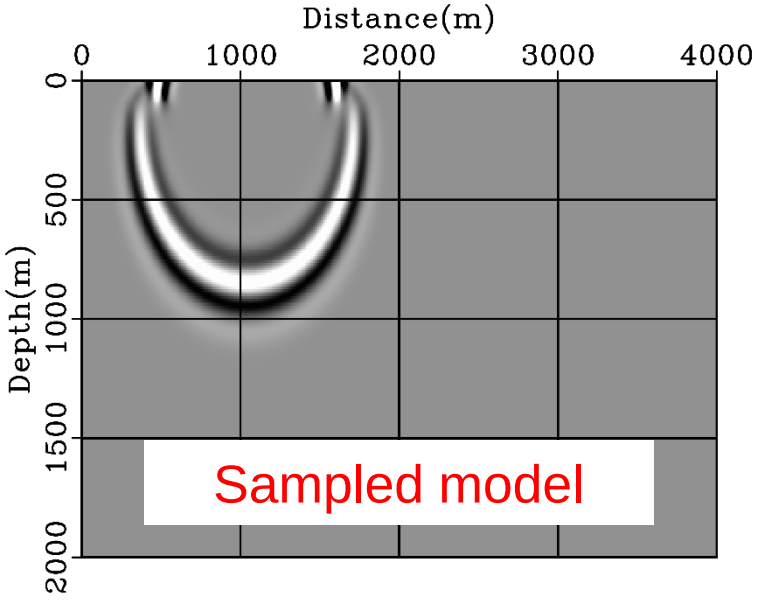
Forward *

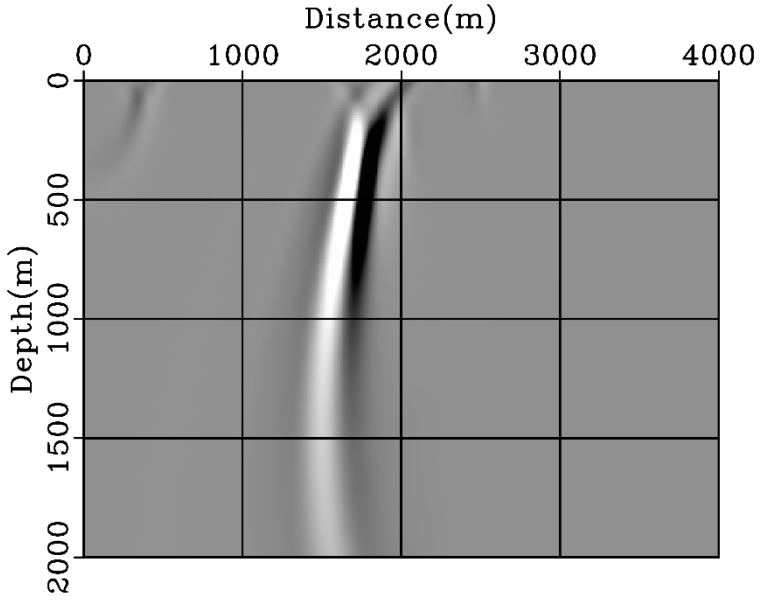


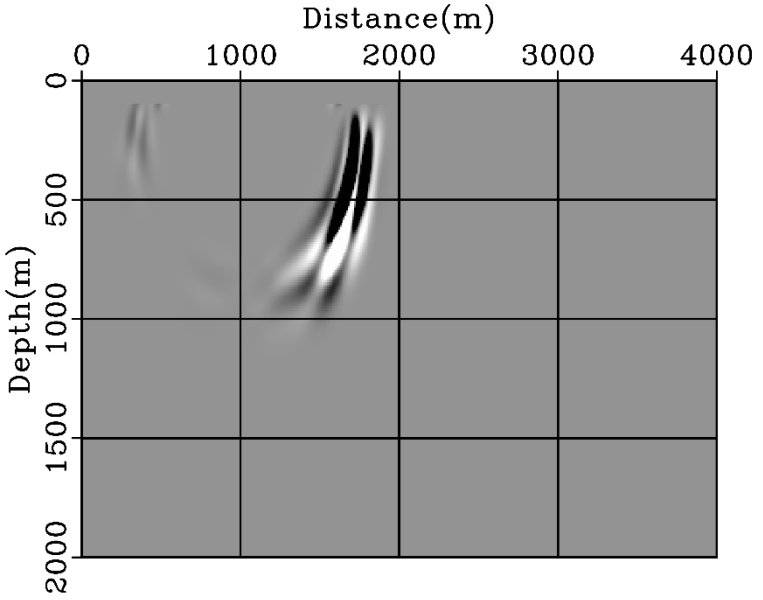
Backward =

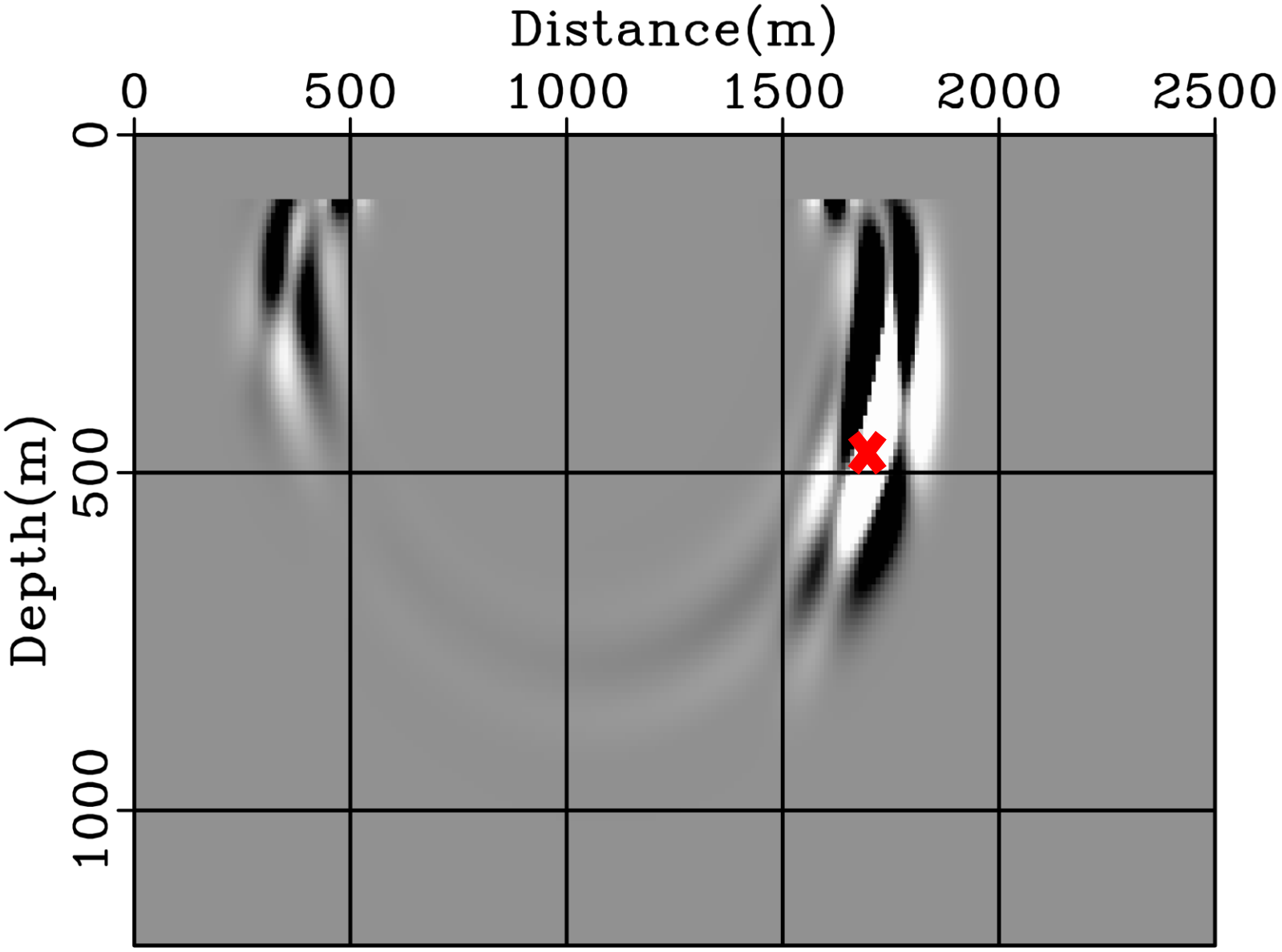


Gradient



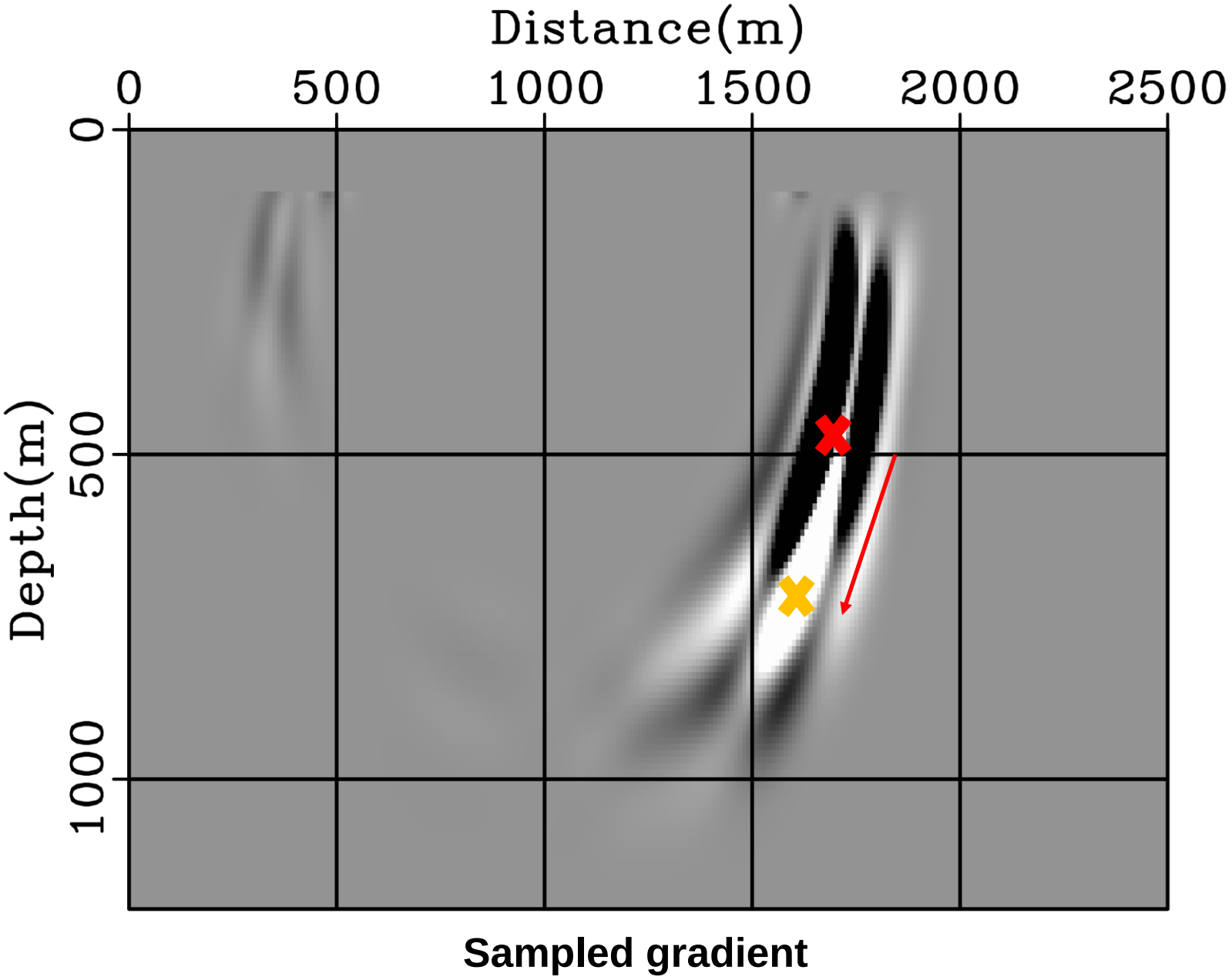






Reference gradient

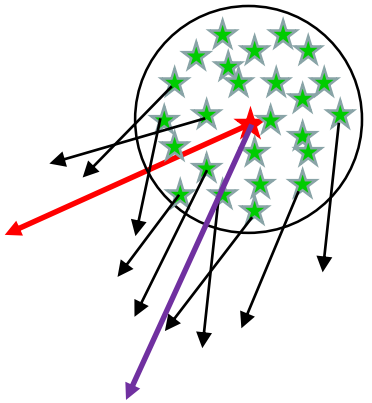
Gradient calculated using the reference velocity model at t=0.5s



Sampled gradient

Gradient calculated using the sampled velocity model at t=0.5s

How can we approximate the sampled gradient in an efficient way?



$$\boldsymbol{g}\left(\boldsymbol{x};\boldsymbol{m}_i\right) \approx \boldsymbol{g}\left(\boldsymbol{x}-\boldsymbol{h}_i;\boldsymbol{m}\right)$$

$$= 2\boldsymbol{m}\left(\boldsymbol{x}-\boldsymbol{h}_i\right)\int U\left(\boldsymbol{x}-\boldsymbol{h}_i,t;\boldsymbol{m}\right)V\left(\boldsymbol{x}-\boldsymbol{h}_i,t;\boldsymbol{m}\right)dt$$

$$\boldsymbol{g}\left(\boldsymbol{x}\right) = \sum \alpha\left(\boldsymbol{h}_i\right)2\boldsymbol{m}\left(\boldsymbol{x}-\boldsymbol{h}_i\right)\int U\left(\boldsymbol{x}-\boldsymbol{h}_i,t;\boldsymbol{m}\right)V\left(\boldsymbol{x}-\boldsymbol{h}_i,t;\boldsymbol{m}\right)dt \qquad \alpha_i > 0, \sum_{i=0}^{N}\alpha_i = 1$$

$$\boldsymbol{g}\left(\boldsymbol{x}\right) \approx 2\int \boldsymbol{m}\left(\boldsymbol{x}-\boldsymbol{h}_t\right)U\left(\boldsymbol{x}-\boldsymbol{h}_t,t;\boldsymbol{m}\right)V\left(\boldsymbol{x}-\boldsymbol{h}_t,t;\boldsymbol{m}\right)dt$$

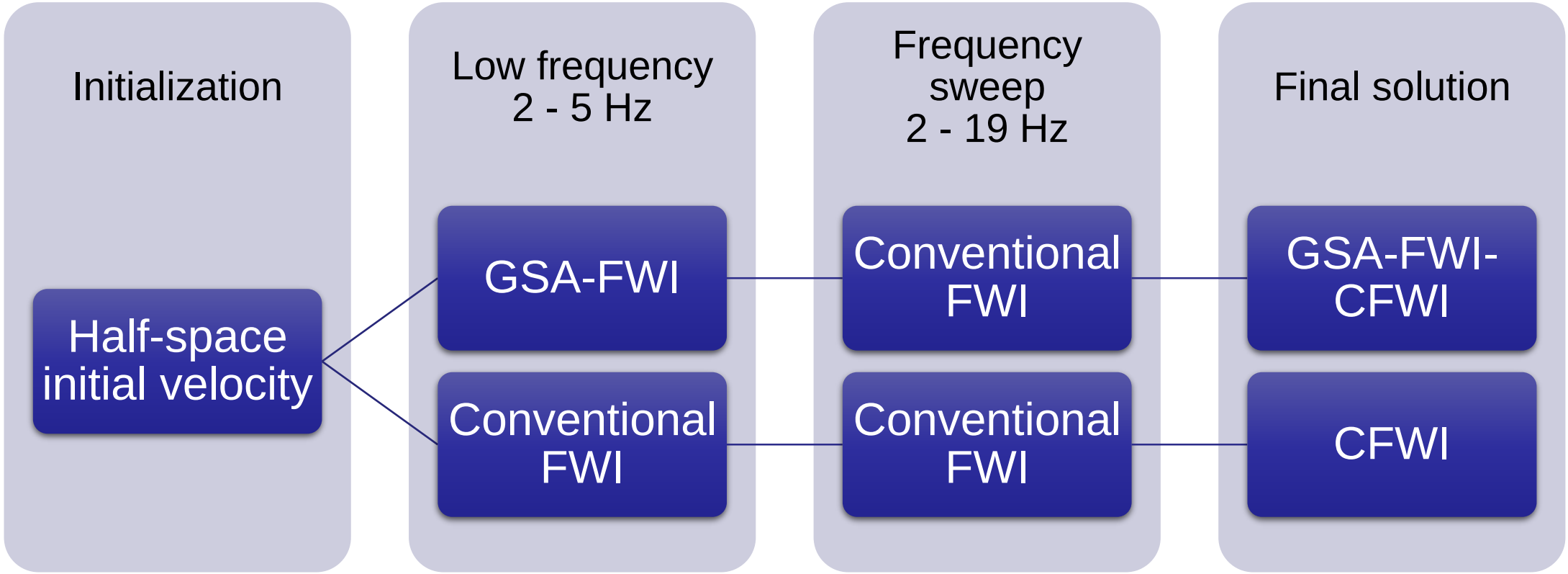
Randomly sample one h within the First Fresnel Zone at each time step!

	Gradient	Computational Cost
Conventional FWI	$g(\boldsymbol{x};\boldsymbol{m}) = 2\boldsymbol{m}(\boldsymbol{x})\int U(\boldsymbol{x},t;\boldsymbol{m})V(\boldsymbol{x},t;\boldsymbol{m})dt$	**2Ns**
GSA-FWI	$g(\boldsymbol{x};\boldsymbol{m}_i) = 2\boldsymbol{m}_i(\boldsymbol{x})\int U(\boldsymbol{x},t;\boldsymbol{m}_i)V(\boldsymbol{x},t;\boldsymbol{m}_i)dt$ $g = \sum_{i=0}^{N}\alpha_i g_{mi} \qquad \alpha_i > 0, \sum_{i=0}^{N}\alpha_i = 1$	**2Ns*(N+1)**
GSA-FWI (Our method)	$\boldsymbol{g}(\boldsymbol{x}) \approx 2\int \boldsymbol{m}(\boldsymbol{x}-\boldsymbol{h}_t)U(\boldsymbol{x}-\boldsymbol{h}_t,t;\boldsymbol{m})V(\boldsymbol{x}-\boldsymbol{h}_t,t;\boldsymbol{m})dt$	**2Ns**

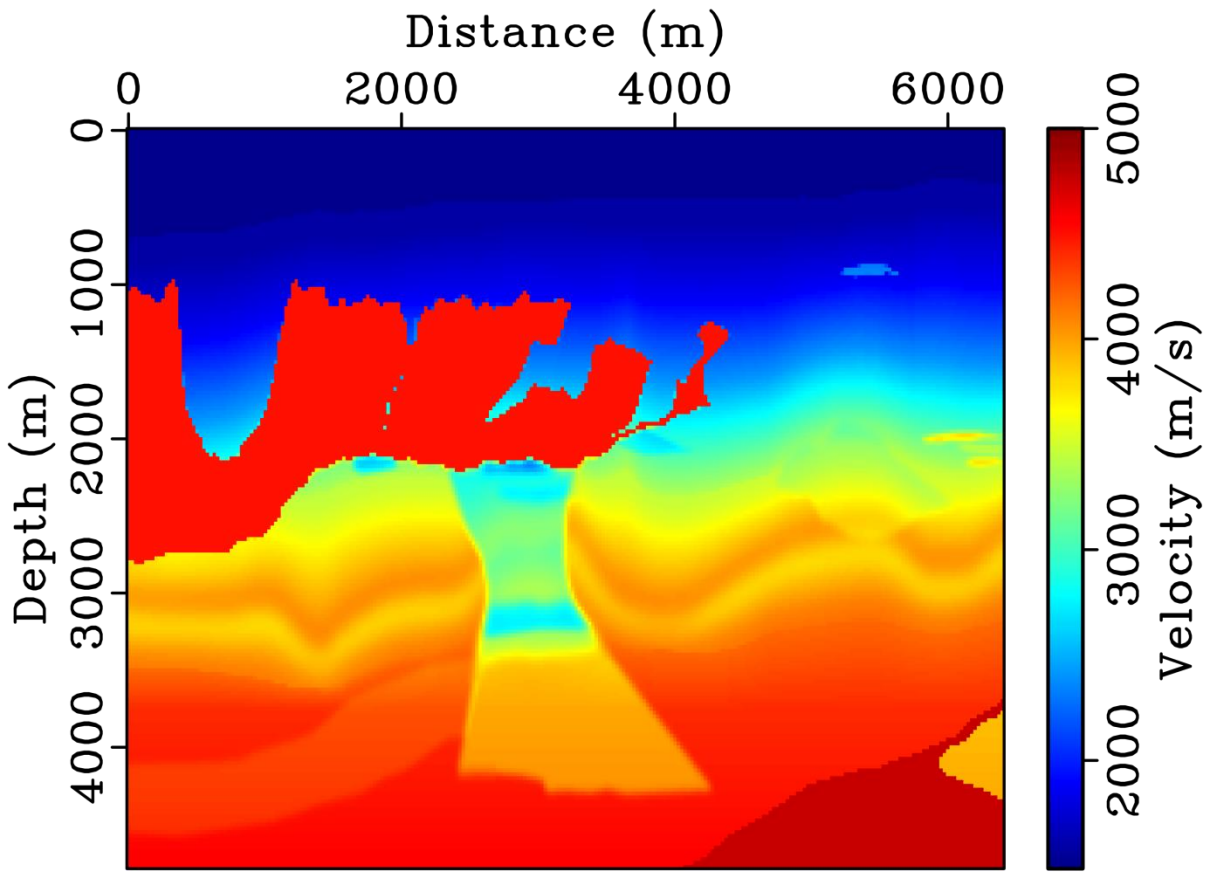
Outline

- Introduction
- Methodology
- Numerical Examples
  - 2004 BP model
- Conclusions

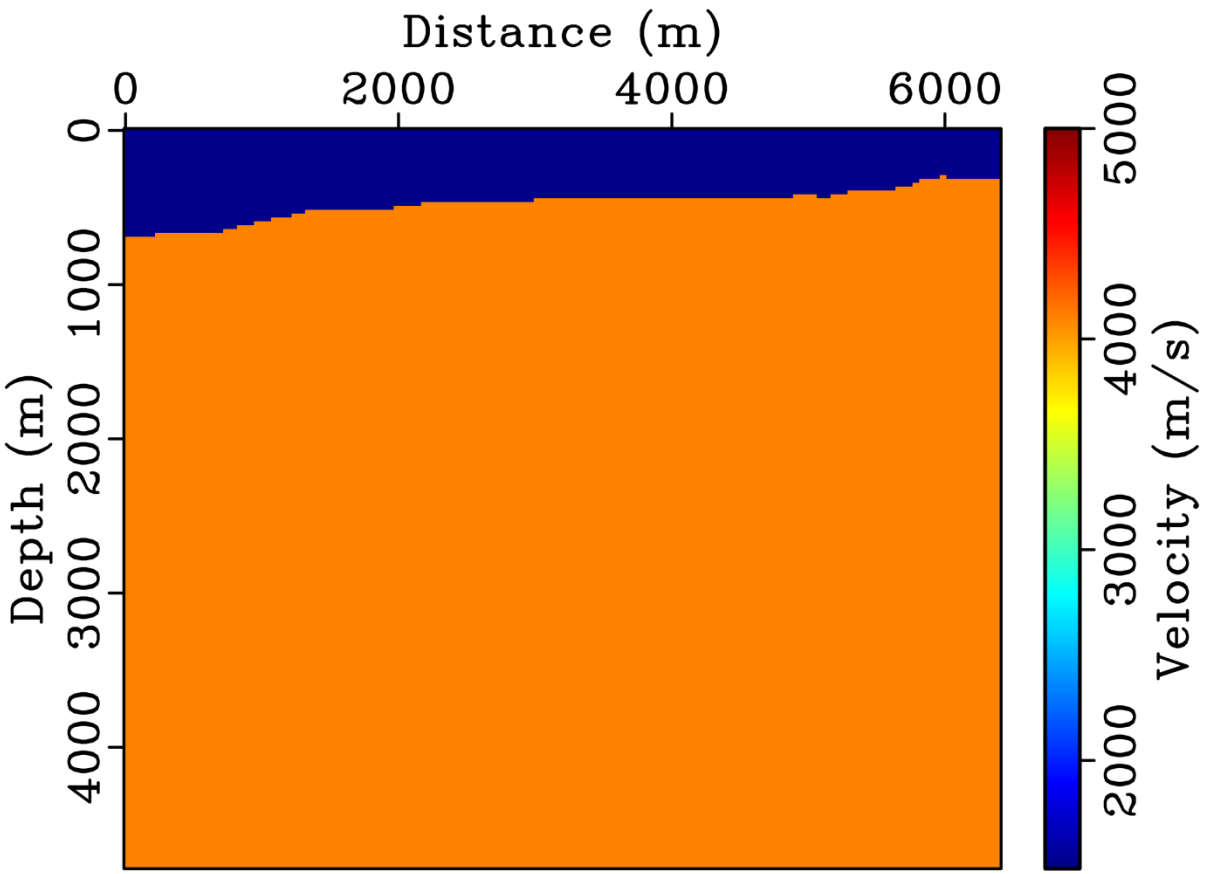
Workflow


Initialization
Half-space initial velocity
Low frequency 2 - 5 Hz
GSA-FWI
Conventional FWI
Frequency sweep 2 - 19 Hz
Conventional FWI
Conventional FWI
Final solution
GSA-FWI-CFWI
CFWI

True velocity

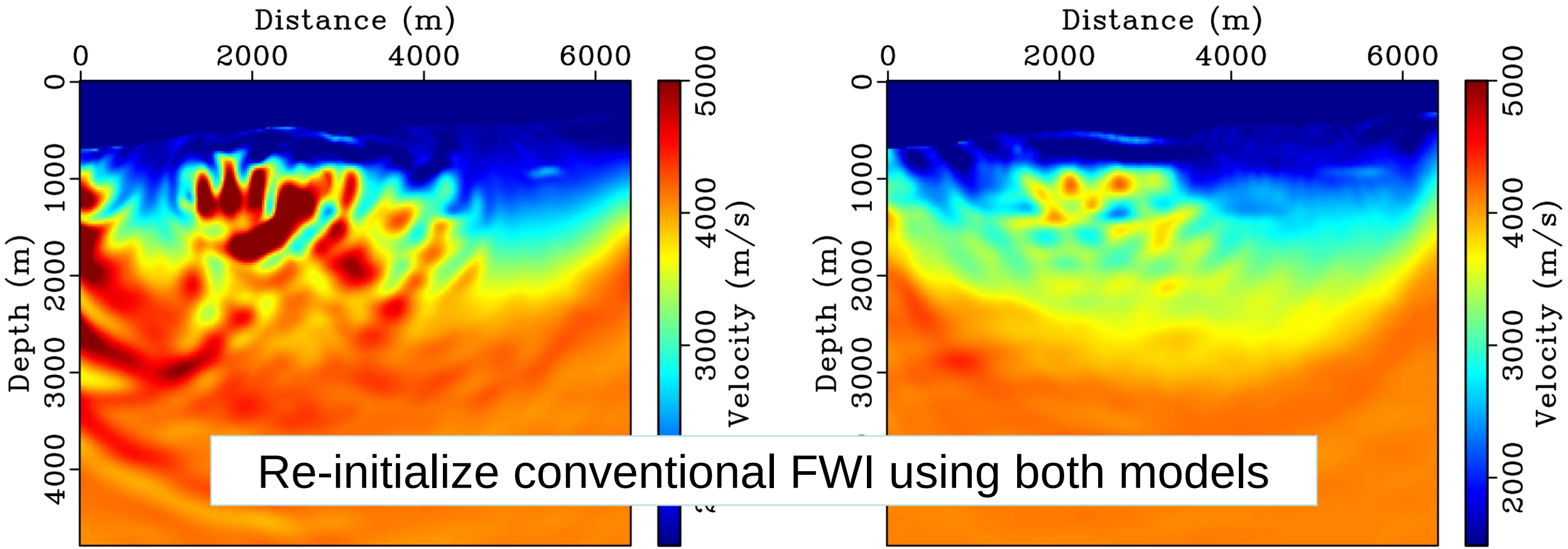


Initial velocity



Conventional FWI

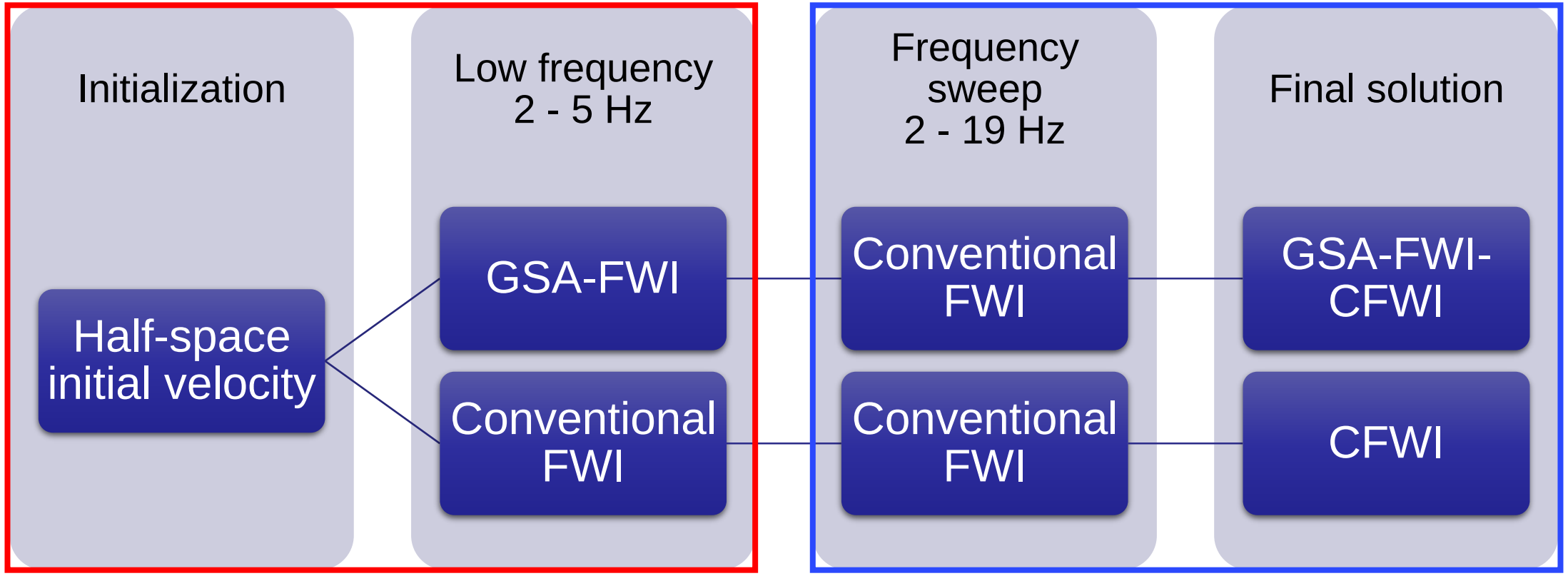
GSA-FWI



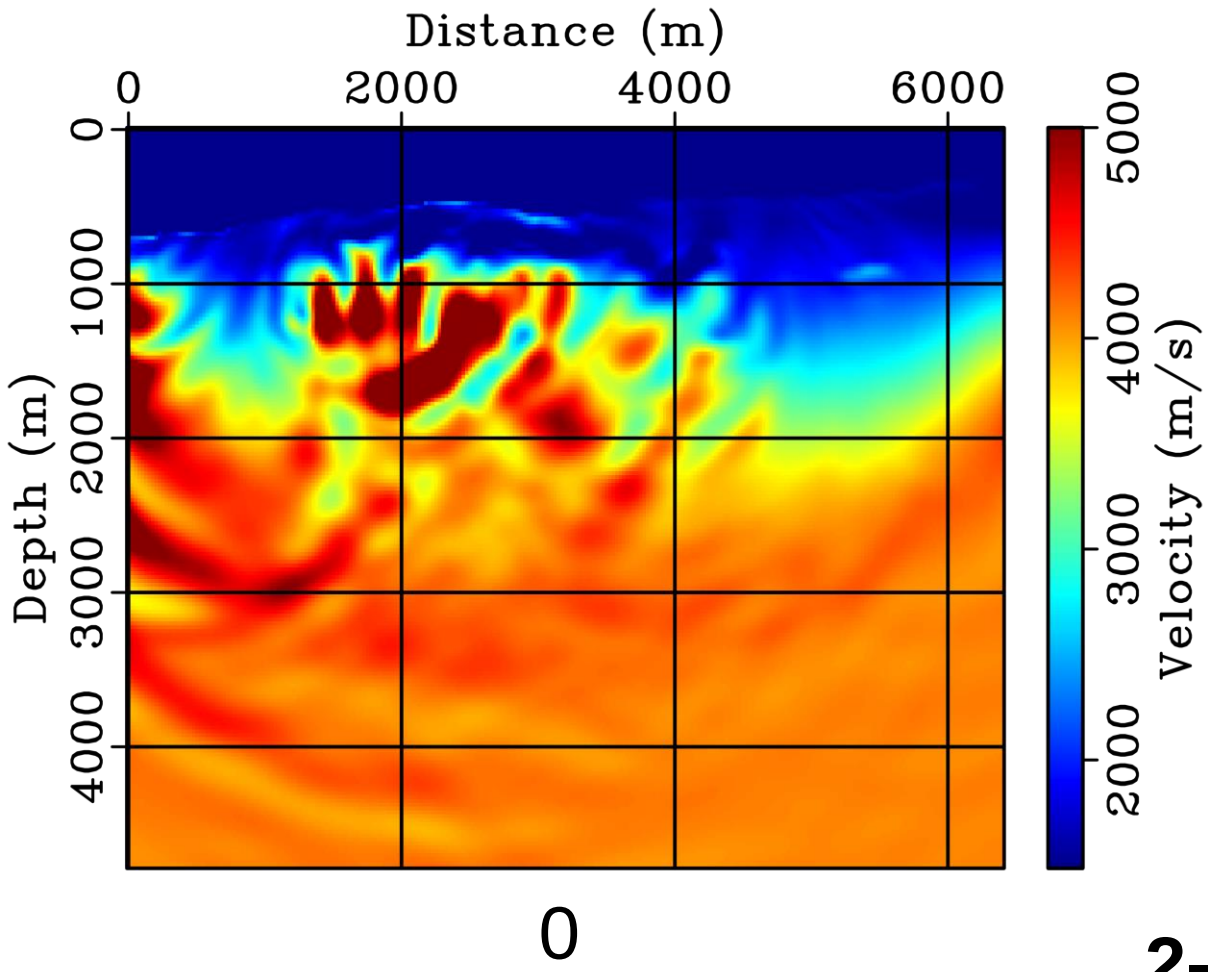
Re-initialize conventional FWI using both models

2-5 Hz

Workflow

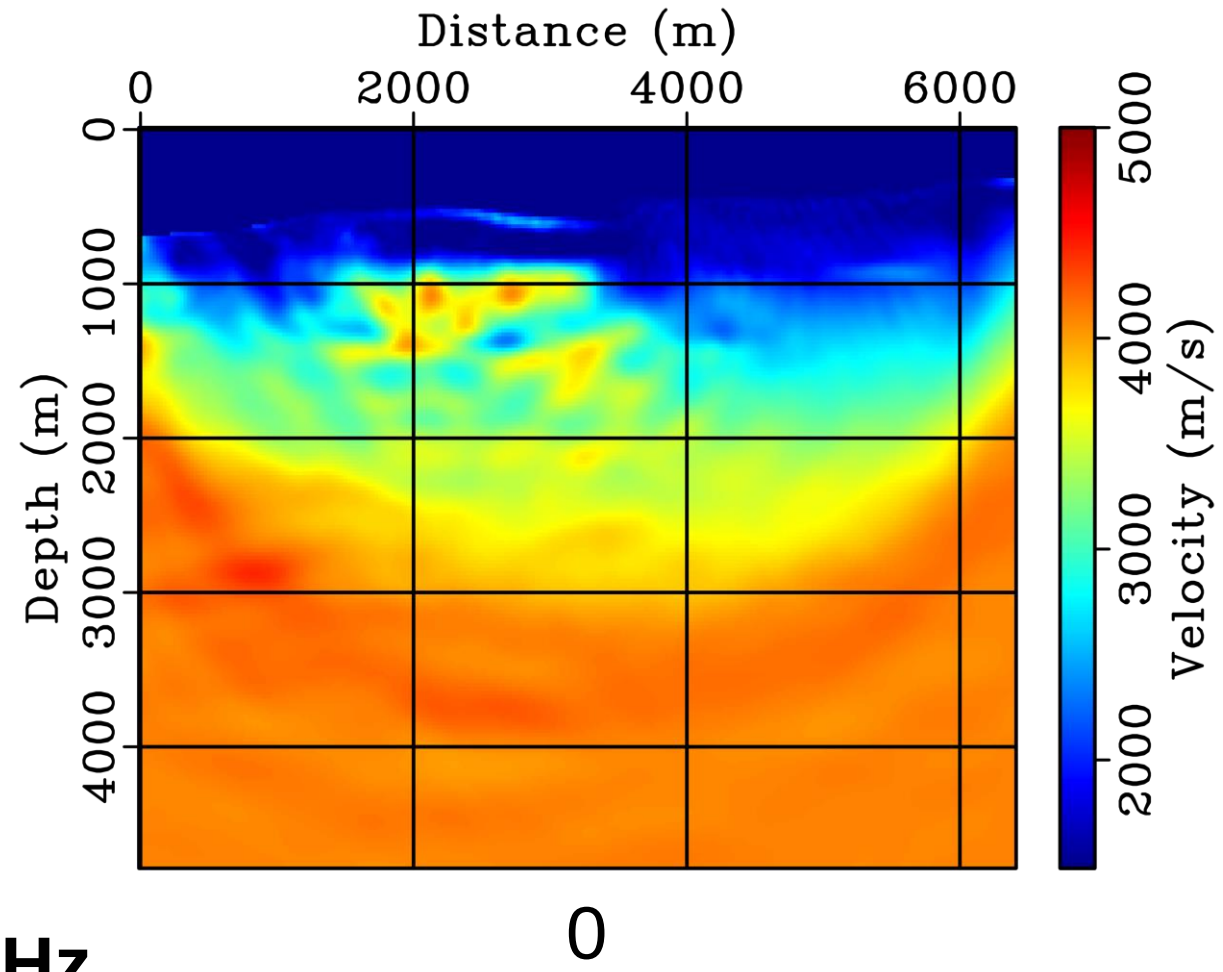

Initialization
Half-space initial velocity
Low frequency 2 - 5 Hz
GSA-FWI
Conventional FWI
Frequency sweep 2 - 19 Hz
Conventional FWI
Conventional FWI
Final solution
GSA-FWI-CFWI
CFWI

Conventional FWI



0

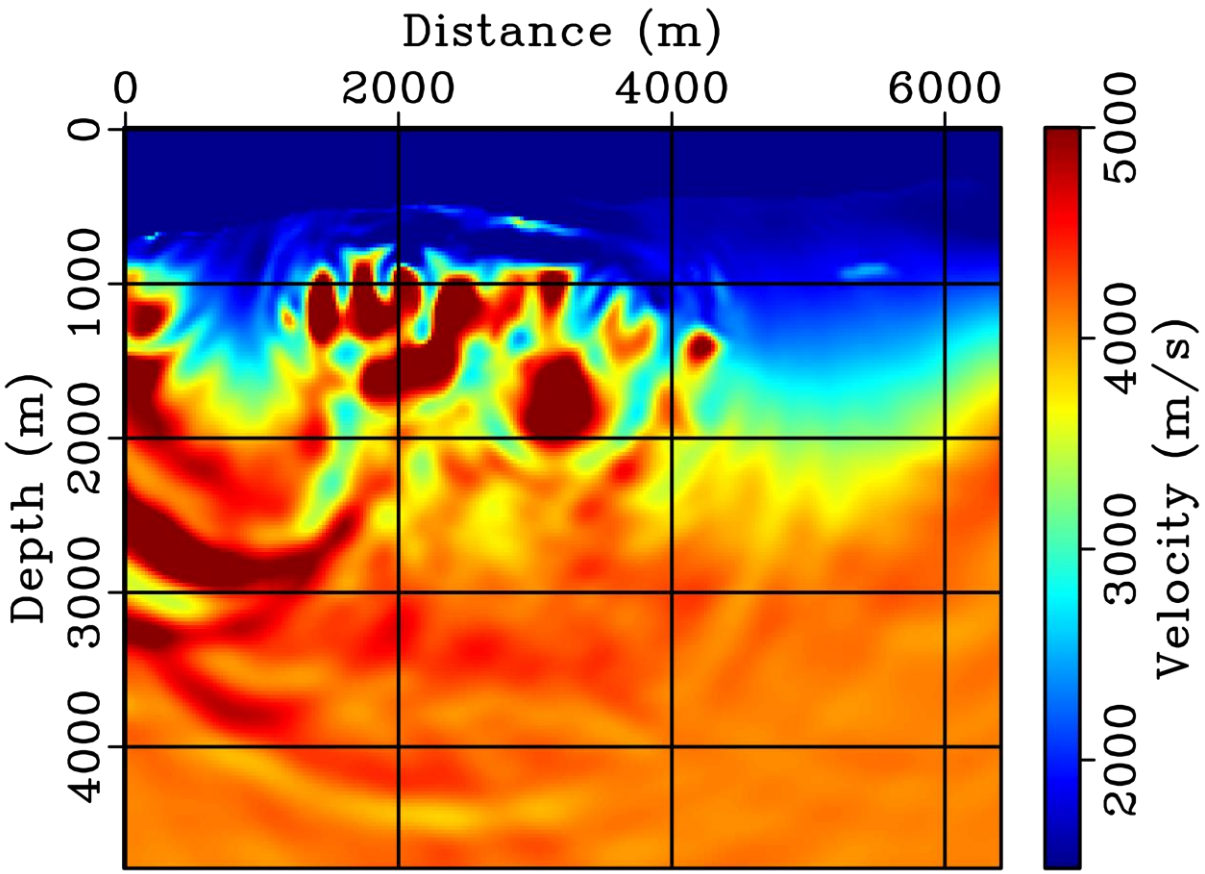
Conventional FWI initialized with GSA-FWI



0

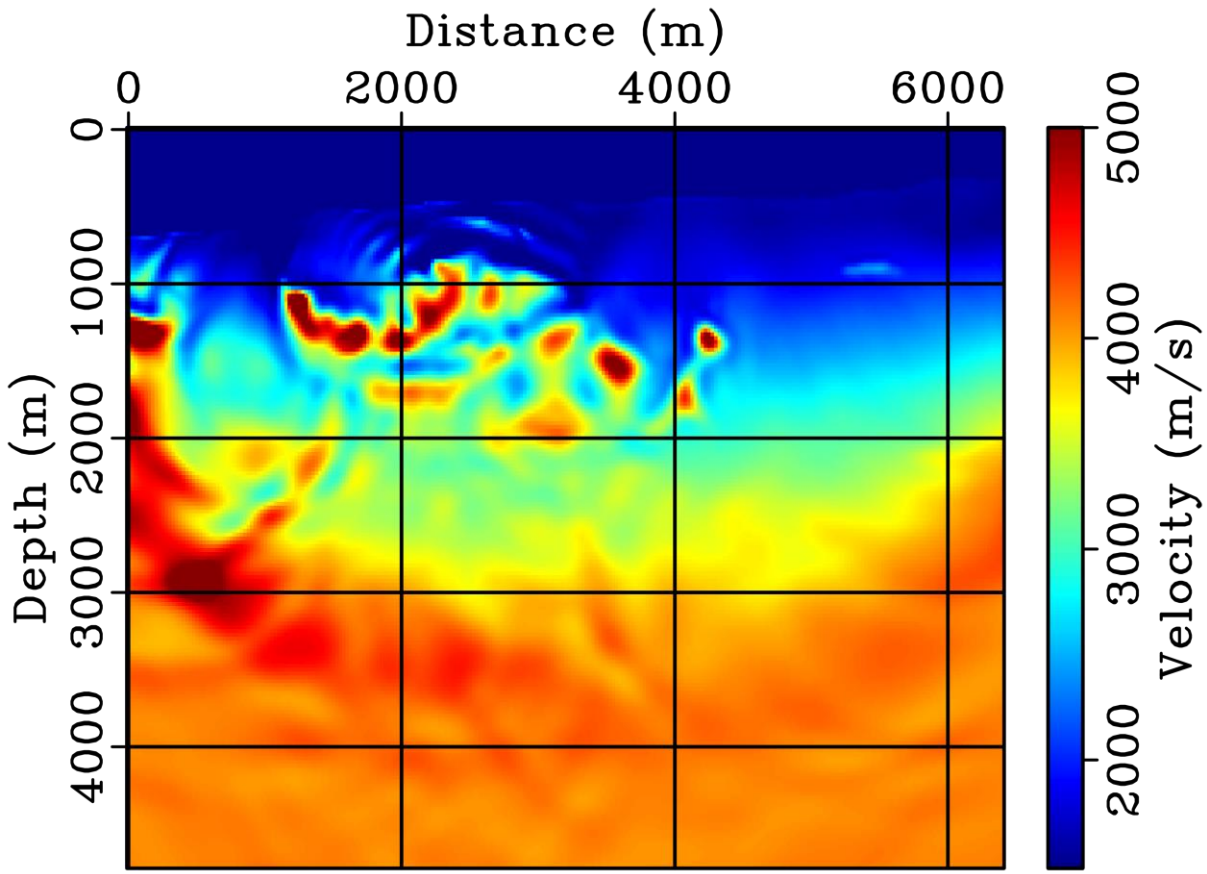
2-5 Hz

Conventional FWI



100

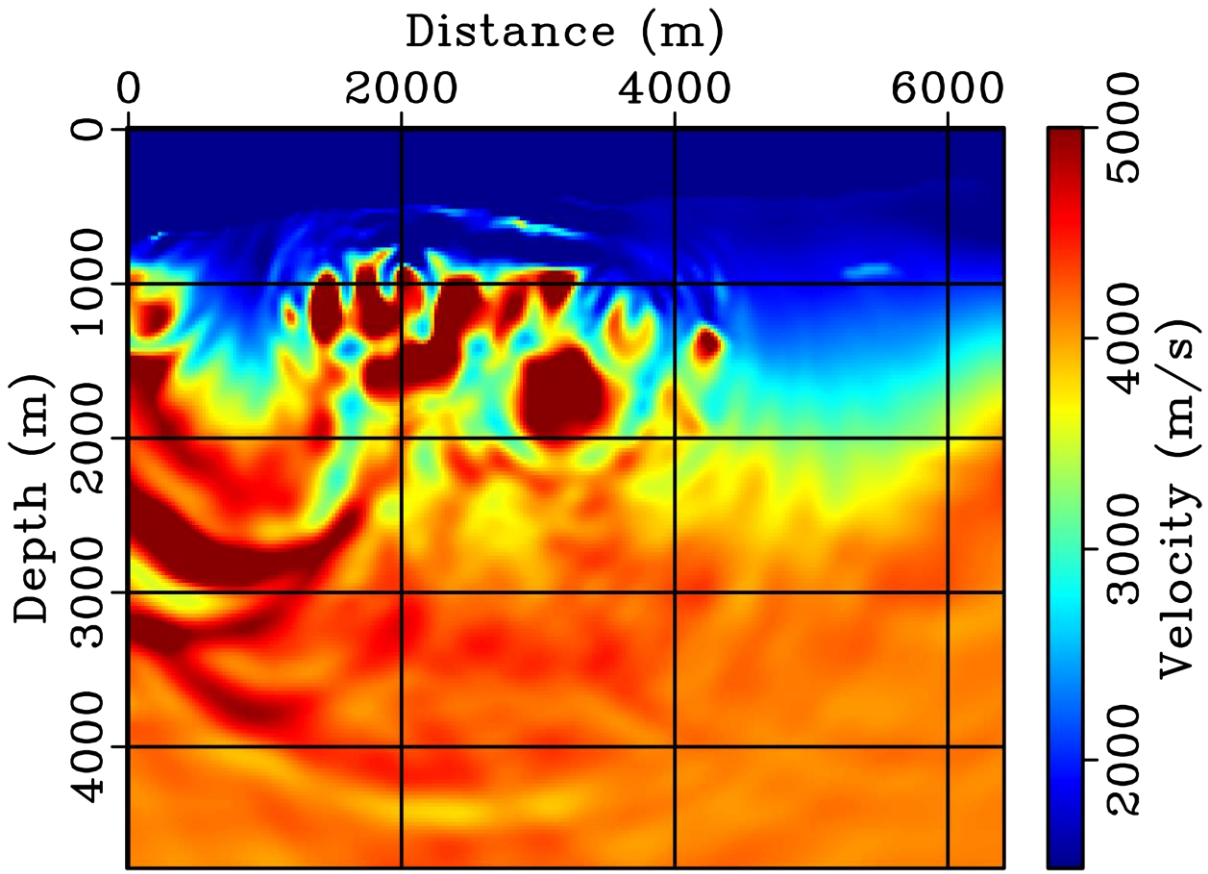
Conventional FWI initialized with GSA-FWI



100

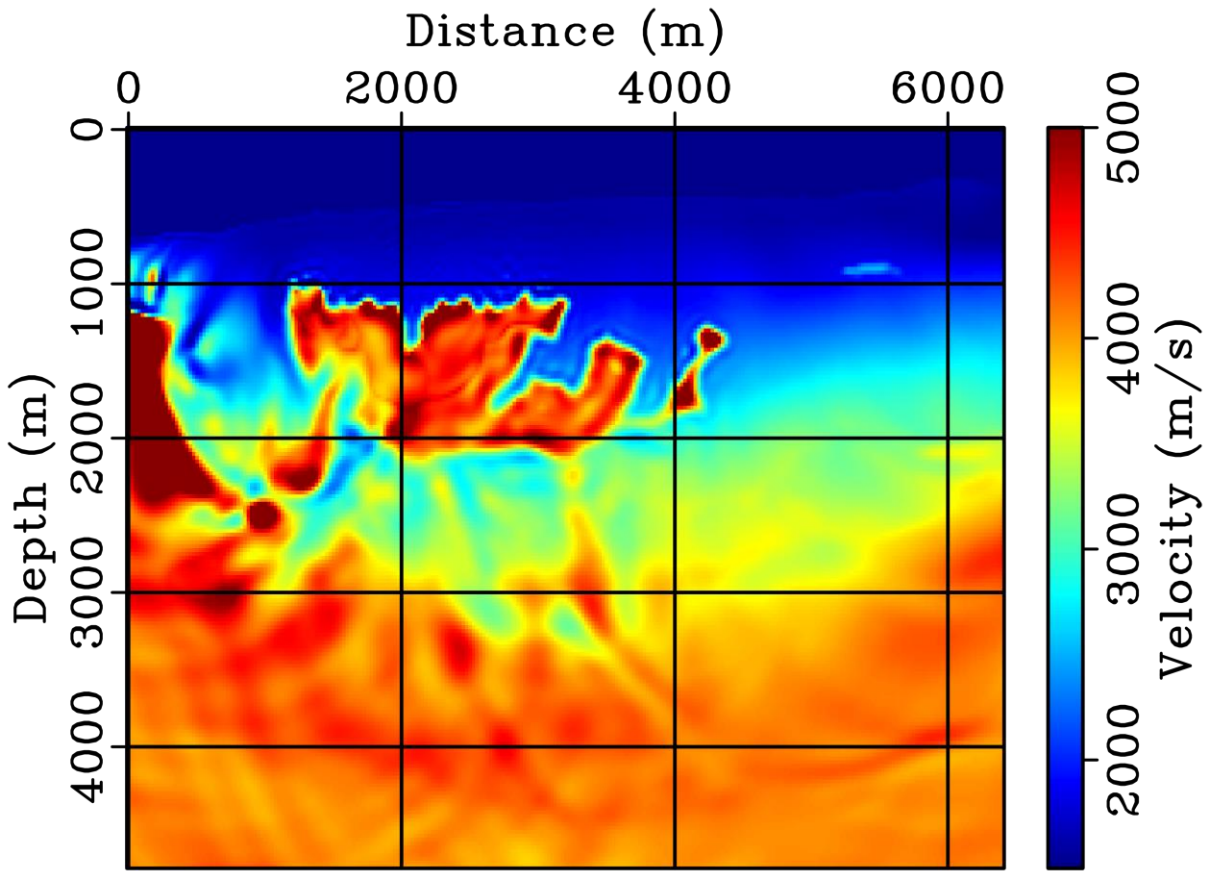
2-5 Hz

Conventional FWI



200

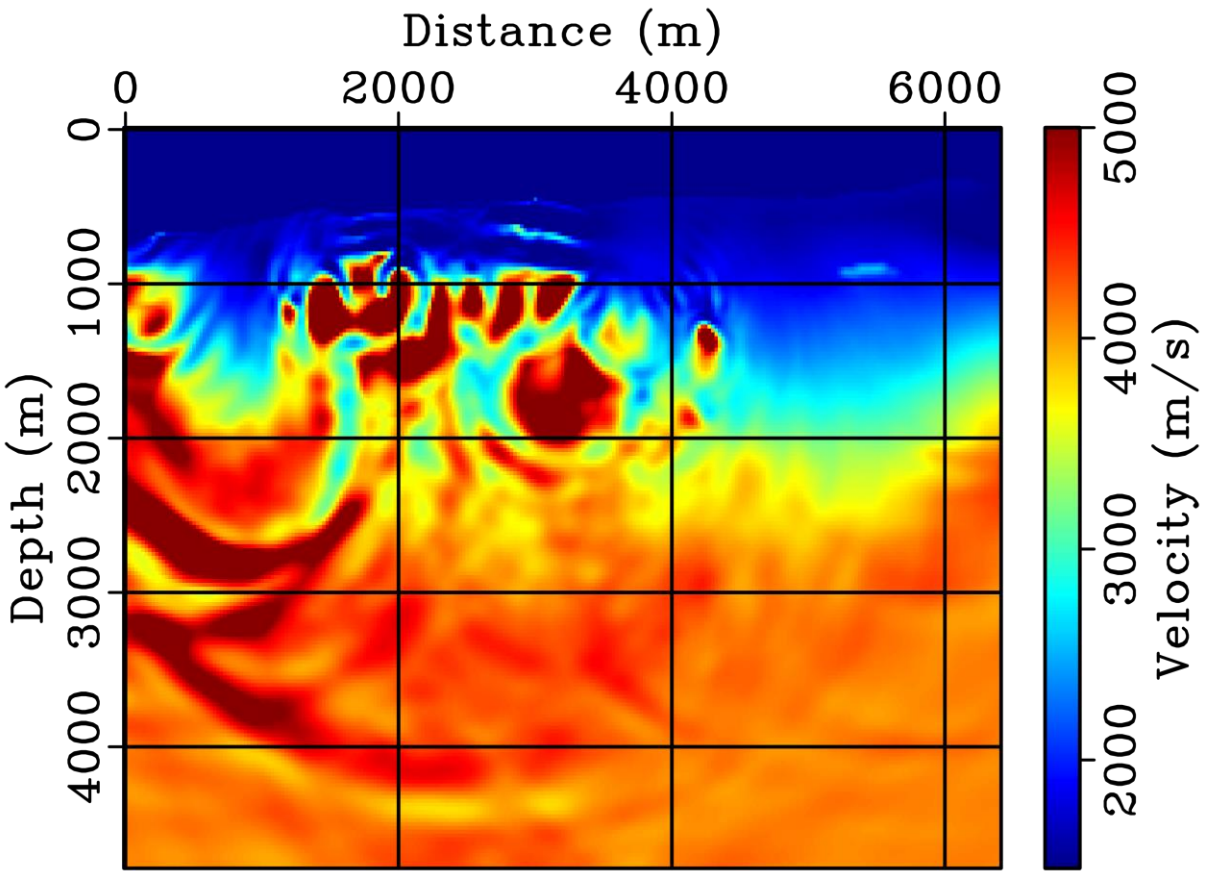
Conventional FWI initialized with GSA-FWI



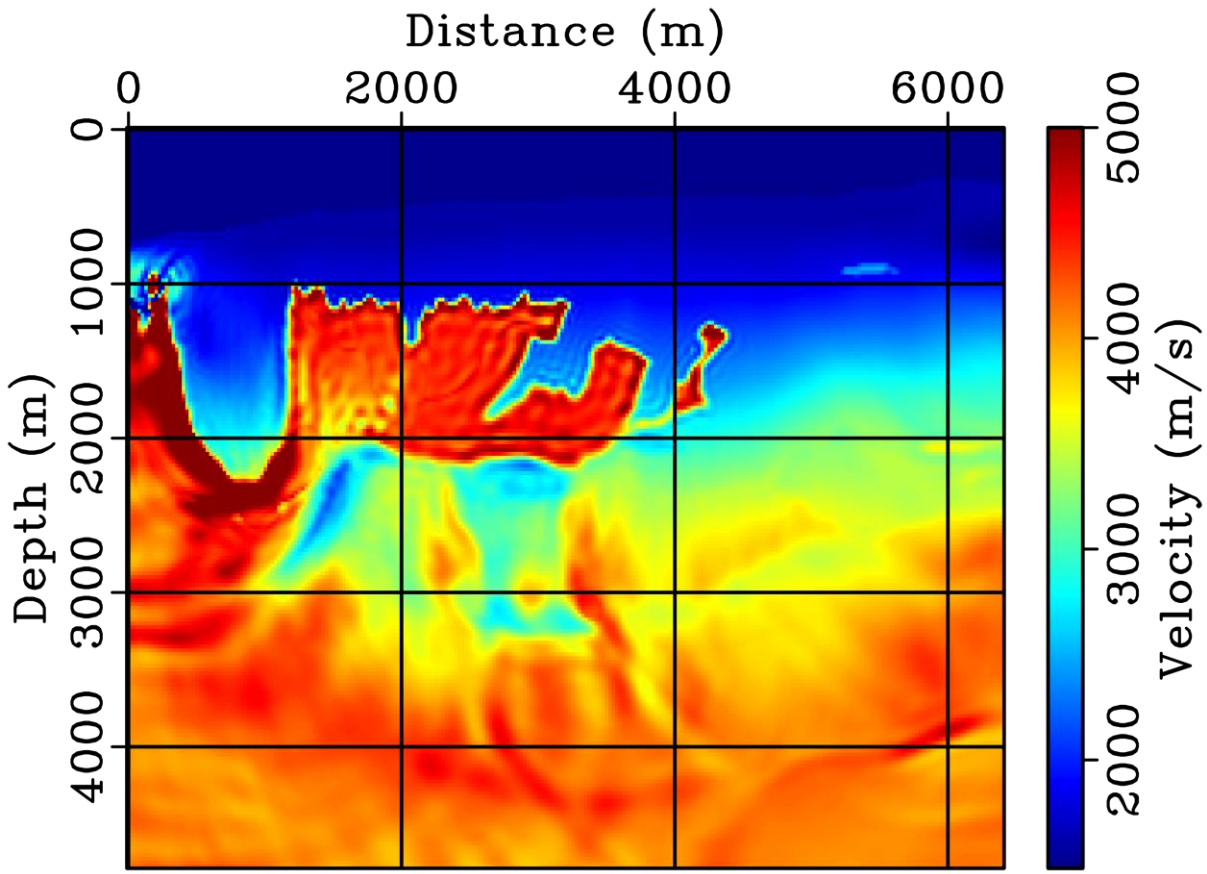
200

2-5 Hz

Conventional FWI

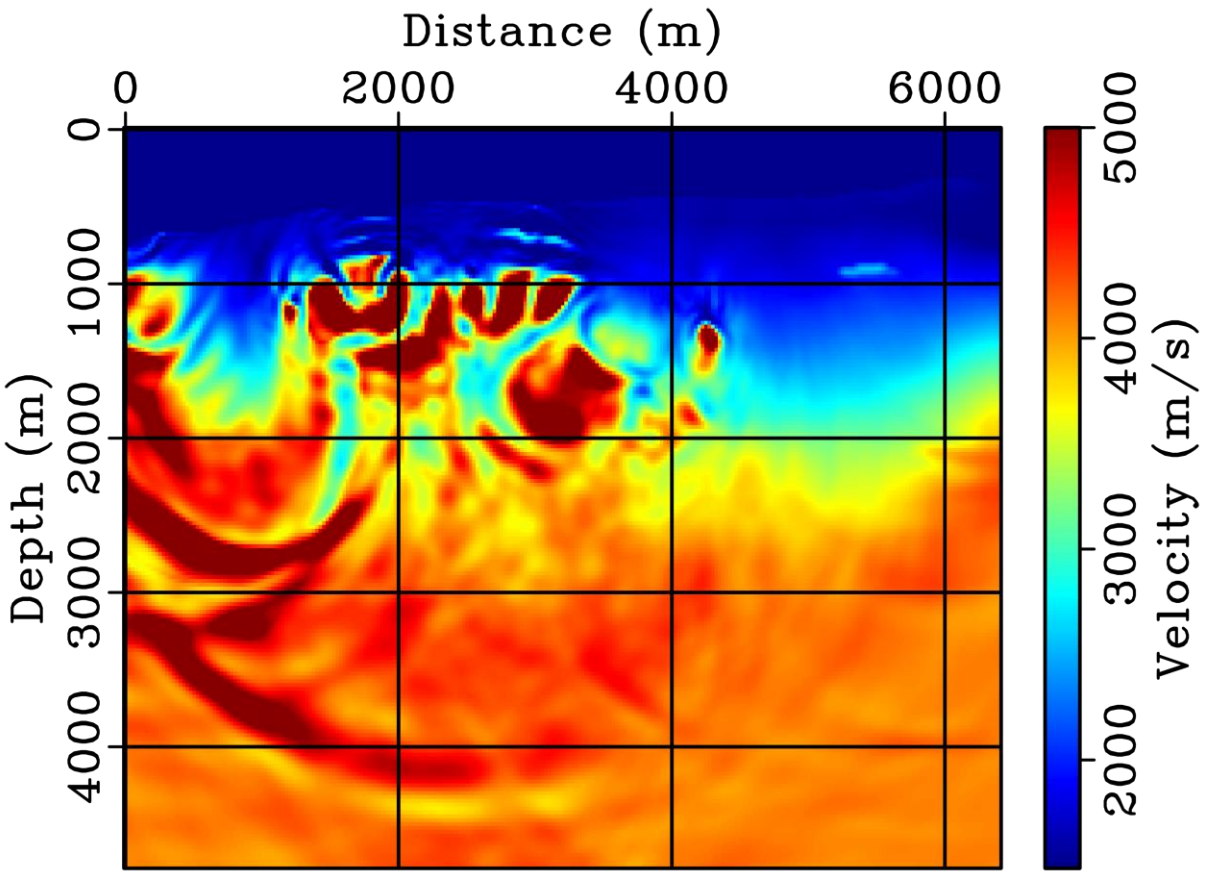


Conventional FWI initialized with GSA-FWI

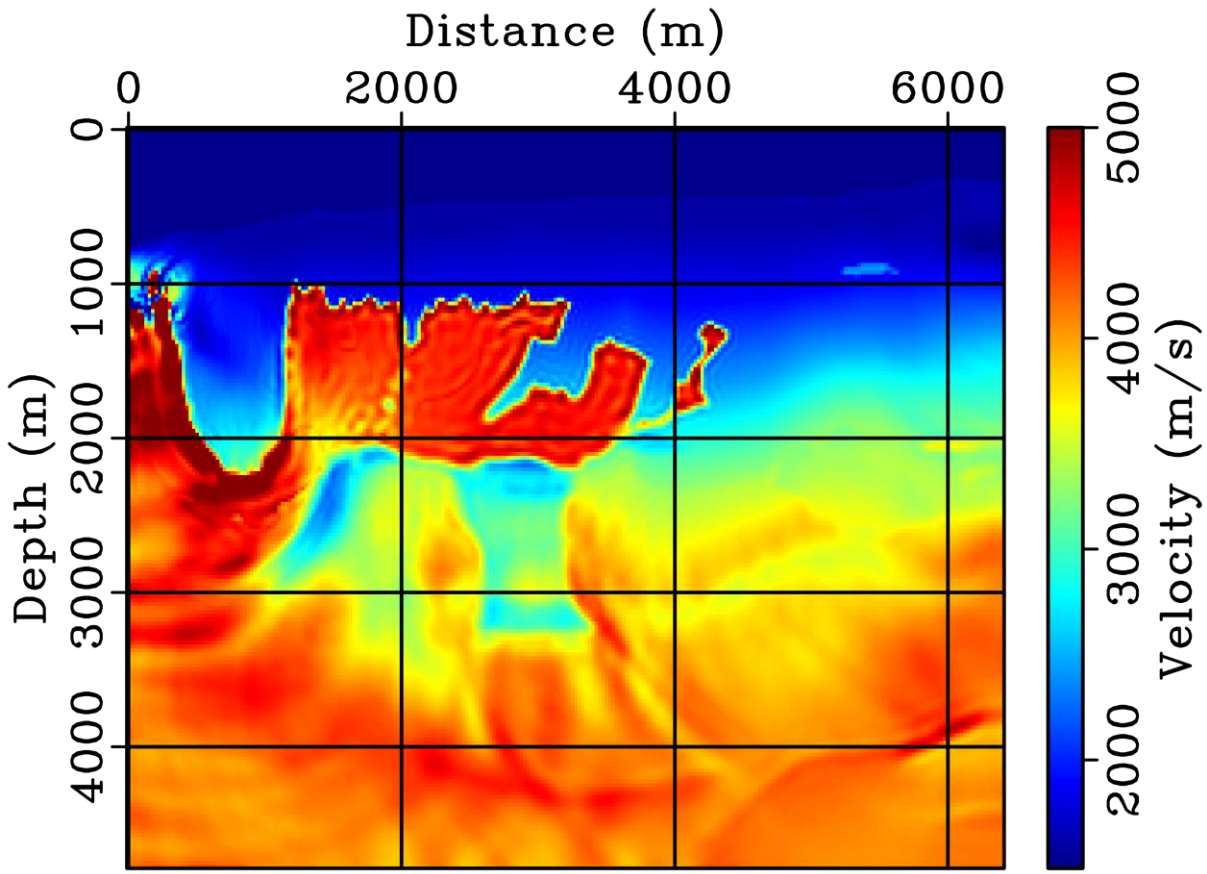


2-7 Hz

Conventional FWI

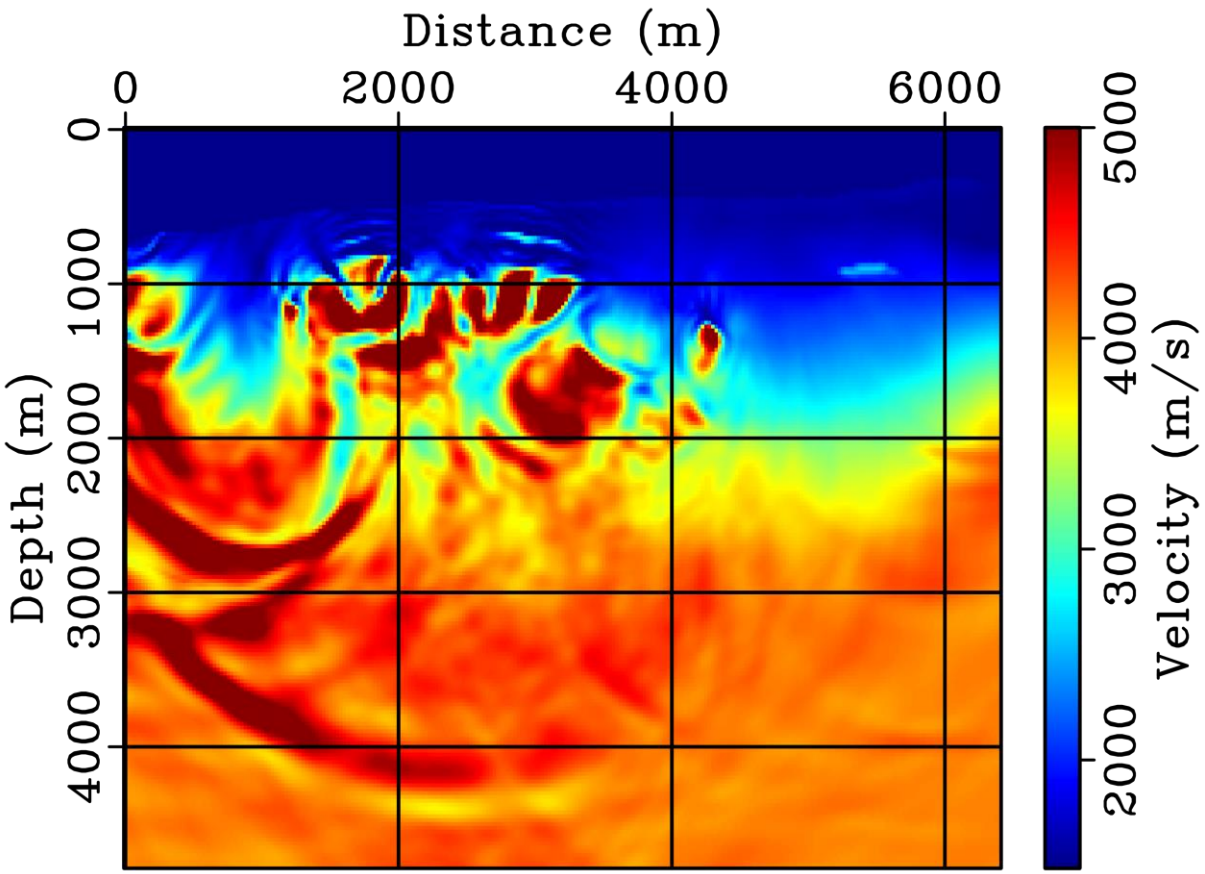


Conventional FWI initialized with GSA-FWI

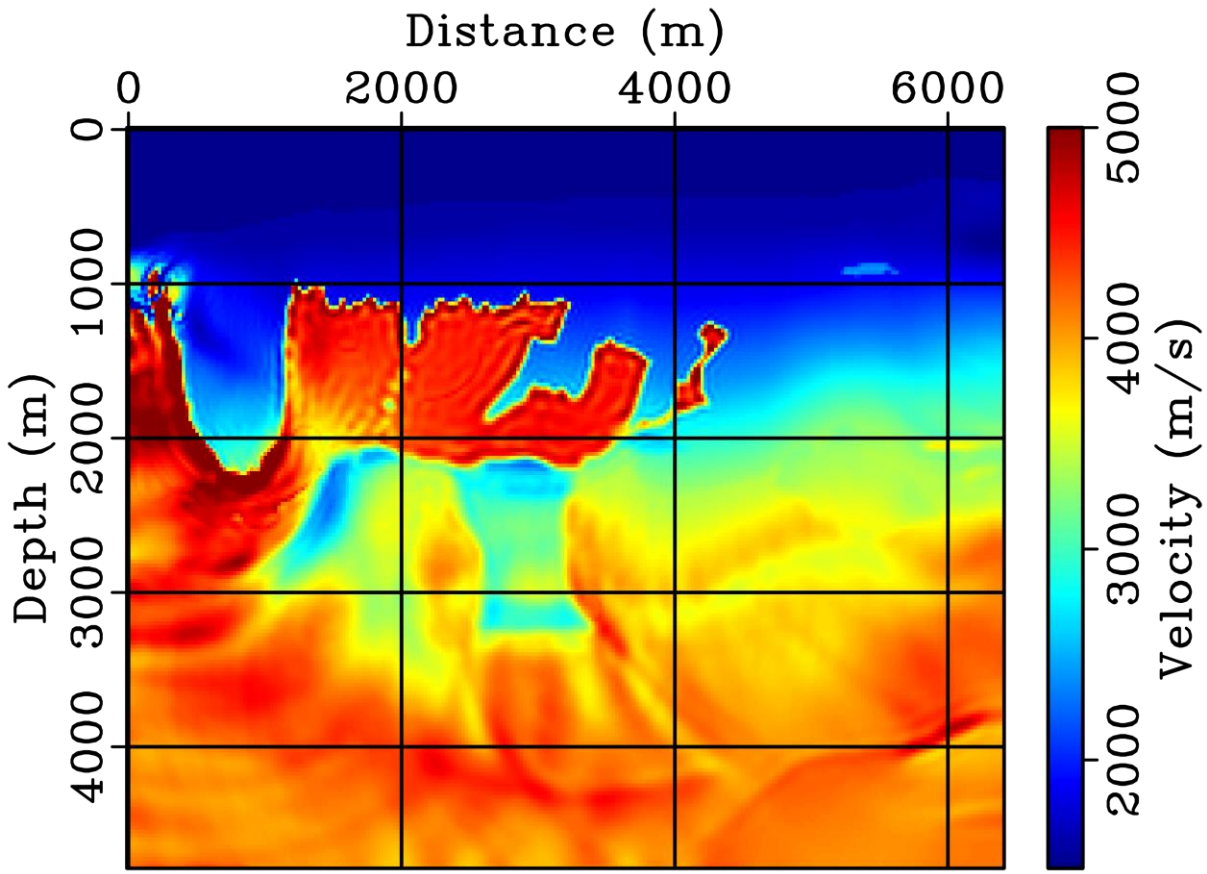


2-11 Hz

Conventional FWI

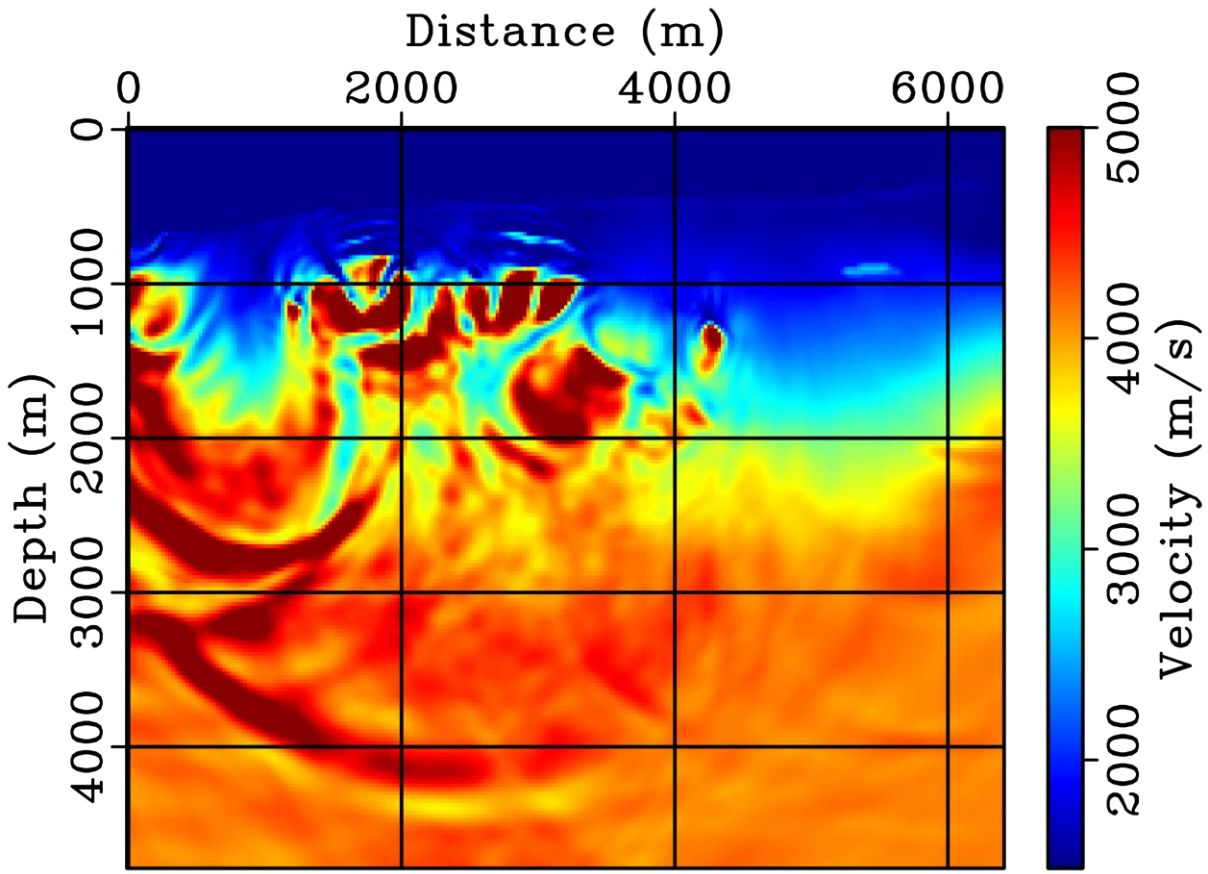


Conventional FWI initialized with GSA-FWI

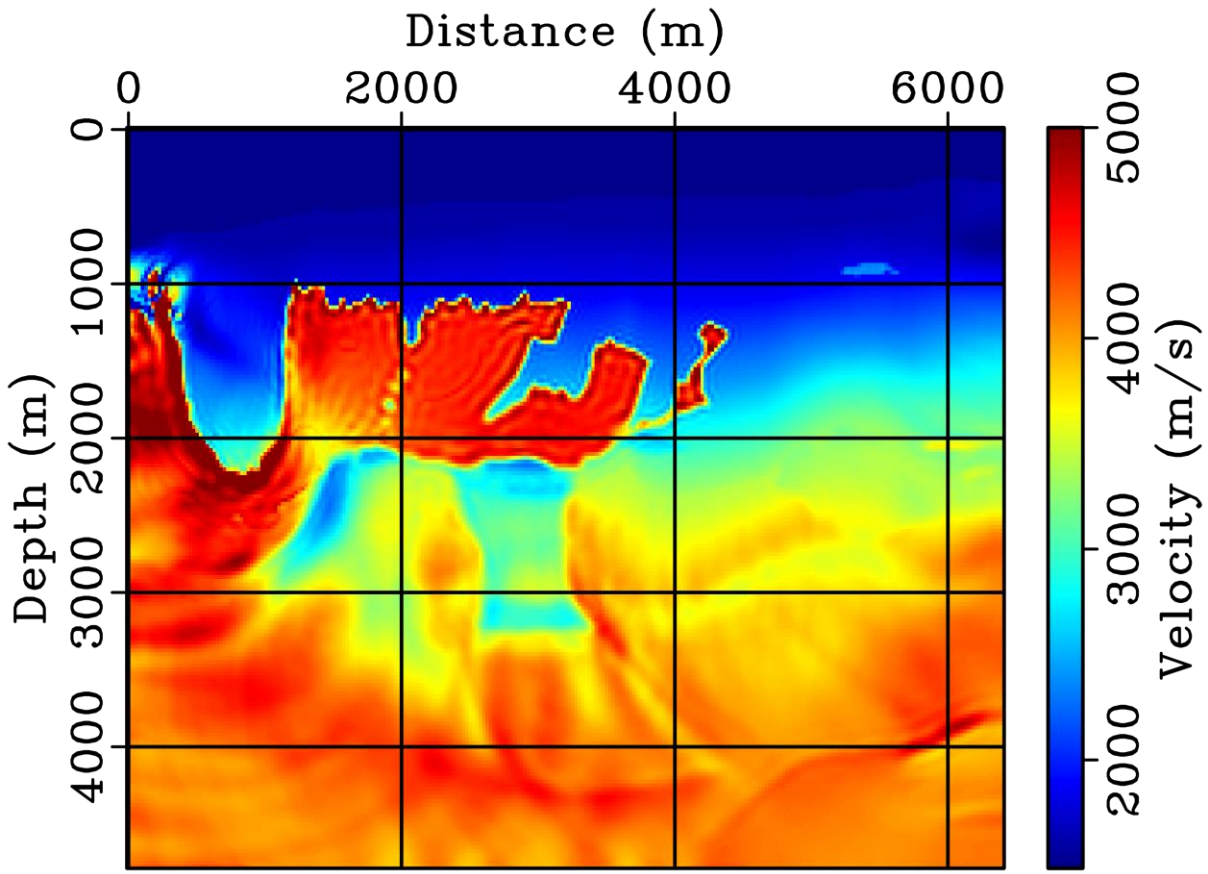


2-15 Hz

Conventional FWI

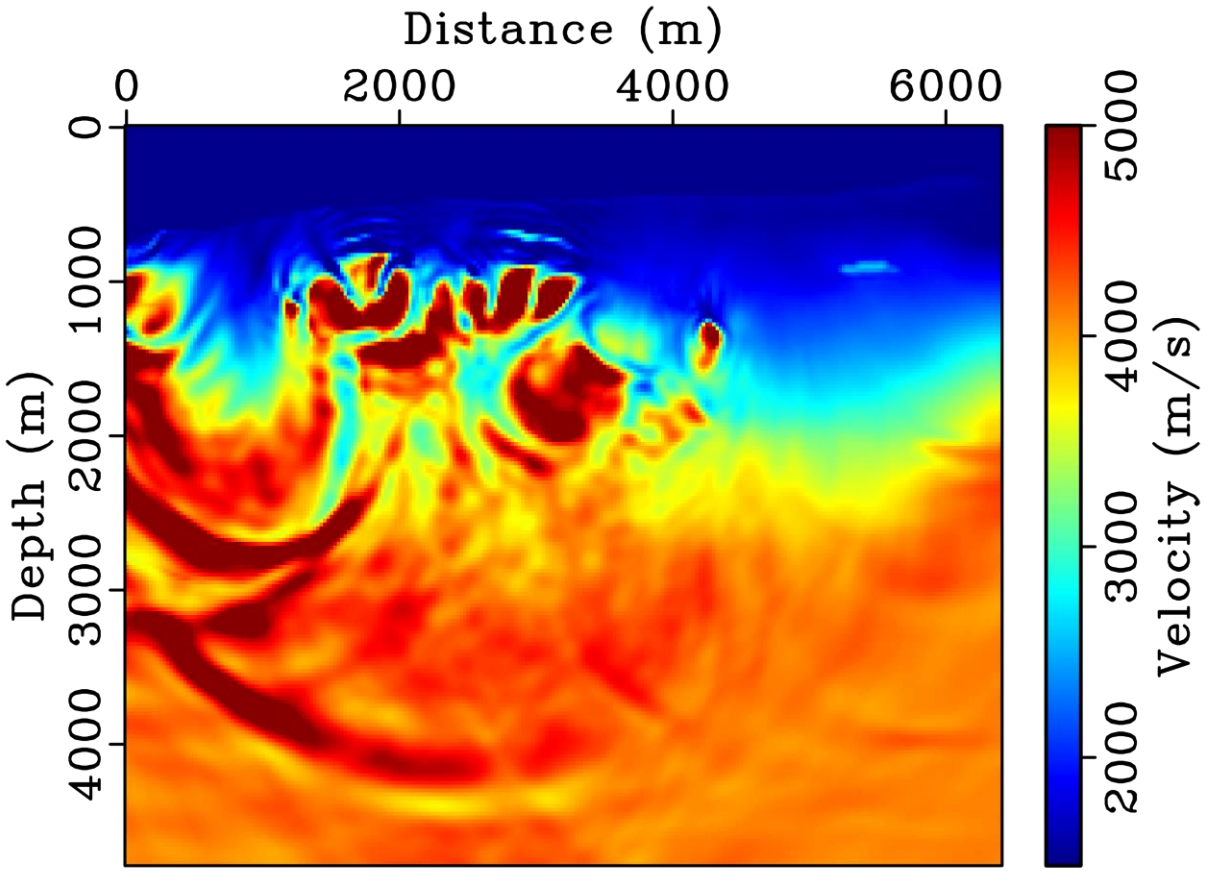


Conventional FWI initialized with GSA-FWI



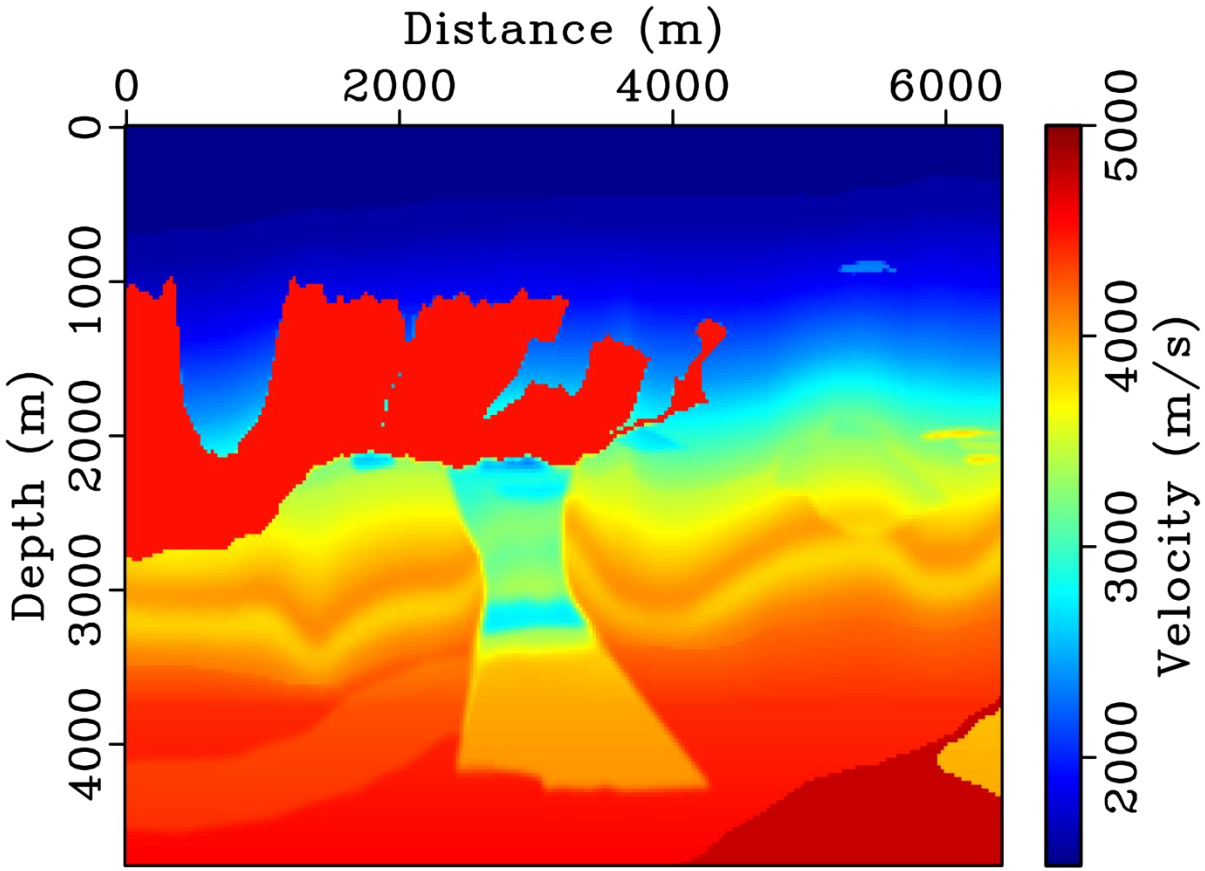
2-19 Hz

Conventional FWI

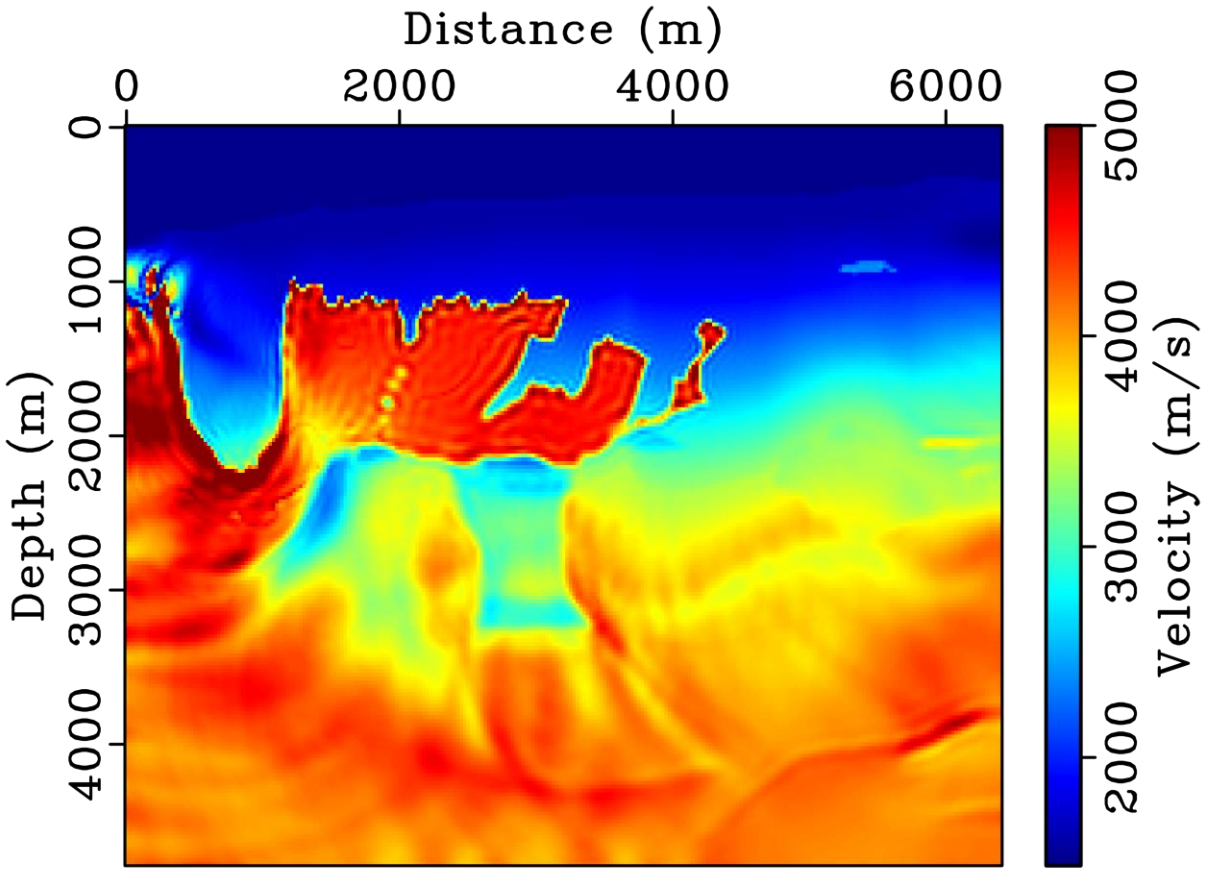


All frequency
2-19 Hz

True velocity

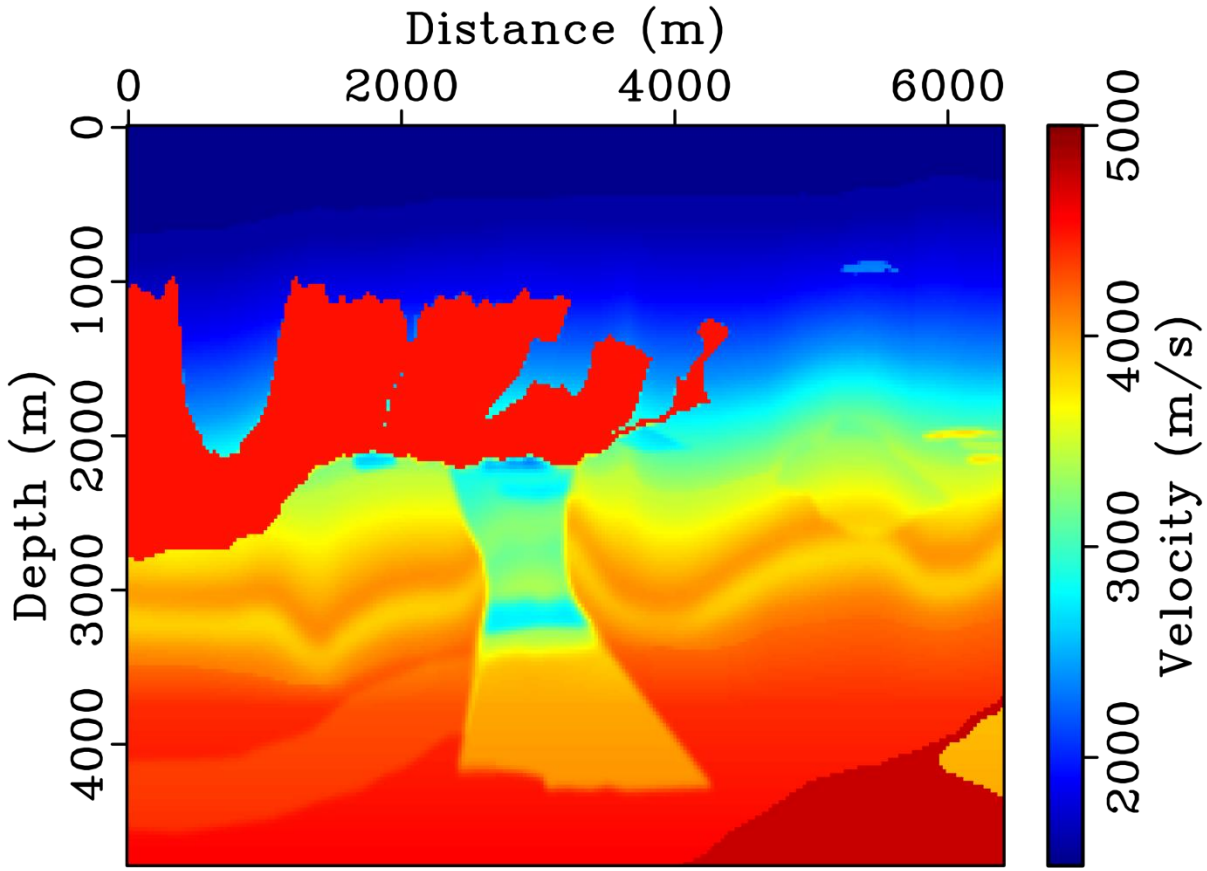


Conventional FWI initialized with GSA-FWI



All frequency
2-19 Hz

True velocity



Outline

- Introduction
- Methodology
- Numerical Examples
  - 2004 BP model
- Conclusions

Conclusions

- GSA-based FWI speeds up the convergence, and is hence less sensitive to the cycle-skipping problem
- We proposed an approximated sampling scheme, which makes the resulting GSA-based FWI as efficient as conventional FWI
- The proposed GSA-based FWI is robust when the seismic data contains abundant diving waves and refraction waves

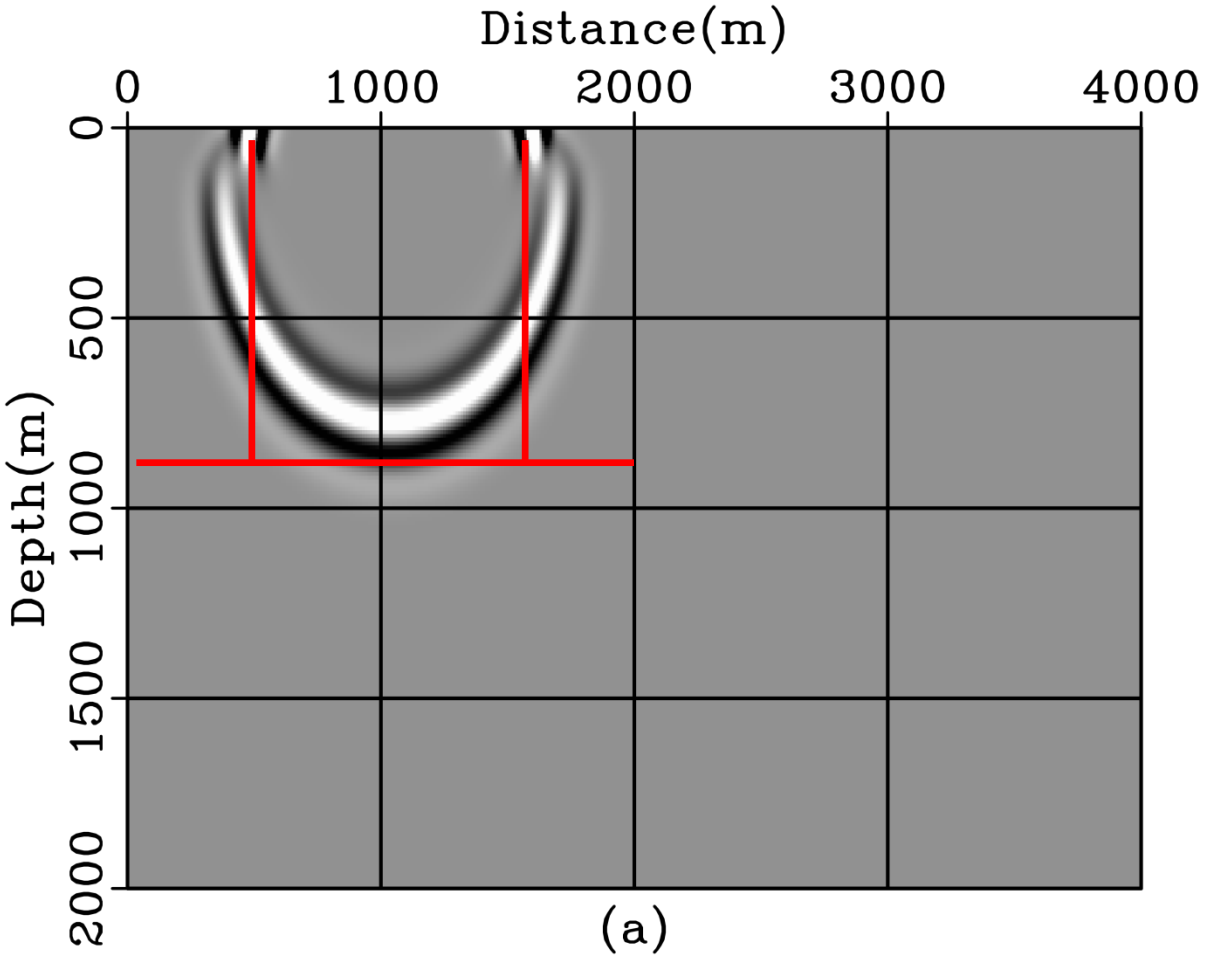
Discussion

- **GSA-based FWI did not overcome the non-convexity of FWI**
  - ➢ Numerical tests starting from 5Hz did not yield significant results
- **Further study is needed on an optimal strategy of space shift**
  - ➢ Numerical tests with different shift strategies show similar, yet different results
- **Applicability of GSA-based FWI to reflection-dominated seismic data**
  - ➢ Numerical tests on transmission-dominated data show better results

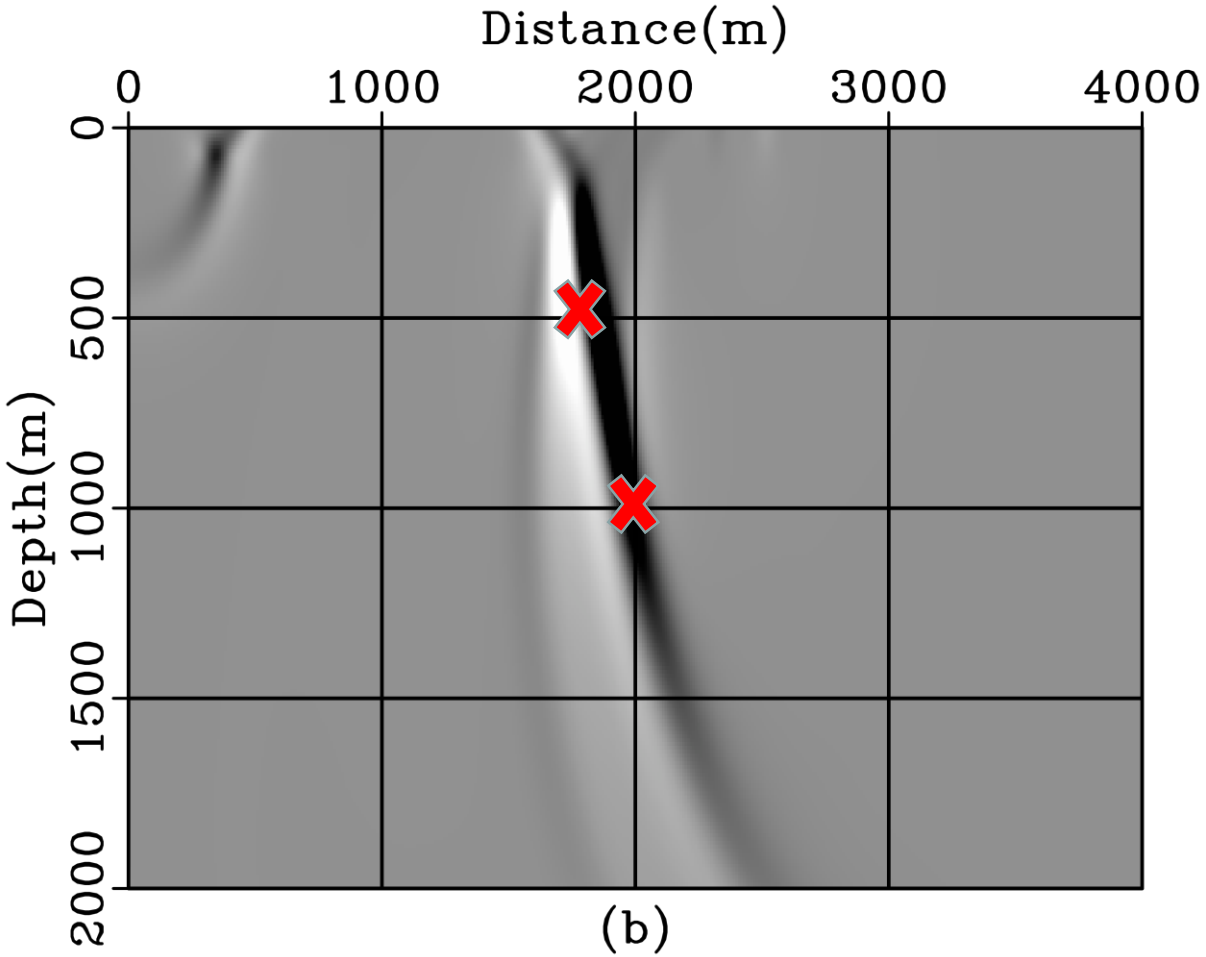
NUS National University of Singapore

THANKS FOR YOUR ATTENTION !

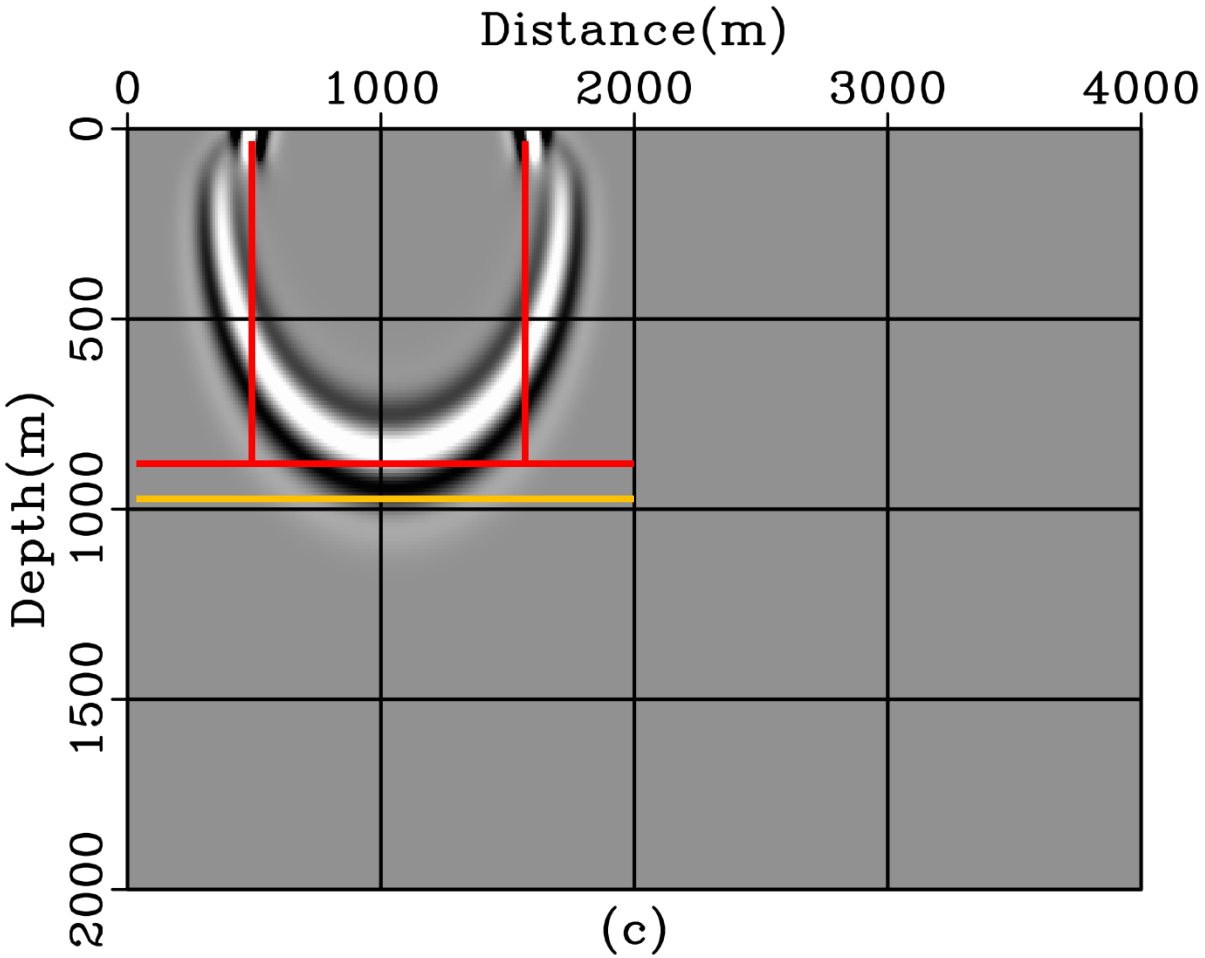
Forward



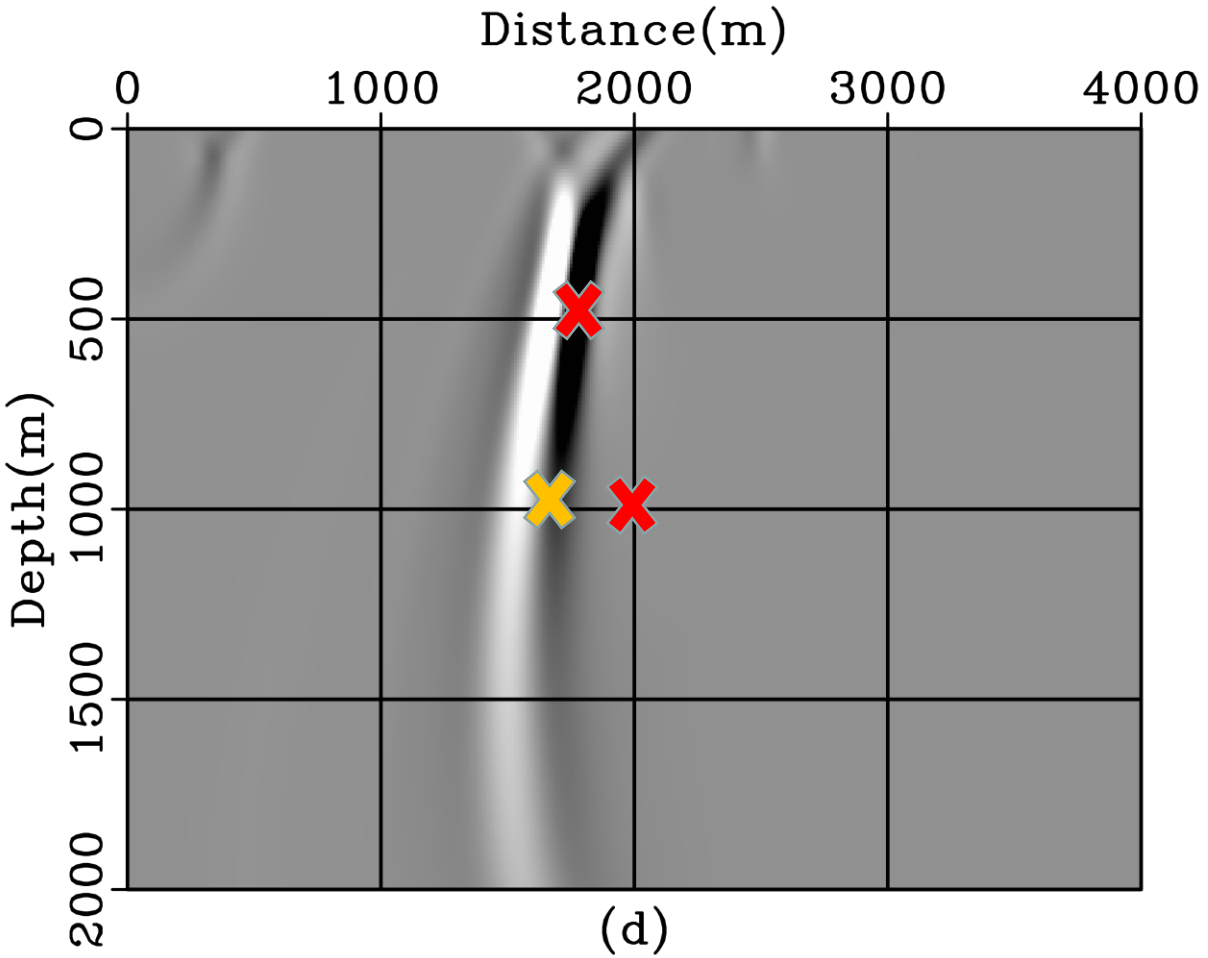
Backward

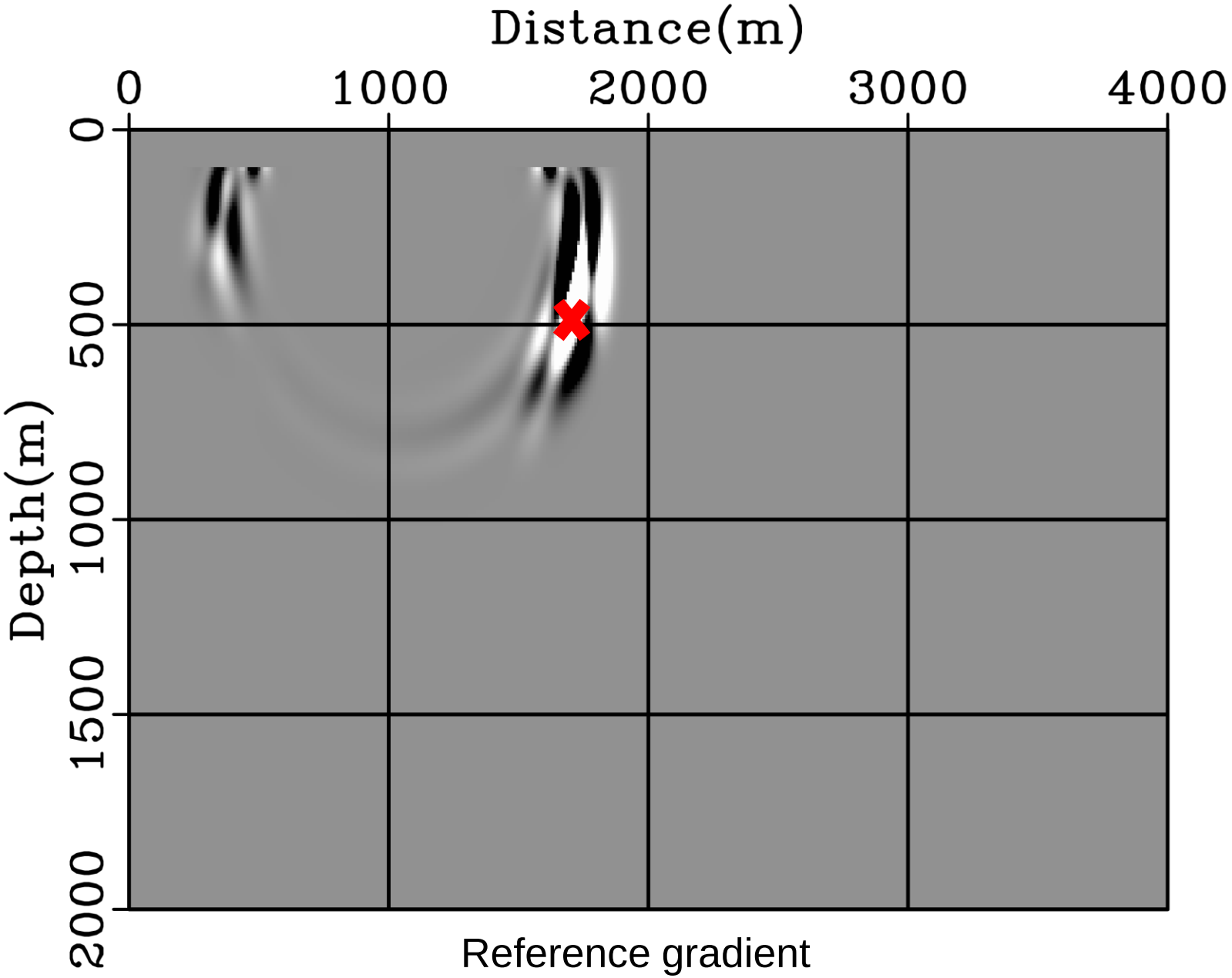


Forward

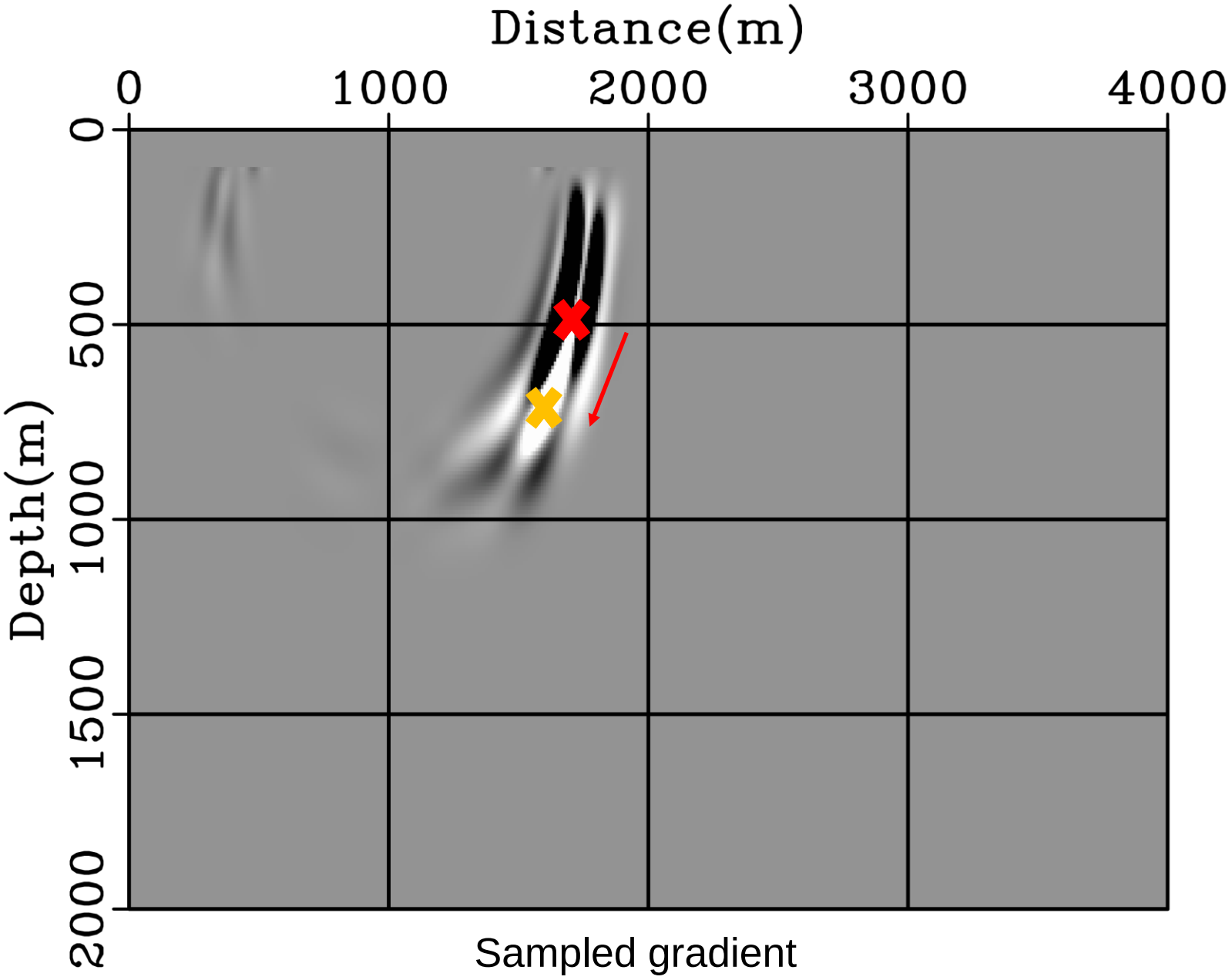


Backward



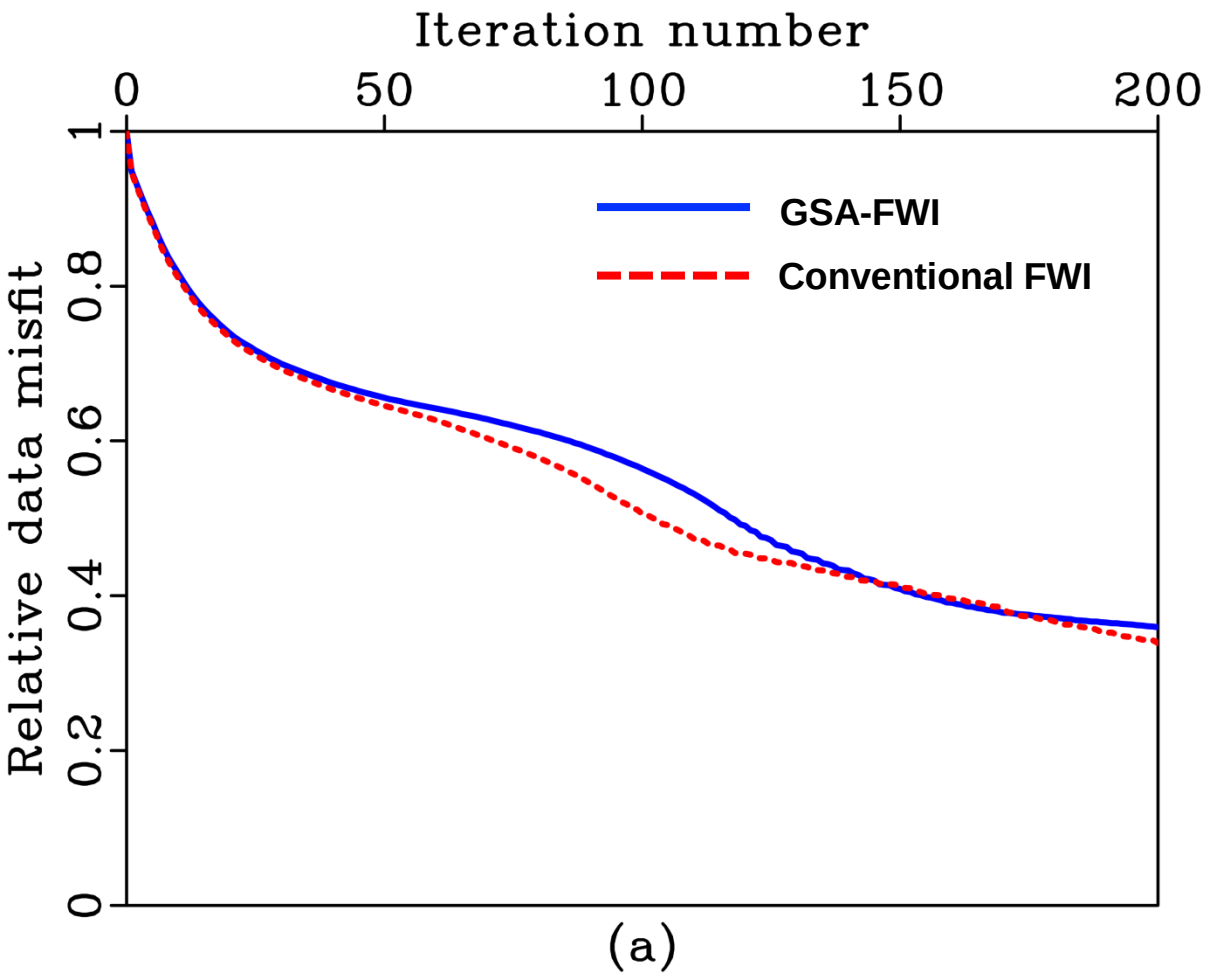


Gradient calculated using the reference velocity model at t=0.5s

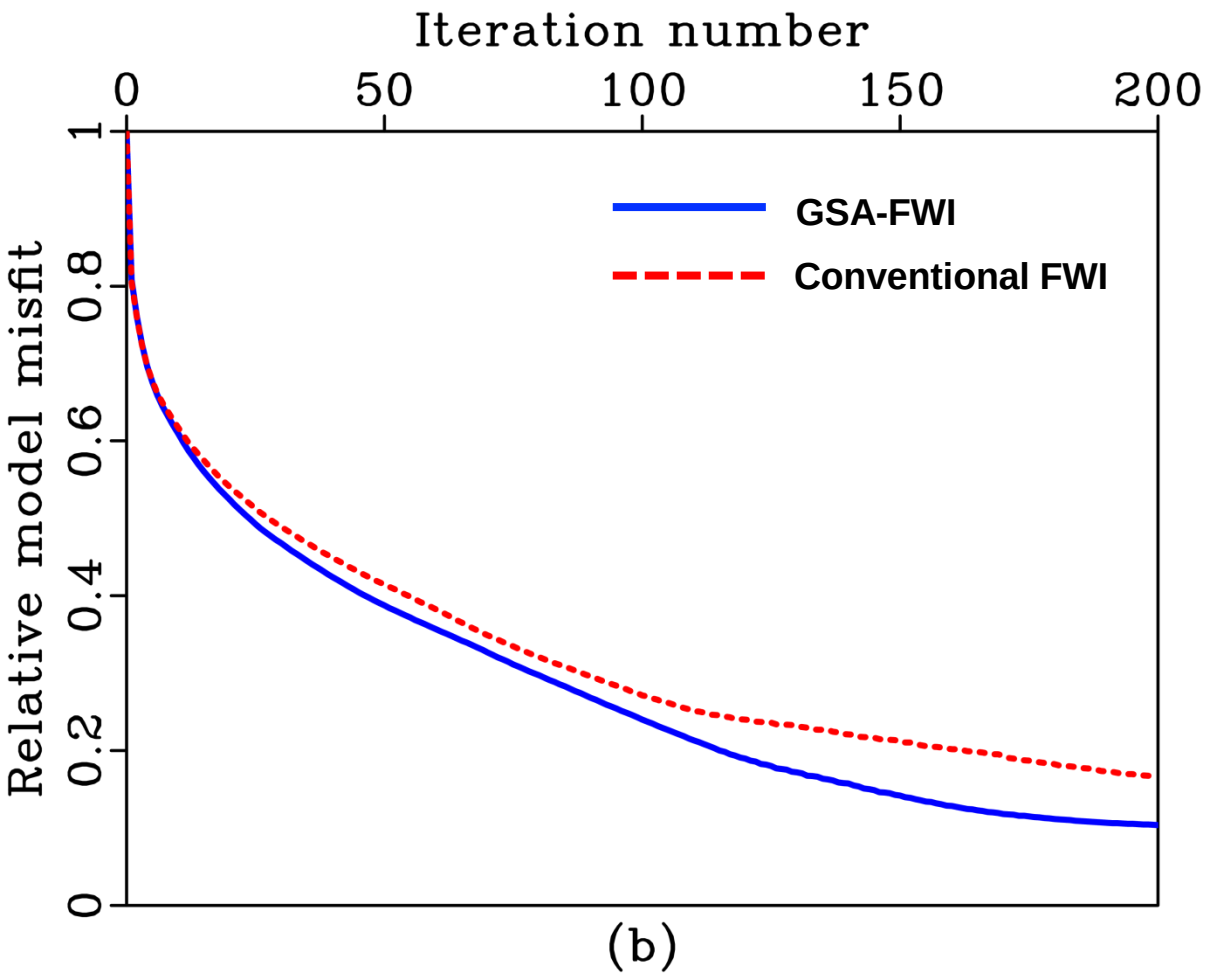


Gradient calculated using the sampled velocity model at t=0.5s

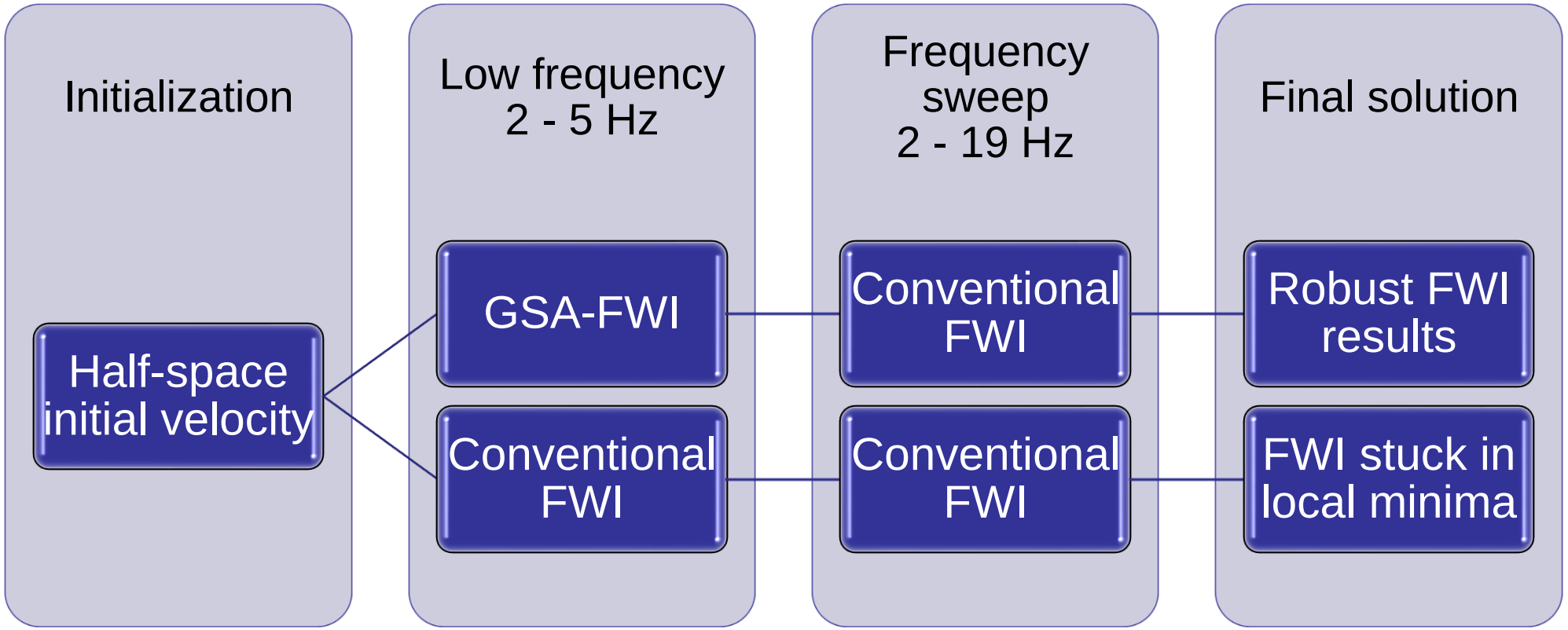
Data misfit



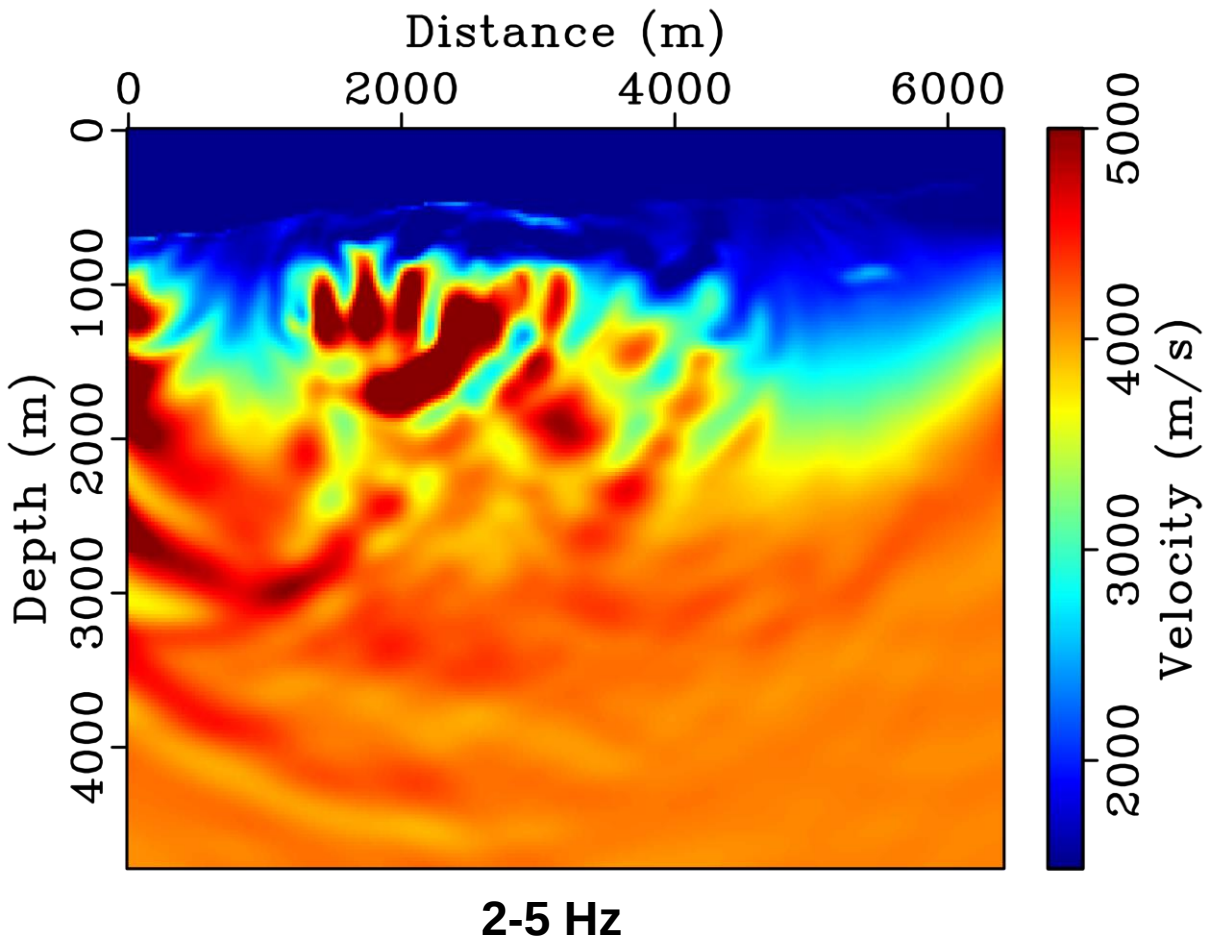
Model misfit



Workflow

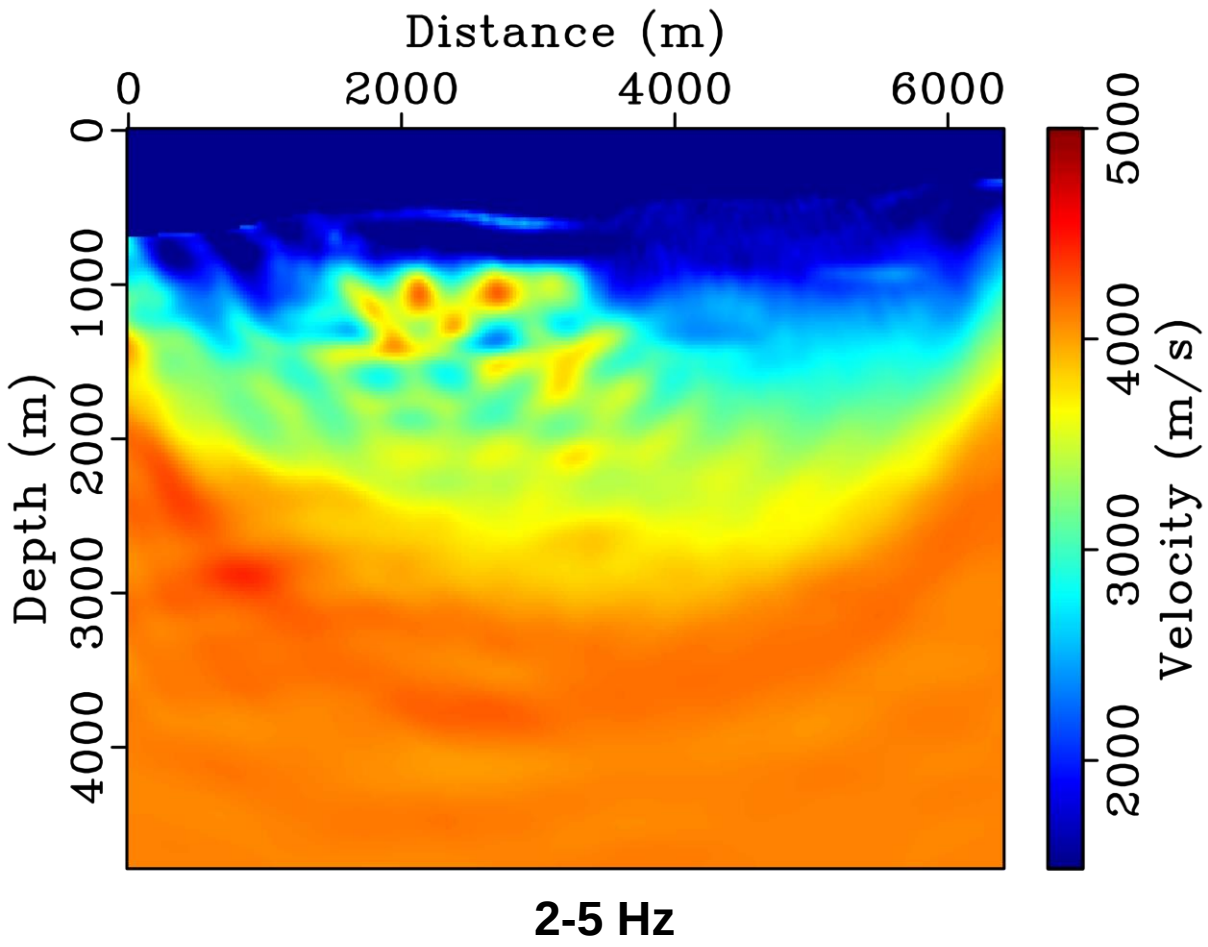

Initialization
Half-space initial velocity
Low frequency 2 - 5 Hz
GSA-FWI
Conventional FWI
Frequency sweep 2 - 19 Hz
Conventional FWI
Conventional FWI
Final solution
Robust FWI results
FWI stuck in local minima

Conventional FWI



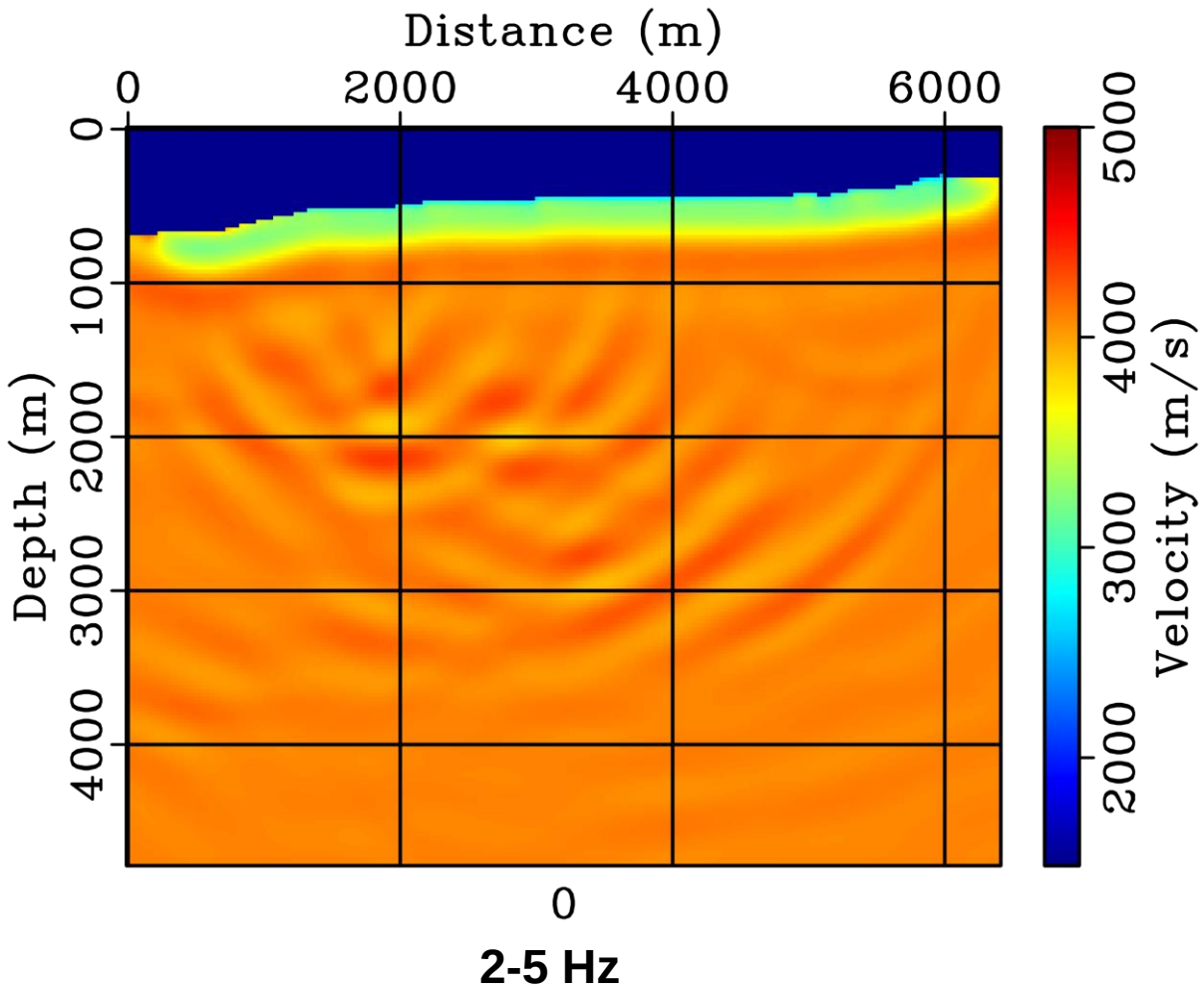
2-5 Hz

GSA-FWI

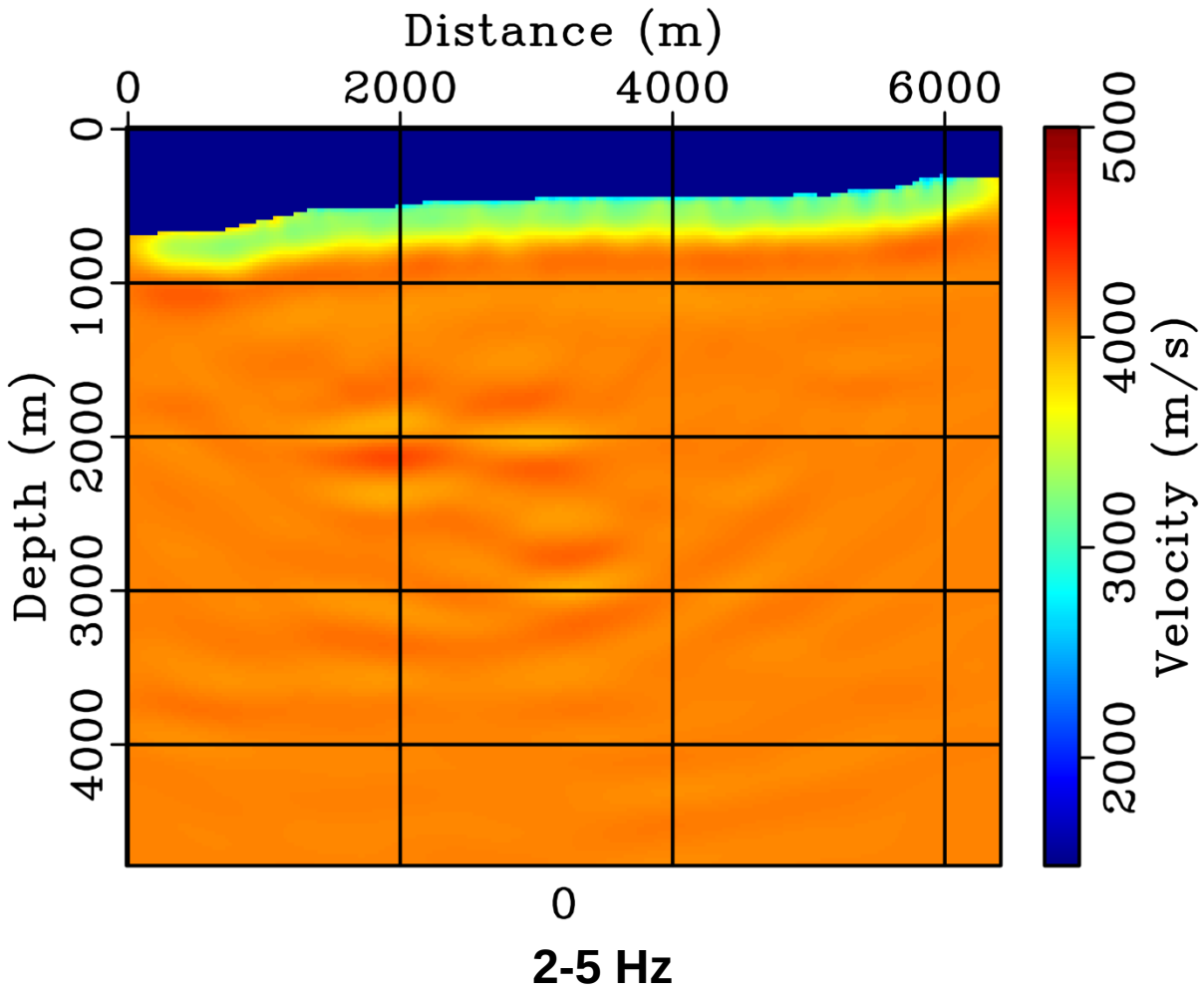


2-5 Hz

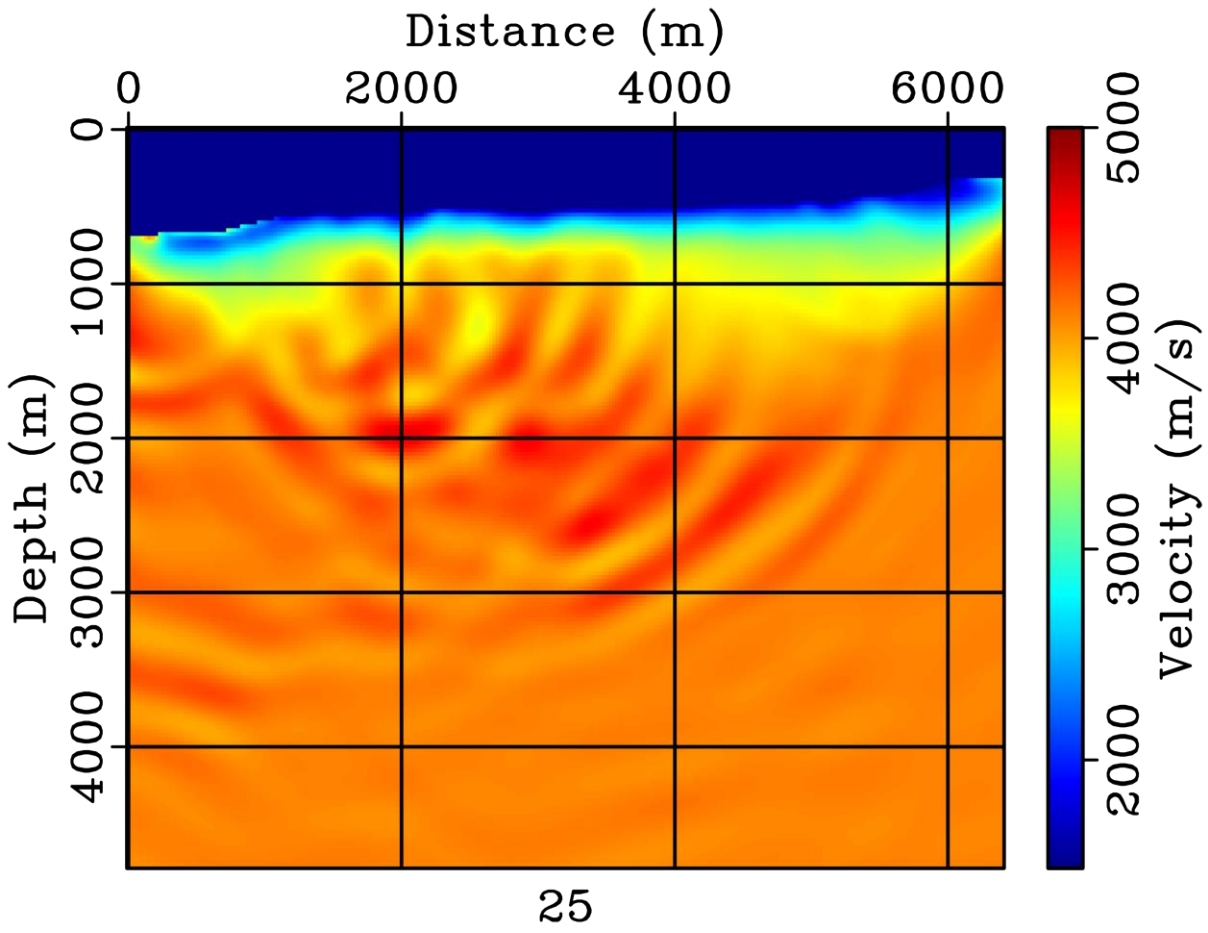
Conventional FWI



GSA-FWI

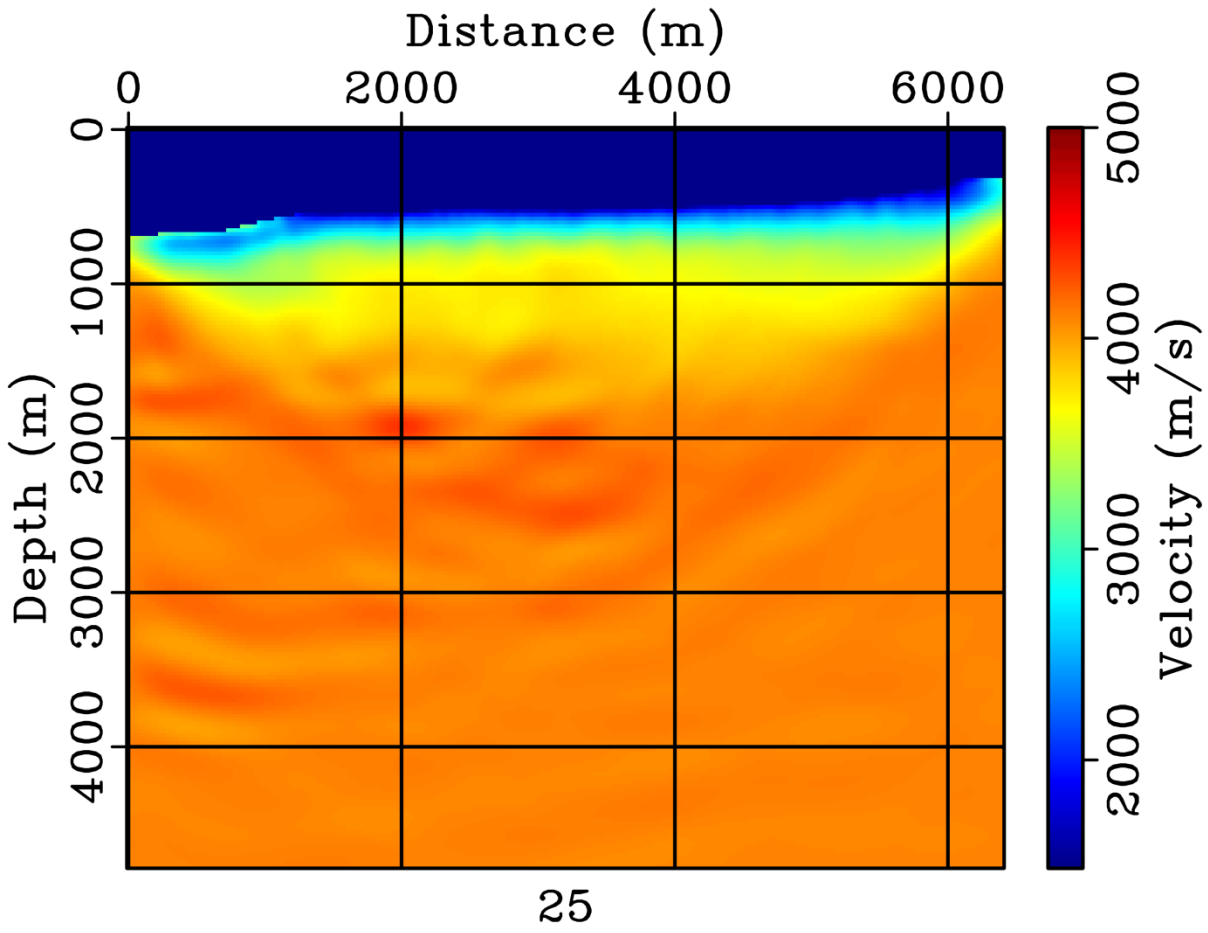


Conventional FWI



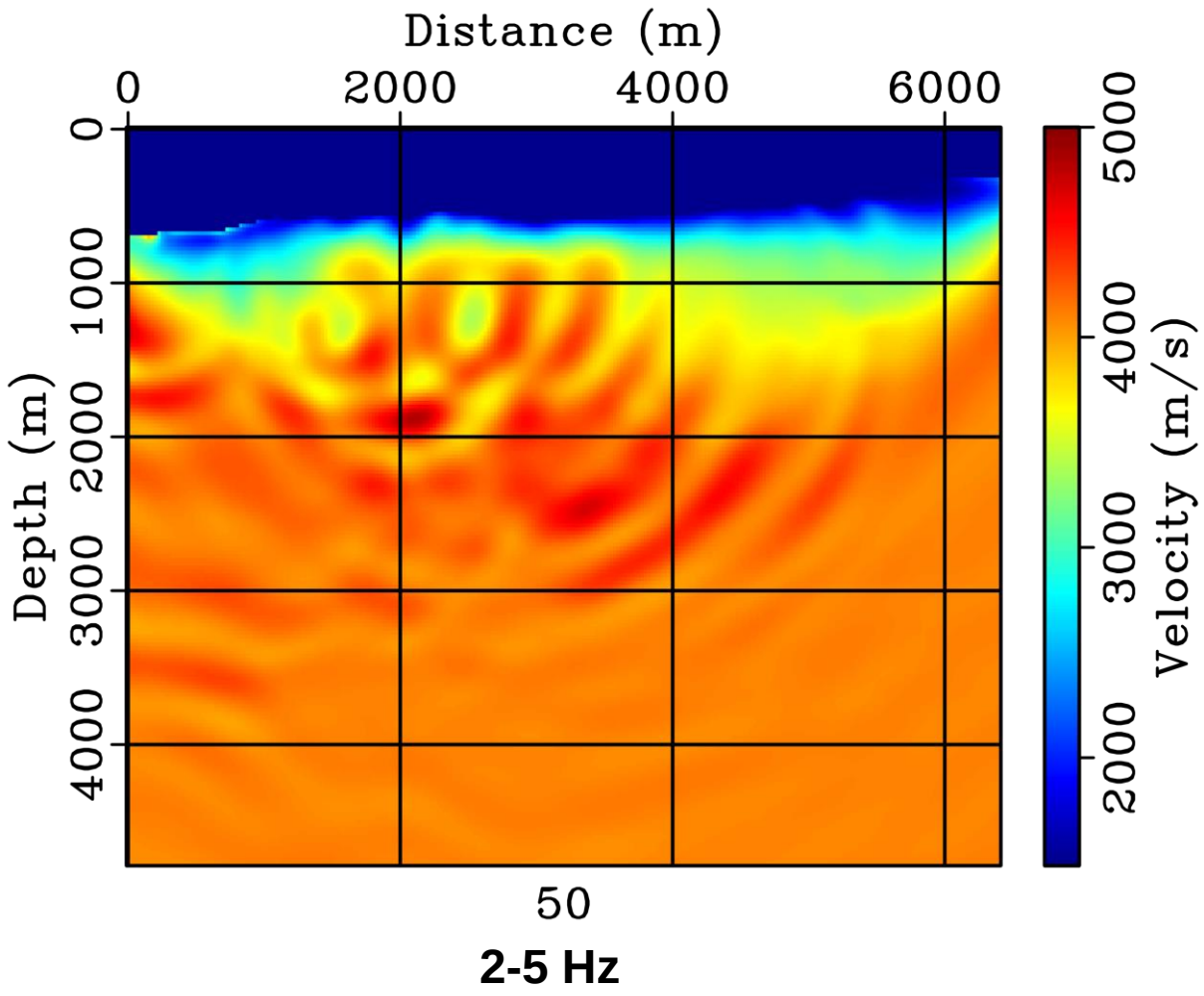
Low frequency
2-5 Hz

GSA-FWI



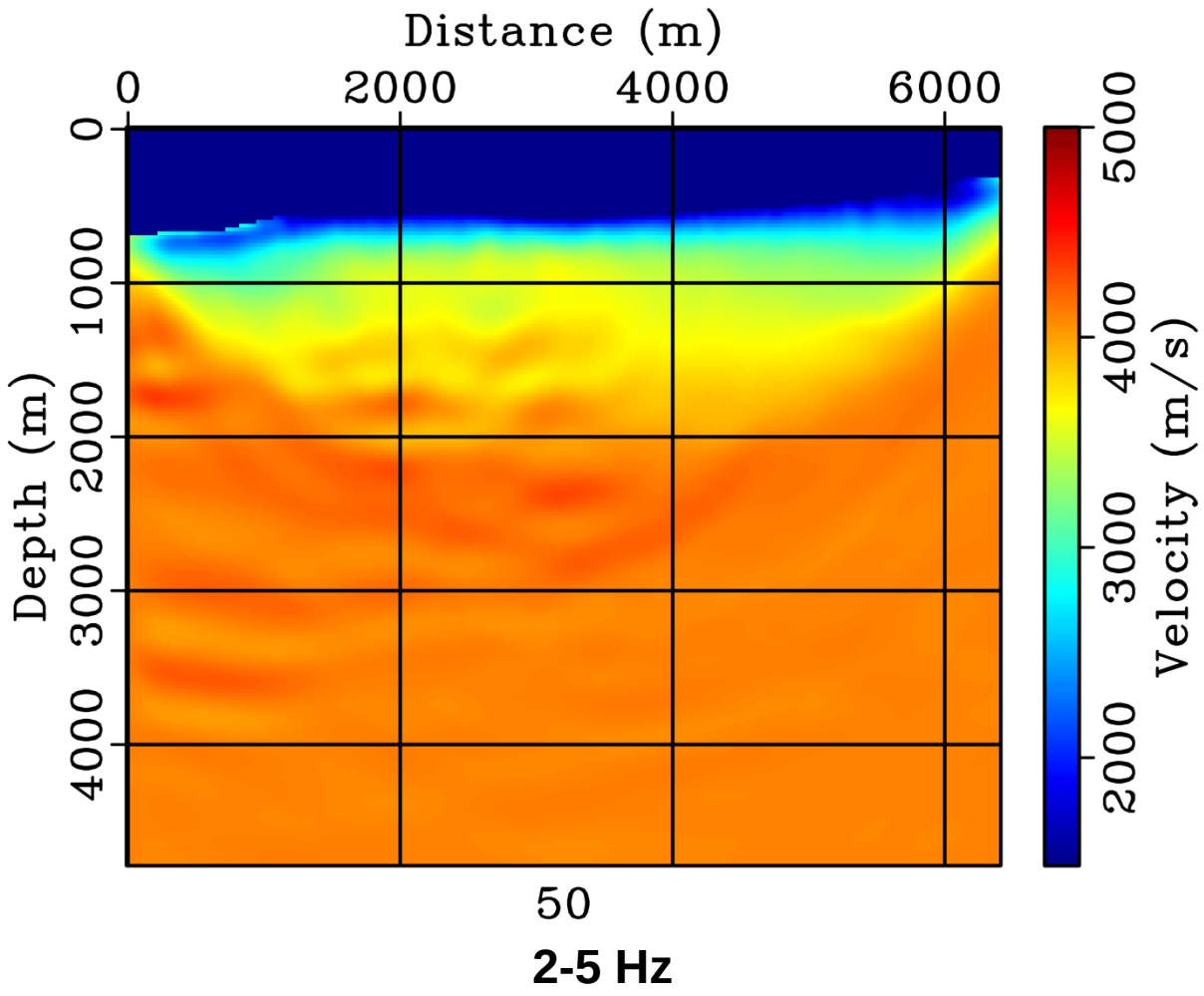
Low frequency
2-5 Hz

Conventional FWI



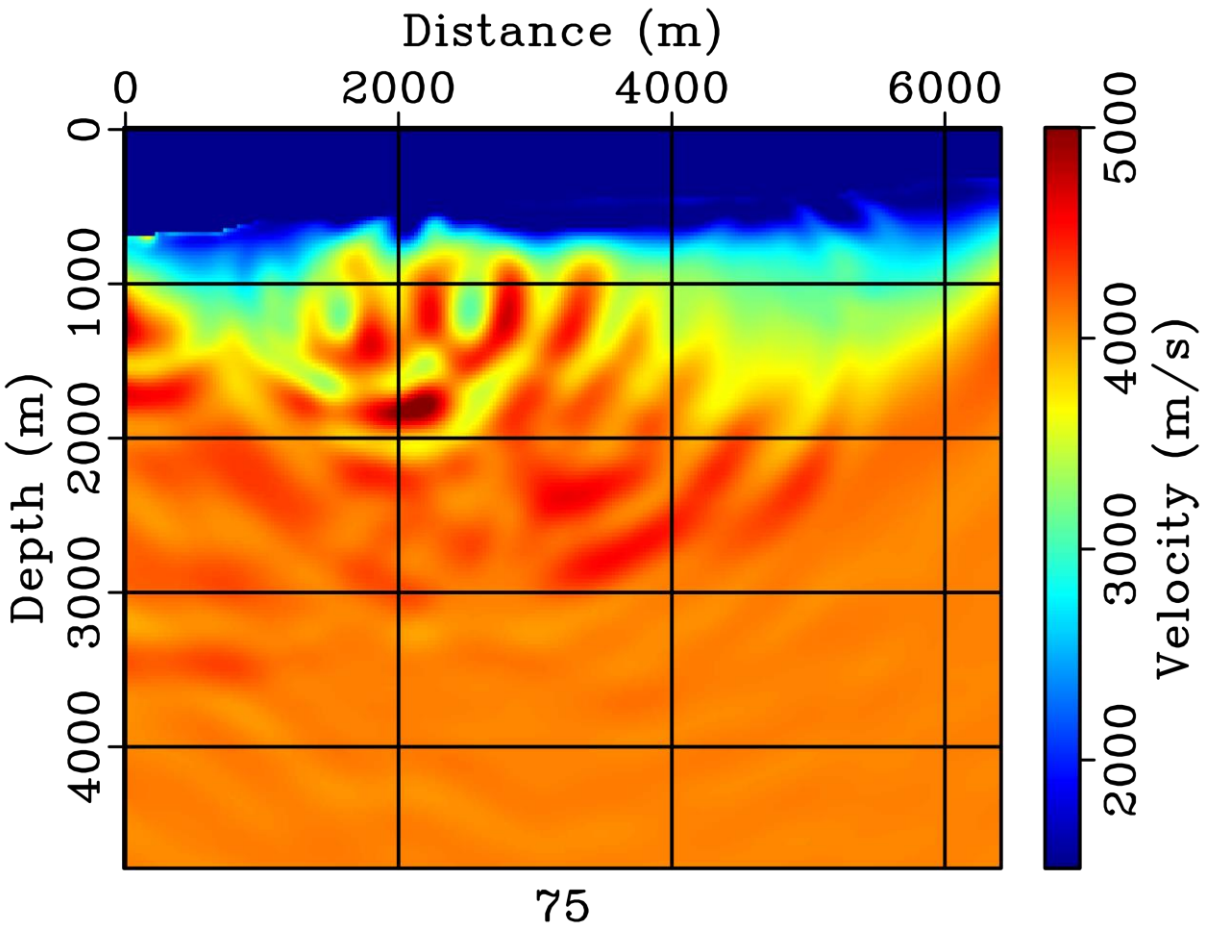
2-5 Hz

GSA-FWI



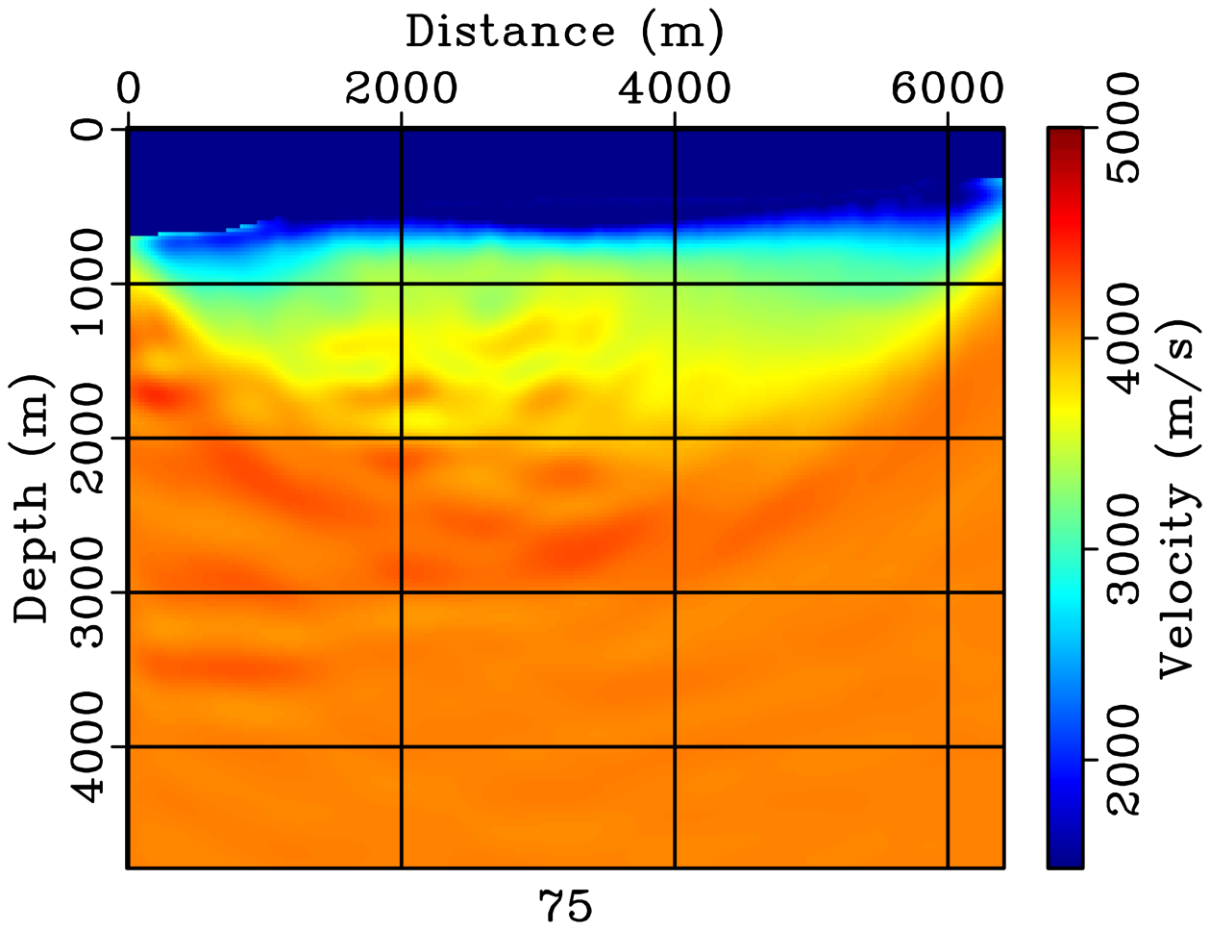
2-5 Hz

Conventional FWI



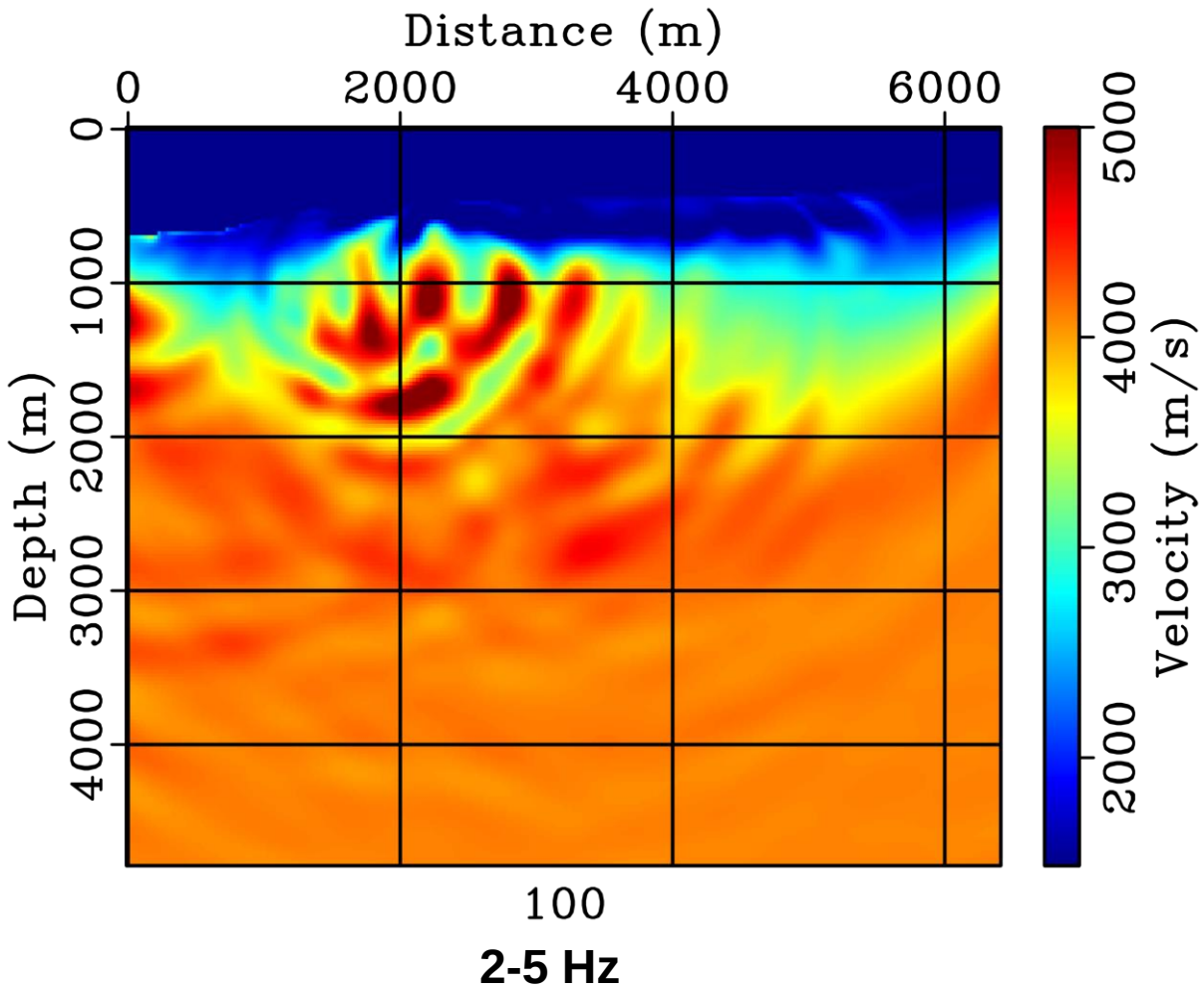
Low frequency
2-5 Hz

GSA-FWI



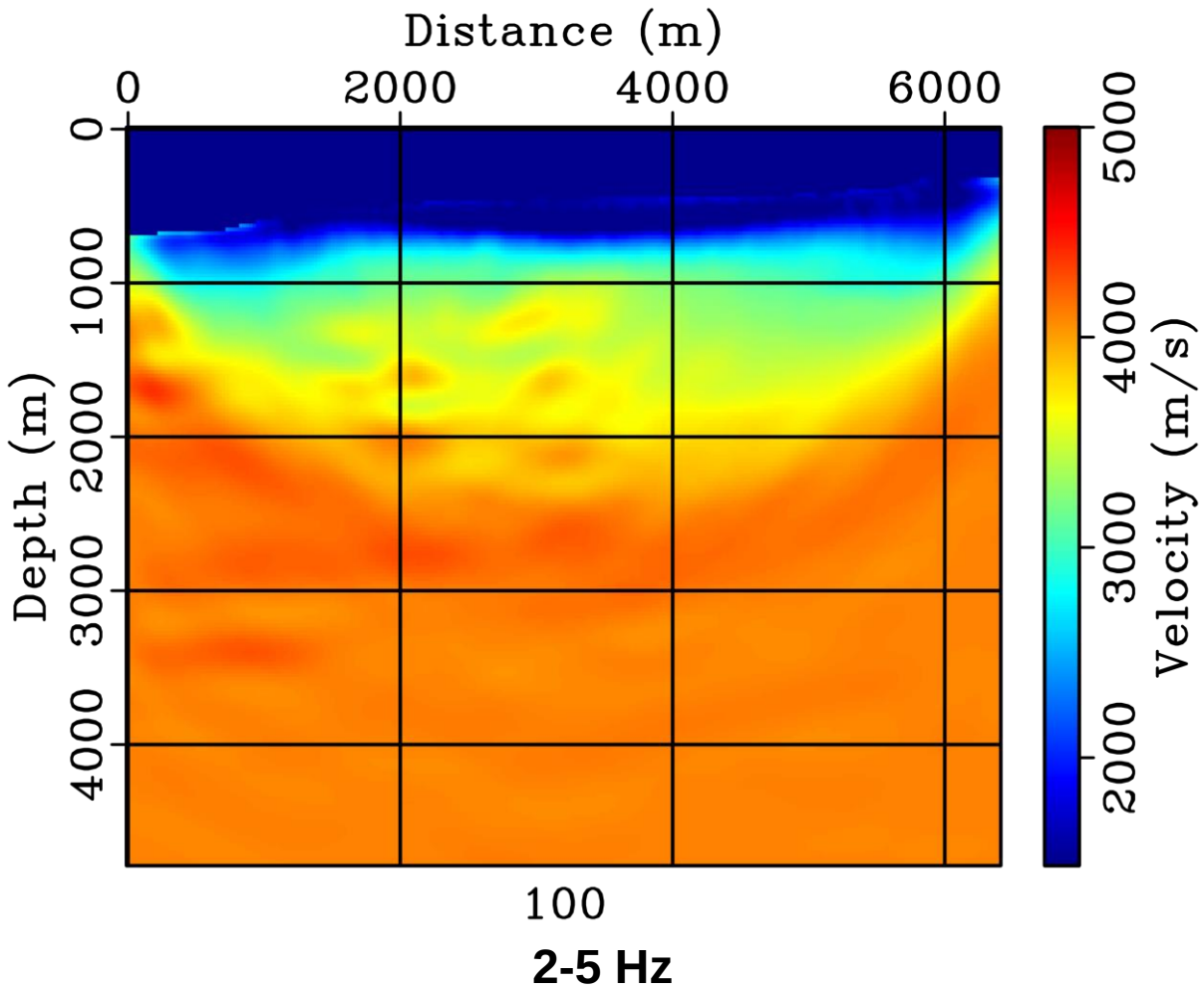
Low frequency
2-5 Hz

Conventional FWI



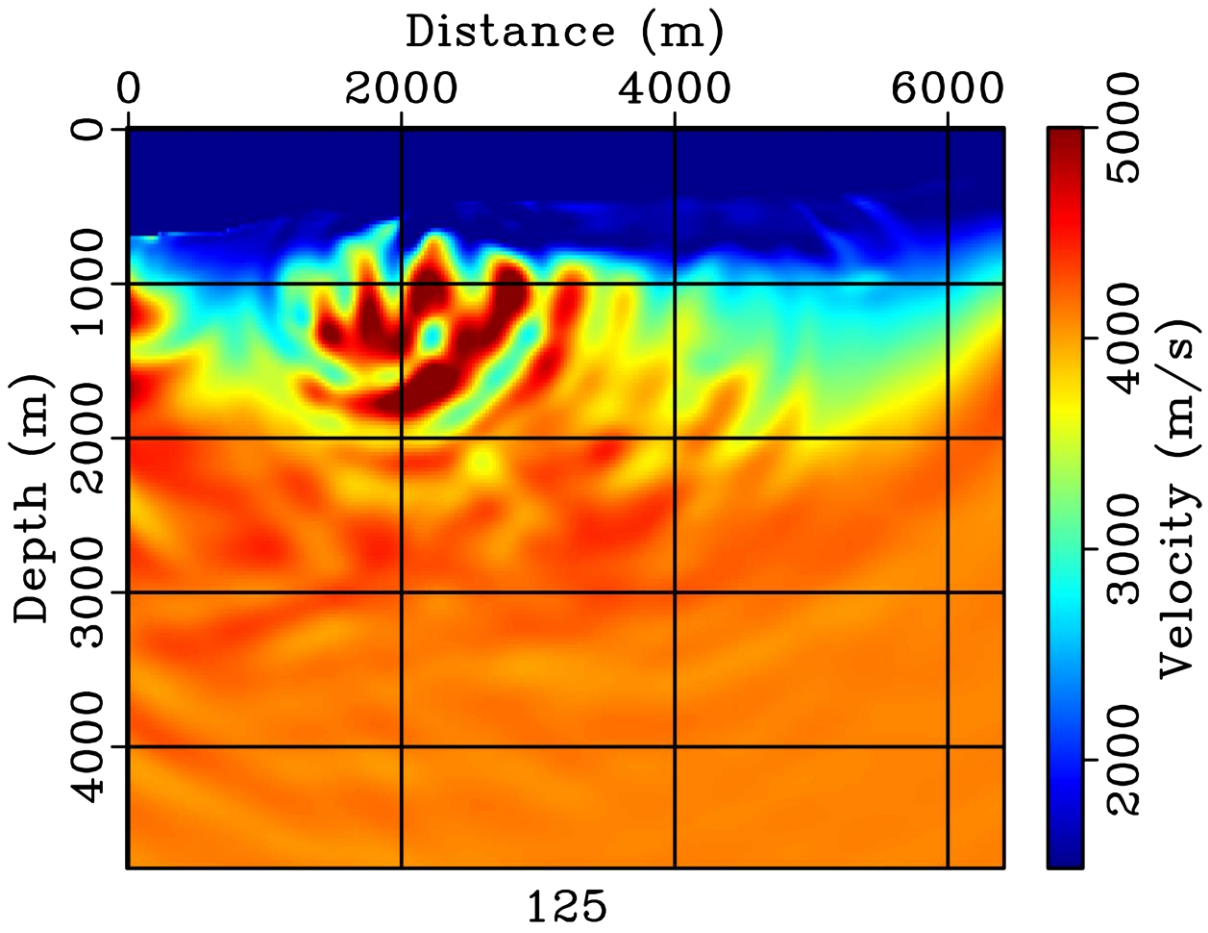
2-5 Hz

GSA-FWI



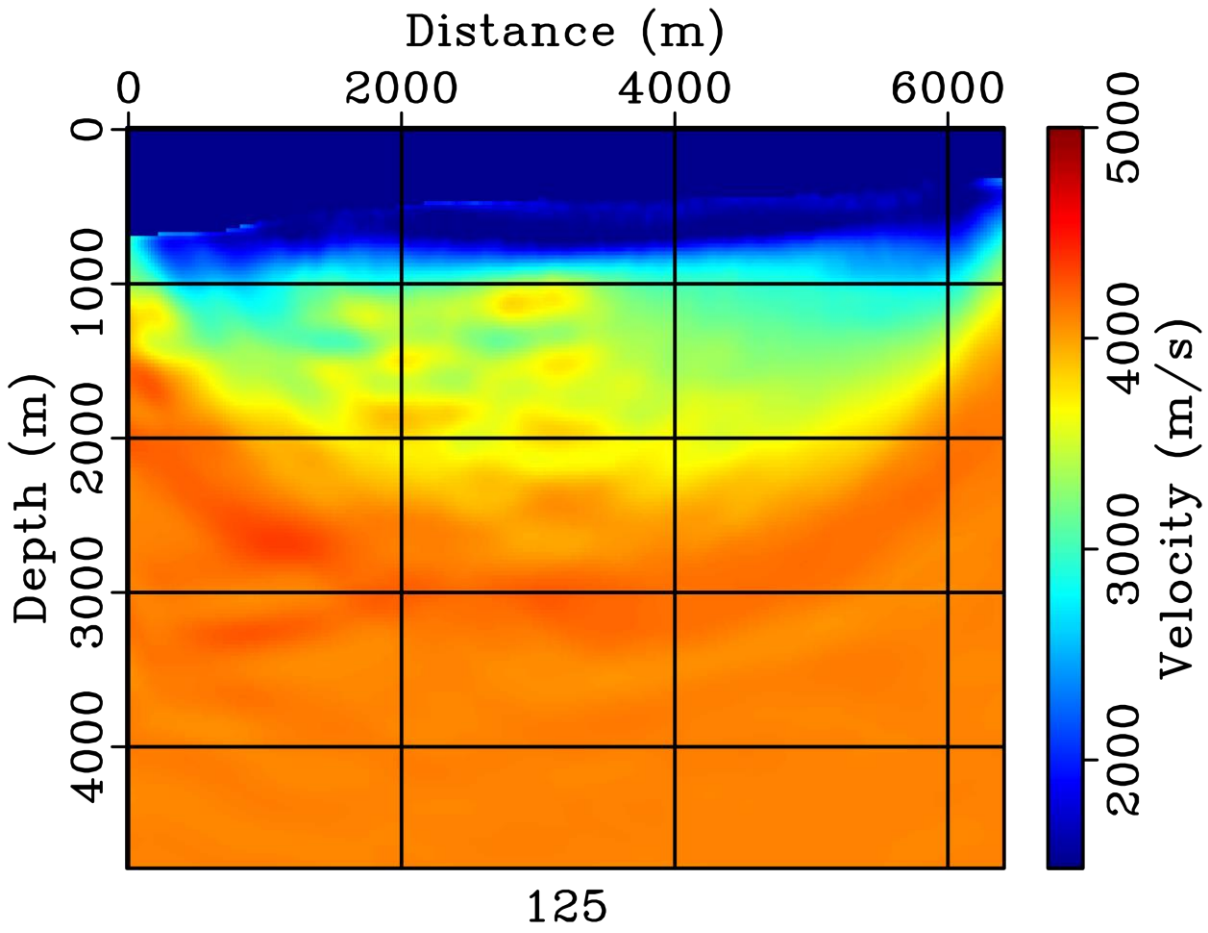
2-5 Hz

Conventional FWI



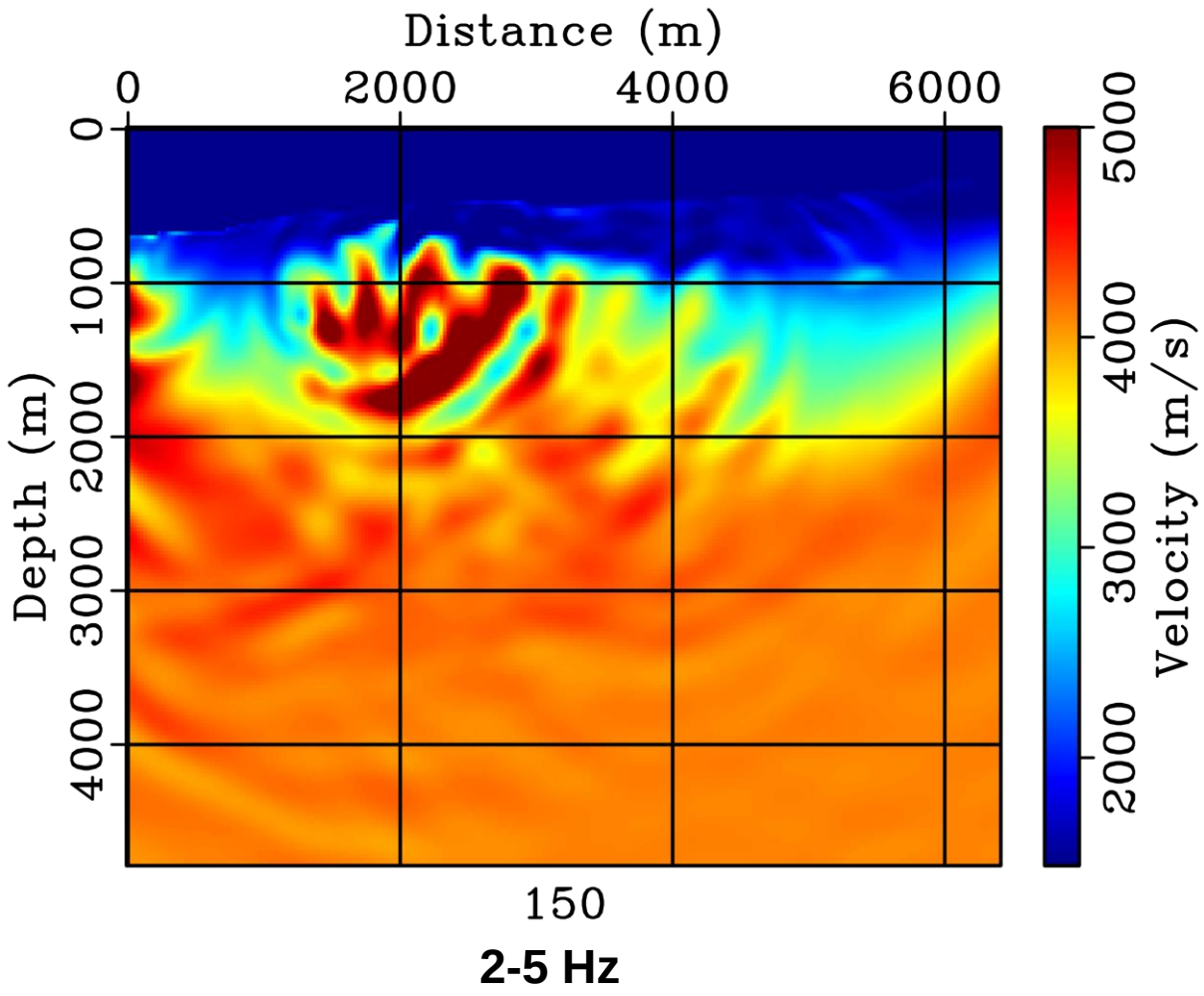
Low frequency
2-5 Hz

GSA-FWI



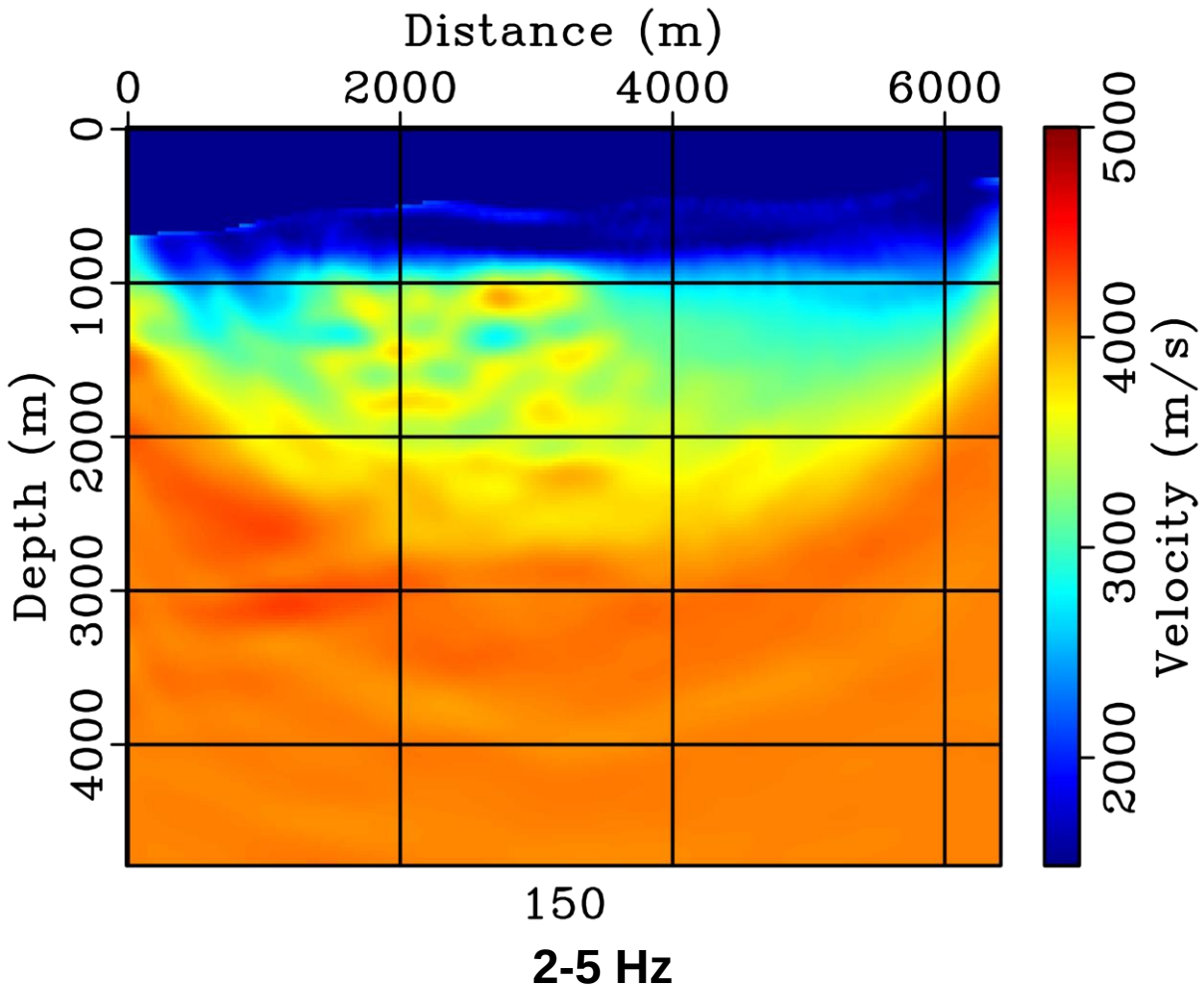
Low frequency
2-5 Hz

Conventional FWI



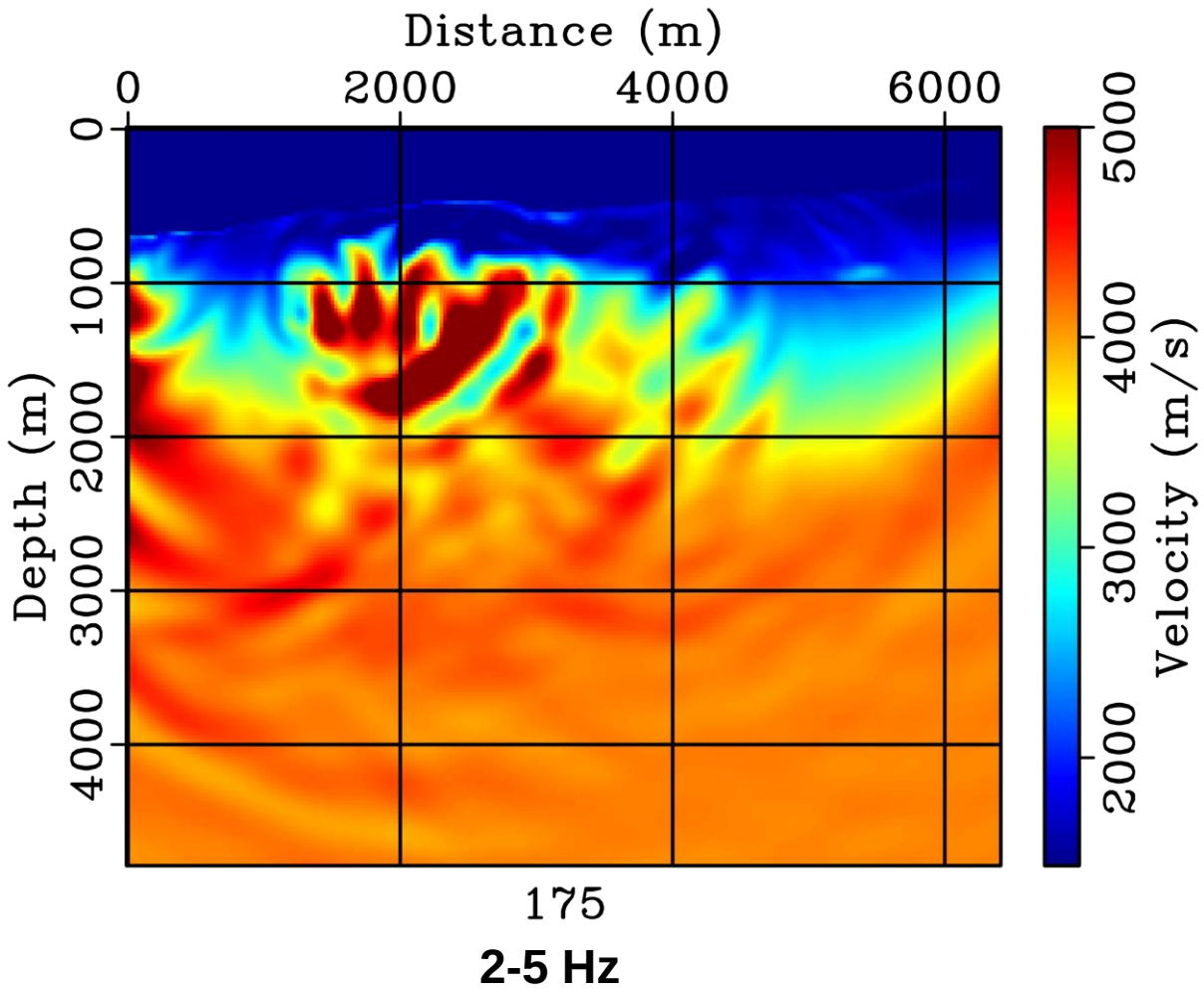
2-5 Hz

GSA-FWI



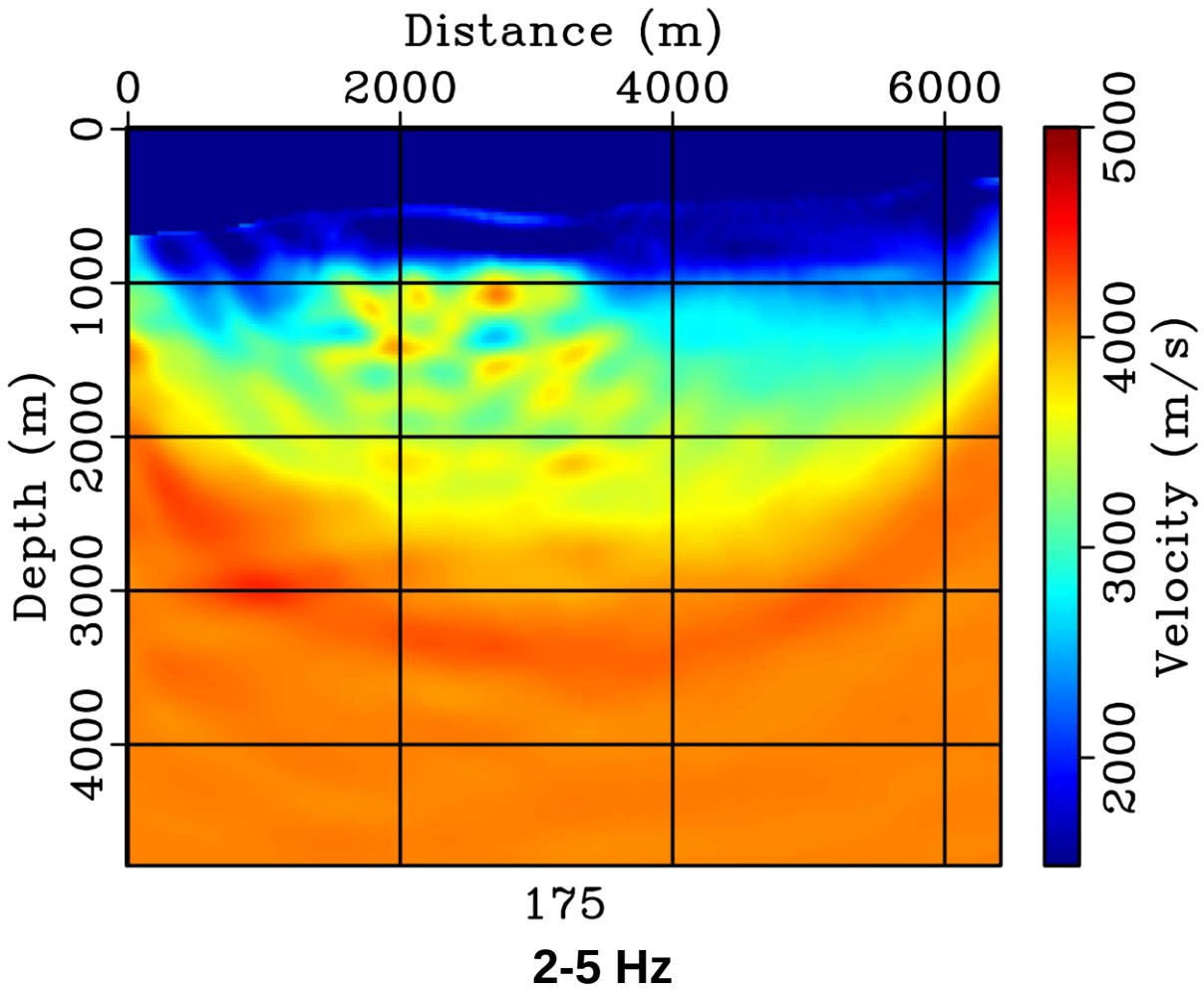
2-5 Hz

Conventional FWI



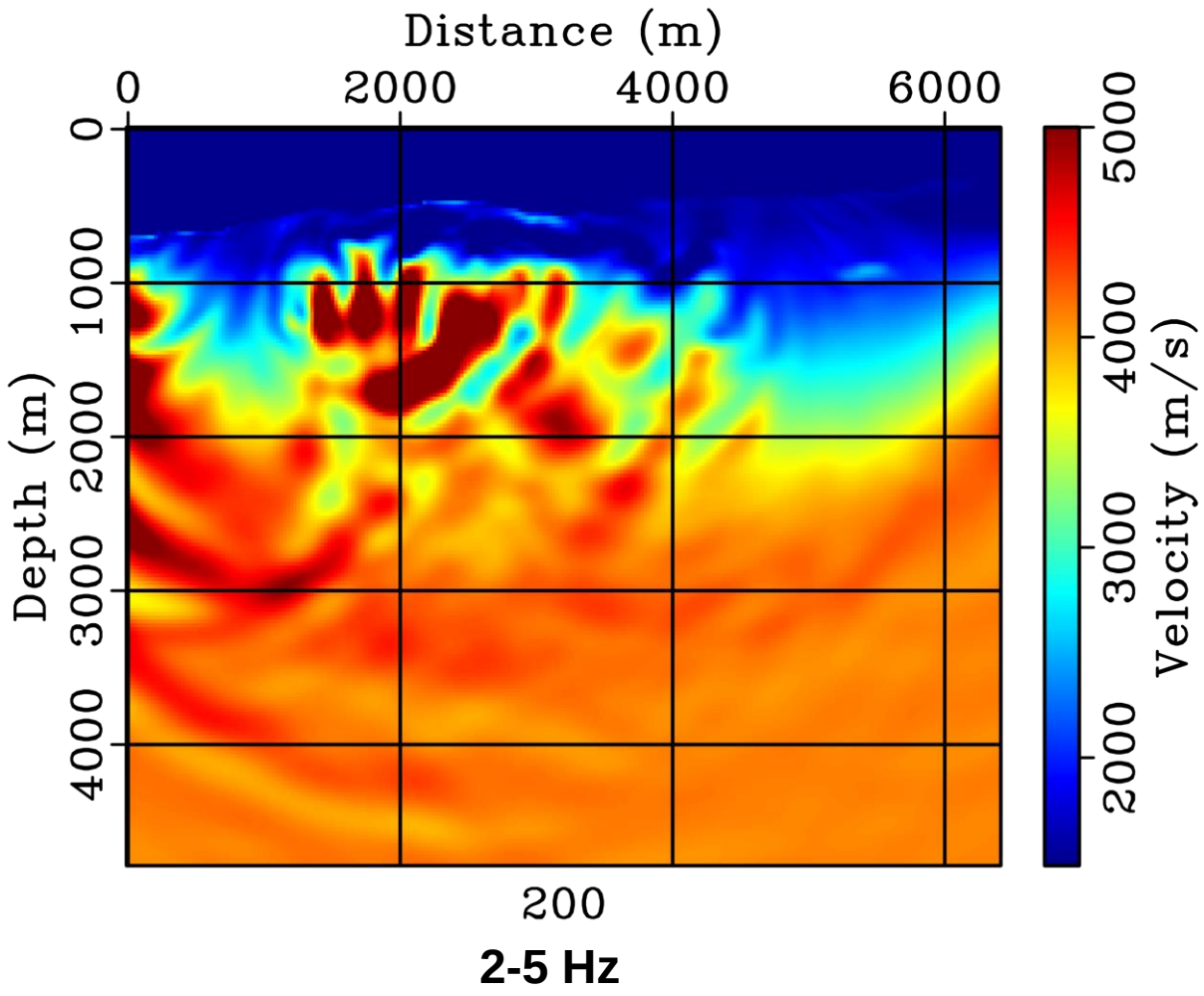
2-5 Hz

GSA-FWI



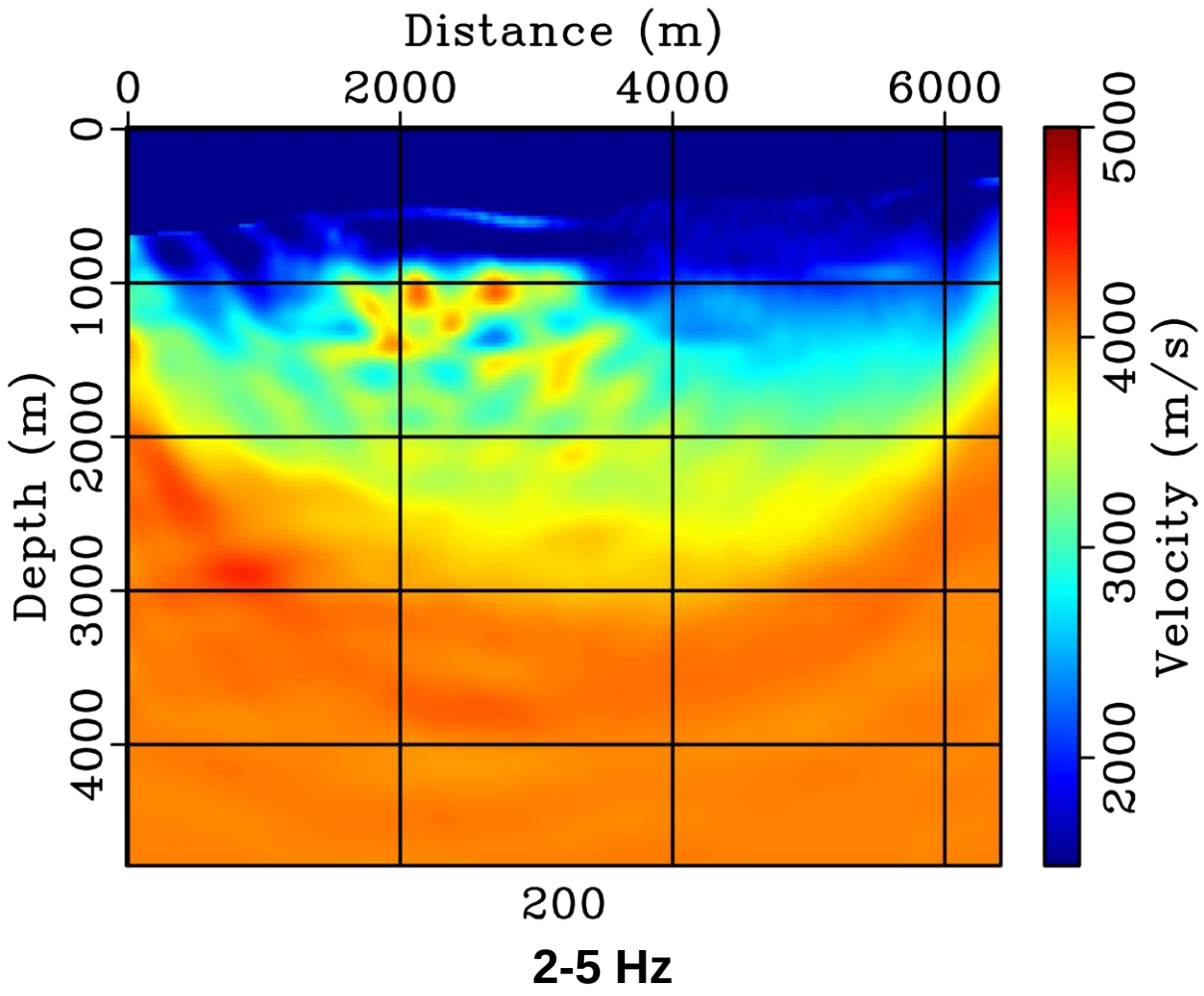
2-5 Hz

Conventional FWI



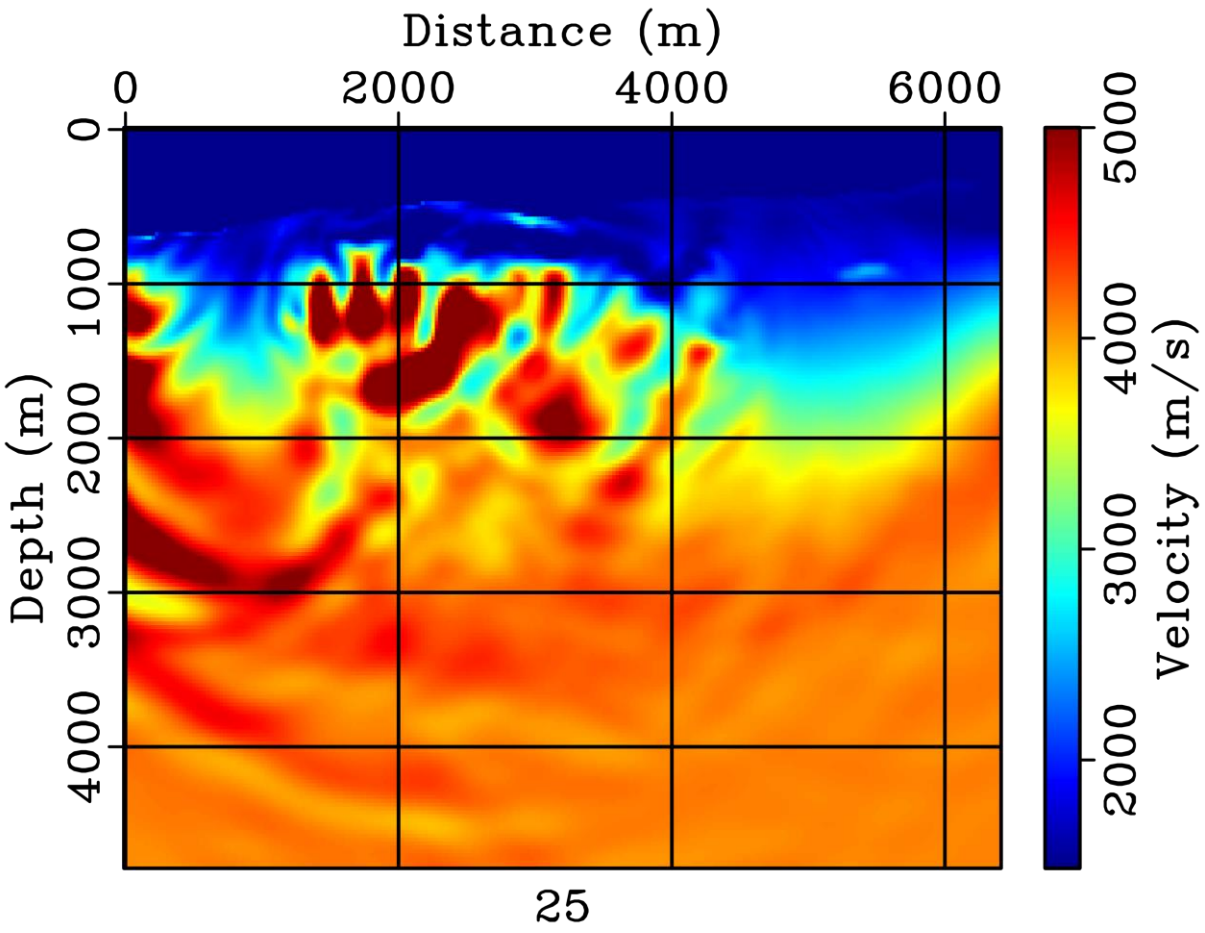
2-5 Hz

GSA-FWI



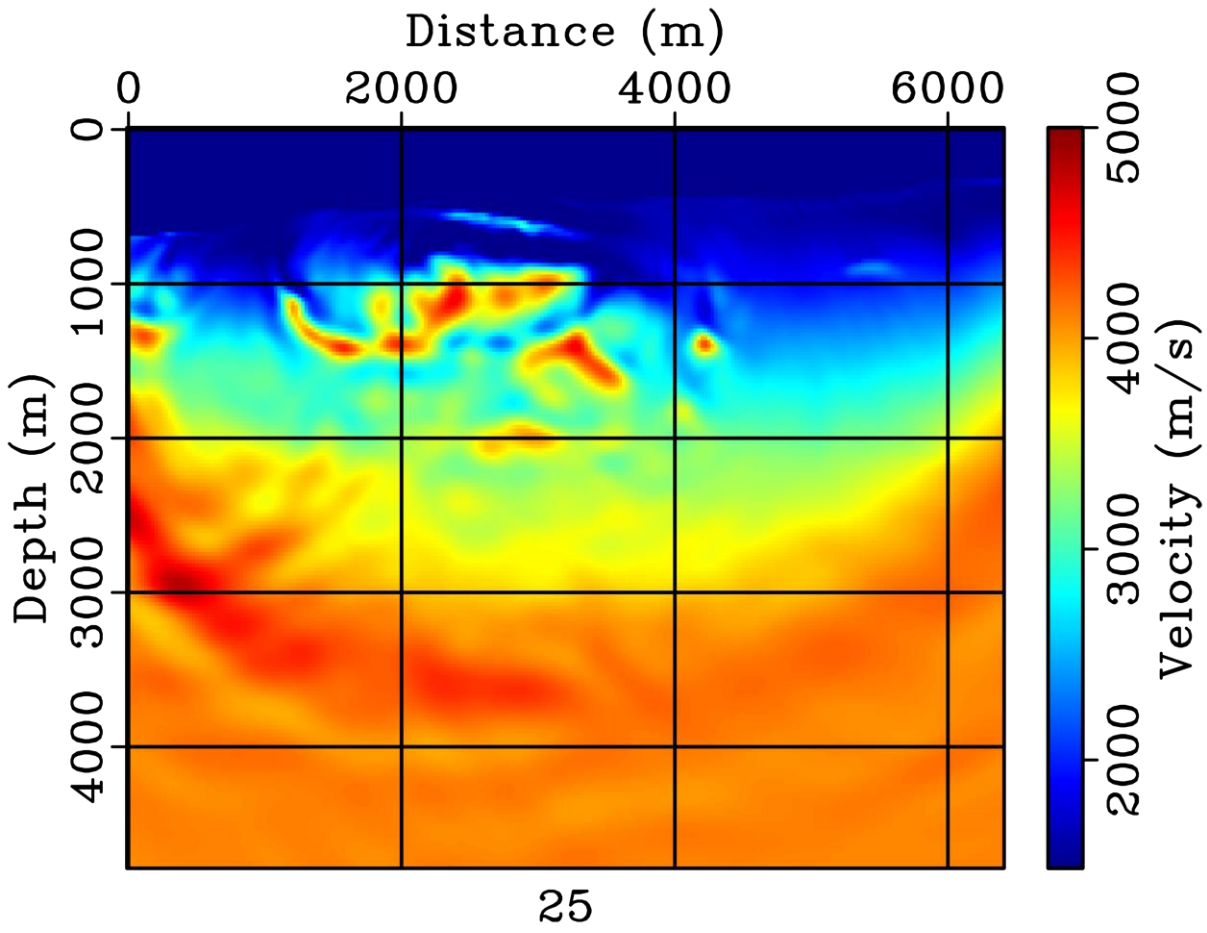
2-5 Hz

Conventional FWI



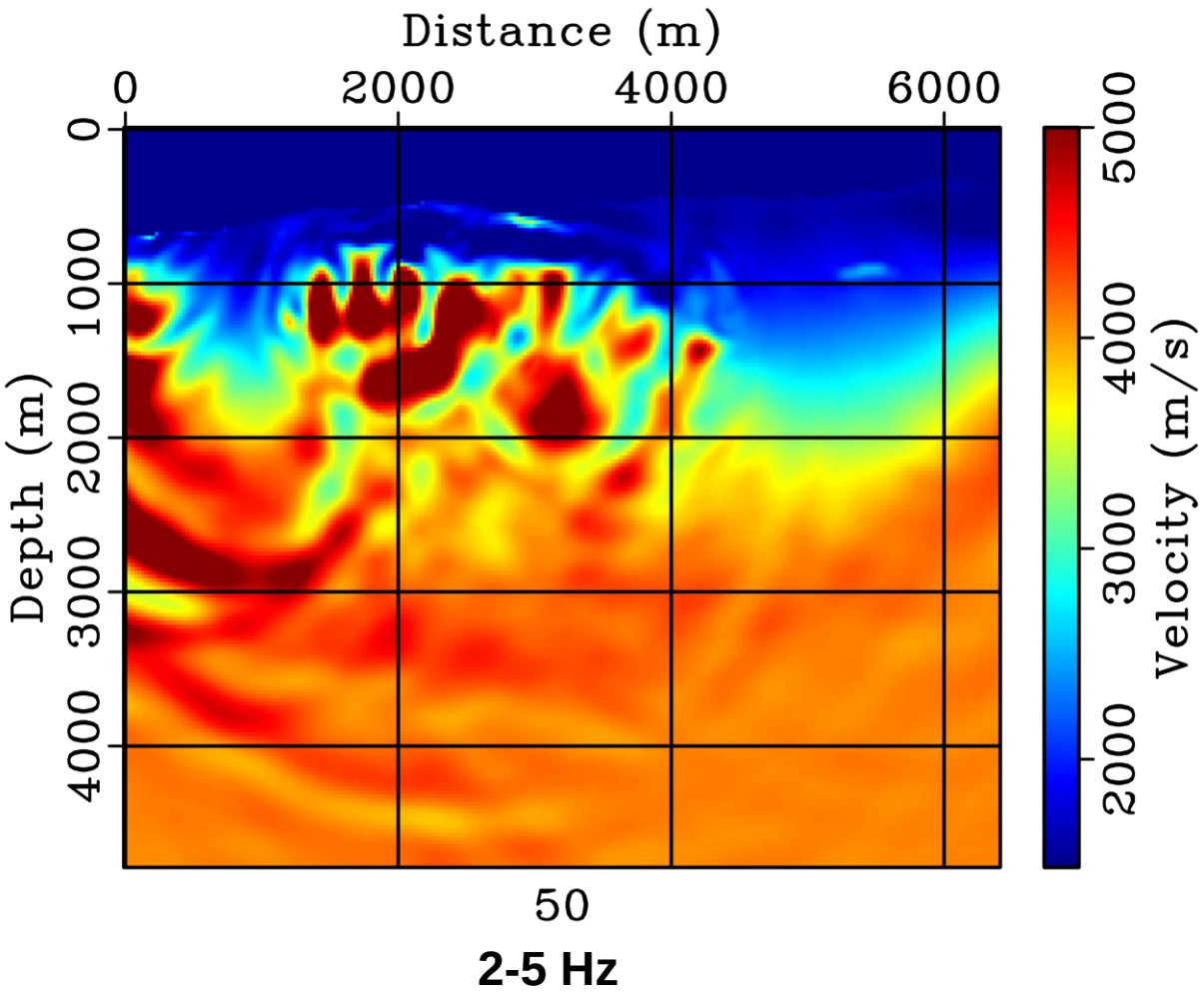
Low frequency
2-5 Hz

Conventional FWI
initialized with GSA-FWI



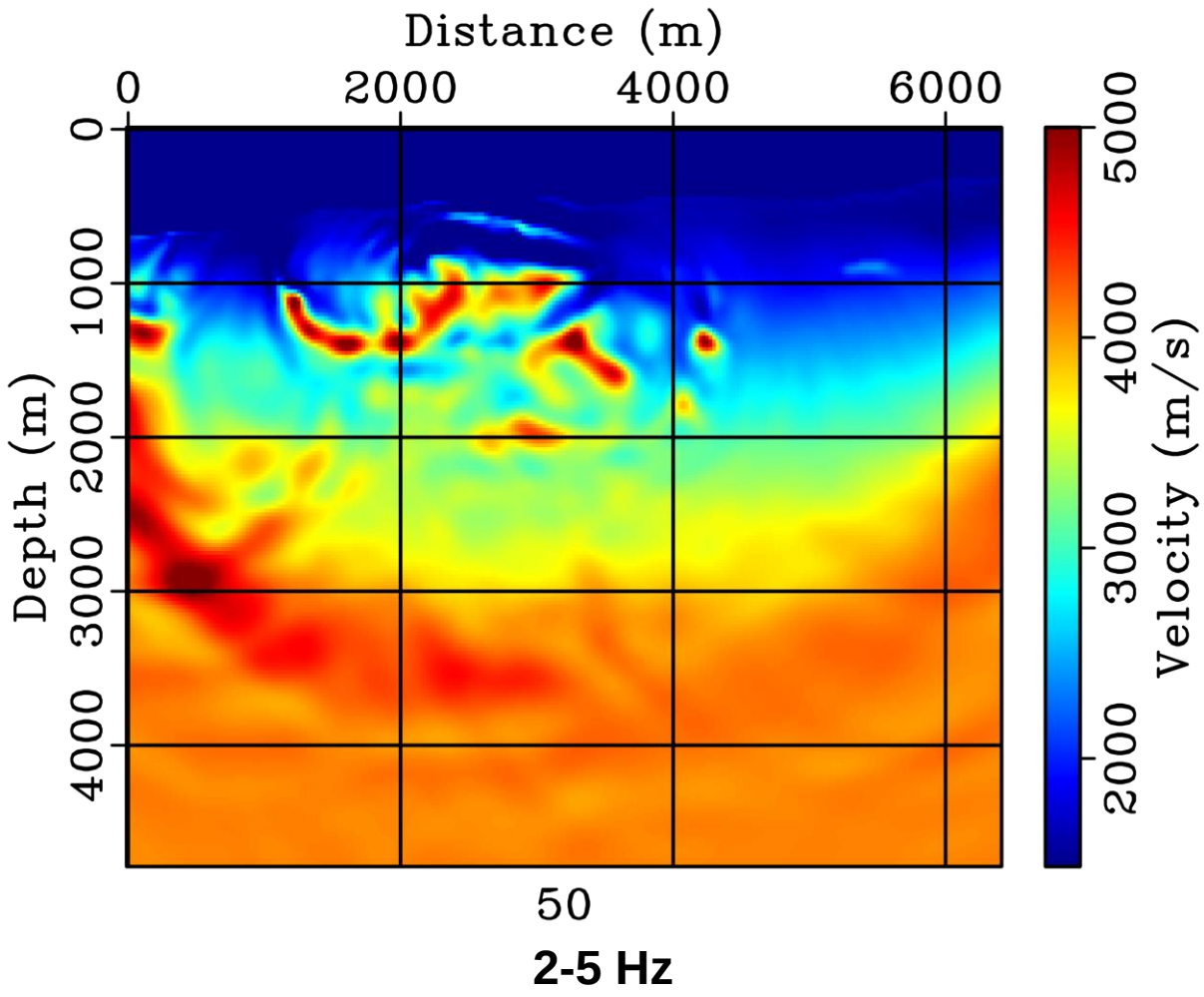
Low frequency
2-5 Hz

Conventional FWI



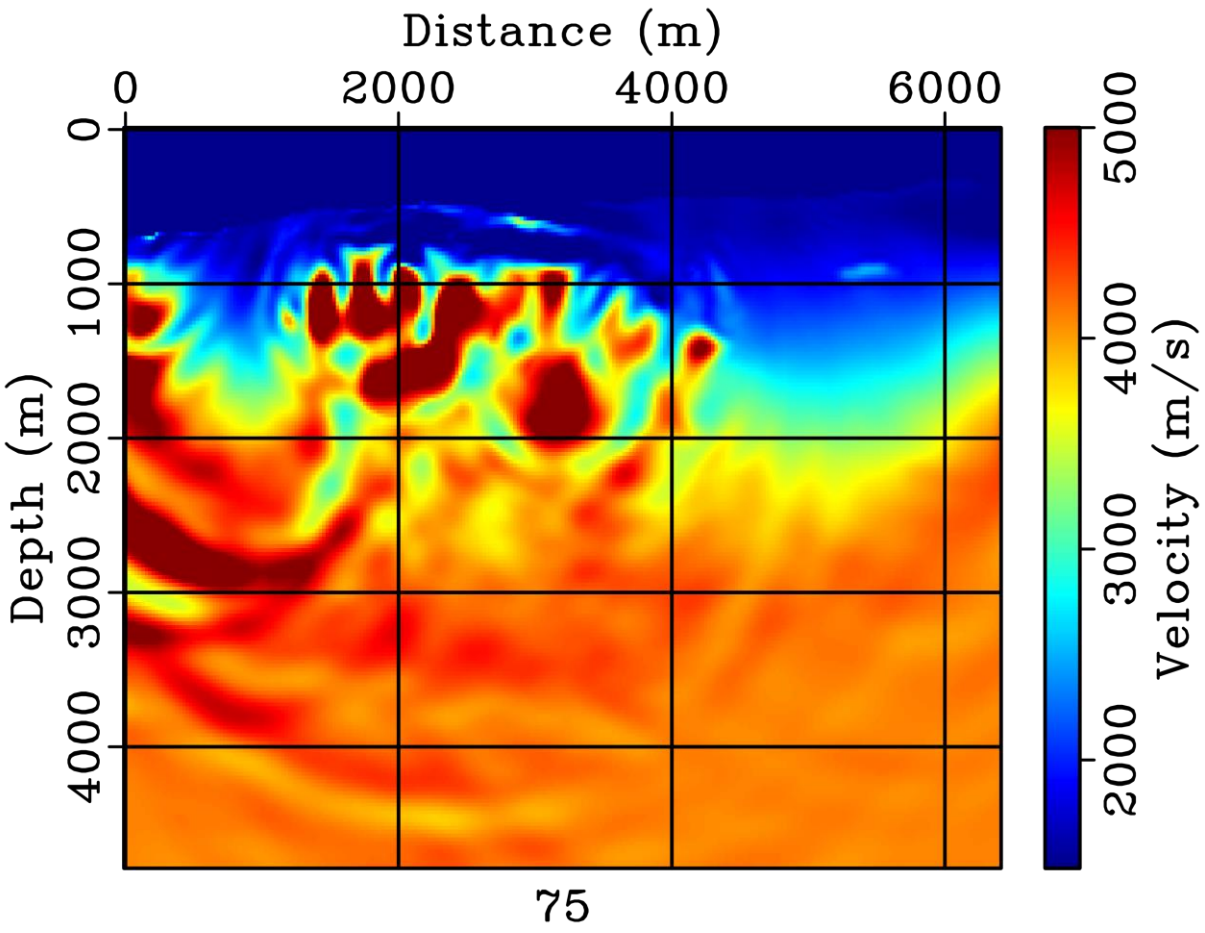
2-5 Hz

Conventional FWI initialized with GSA-FWI



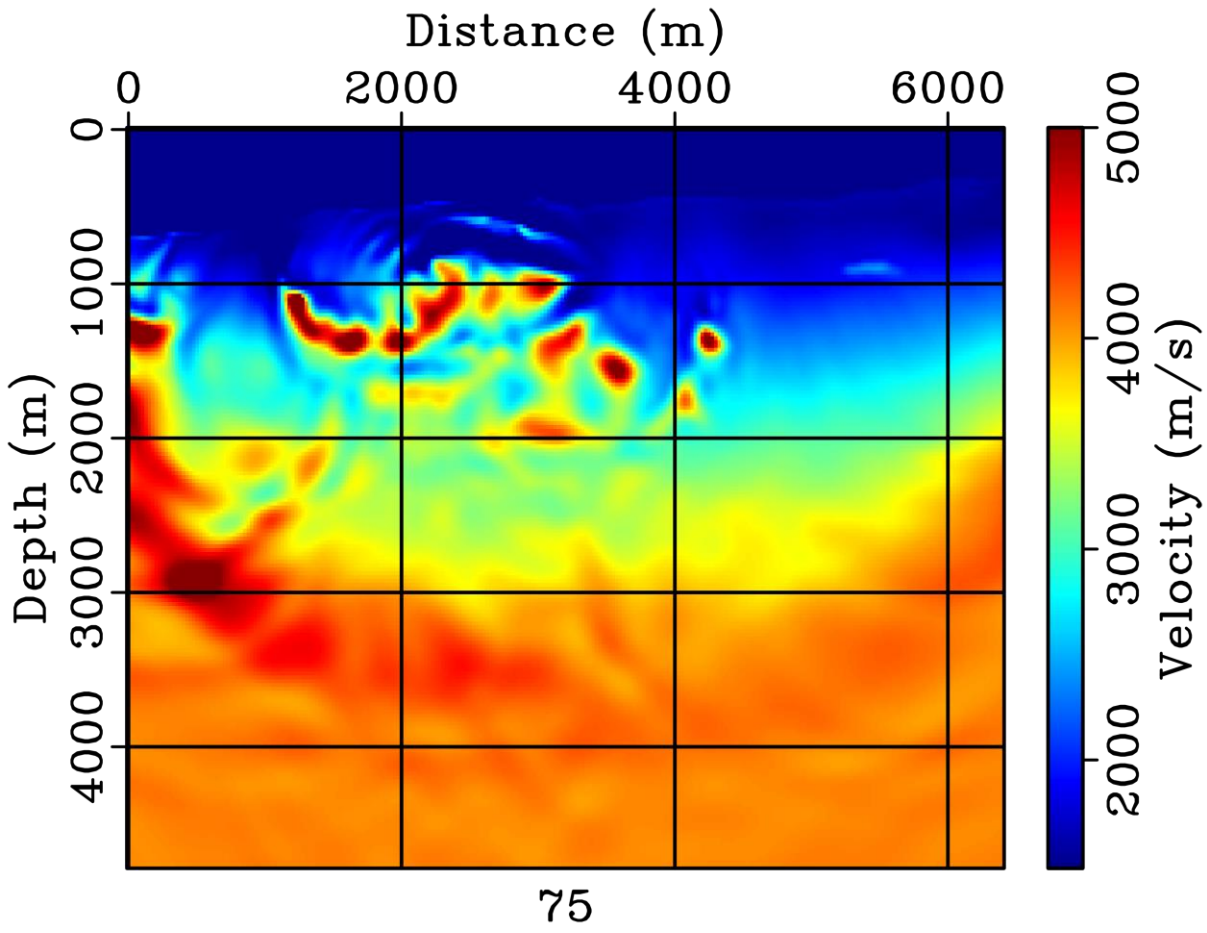
2-5 Hz

Conventional FWI



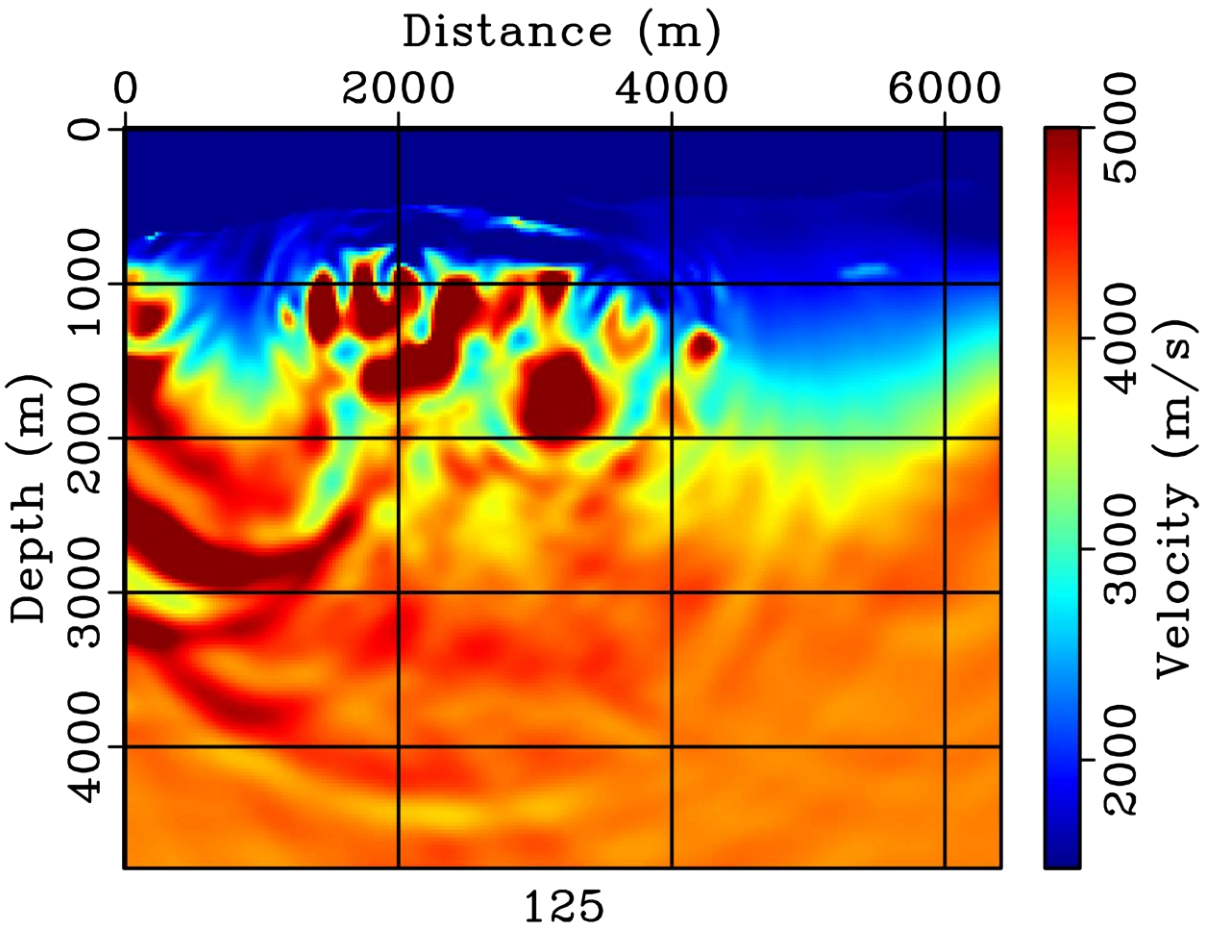
Low frequency
2-5 Hz

Conventional FWI
initialized with GSA-FWI



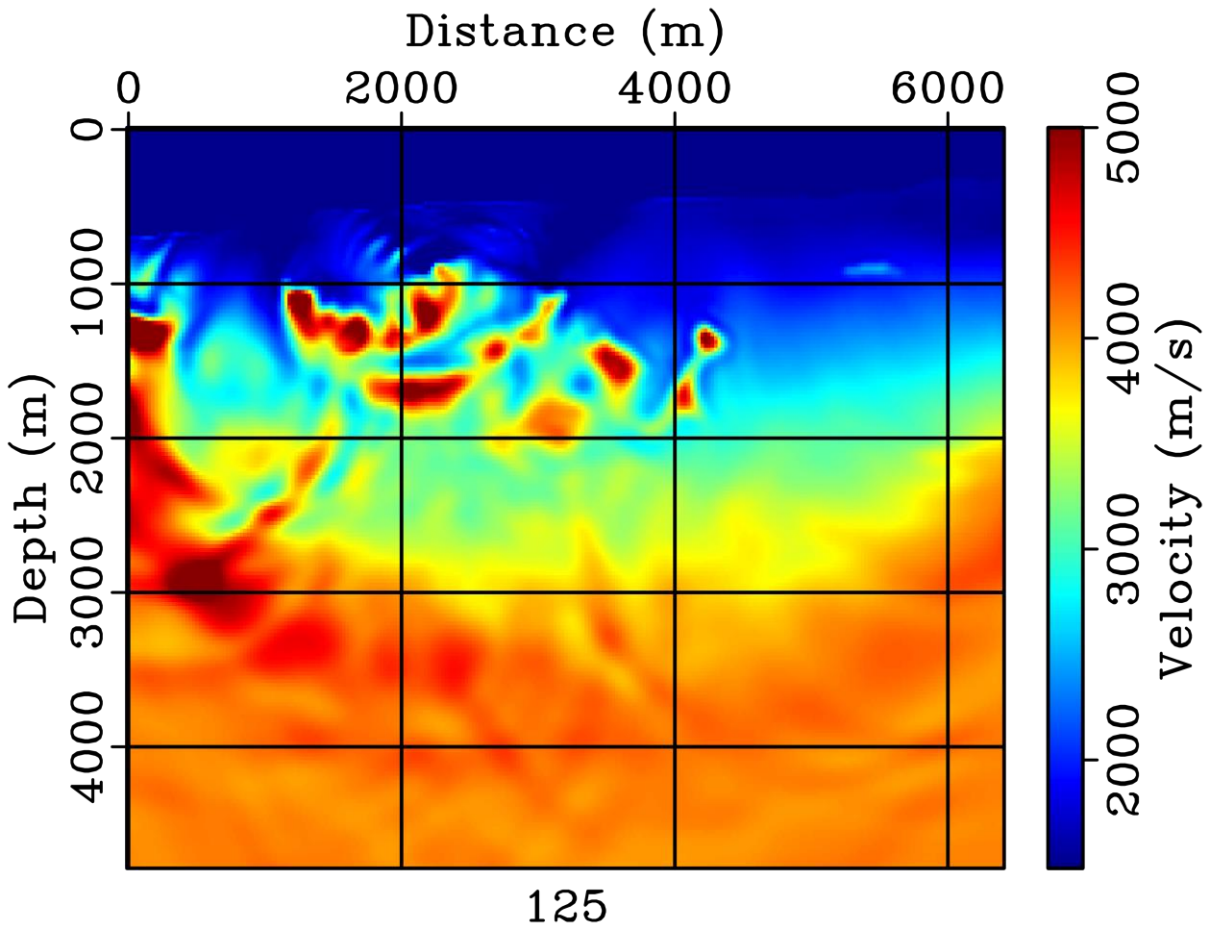
Low frequency
2-5 Hz

Conventional FWI



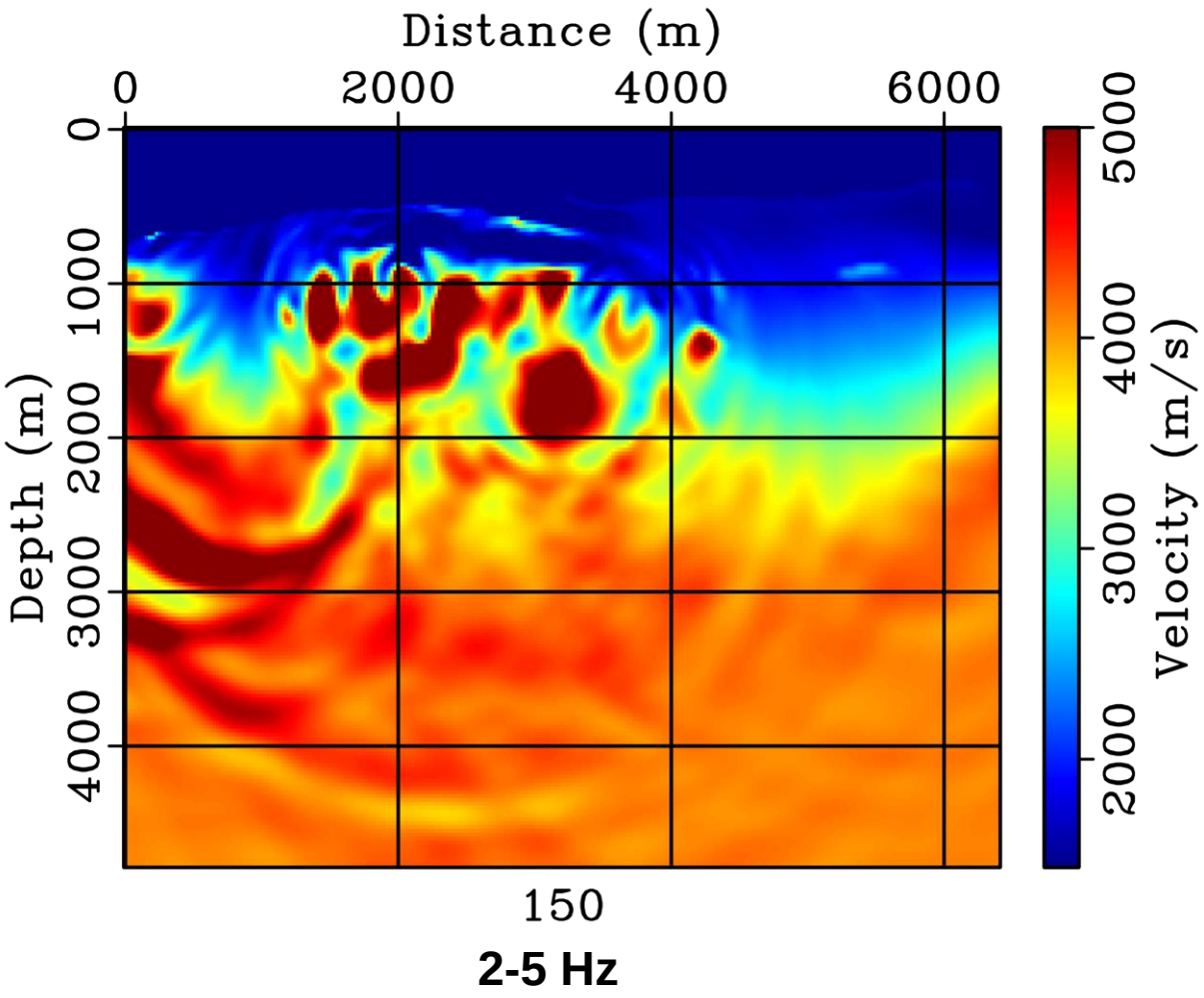
Low frequency
2-5 Hz

Conventional FWI
initialized with GSA-FWI



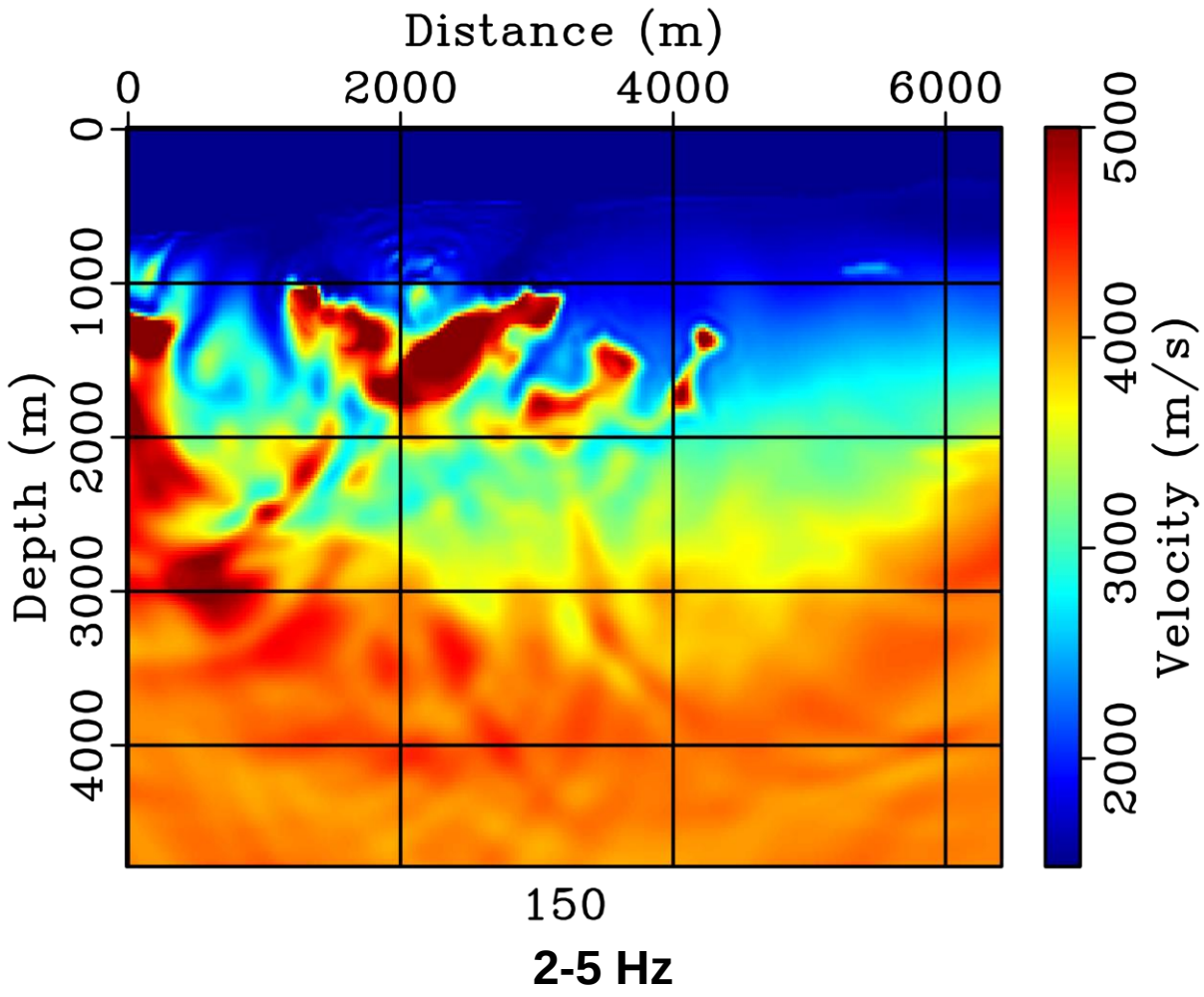
Low frequency
2-5 Hz

Conventional FWI



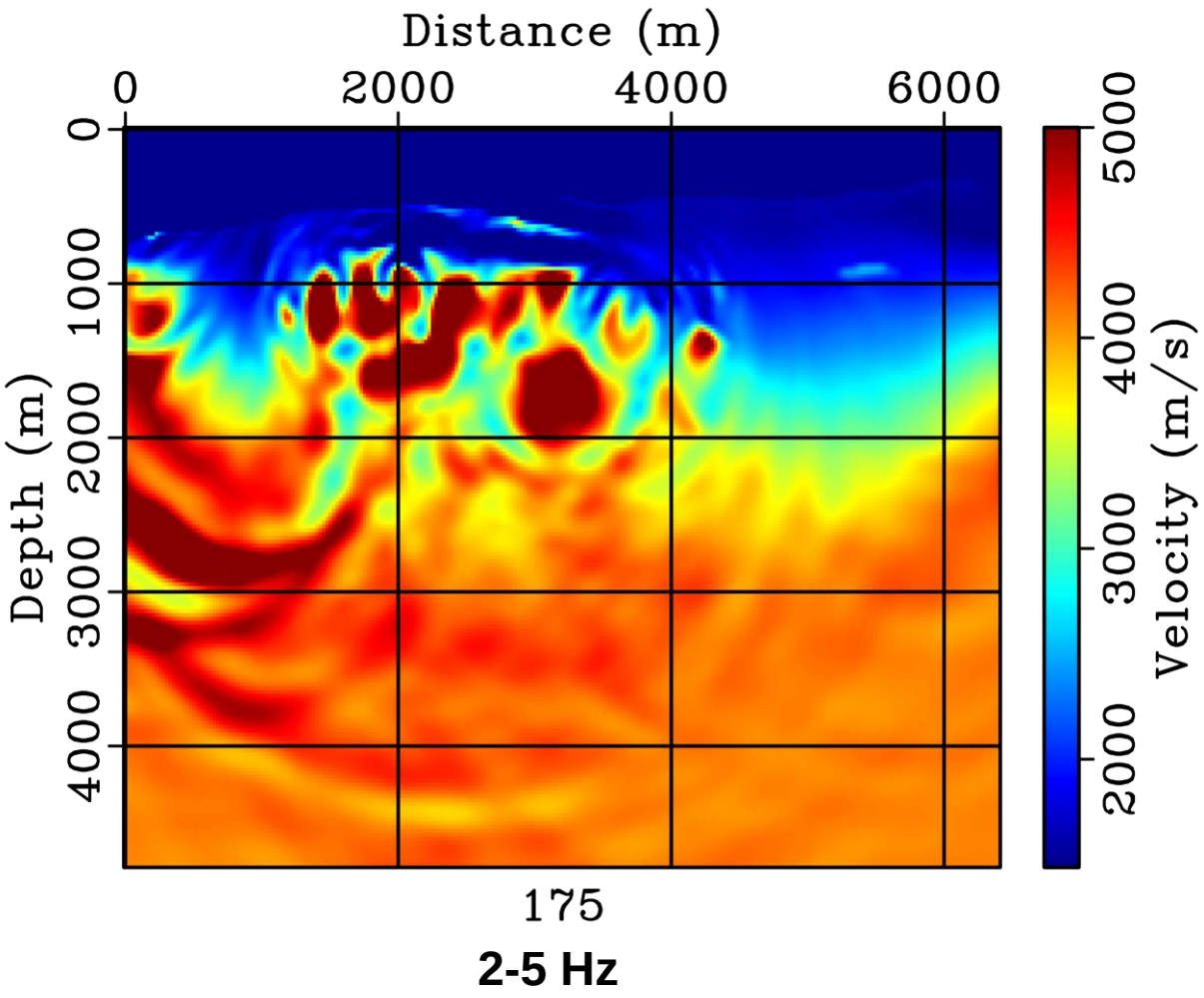
2-5 Hz

Conventional FWI initialized with GSA-FWI



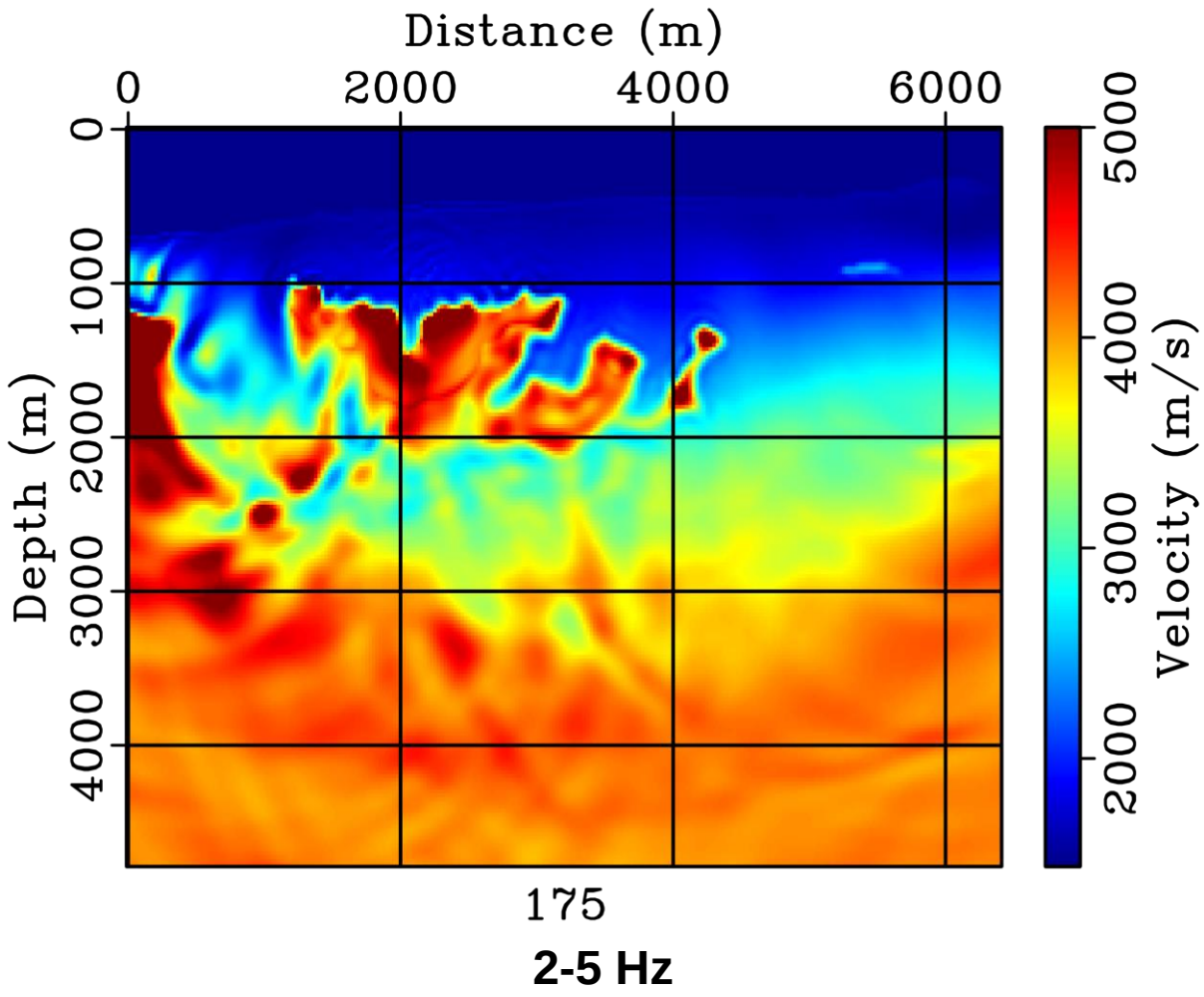
2-5 Hz

Conventional FWI



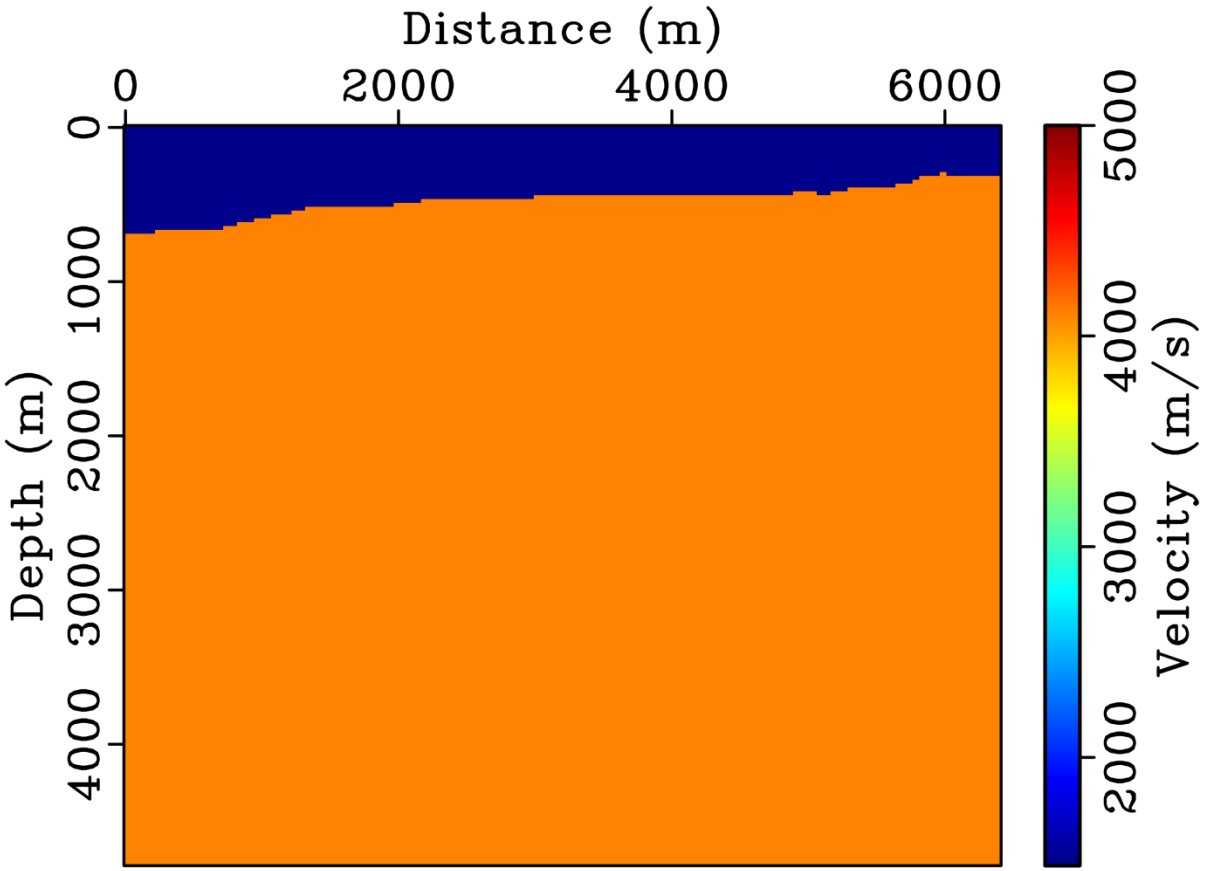
2-5 Hz

Conventional FWI initialized with GSA-FWI

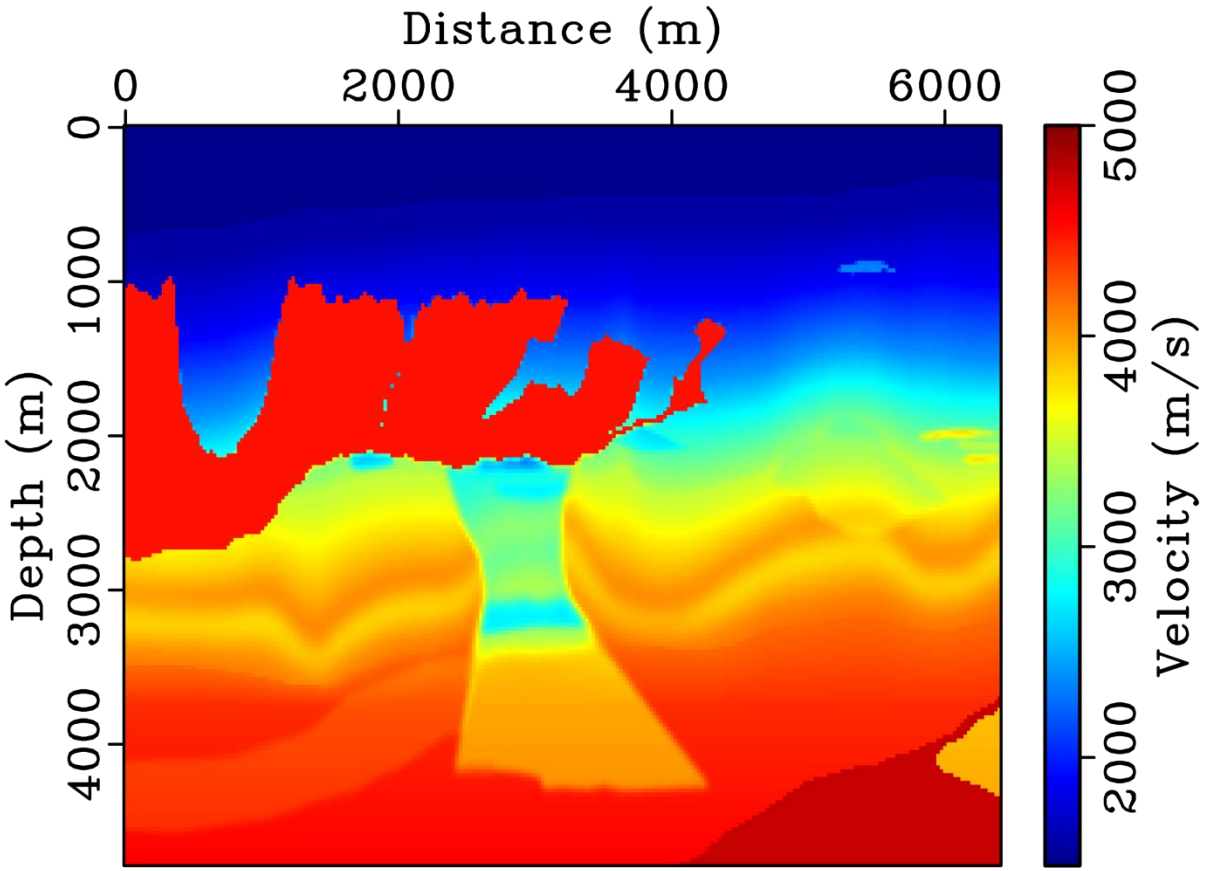


2-5 Hz

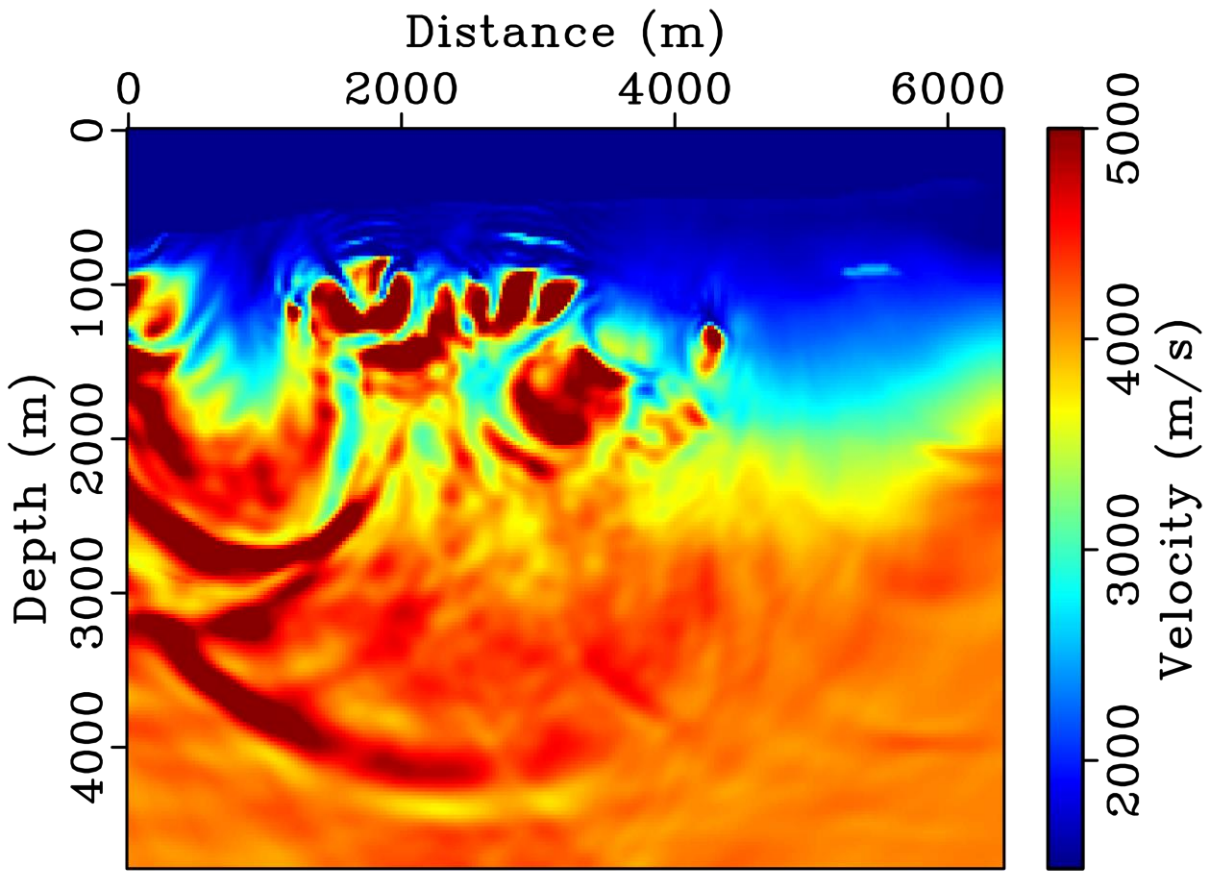
Initial velocity



True velocity

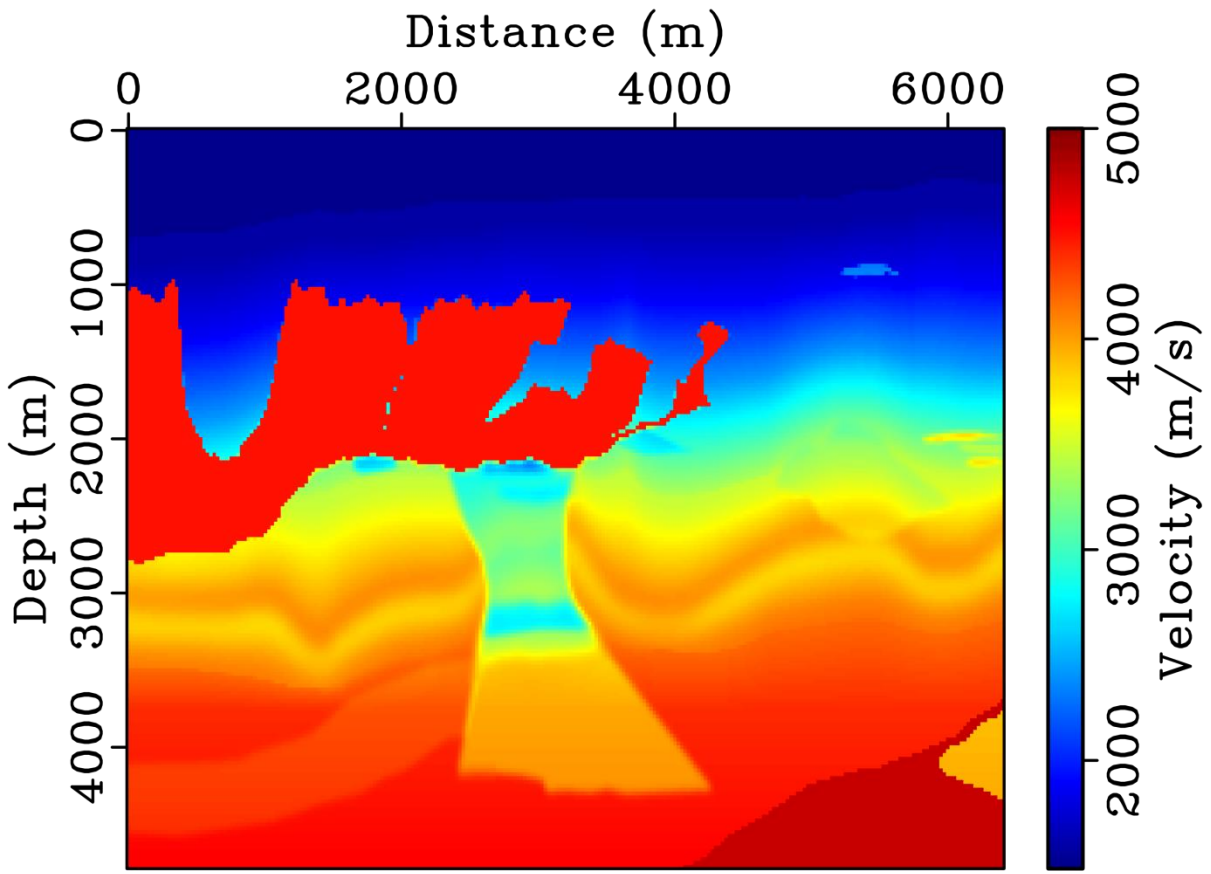


Conventional FWI

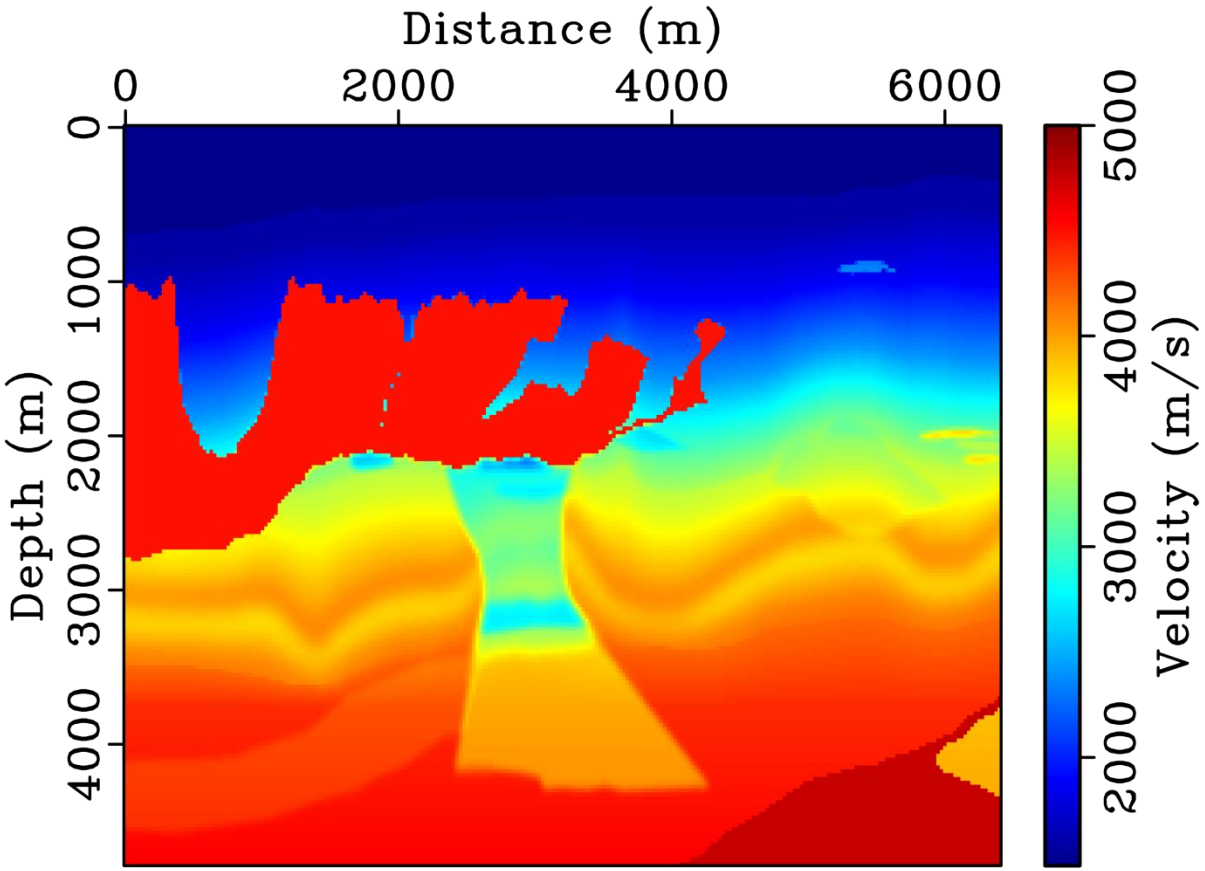


All frequency
2-19 Hz

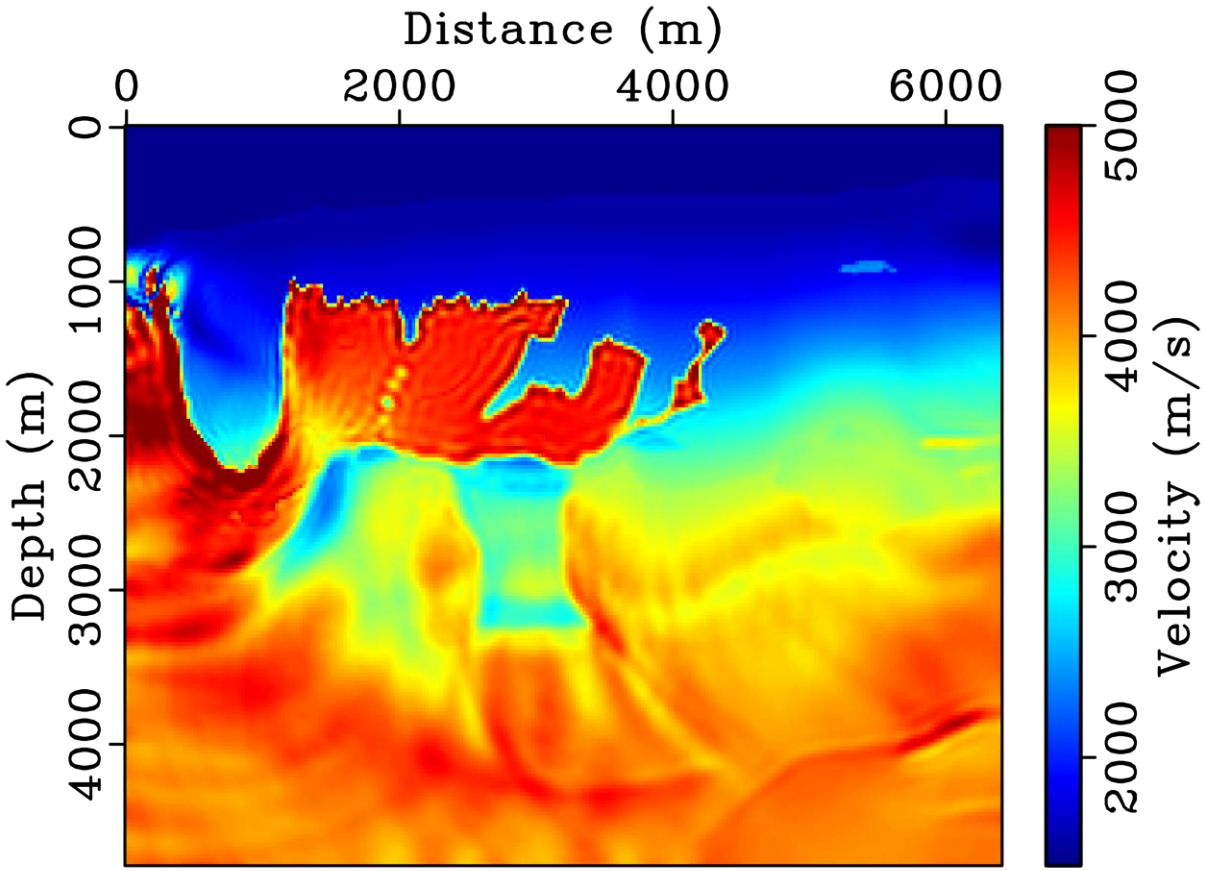
True velocity



True velocity



Conventional FWI initialized with GSA-FWI



All frequency
2-19 Hz