Wave-equation based inversion with the penalty method: adjoint-state versus wavefield-reconstruction inversion

Bas Peters, Felix J. Herrmann & Tristan van Leeuwen

76th EAGE Conference & Exhibition 2014

SLIM
University of British Columbia

Motivation

- We would like to do seismic waveform inversion with more inaccurate initial models and higher starting frequencies.
- There are indications that Wavefield Reconstruction Inversion (WRI) mitigates the non-linearity of the problem to some extend.

Conventional FWI

Least-squares objective:

$$\phi_{\mathrm{red}}(\mathbf{m}) = \frac{1}{2}\sum_{kl}\|PH_k(\mathbf{m})^{-1}\mathbf{q}_{kl} - \mathbf{d}_{kl}\|_2^2 = \frac{1}{2}\|\mathbf{d}_{\mathrm{pred}} - \mathbf{d}_{\mathrm{obs}}\|_2^2$$

$\mathbf{m}$: model
P : Restriction to receiver locations
k, l : frequency and source index
H_k : discrete Helmholtz system
$\mathbf{q}_{kl}$: source term
$\mathbf{d}_{kl}$: observed data

Conventional FWI

Least-squares objective:

$$\phi_{\text{red}}(\mathbf{m}) = \frac{1}{2}\sum_{kl}\|PH_k(\mathbf{m})^{-1}\mathbf{q}_{kl} - \mathbf{d}_{kl}\|_2^2 = \frac{1}{2}\|\mathbf{d}_{\text{pred}} - \mathbf{d}_{\text{obs}}\|_2^2$$

with the gradient (via the adjoint-state method):

$$\nabla_{\mathbf{m}}\phi_{\text{red}} = \sum_{kl} G_{kl}^{*}\mathbf{v}_{kl}$$

where

G_{kl}^{*} is the partial derivative of the discrete Helmholtz system

$\mathbf{v}_{kl}$ is the adjoint field/back propagated data residue

Wavefield Reconstruction Inversion [T. van Leeuwen & F.J. Herrmann, 2013]

Objective:

Data-misfit ↓ PDE-misfit ↓

$$\bar{\phi}_\lambda(\mathbf{m}) = \frac{1}{2}\sum_{kl}\|P\bar{\mathbf{u}}_{kl} - \mathbf{d}_{kl}\|_2^2 + \frac{\lambda^2}{2}\|H_k(\mathbf{m})\bar{\mathbf{u}}_{kl} - \mathbf{q}_{kl}\|_2^2$$

where $\bar{\mathbf{u}}_{kl} = \arg\min_{\mathbf{u}_{kl}} \left\| \begin{pmatrix} \lambda H_k(\mathbf{m}) \\ P \end{pmatrix} \mathbf{u}_{kl} - \begin{pmatrix} \lambda \mathbf{q}_{kl} \\ \mathbf{d}_{kl} \end{pmatrix} \right\|_2$

and λ is a tradeoff parameter between PDE-fit and data-fit

Wavefield Reconstruction Inversion [T. van Leeuwen & F.J. Herrmann, 2013]

Objective:

Data-misfit ↓ PDE-misfit ↓

$$\bar{\phi}_\lambda(\mathbf{m}) = \frac{1}{2}\sum_{kl} \|P\bar{\mathbf{u}}_{kl} - \mathbf{d}_{kl}\|_2^2 + \frac{\lambda^2}{2}\|H_k(\mathbf{m})\bar{\mathbf{u}}_{kl} - \mathbf{q}_{kl}\|_2^2$$

with gradient:

$$\nabla_{\mathbf{m}}\bar{\phi}_\lambda = \sum_{kl} \lambda^2 G_{kl}(\mathbf{m}, \bar{\mathbf{u}}_{kl})^* \left(H_k(\mathbf{m})\bar{\mathbf{u}}_{kl} - \mathbf{q}_{kl}\right)$$

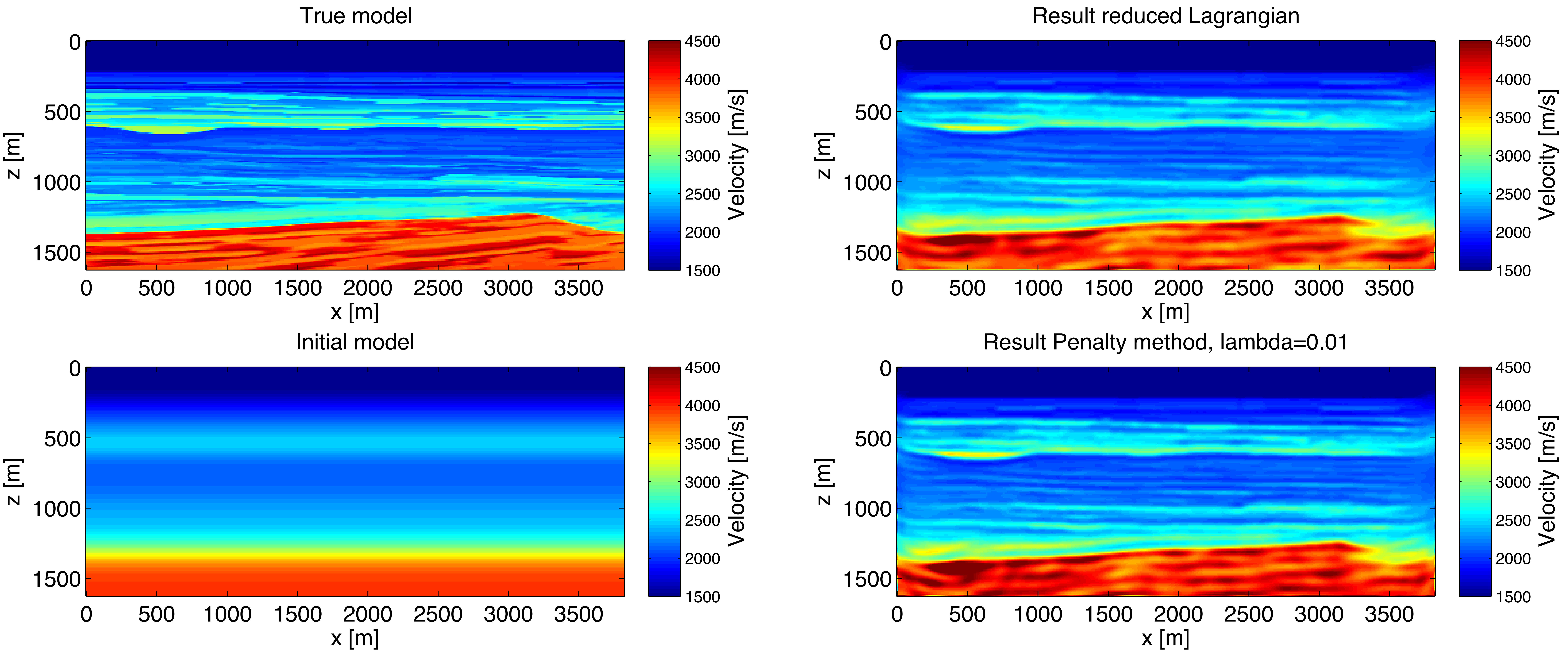
Non-linear waveform inversion

Example 1a (easy):

- Used the L-BFGS algorithm
- 64 equally distributed sources and receivers near the surface
- 18 frequency batches (10 iterations each) as {2 3}, {3 4} ,... ,{19 20} Hertz
- No noise

- Solve least-squares problem using SuiteSparseQR. **[T.A. Davis, 2011]**

True, initial and final models



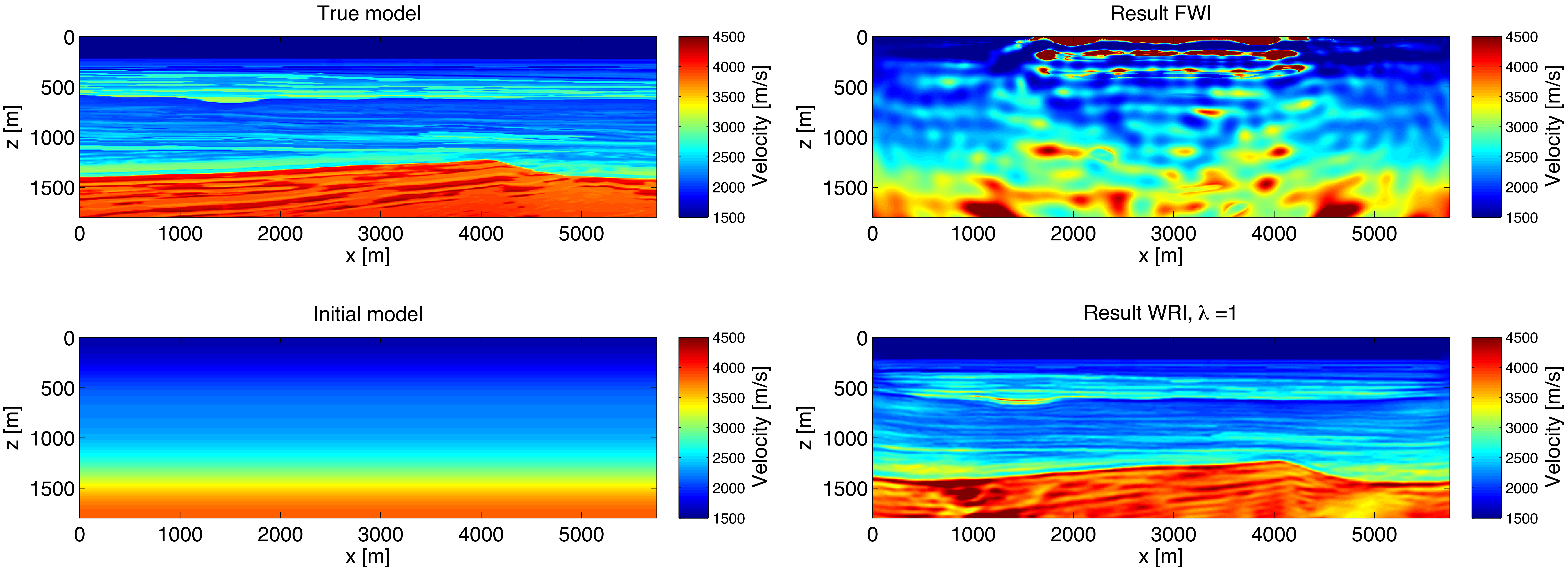
Non-linear waveform inversion

Example 1b (difficult):

- Lots of low frequencies missing, 24 frequency batches (15 iterations each) with intervals {5 6} ,{6 7},... ,{28 29} Hertz. Each interval contains 5 frequencies.
- We use 2 cycles through the batches: {5 6} ,{6 7},... ,{28 29}, {5 6} ,{6 7},... ,{28 29}
- Inaccurate initial model
- 103 sources and receivers near the surface, spread over the whole domain (6km).
- Source & receiver interval: 55m. Max. offset 6km.
- Shortest wavelength: 290m @ 5Hz. and 50m @ 29 Hz.
- Used Two-metric projection with L-BFGS Hessian for optimization with bound-constraints. **[Bertsekas, 1982 ; Gafni & Bertsekas, 1982 ; Schmidt, Kim & Sra, 2009]**

True, initial and final models

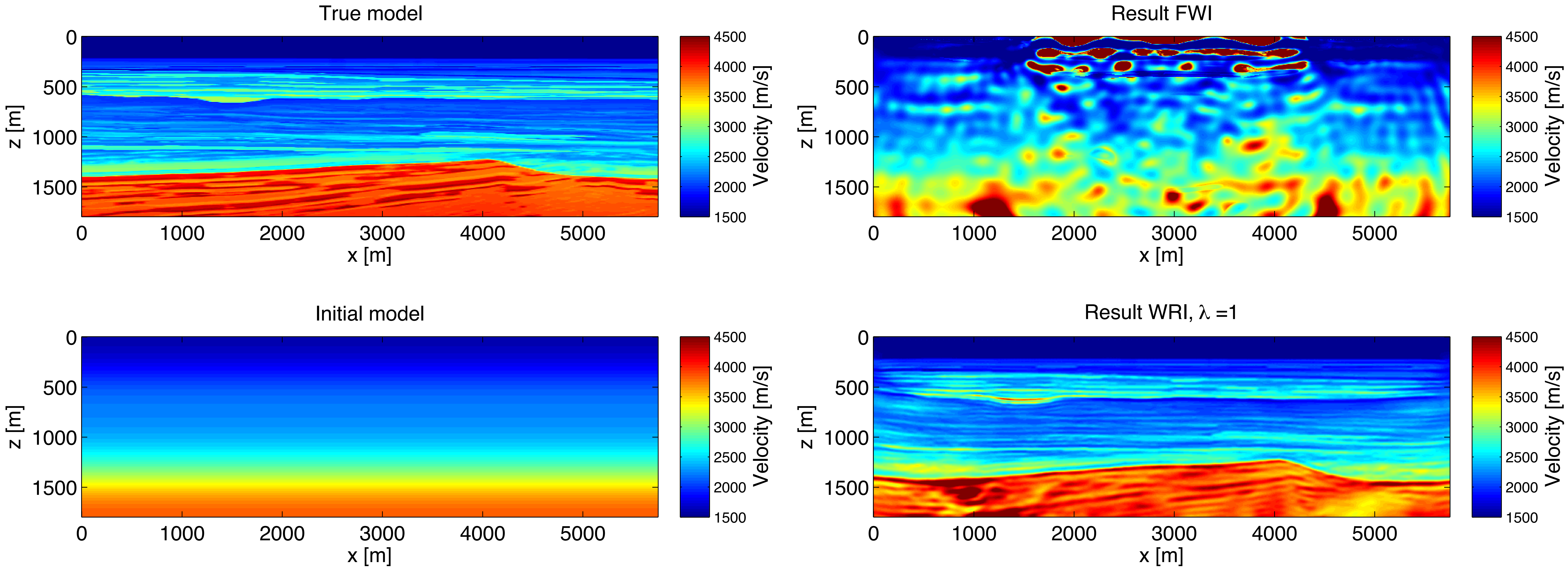
after 1st frequency batch



final result

True, initial and final models

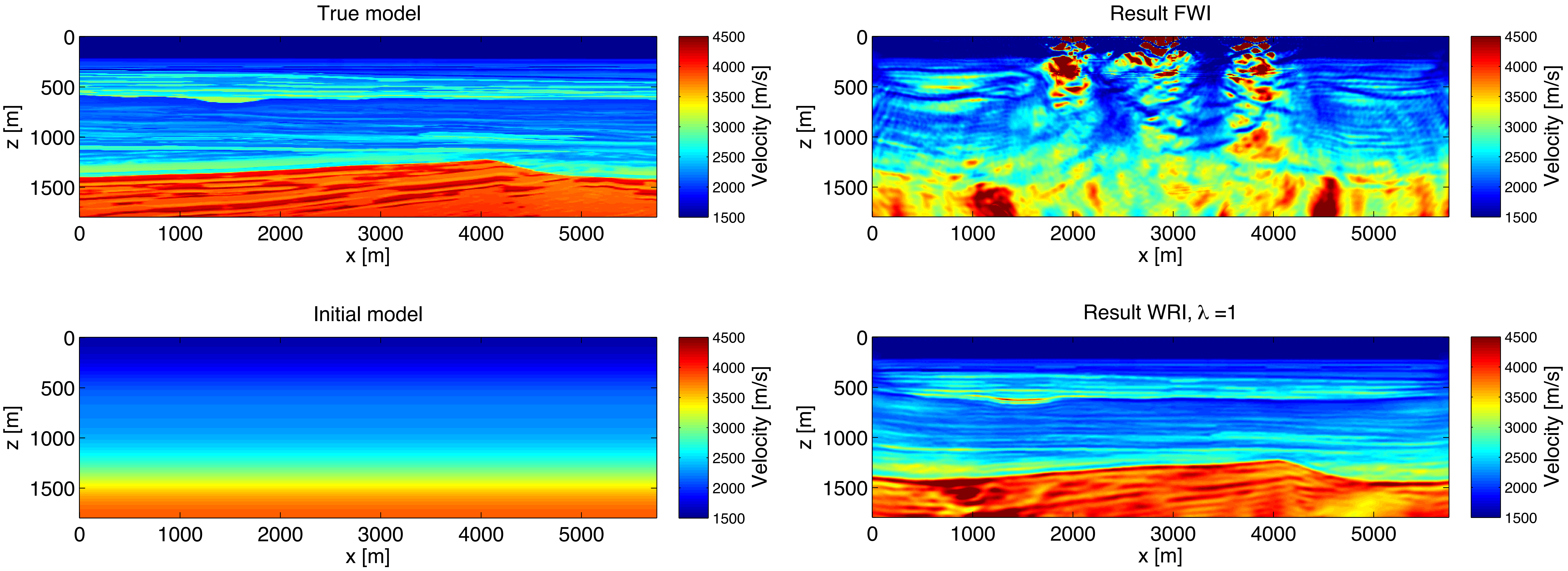
after 2nd frequency batch



final result

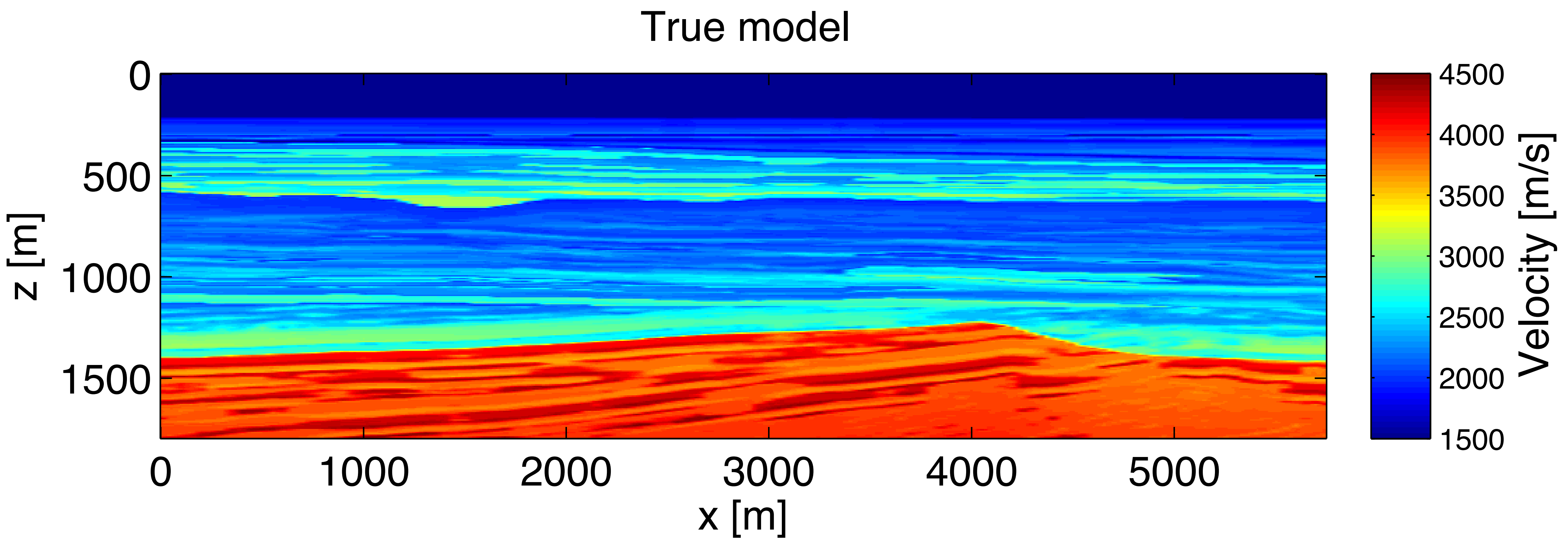
True, initial and final models

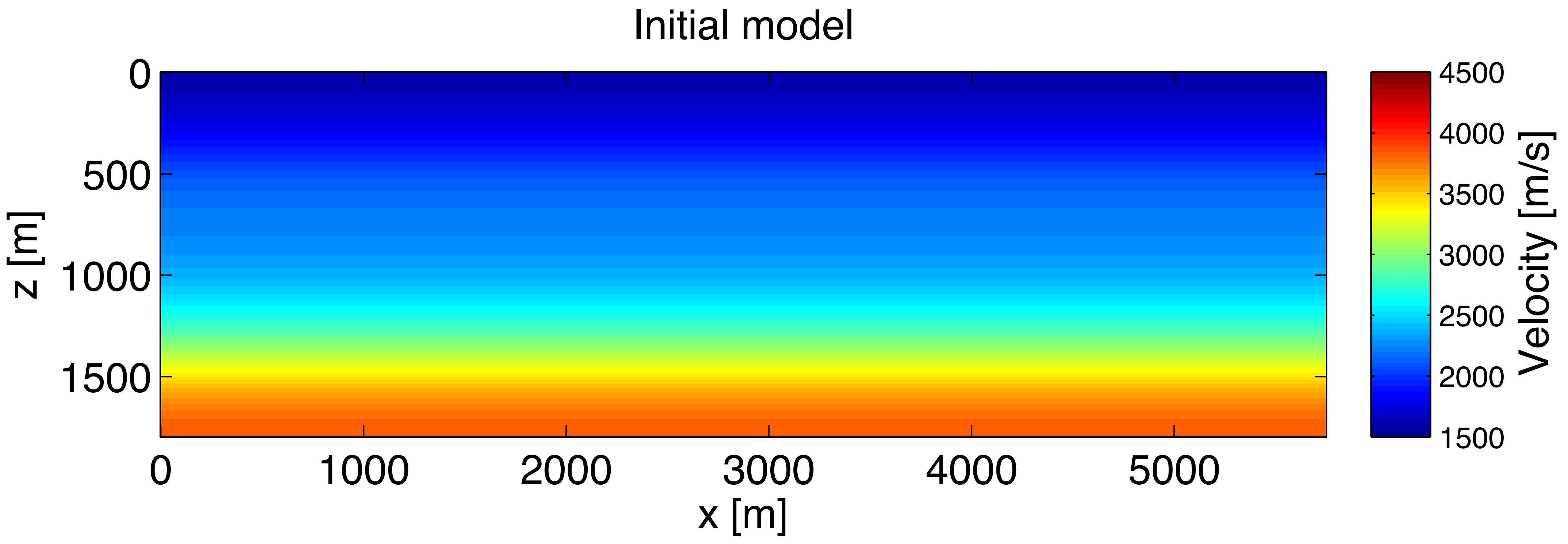
after last frequency batch



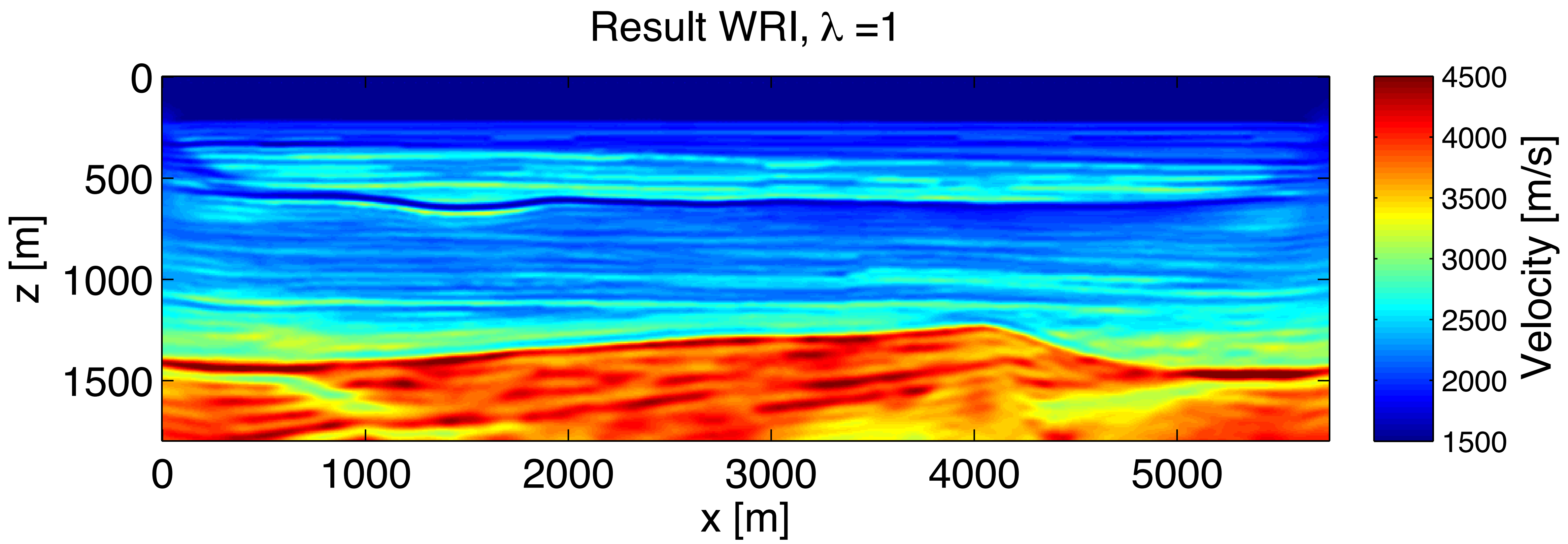
final result

True, initial and final models



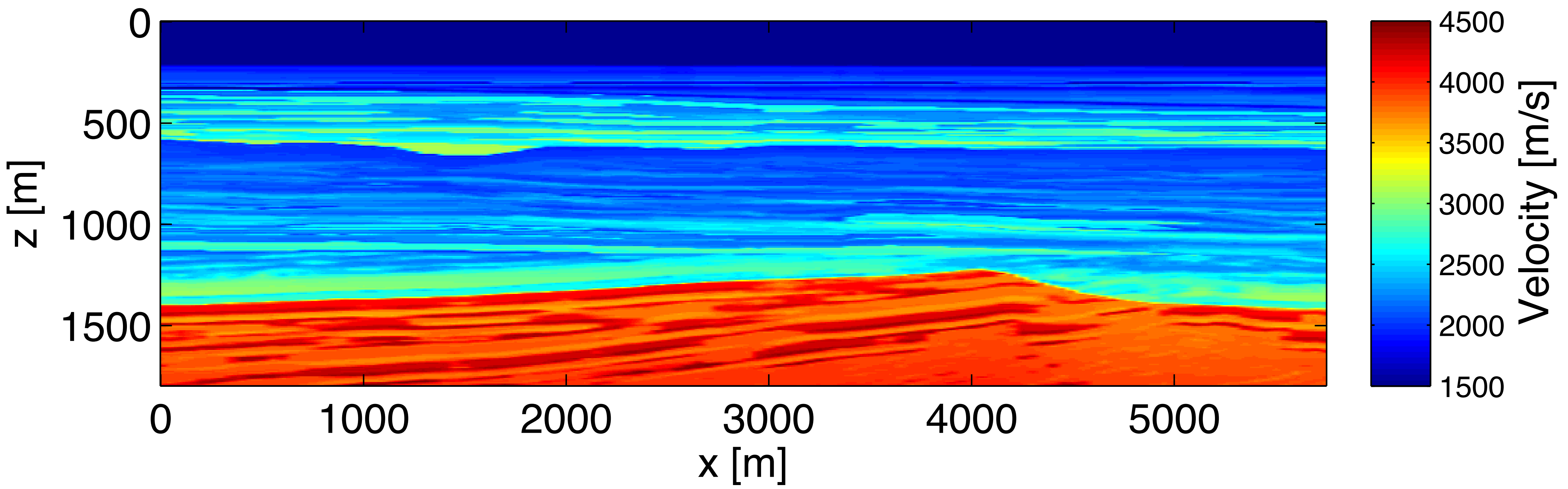


After 1st cycle

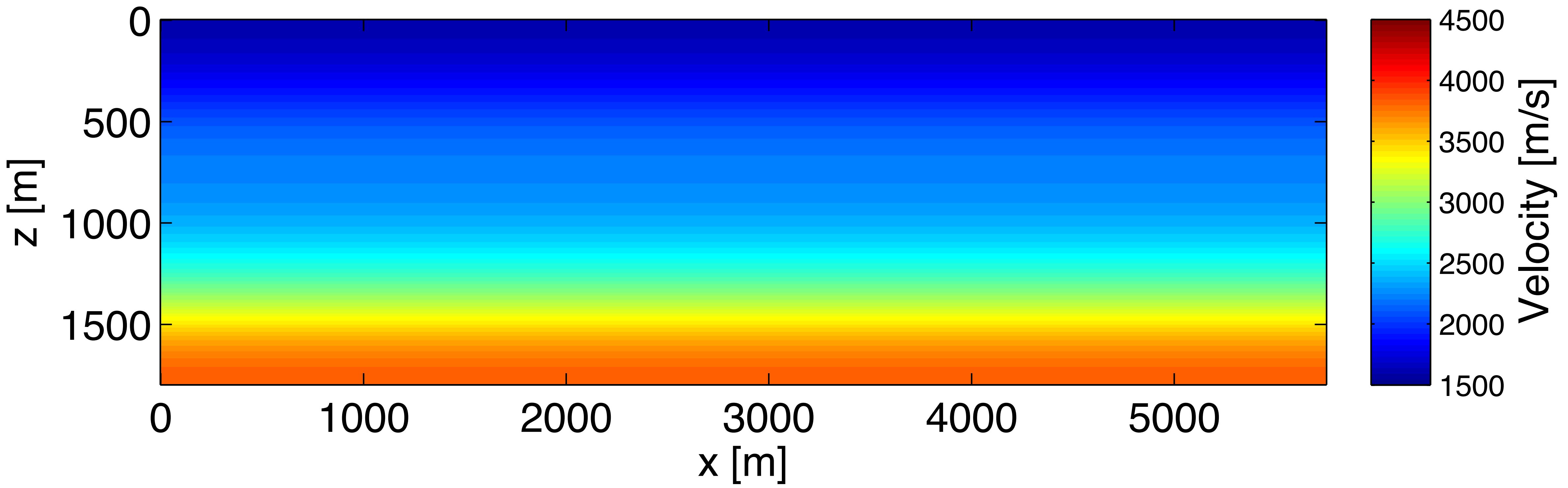


True, initial and final models

True model

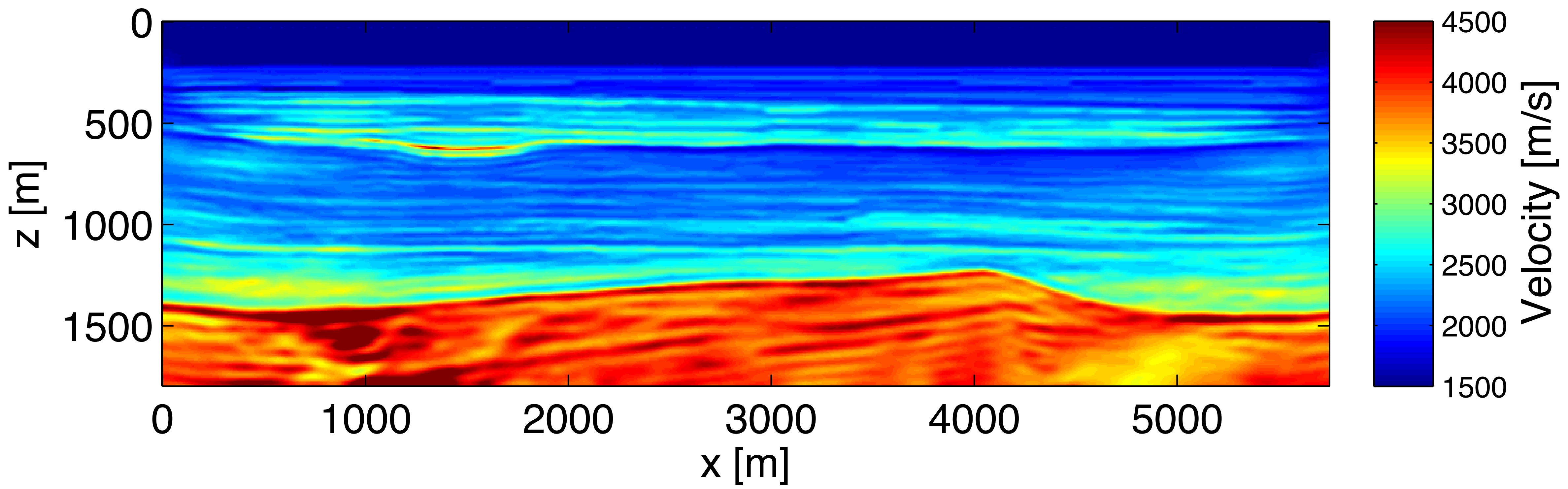


Initial model

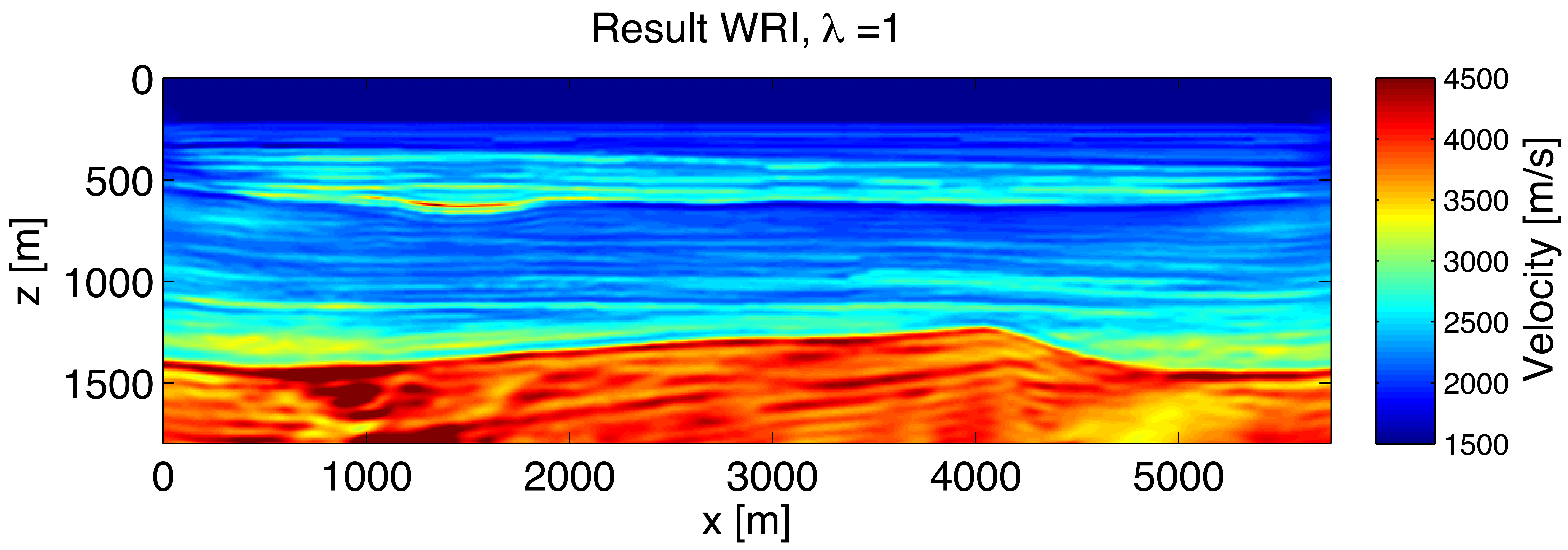


After 2nd cycle

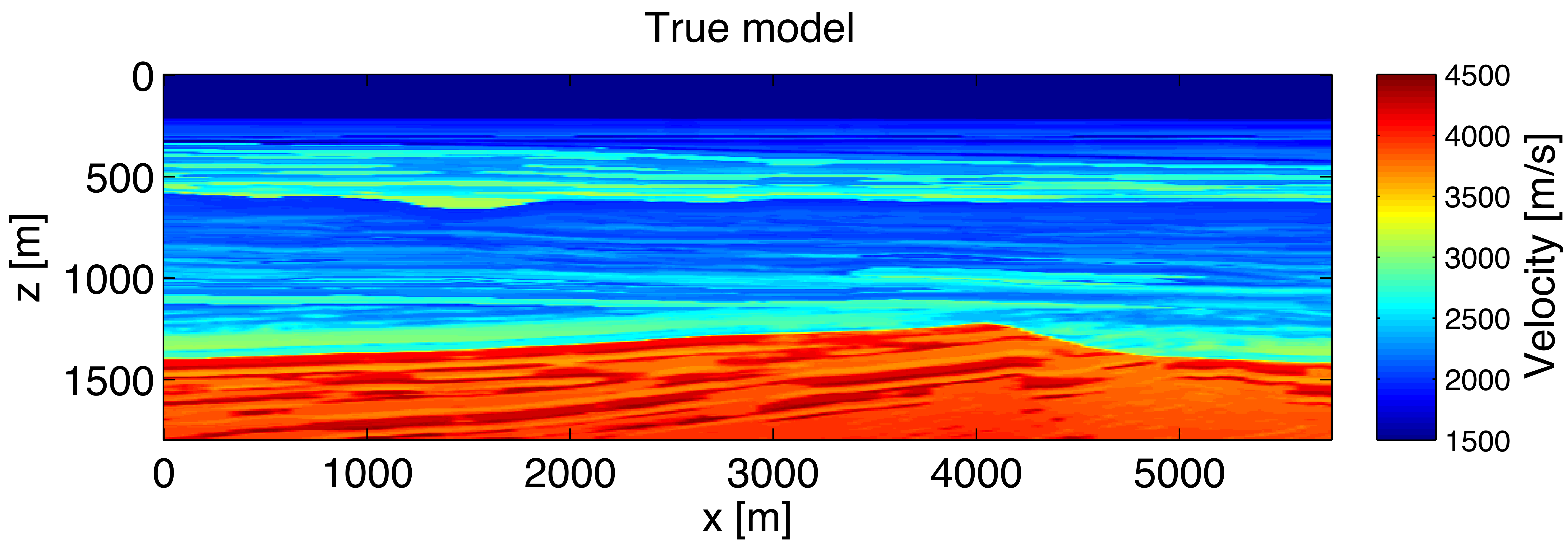
Result WRI, λ =1



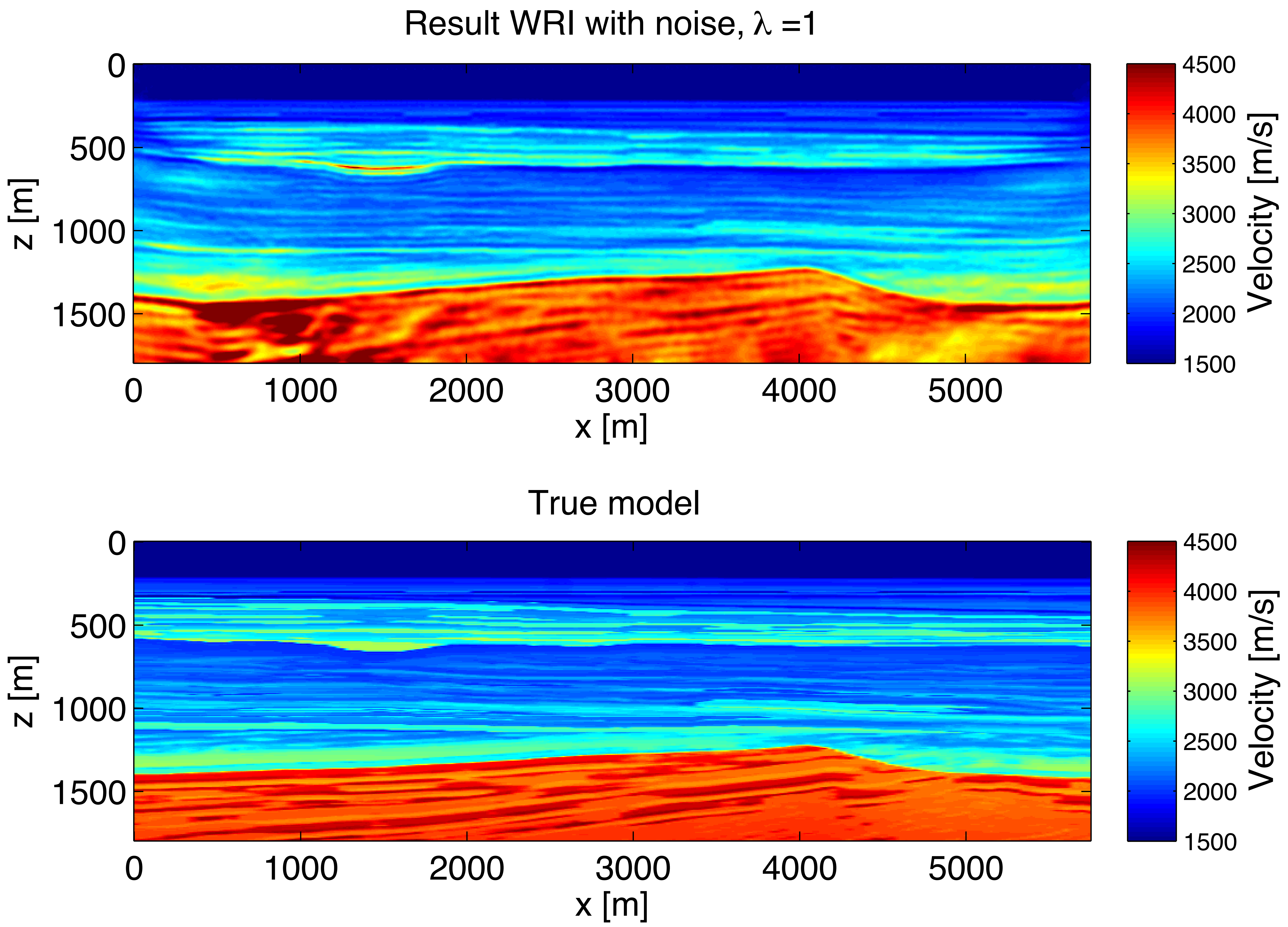
True and final models



No noise



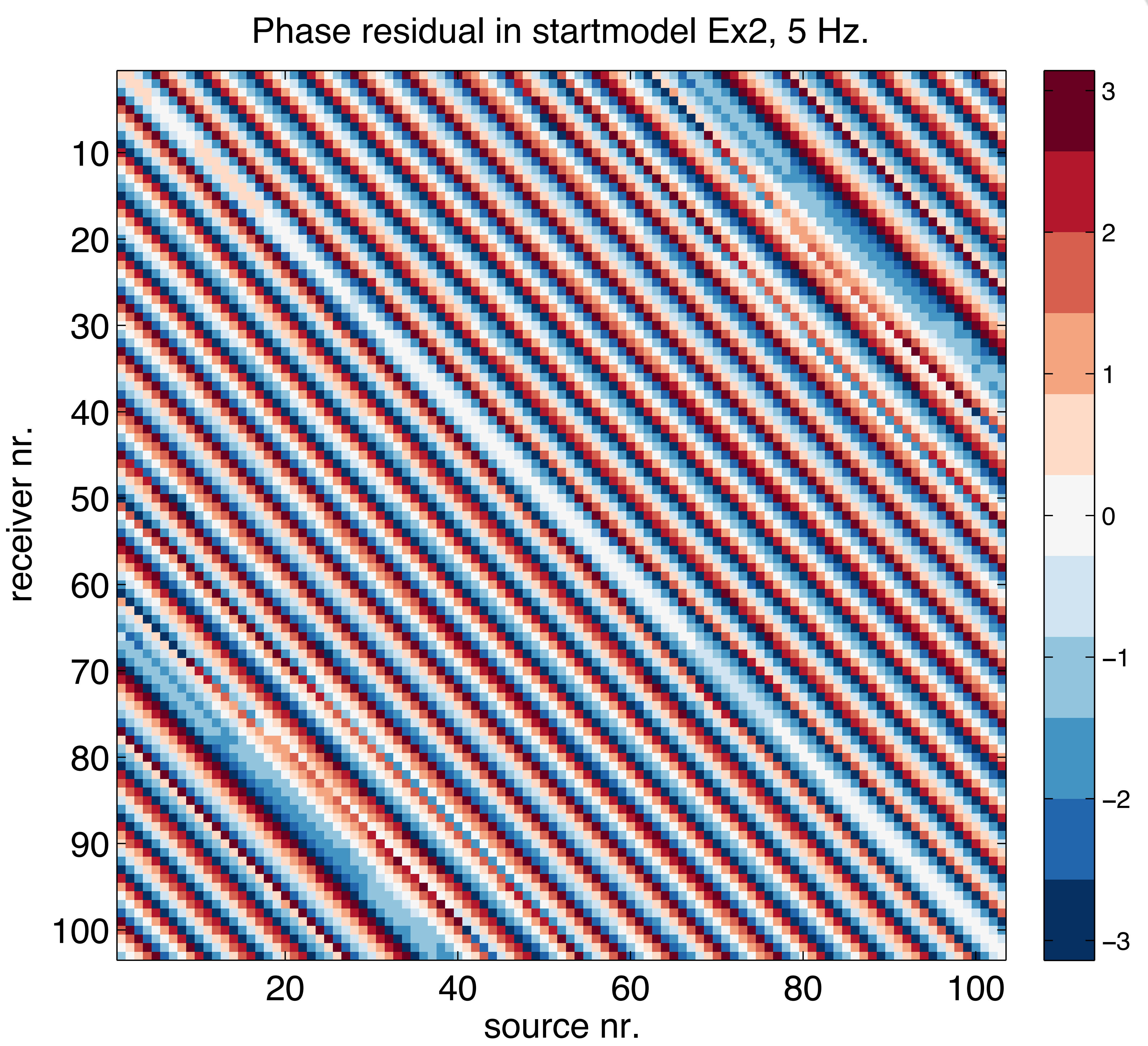
True and final models



$$\frac{\|\text{noise}\|}{\|\text{data}\|} = 0.04$$

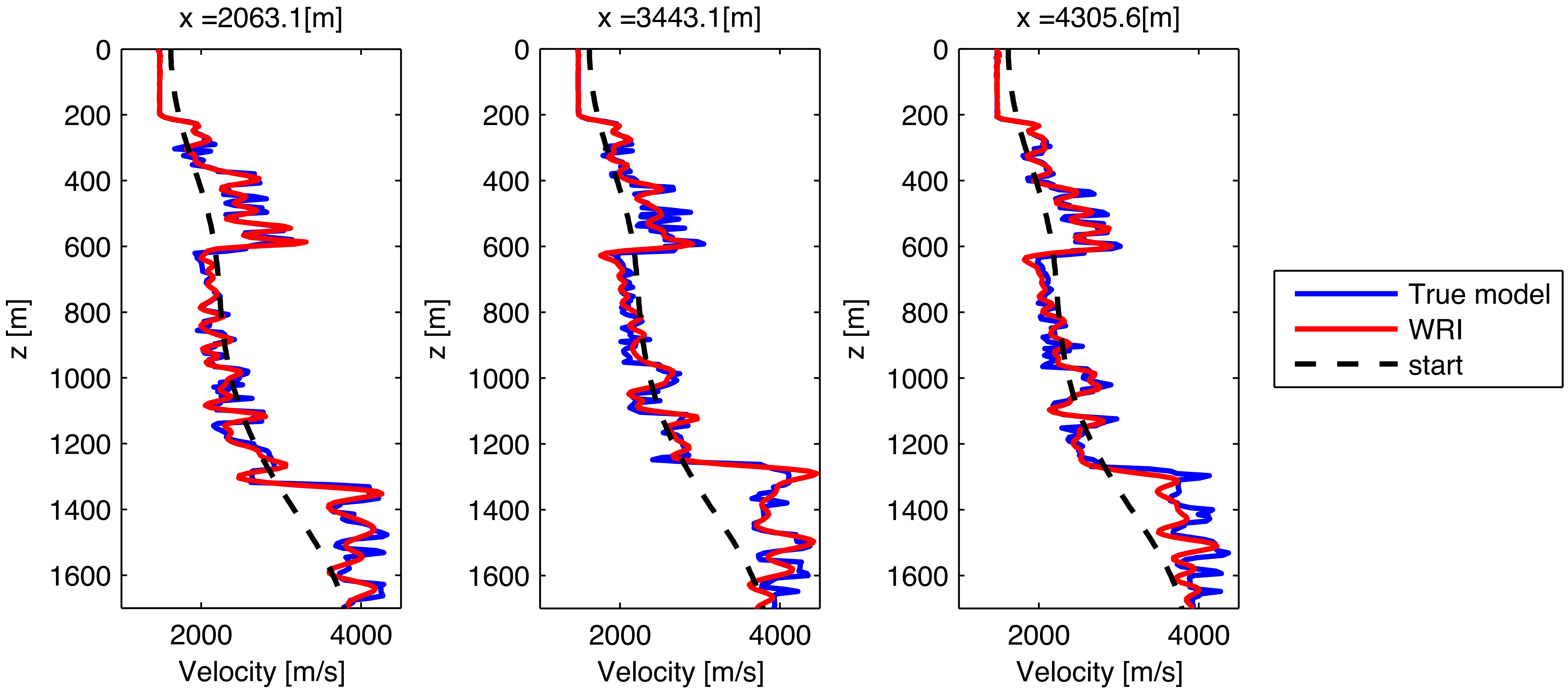
Initial phase-residuals

- Phase residuals computed using the Helmholtz equation in the start model
- WRI does not work with exact wavefields
- WRI uses the 'data-augmented' wavefield
- for λ small enough, the phase residual will be 0.

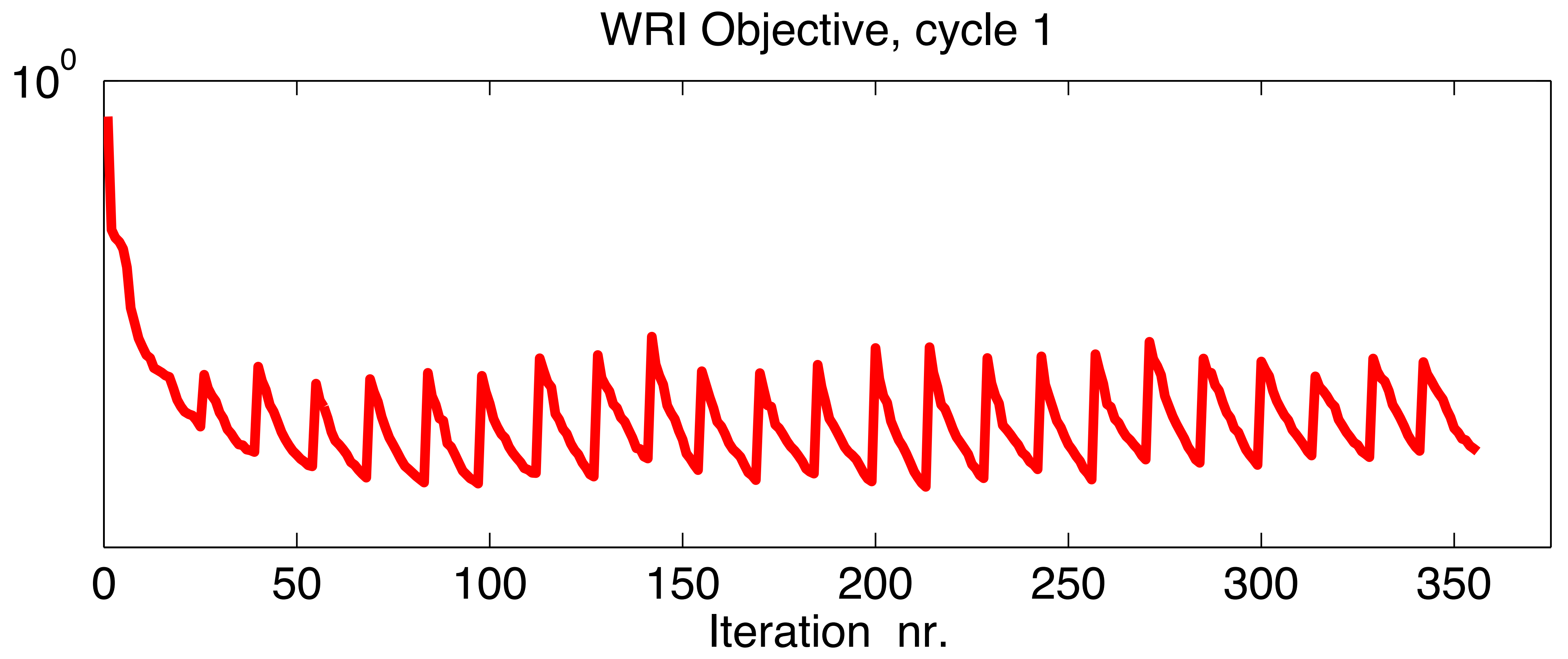


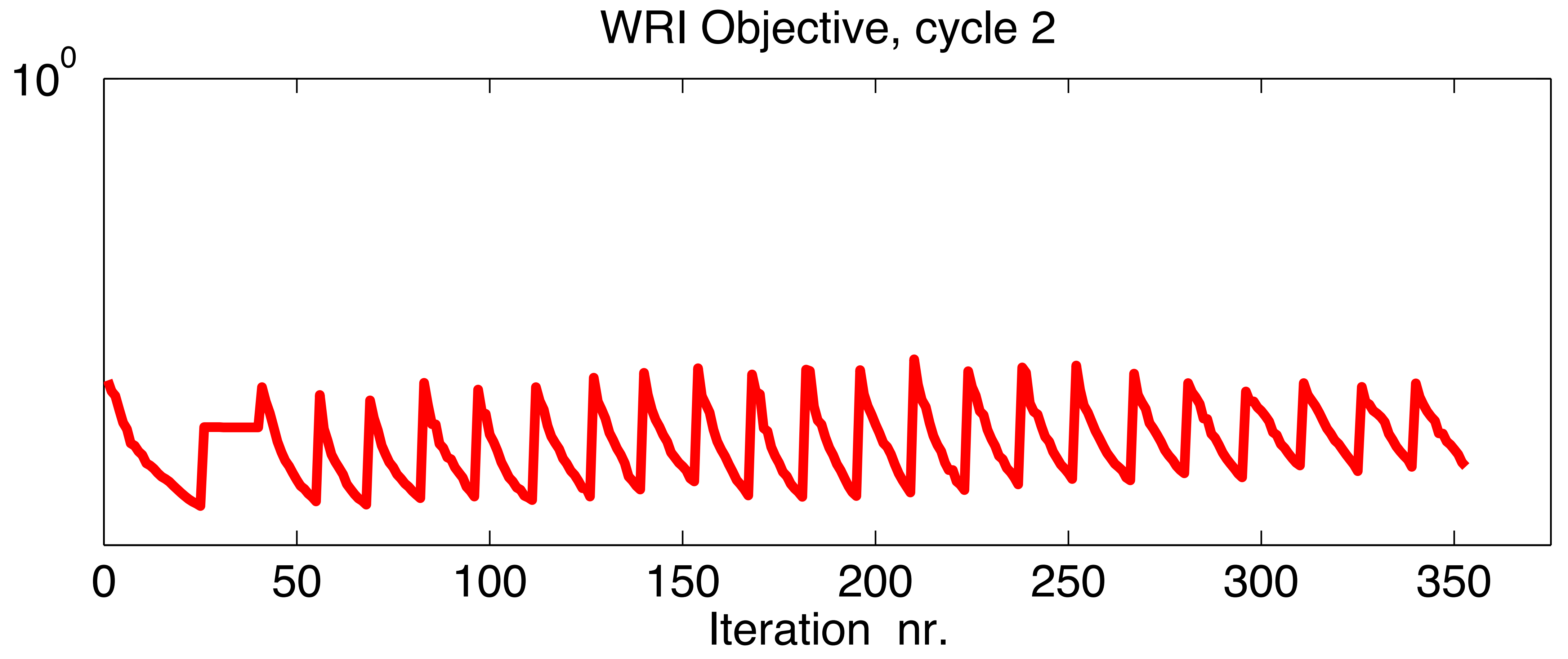
$$\bar{\mathbf{u}}_{kl} = \arg\min_{\mathbf{u}_{kl}} \left\| \begin{pmatrix} \lambda H_k(\mathbf{m}) \\ P \end{pmatrix} \mathbf{u}_{kl} - \begin{pmatrix} \lambda \mathbf{q}_{kl} \\ \mathbf{d}_{kl} \end{pmatrix} \right\|_2$$

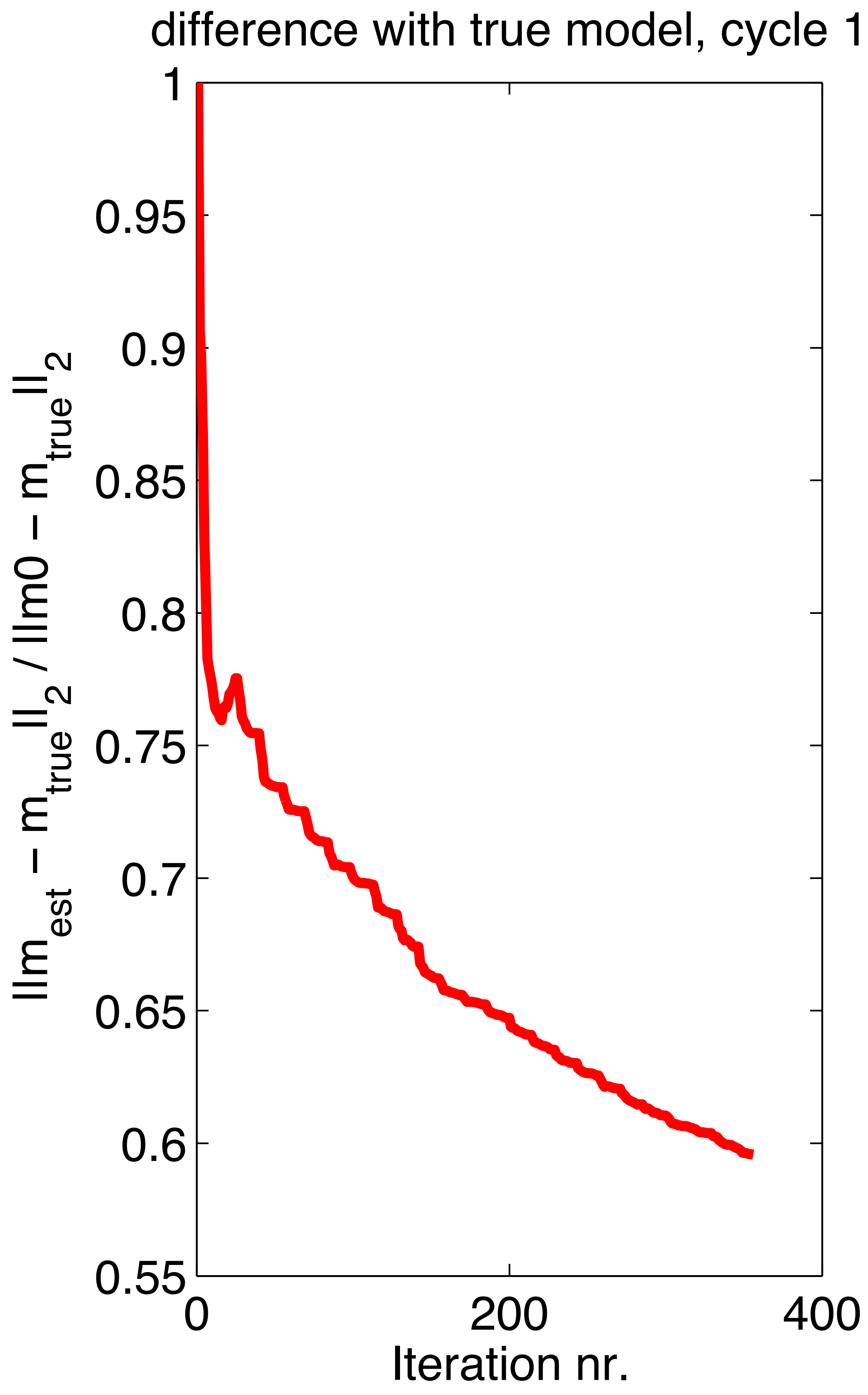
Cross sections

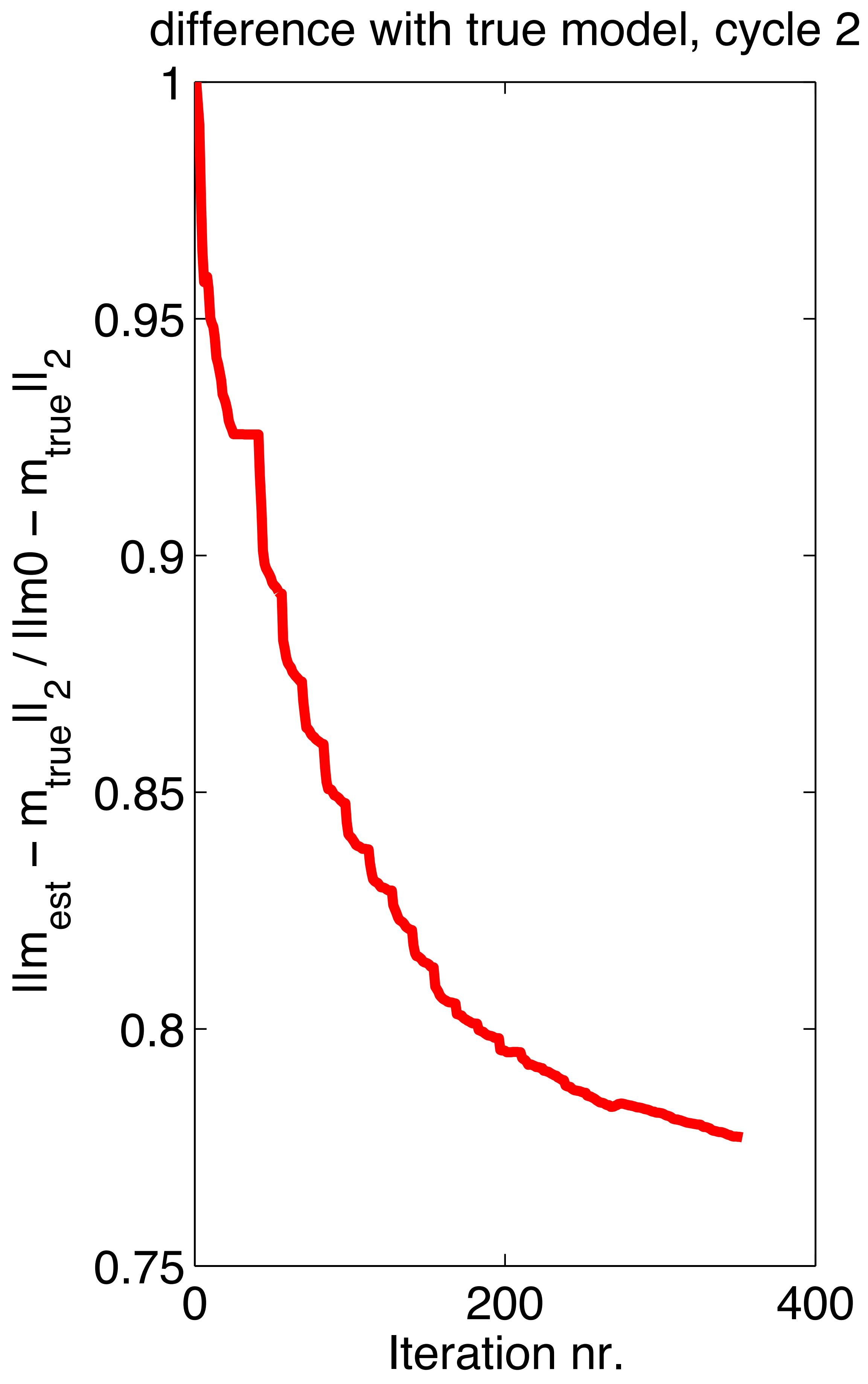


Objective and model error









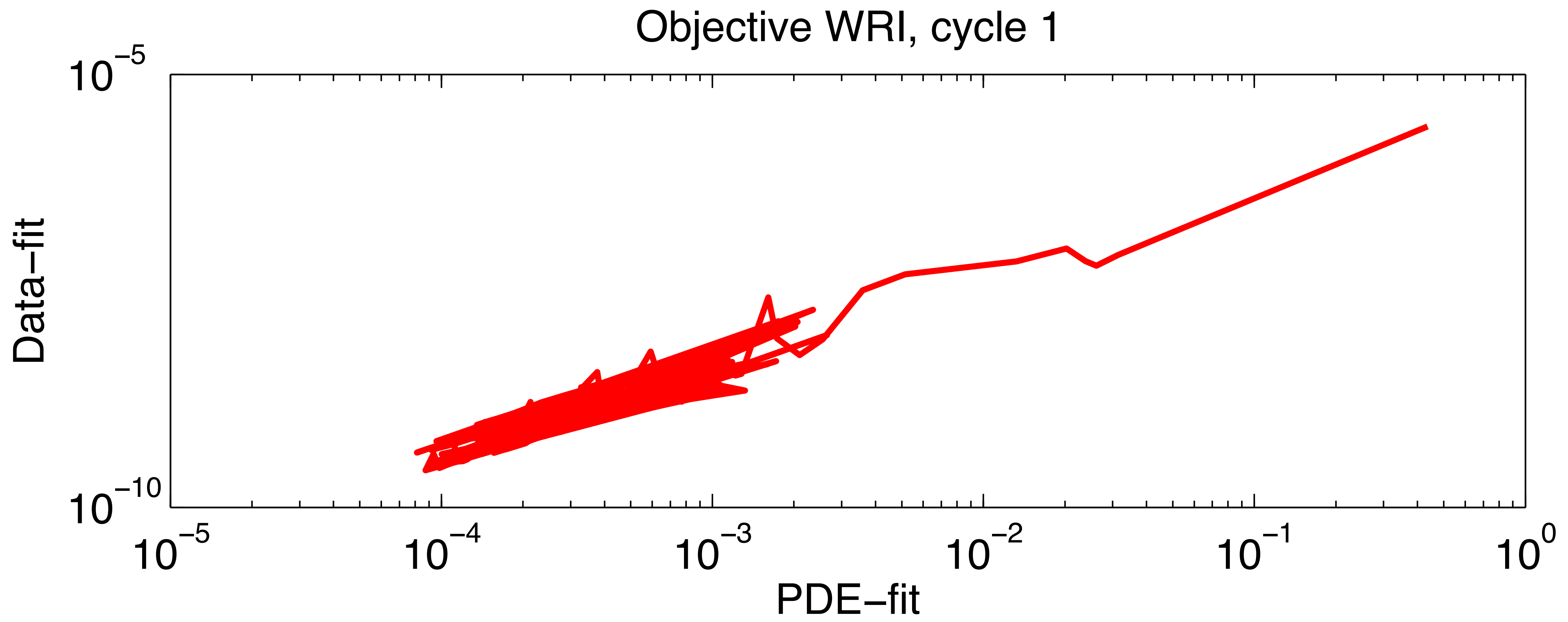
Objective

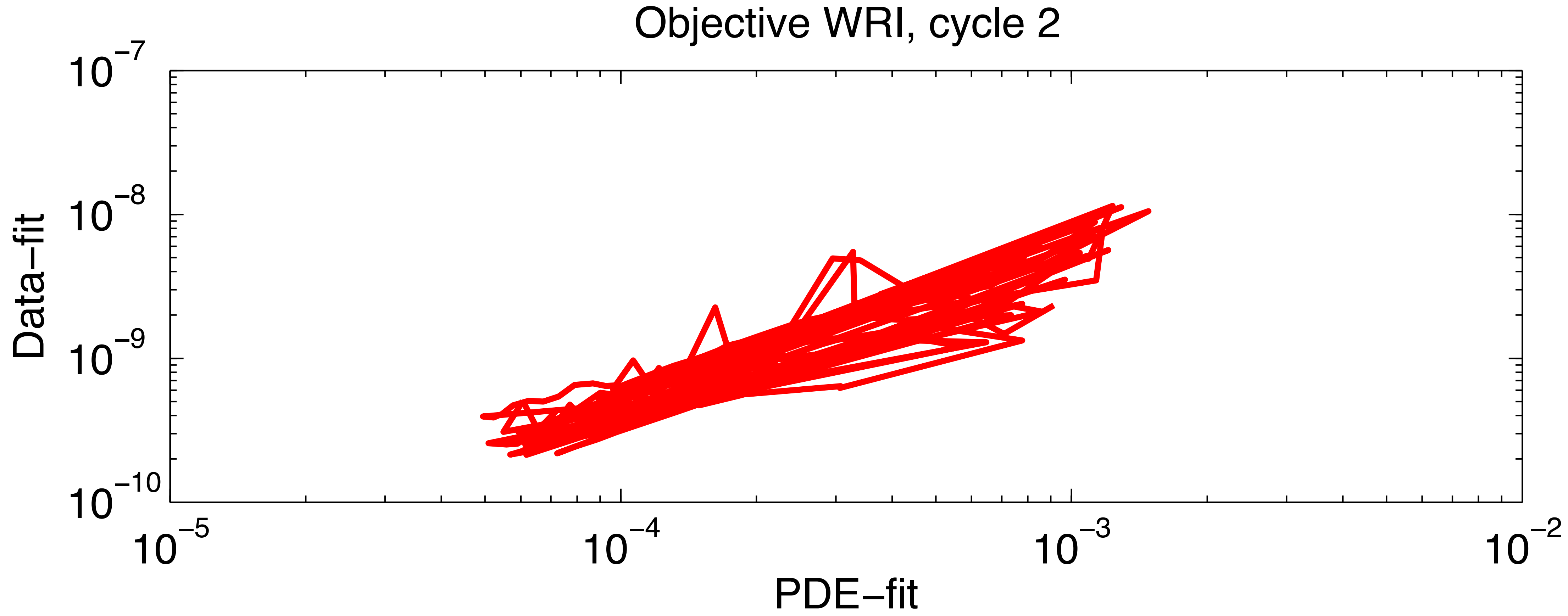
Data-misfit

PDE-misfit

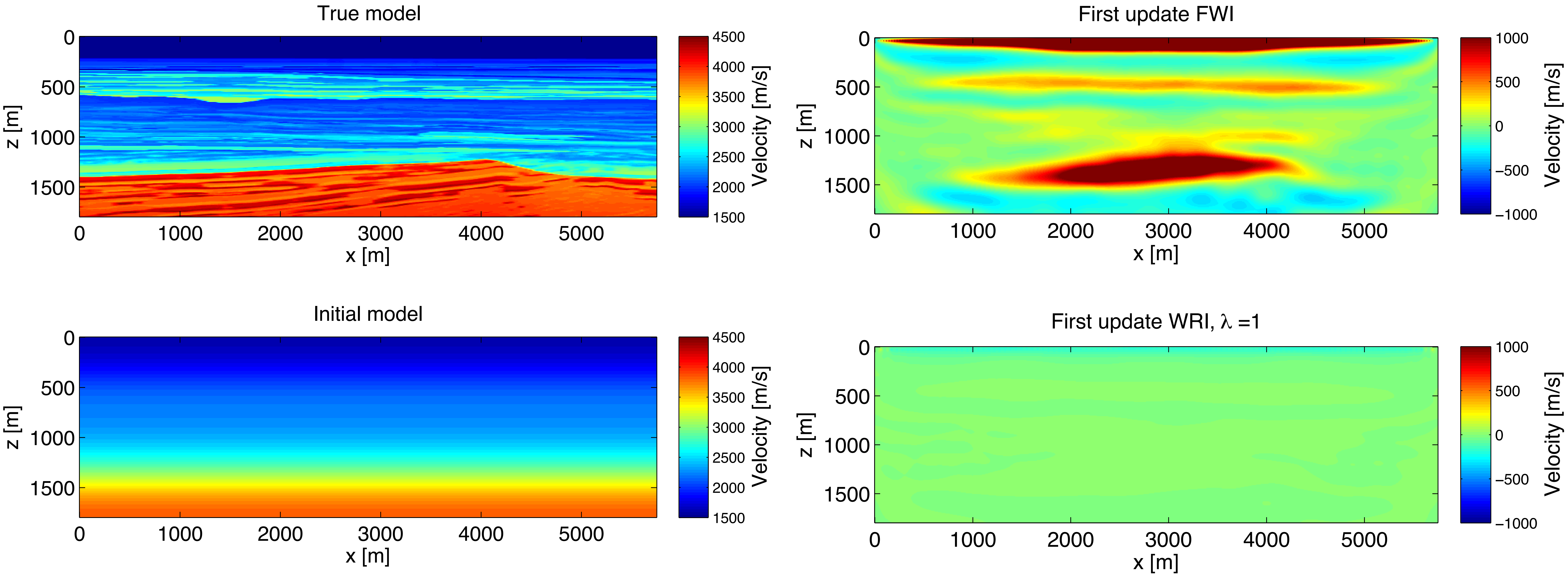
$$\bar{\phi}_\lambda(\mathbf{m}) = \frac{1}{2}\sum_{kl} \|P\bar{\mathbf{u}}_{kl} - \mathbf{d}_{kl}\|_2^2 + \frac{\lambda^2}{2}\|H_k(\mathbf{m})\bar{\mathbf{u}}_{kl} - \mathbf{q}_{kl}\|_2^2$$

- We can take a look at each part separately:
- Data-misfit can go up while iterating!

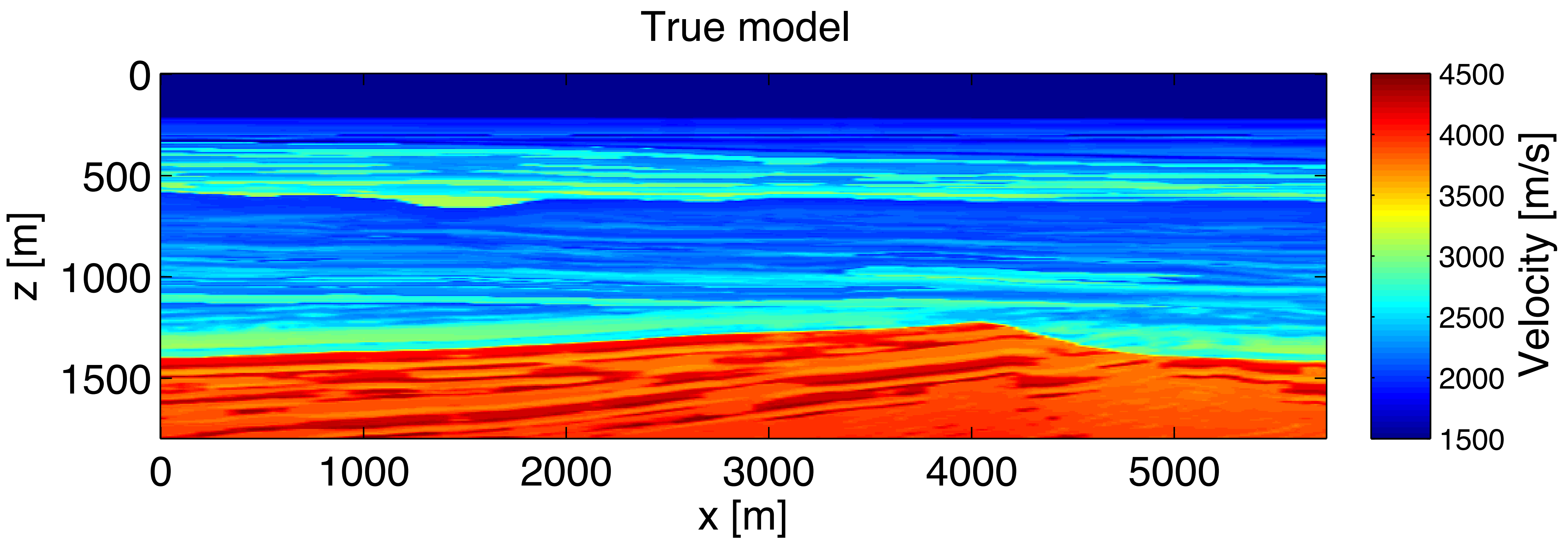


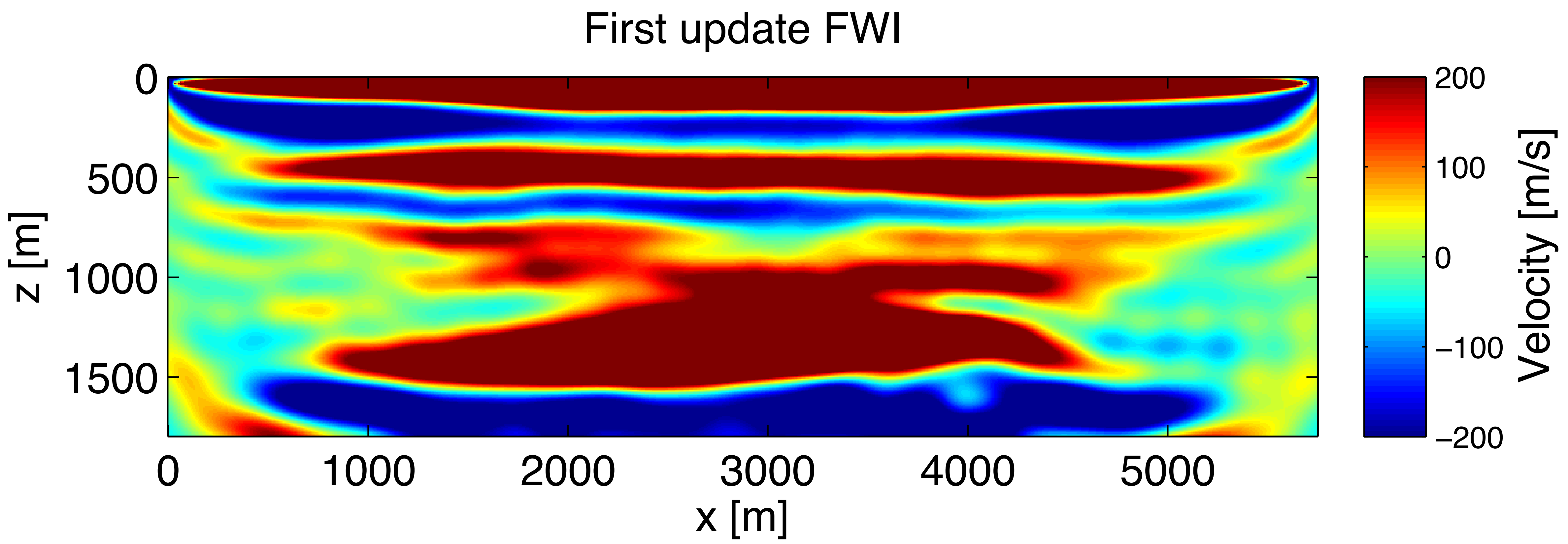


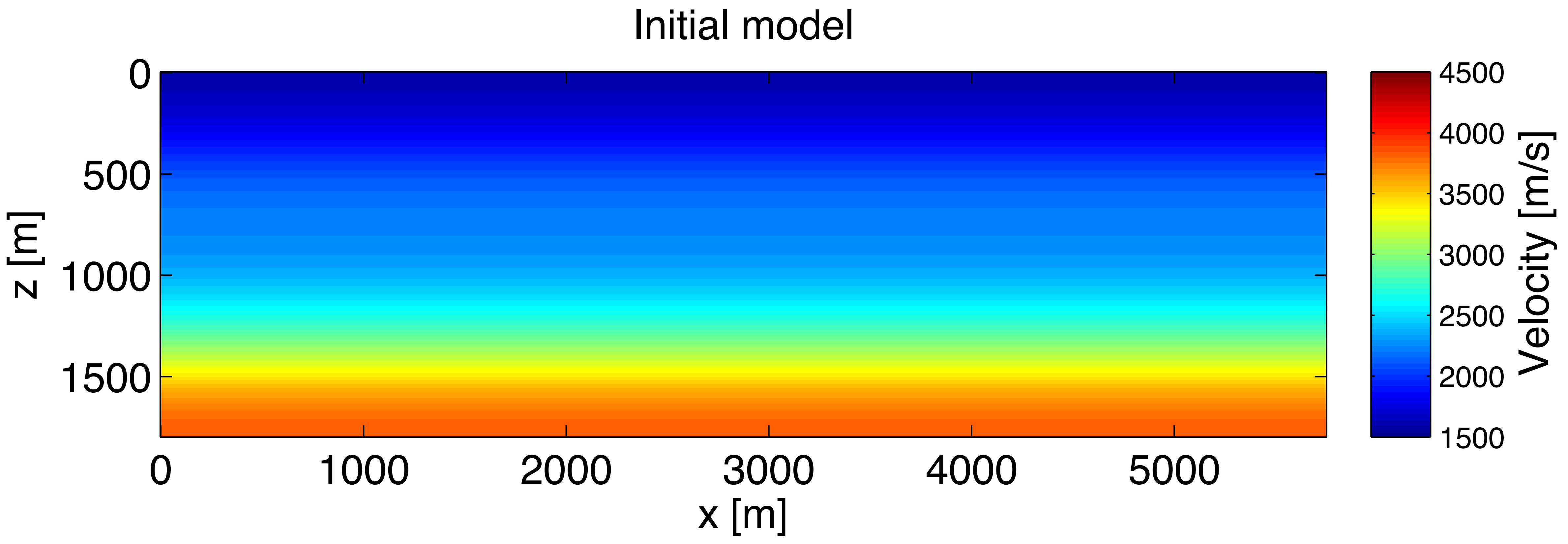
A look at the first update...

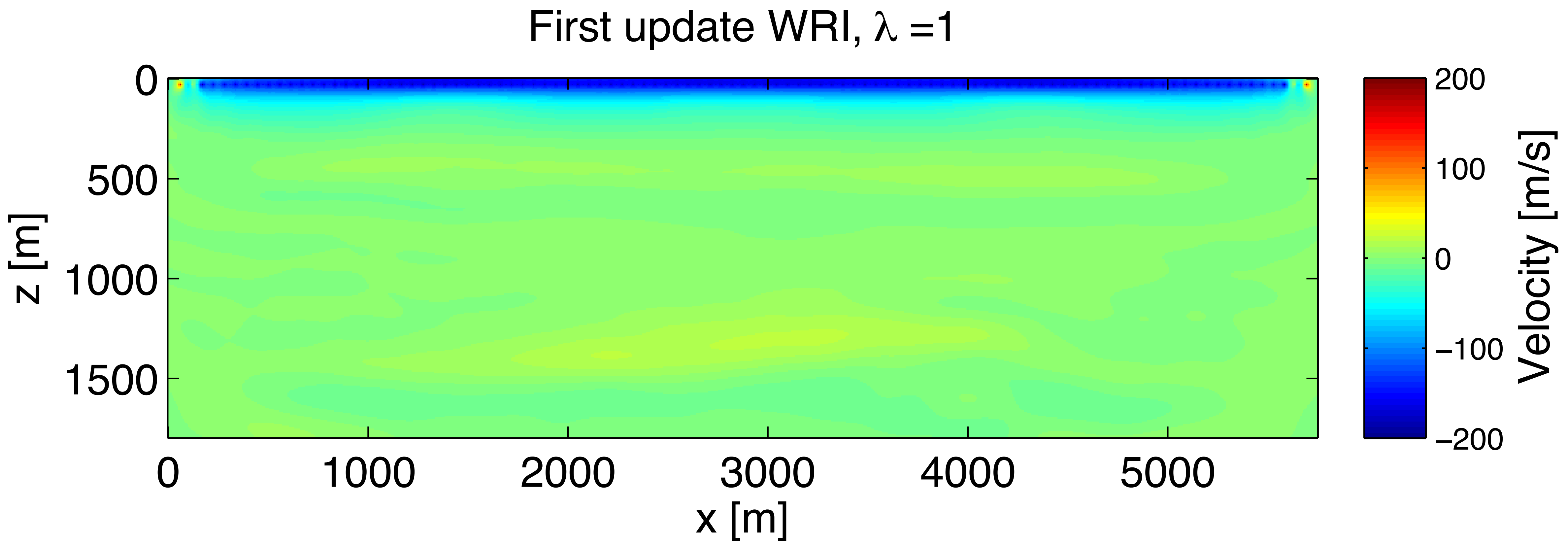


A look at the first update...









Different color scale

Observations about waveform inversion

- WRI performs much better for difficult problems
- WRI performs similar to FWI for not so difficult problems.

- Even for more difficult problems, only frequency continuation is required.
- No penalty parameter continuation was used, which can potentially increase quality and decrease the number of iterations.

Conclusions

WRI vs. FWI

- WRI:
  - Much better waveform inversion results for some difficult problems.
  - Less sensitive to missing low frequencies and poor start models.
  - Similar results for easy problems.
  - Not more sensitive to noise.
  - No hidden parameters or fine-tuning of settings, just choose λ.
  - Passing though the data twice can be beneficial for this method.

  - 1 least-squares problem instead of the usual 2 PDE solves.

Work in progress..

- Application of WRI to a land data set, results look quite promising and consistent with the FWI results.
- Assess the added value to imaging.

Outlook

- Find most efficient ways to solve the least-squares problem (also using iterative methods
- Find ‘optimal’ combination of tradeoff parameter λ and inversion set up. (Results in this talk may not be the best possible)
- The WRI objective function offers some interesting possibilities for multi-parameter inversion which Lagrangian based methods do not offer. (Will be presented at the SEG meeting in Denver later this year)

Acknowledgements

The SLIM students & postdocs

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, BP, CGG, Chevron, ConocoPhillips, ION, Petrobras, PGS, Total SA, WesternGeco, and Woodside.