MCMC-based time-lapse FWI

Xin Fu* and Kris Innanen

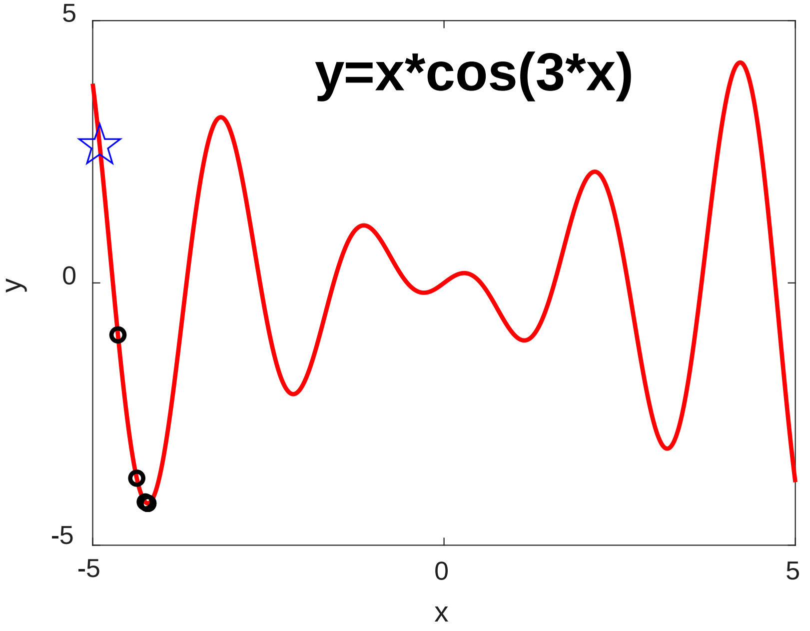
University of Calgary, CREWES Project

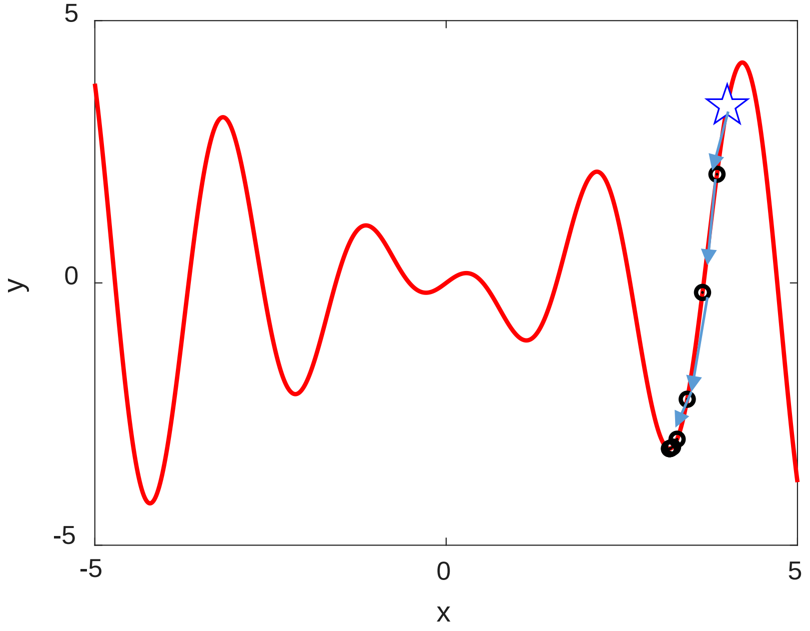
Dec 3, 2020

Outline

- Introduction
- Deterministic optimization (DO) Full-waveform inversion (FWI)
- Bayesian inference based on Markov chain Monte Carlo (MCMC)
- Inversion Strategies
- Numerical examples
- Conclusions

Introduction



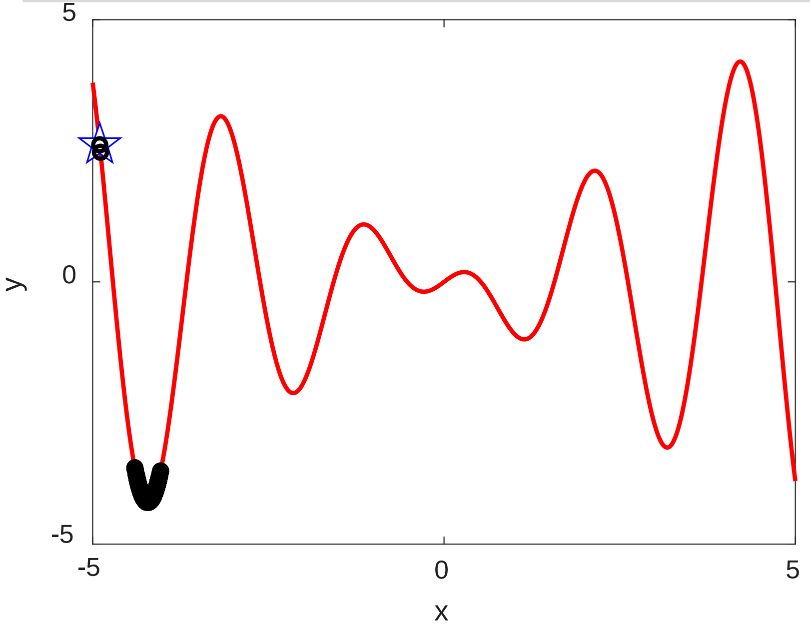


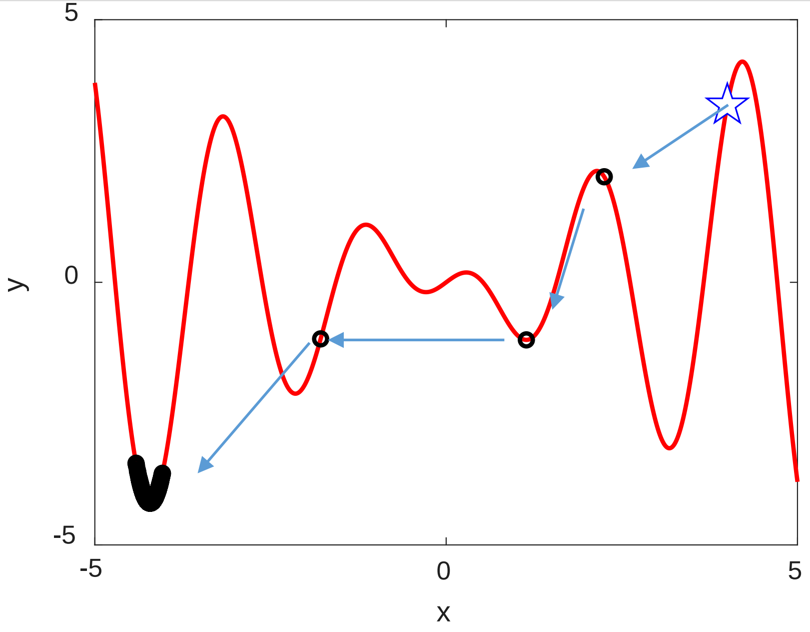
Deterministic Optimization:

Fast

Large model size

Model-dependent





Stochastic Global Optimization:

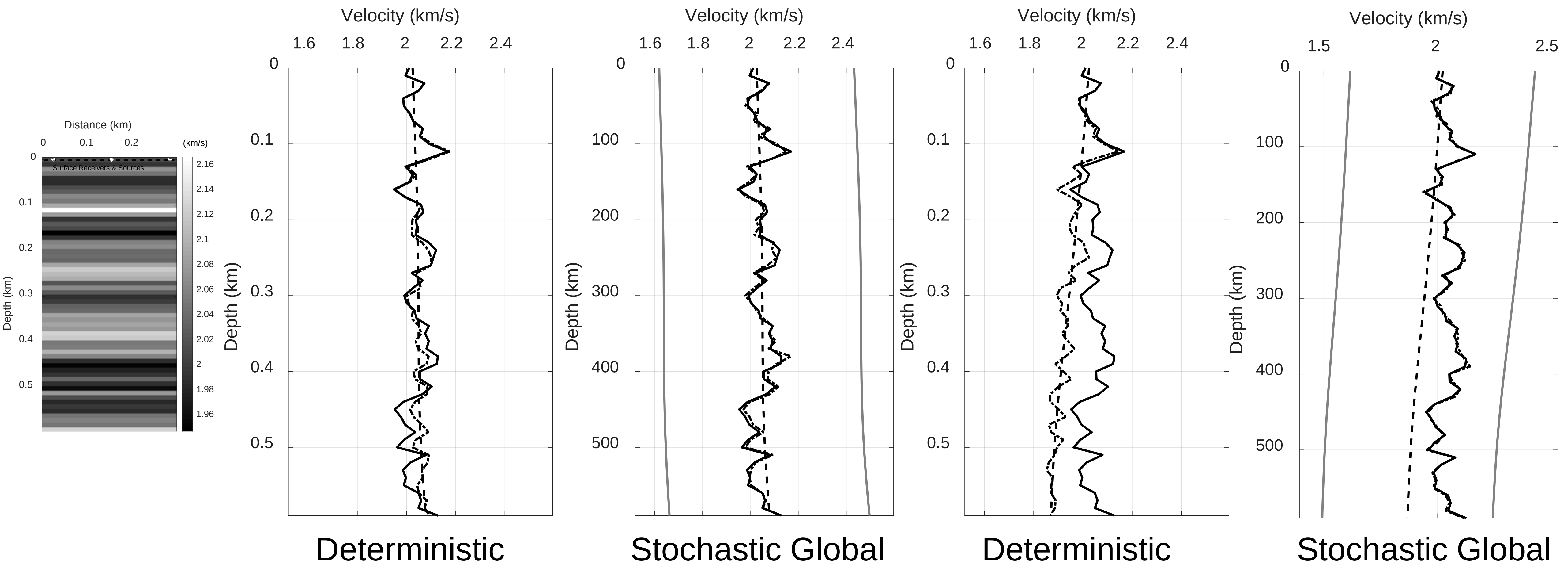
Slow

Dimension-limited

Not model-dependent

Uncertainty qualification

1.5D model examples

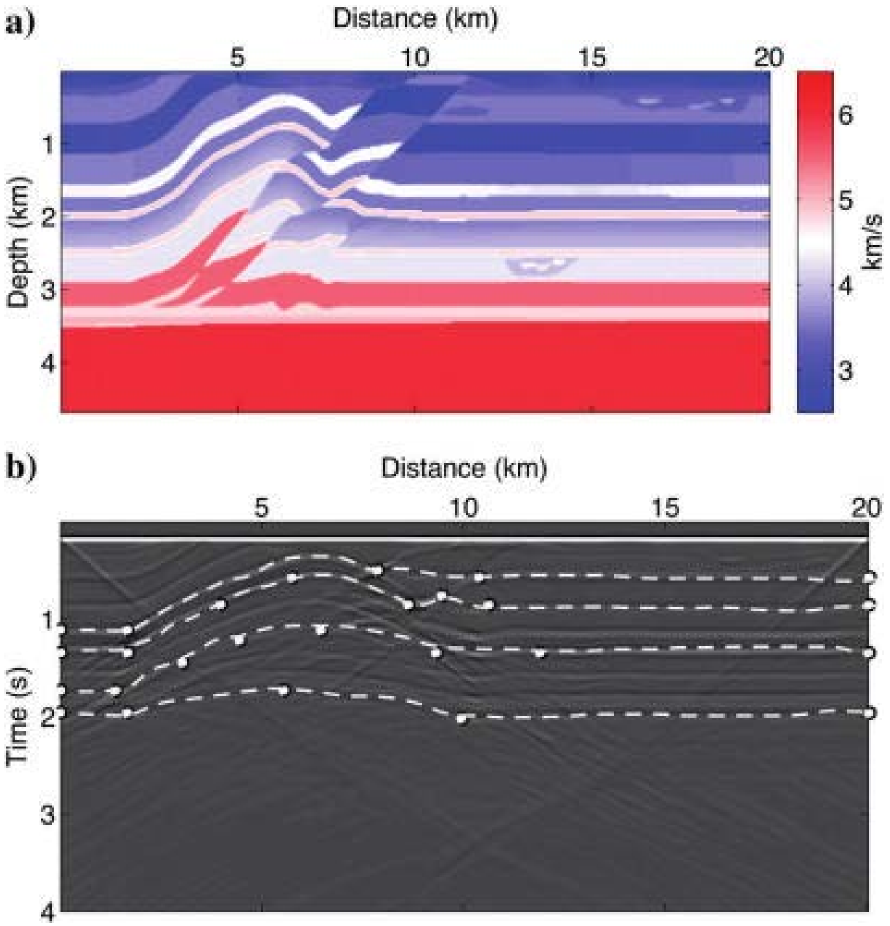


Good initial model

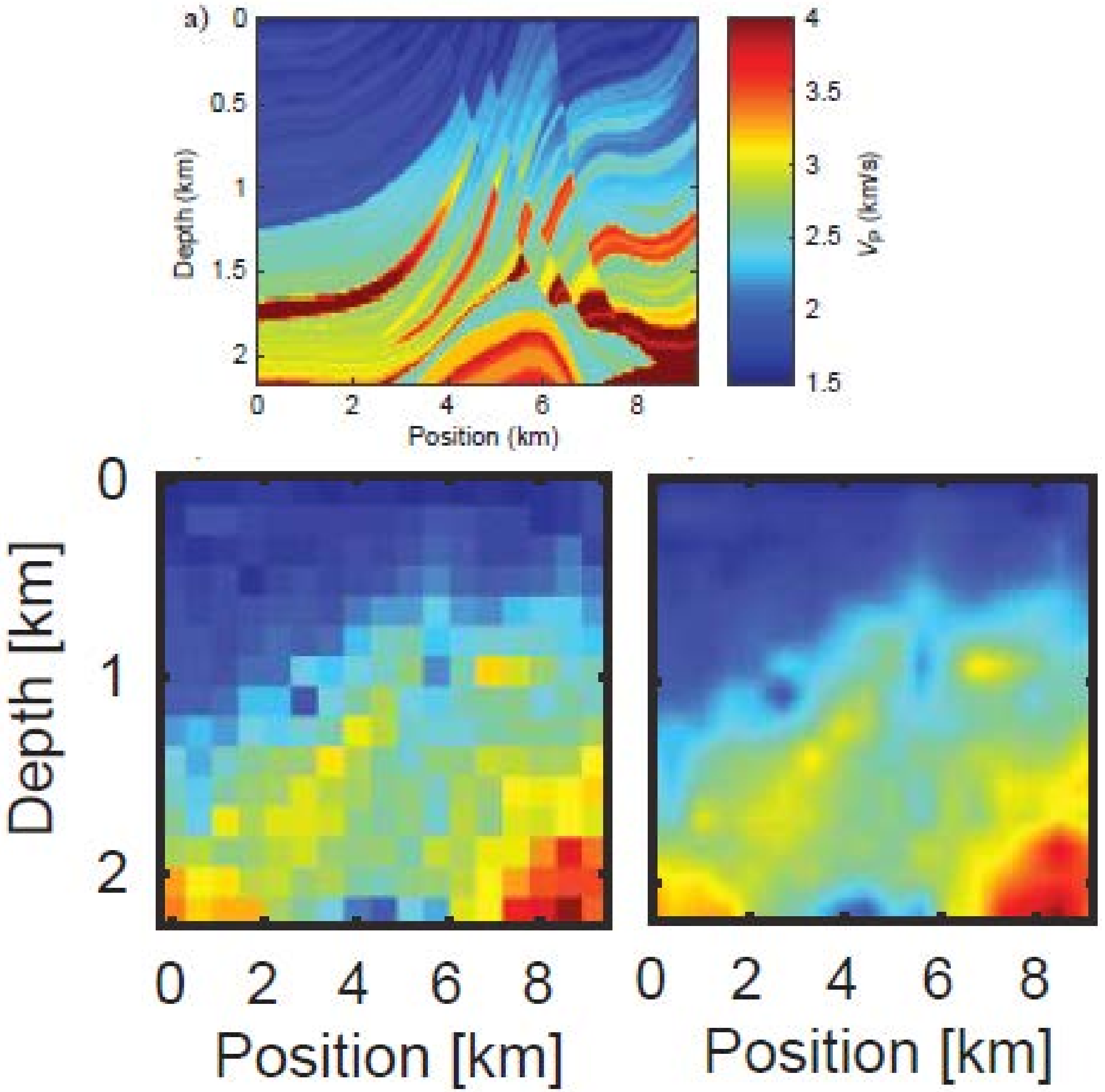
Bad initial model

Fu, X., and Innanen, K., 2020, A new parallel simulated annealing algorithm for 1.5D acoustic full-waveform inversion: the 32nd Annual Research Report of the CREWES Project.

2D model examples: sparse parameterization



Datta and Sen, 2016



Sajeva et al., 2016

Outline

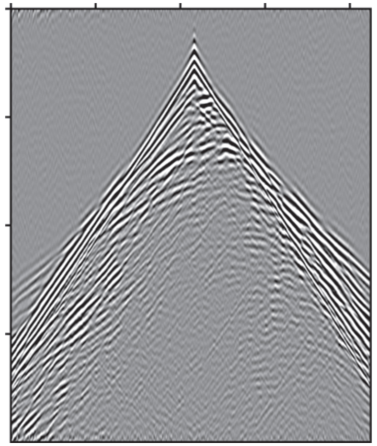
- Introduction
- Deterministic optimization (DO) Full-waveform inversion (FWI)
- Bayesian inference based on Markov chain Monte Carlo (MCMC)
- Inversion Strategies
- Numerical examples
- Conclusions

DO FWI

Objective function: $$E(\mathbf{m})=\frac{1}{2}(\mathbf{d}_{syn}(\mathbf{m})-\mathbf{d}_{obs})^T(\mathbf{d}_{syn}(\mathbf{m})-\mathbf{d}_{obs})$$

$$\frac{1}{v^2(\mathbf{x})}\frac{\partial^2 P(\mathbf{x},t)}{\partial t^2}-\nabla^2 P(\mathbf{x},t)=s(t)\delta(\mathbf{x}-\mathbf{x}_s)$$

Shot gathers



Model update:
(steepest descent)

$$\mathbf{m}=\mathbf{m}_0+\Delta\mathbf{m}$$
$$\Delta\mathbf{m}=-\mu\mathbf{g}$$

Adjoint method + deconvolution imaging condition:

$$\Delta v(\mathbf{x})=-\mu\sum_{r=1}^{ng}\sum_{i=1}^{ns}\frac{2}{v(\mathbf{x})^3}\frac{\int_0^{t_{max}}dt[\ddot{P}_f(\mathbf{x},t;\mathbf{x}_s)P_b(\mathbf{x},t;\mathbf{x}_r)]}{\int_0^{t_{max}}dt[P_f(\mathbf{x},t;\mathbf{x}_s)P_f(\mathbf{x},t;\mathbf{x}_s)+\lambda I_{max}]}$$

Outline

- Introduction
- Deterministic optimization (DO) Full-waveform inversion (FWI)
- Bayesian inference based on Markov chain Monte Carlo (MCMC)
- Inversion Strategies
- Numerical examples
- Conclusions

Bayesian inference

Bayes' rule:

Posterior probability: $p(\mathbf{m}\,|\,\mathbf{d}_{obs}) = \dfrac{p(\mathbf{d}_{obs}\,|\,\mathbf{m})\,p(\mathbf{m})}{p(\mathbf{d}_{obs})} \propto p(\mathbf{d}_{obs}\,|\,\mathbf{m})\,p(\mathbf{m})$

Likelihood function (data prior):

$$p(\mathbf{d}_{obs}\,|\,\mathbf{m}) \propto \exp\{-\frac{1}{2}(\mathbf{d}_{syn}(\mathbf{m}) - \mathbf{d}_{obs})^T \mathbf{C}_d^{-1} (\mathbf{d}_{syn}(\mathbf{m}) - \mathbf{d}_{obs})\}$$

Model prior:

$$p(\mathbf{m}) \propto \exp\{-\frac{1}{2}(\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0)\}$$

Objective function:

$$p(\mathbf{m}\,|\,\mathbf{d}_{obs}) \propto \exp\{-\chi(\mathbf{m})\}$$

Misfit function:

$$\chi(\mathbf{m}) = \frac{1}{2}(\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0) + \frac{1}{2}(\mathbf{d}_{syn}(\mathbf{m}) - \mathbf{d}_{obs})^T \mathbf{C}_d^{-1} (\mathbf{d}_{syn}(\mathbf{m}) - \mathbf{d}_{obs})$$

Markov chain Monte Carlo (MCMC)

Metropolis-Hastings (MH) MCMC

ingredients

Proposal distribution $T(\mathbf{m}^* | \mathbf{m})$

Posterior distribution $p(\mathbf{m} | \mathbf{d}_{obs})$

Start with $\mathbf{m}_0$

For k=1, 2, 3, …

1, Use $T(\mathbf{m}^* | \mathbf{m})$ to produce the proposal model $\mathbf{m}^*$

2, Calculate the acceptance probability

$$a(\mathbf{m}^*, \mathbf{m}) = min(1, \frac{T(\mathbf{m} | \mathbf{m}^*) p(\mathbf{m}^* | \mathbf{d}_{obs})}{T(\mathbf{m}^* | \mathbf{m}) p(\mathbf{m} | \mathbf{d}_{obs})})$$

3, Accept $\mathbf{m}^*$ according to probability a

end

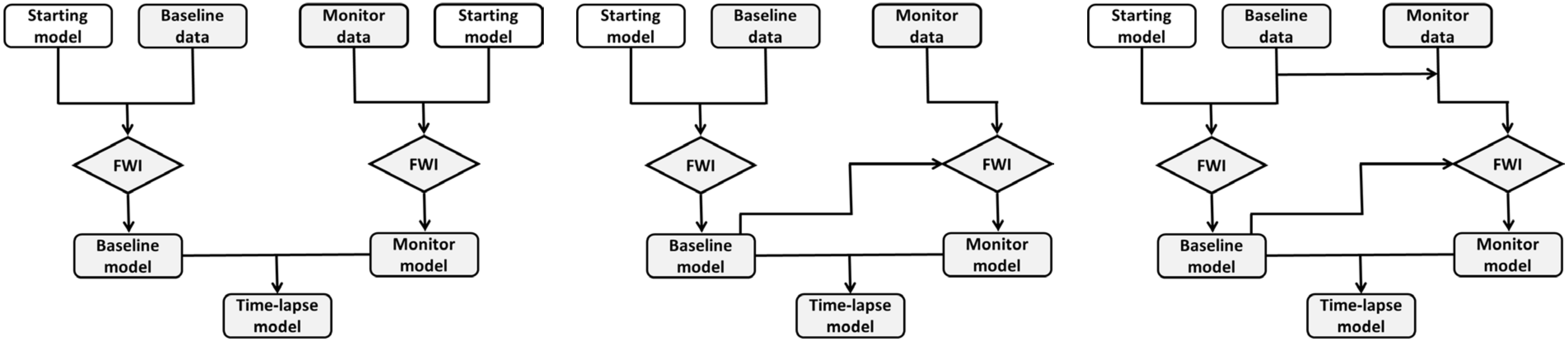
Random Walk sampler $\mathbf{m}^* = \mathbf{m} + \mathbf{X}$

Uniform distribution $T(\mathbf{m}^* | \mathbf{m})$

Outline

- Introduction
- Deterministic optimization (DO) Full-waveform inversion (FWI)
- Bayesian inference based on Markov chain Monte Carlo (MCMC)
- Inversion Strategies
- Numerical examples
- Conclusions

Time-lapse inversion strategies



Parallel difference strategy

Sequential difference strategy

Double-difference strategy (DDWI)

DDWI:

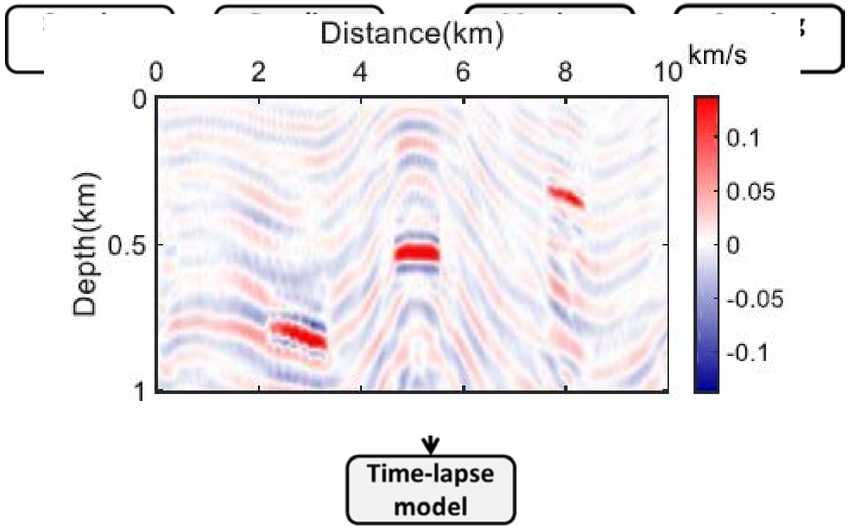
Composited monitoring data:

$$\mathbf{d}_2 = \mathbf{S}_{baseline} + (\mathbf{d}_{monitor} - \mathbf{d}_{baseline})$$

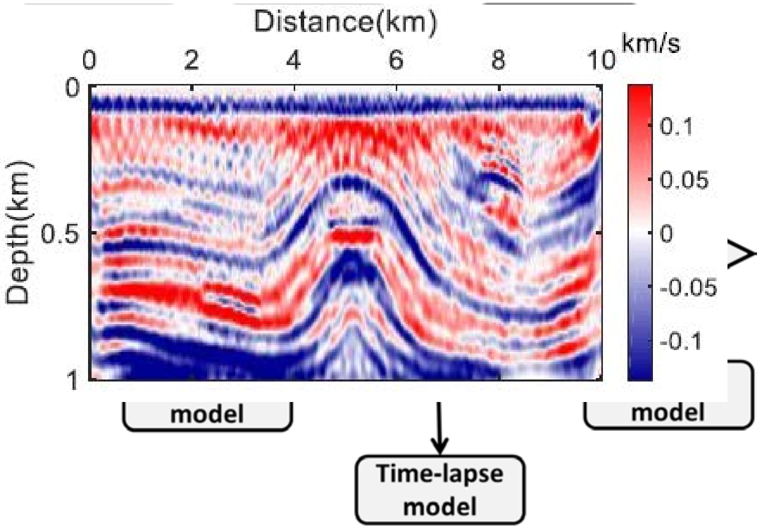
Objective function:

$$E_{DDWI} = \frac{1}{2}\|\mathbf{d}_2 - \mathbf{S}_2\|^2 = \frac{1}{2}\|(\mathbf{d}_{monitor} - \mathbf{S}_{monitor}) - (\mathbf{d}_{baseline} - \mathbf{S}_{baseline})\|^2$$

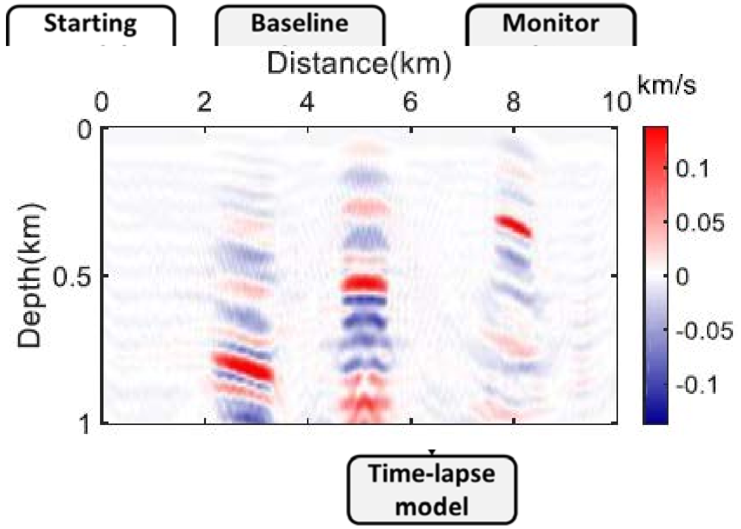
Time-lapse inversion strategies



Parallel difference strategy



Sequential difference strategy



Double-difference strategy (DDWI)

DDWI:

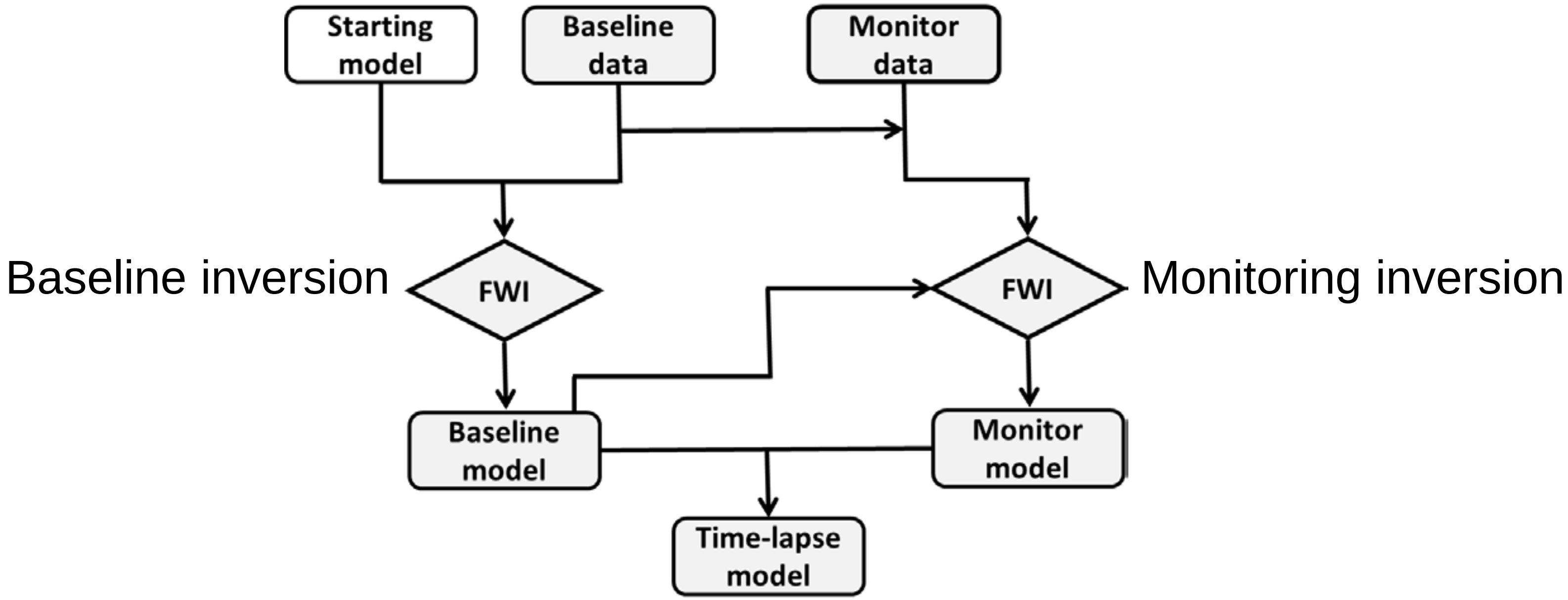
Composited monitoring data:

$$\mathbf{d}_2 = \mathbf{S}_{baseline} + (\mathbf{d}_{monitor} - \mathbf{d}_{baseline})$$

Object function:

$$E_{DDWI} = \frac{1}{2}\|\mathbf{d}_2 - \mathbf{S}_2\|^2 = \frac{1}{2}\|(\mathbf{d}_{monitor} - \mathbf{S}_{monitor}) - (\mathbf{d}_{baseline} - \mathbf{S}_{baseline})\|^2$$

Time-lapse inversion strategies



DO DDWI：

Baseline inversion -> DO FWI

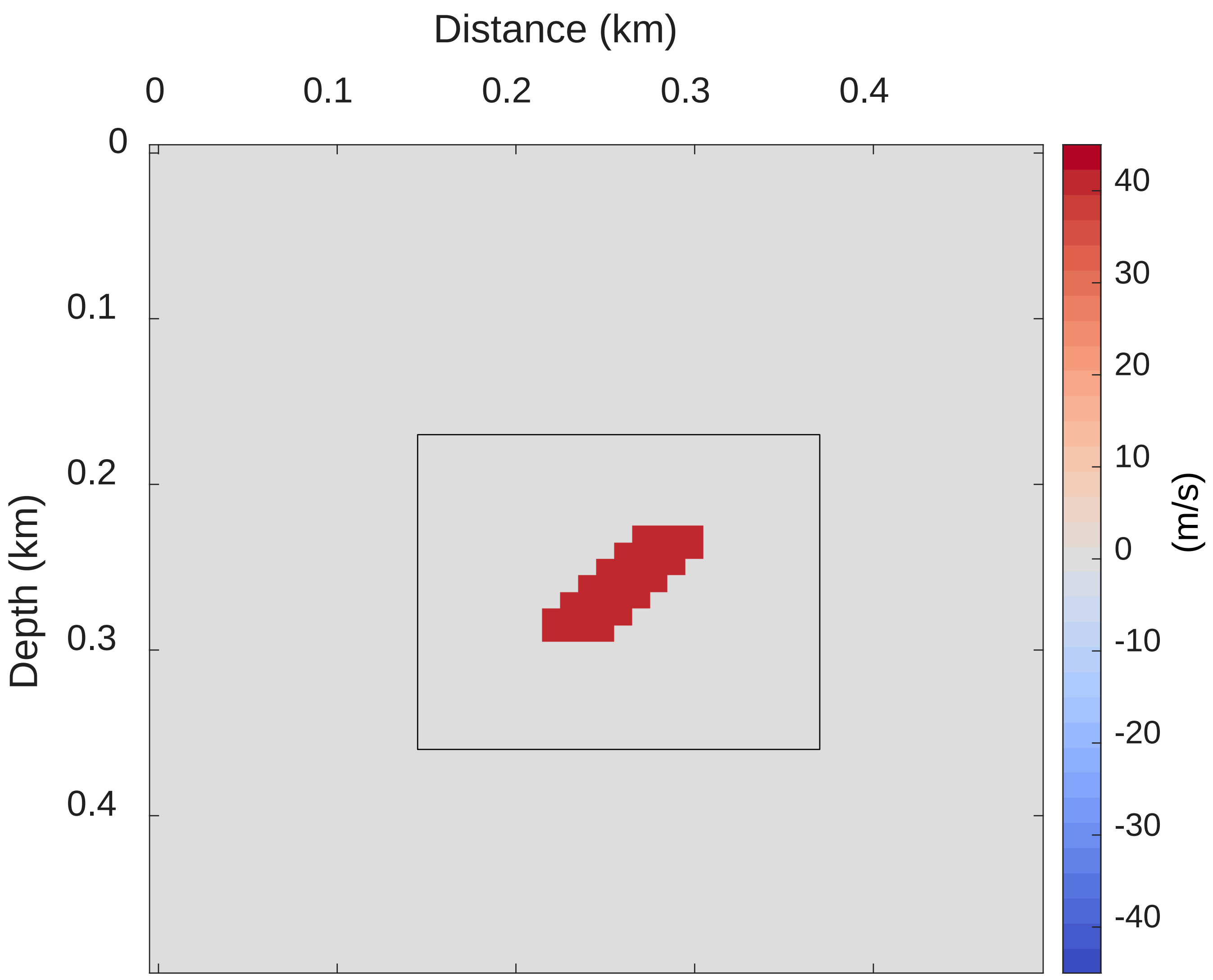
Monitoring inversion -> DO FWI

MCMC DDWI：

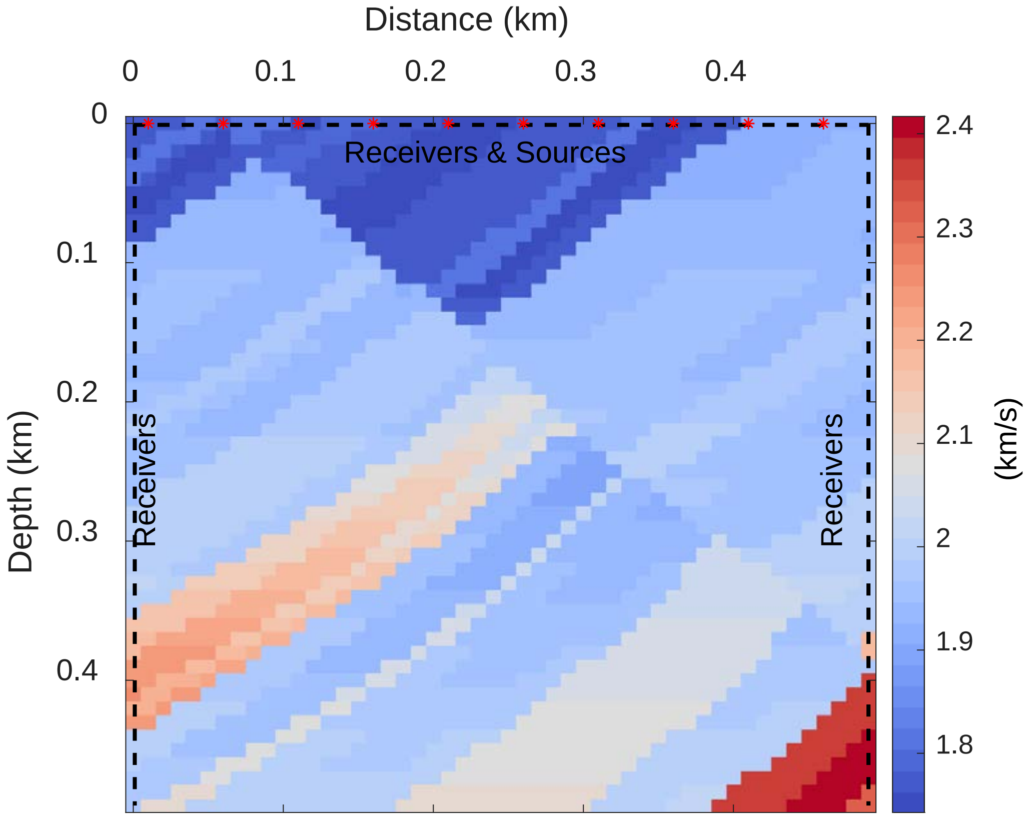
Baseline inversion -> DO FWI

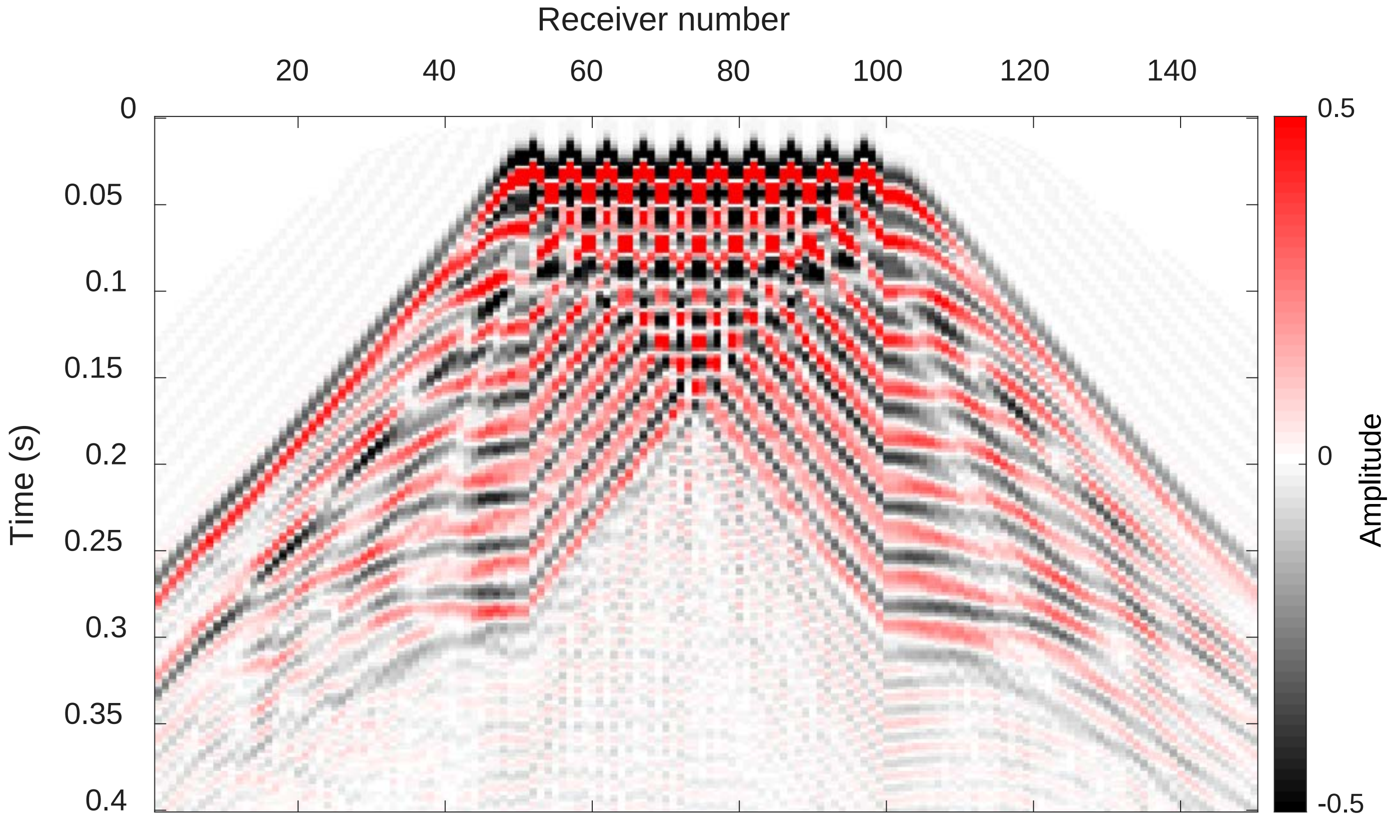
Monitoring inversion -> MCMC FWI

Local-updating target-oriented time-lapse inversion



Multisource waveform inversion



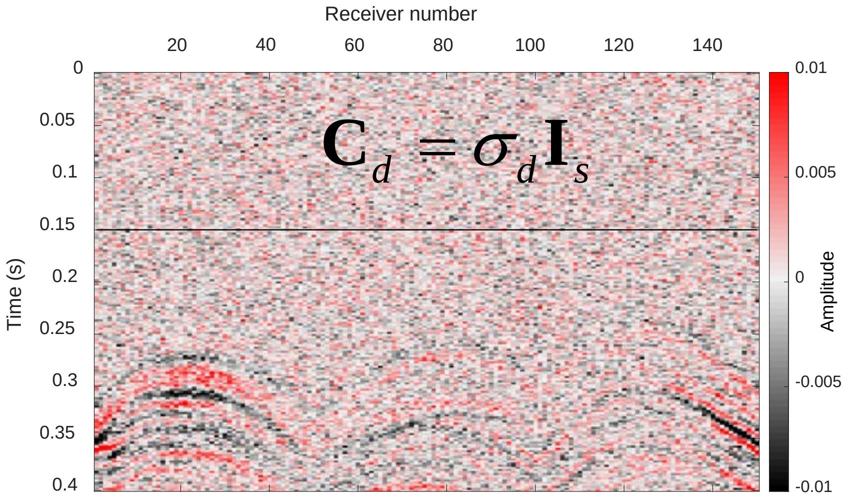


Prior information

Model prior information according to an adaptive Metropolis (AM) algorithm (Gelman et al., 1996, Haario et al., 1999):

$$\mathbf{C}_{k+1} = \frac{k-1}{k}\mathbf{C}_k + \frac{s_m}{k}\left(k\overline{\mathbf{m}}_{k-1}\overline{\mathbf{m}}_{k-1}^T - (k+1)\overline{\mathbf{m}}_k\overline{\mathbf{m}}_k^T + \mathbf{m}_k\mathbf{m}_k^T + \epsilon\mathbf{I}_m\right)$$

New data prior information estimation : $\mathbf{C}_d = \sigma_d \mathbf{I}_s$

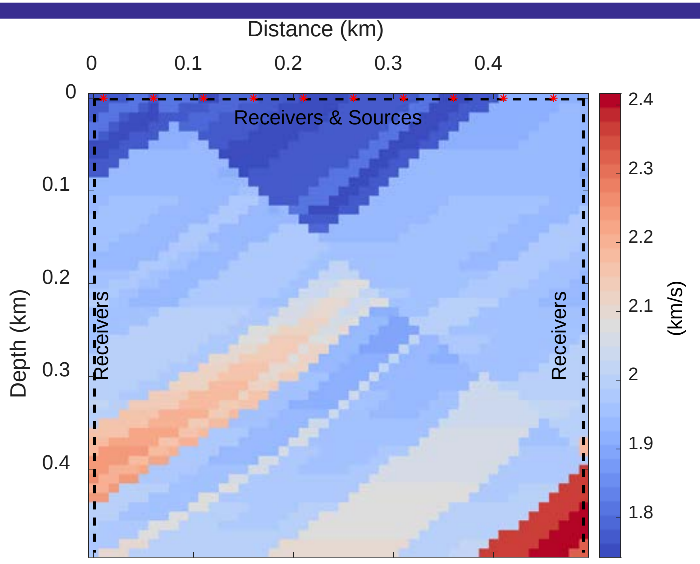


$$\mathbf{d}_{monitor} - \mathbf{d}_{baseline}$$

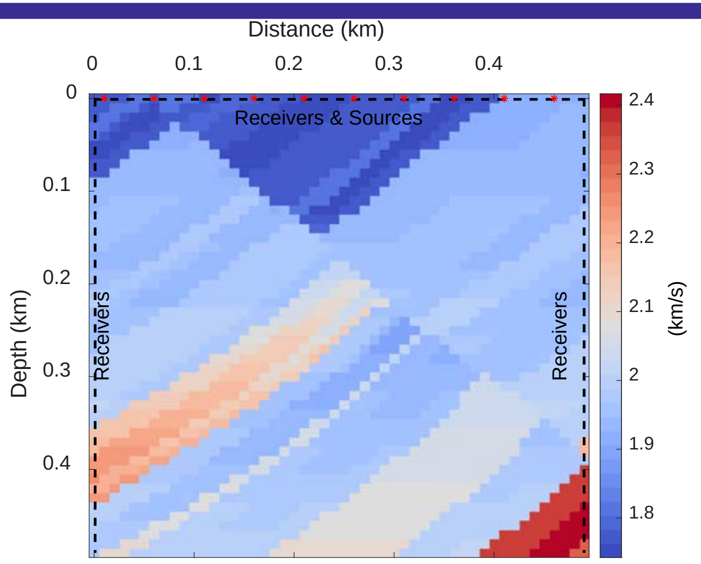
Outline

- Introduction
- Deterministic optimization (DO) Full-waveform inversion (FWI)
- Bayesian inference based on Markov chain Monte Carlo (MCMC)
- Inversion Strategies
- Numerical examples
- Conclusions

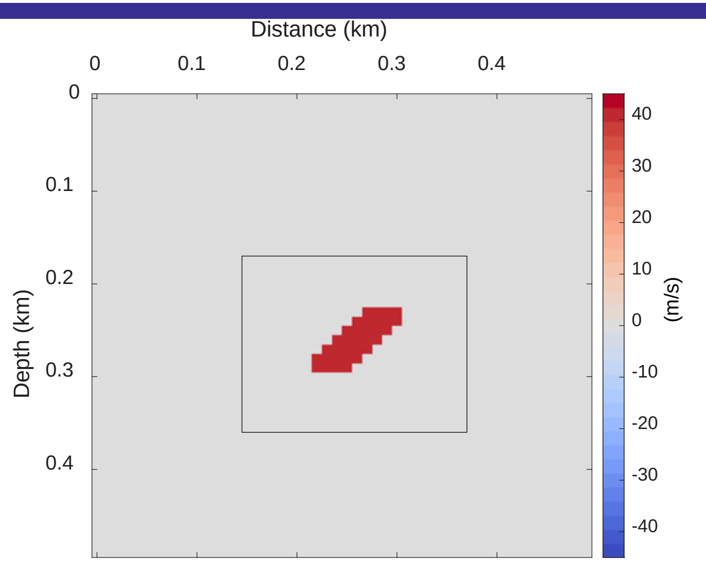
Numerical examples



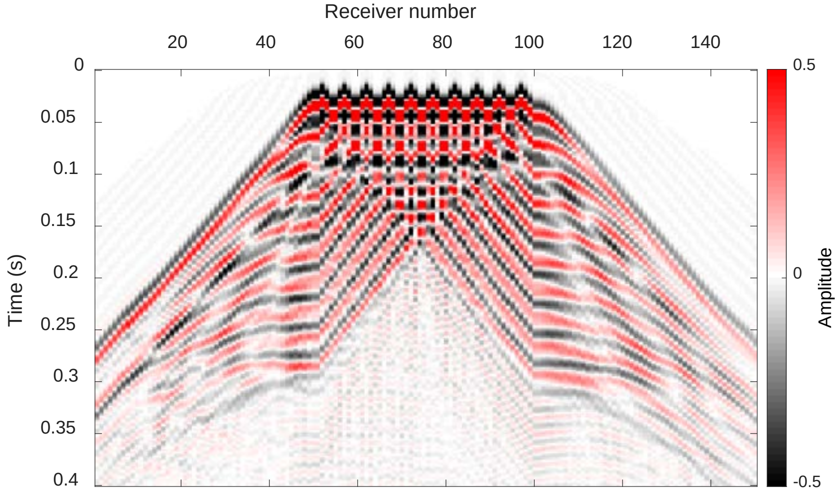
Baseline model



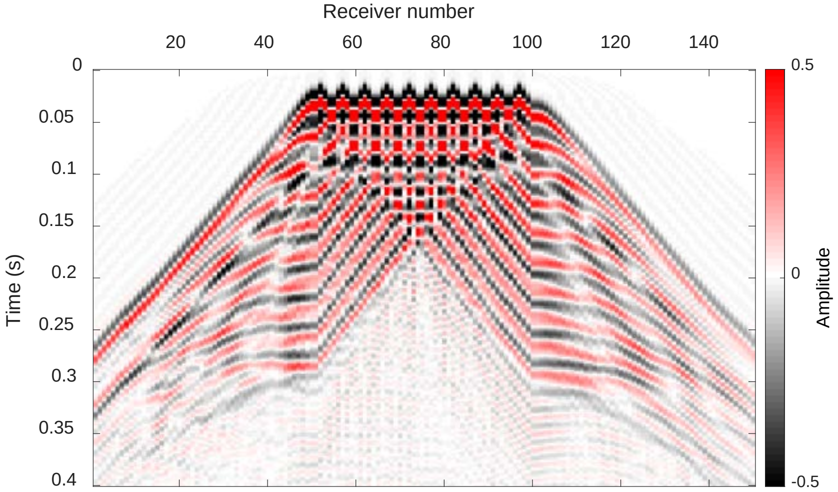
Monitoring model



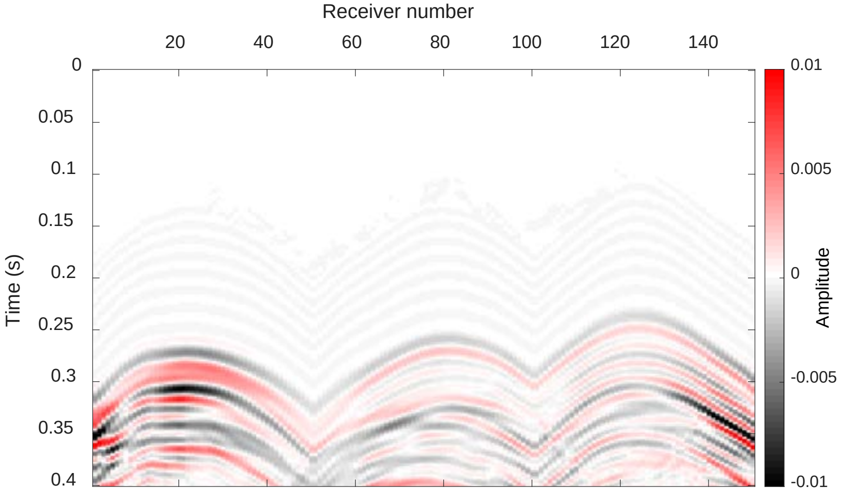
Time-lapse model



Baseline data

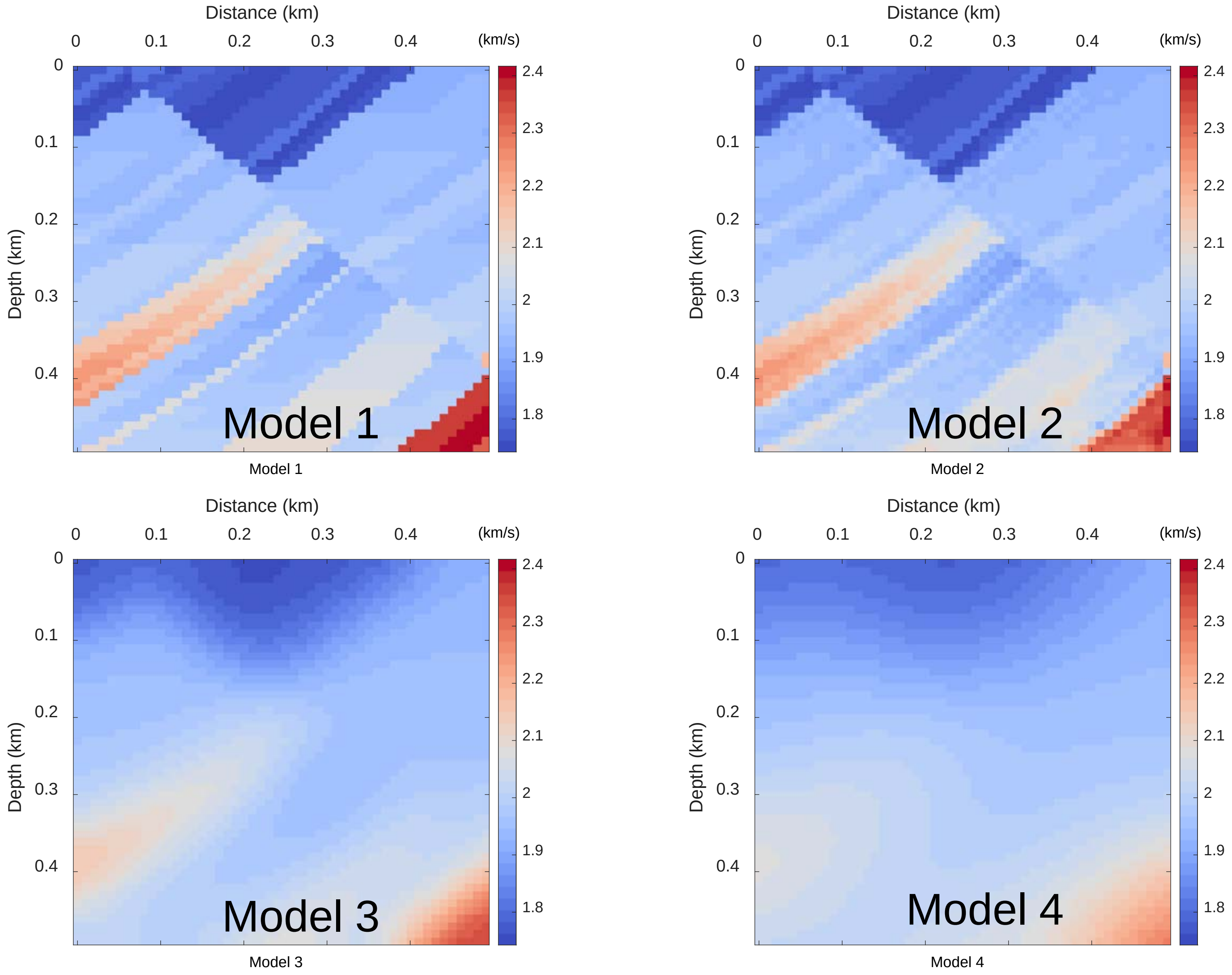


Monitoring data

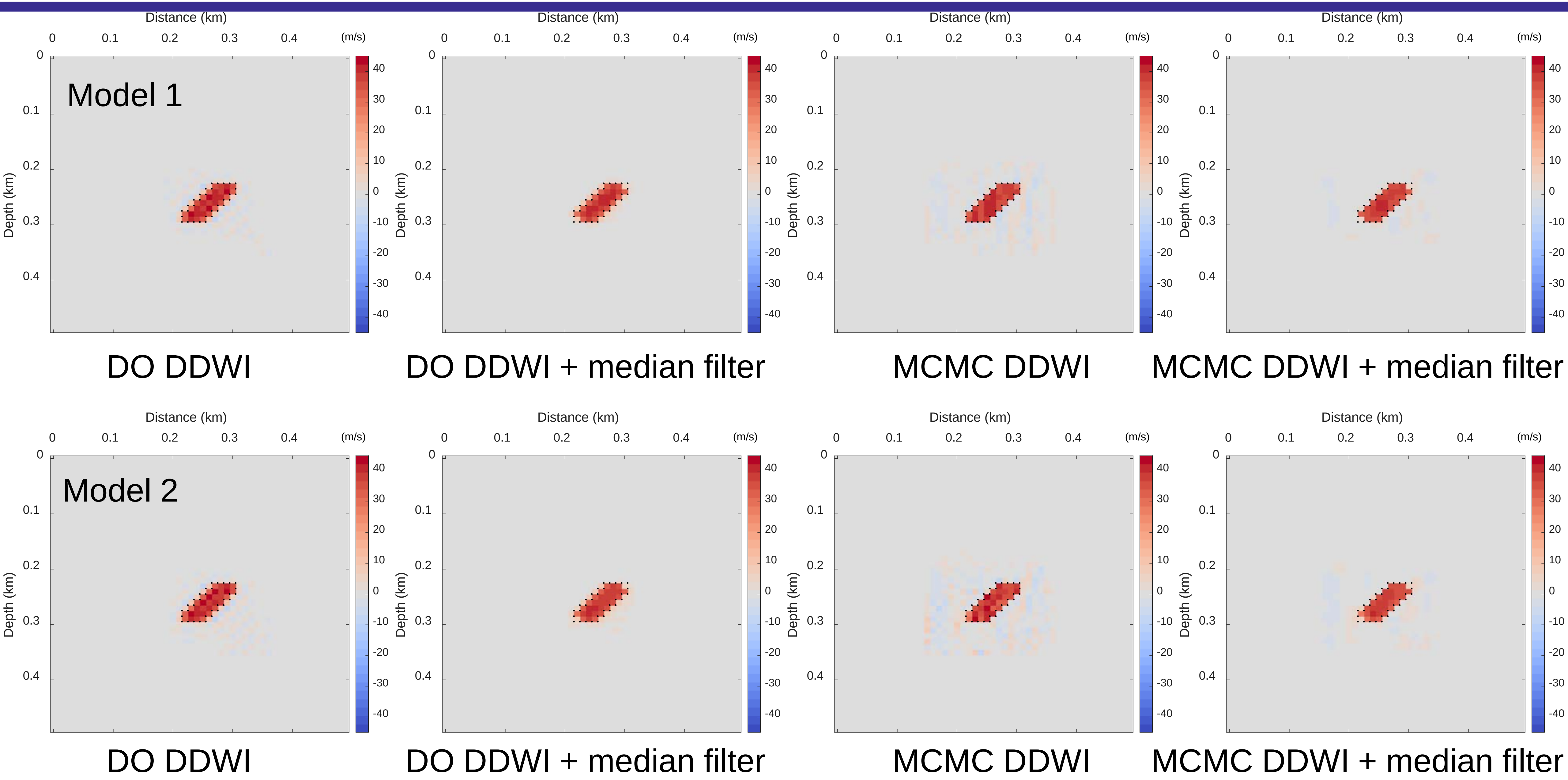


Difference data

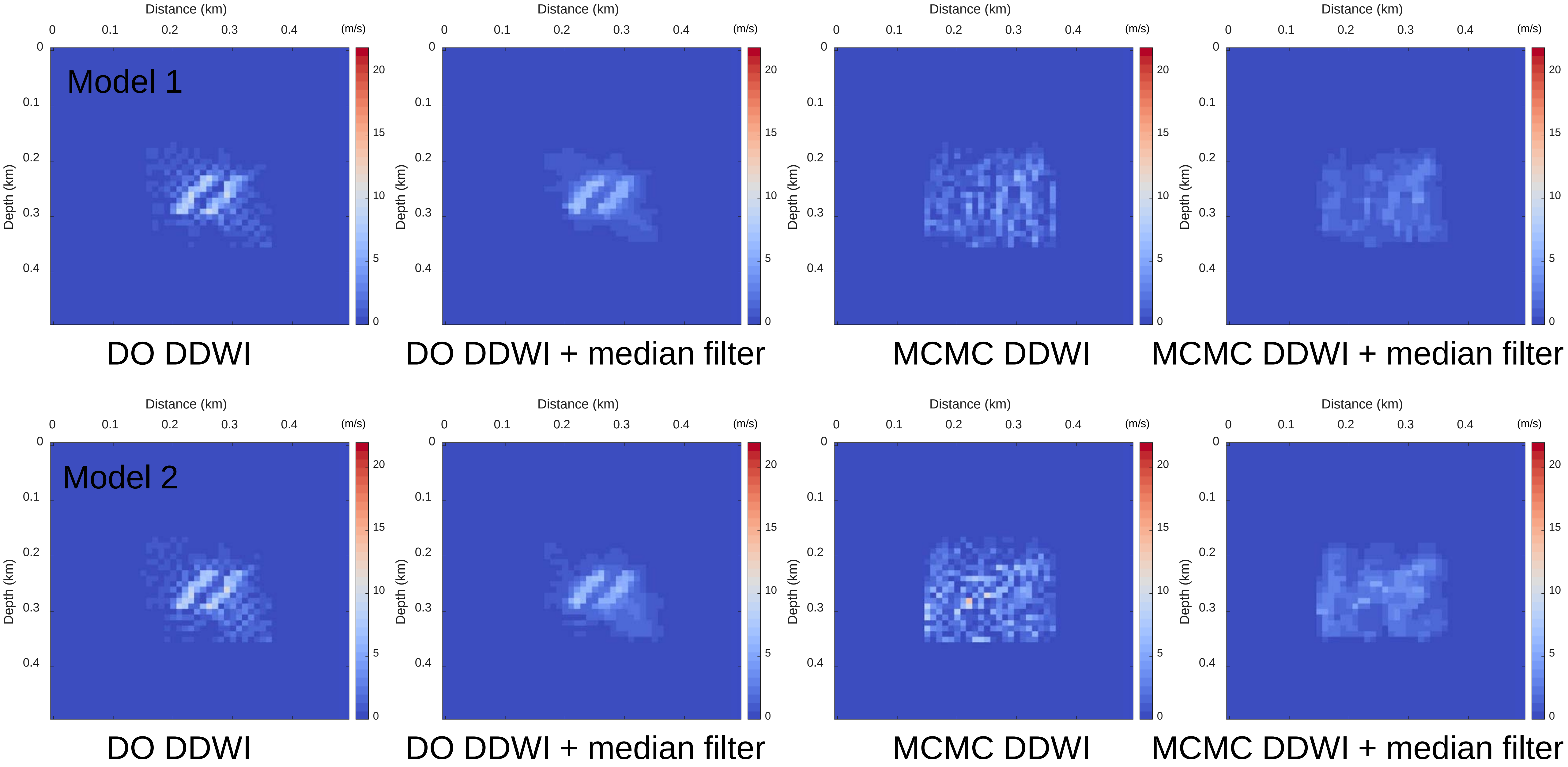
DO DDWI vs MCMC DDWI: noise-free data tests



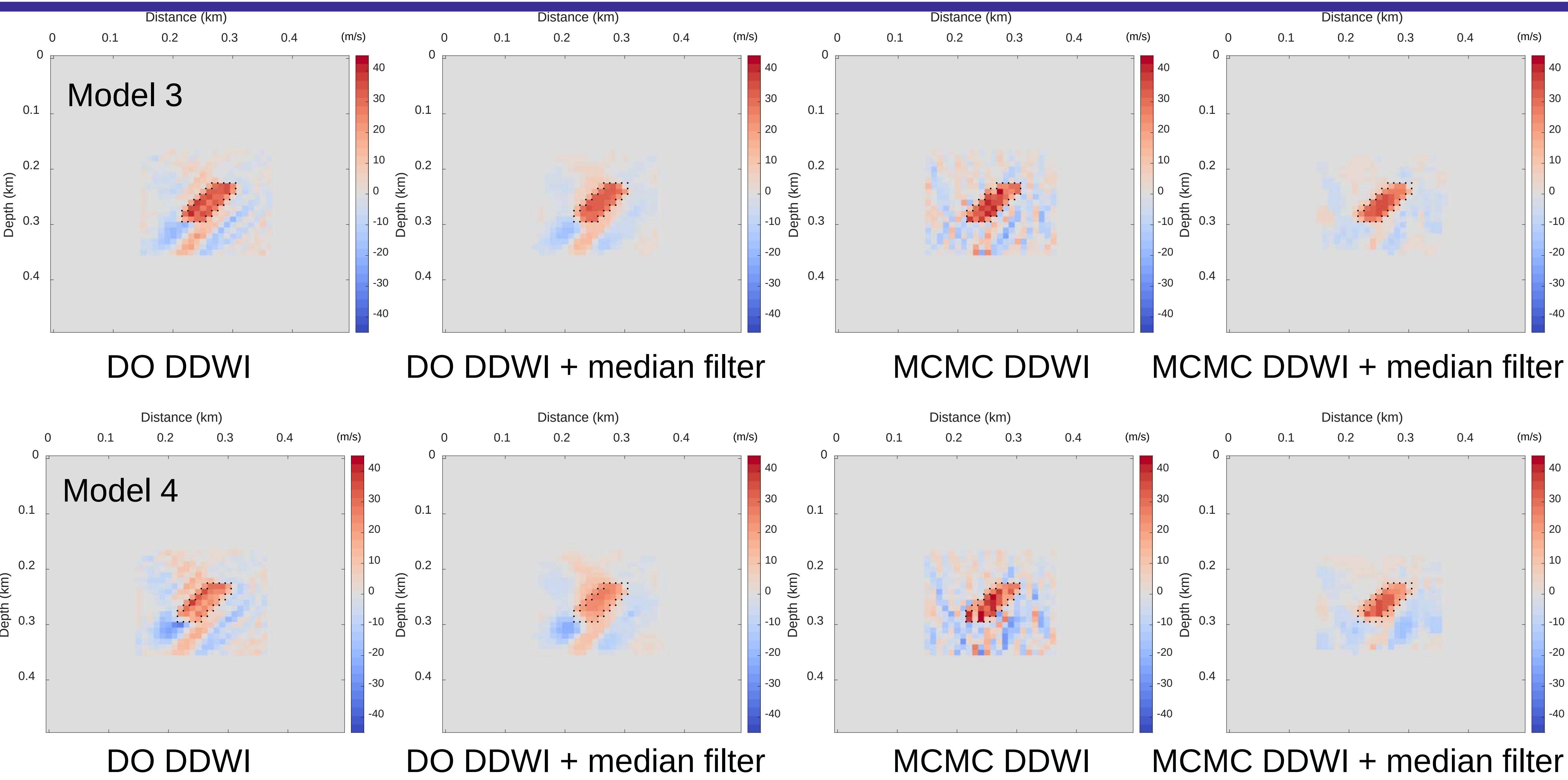
Inverted time-lapse models



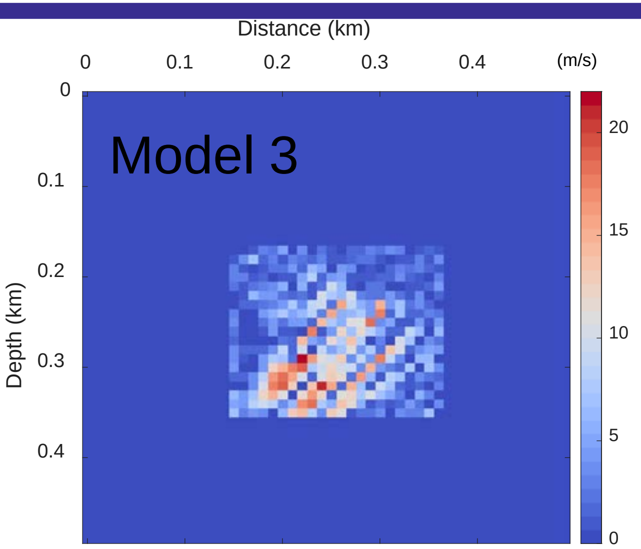
Model errors



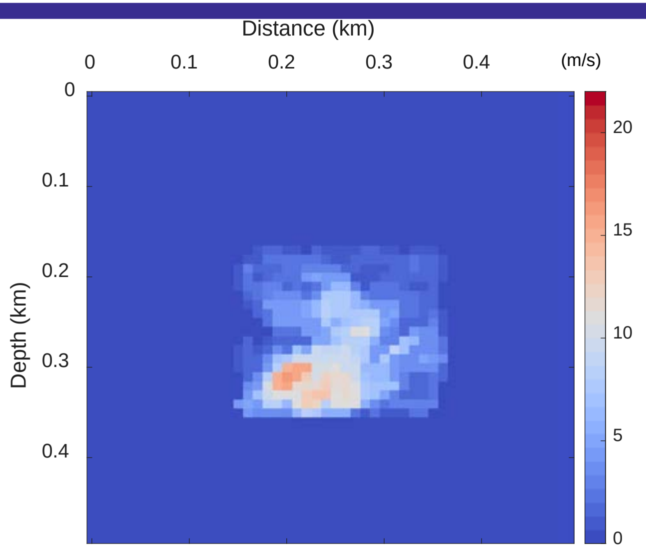
Inverted time-lapse models



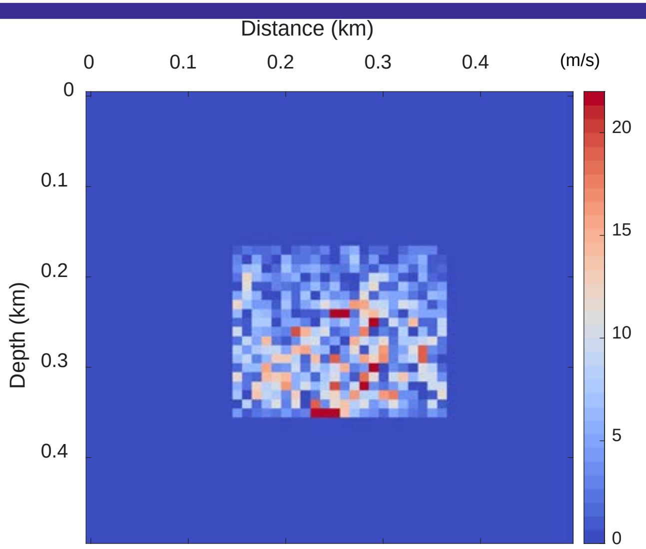
Model errors



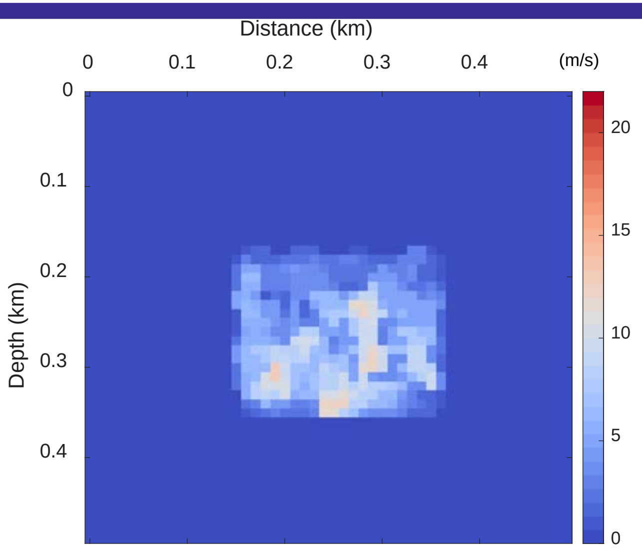
DO DDWI



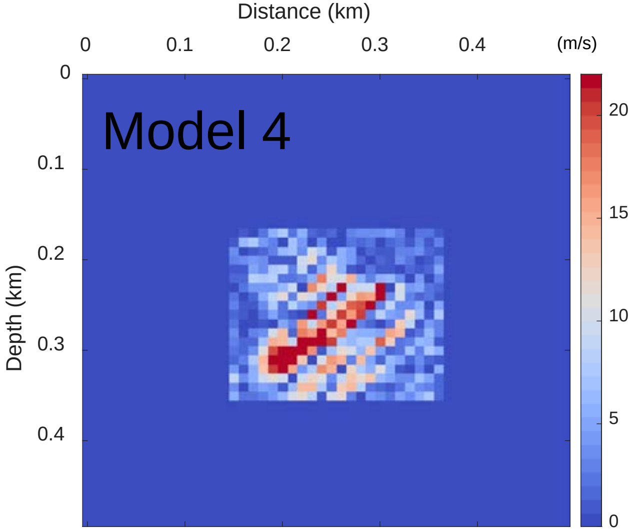
DO DDWI + median filter



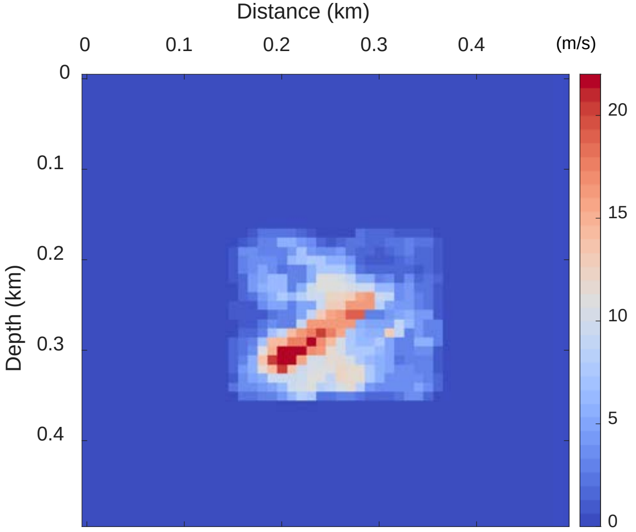
MCMC DDWI



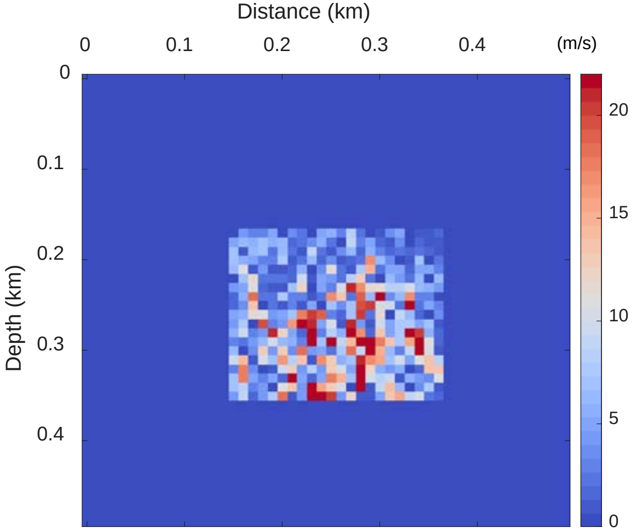
MCMC DDWI + median filter



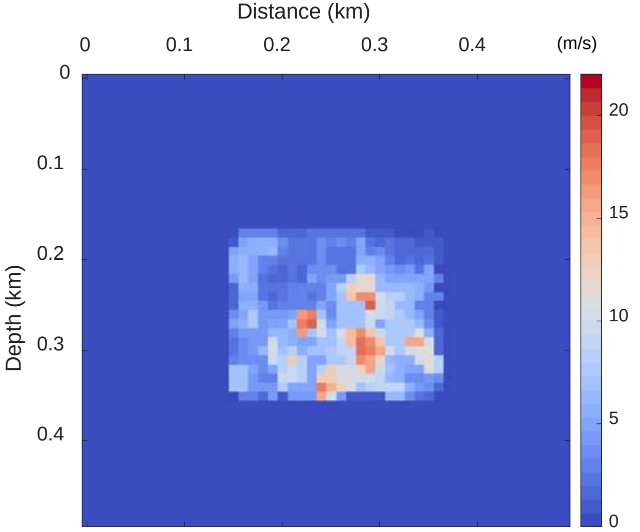
DO DDWI



DO DDWI + median filter

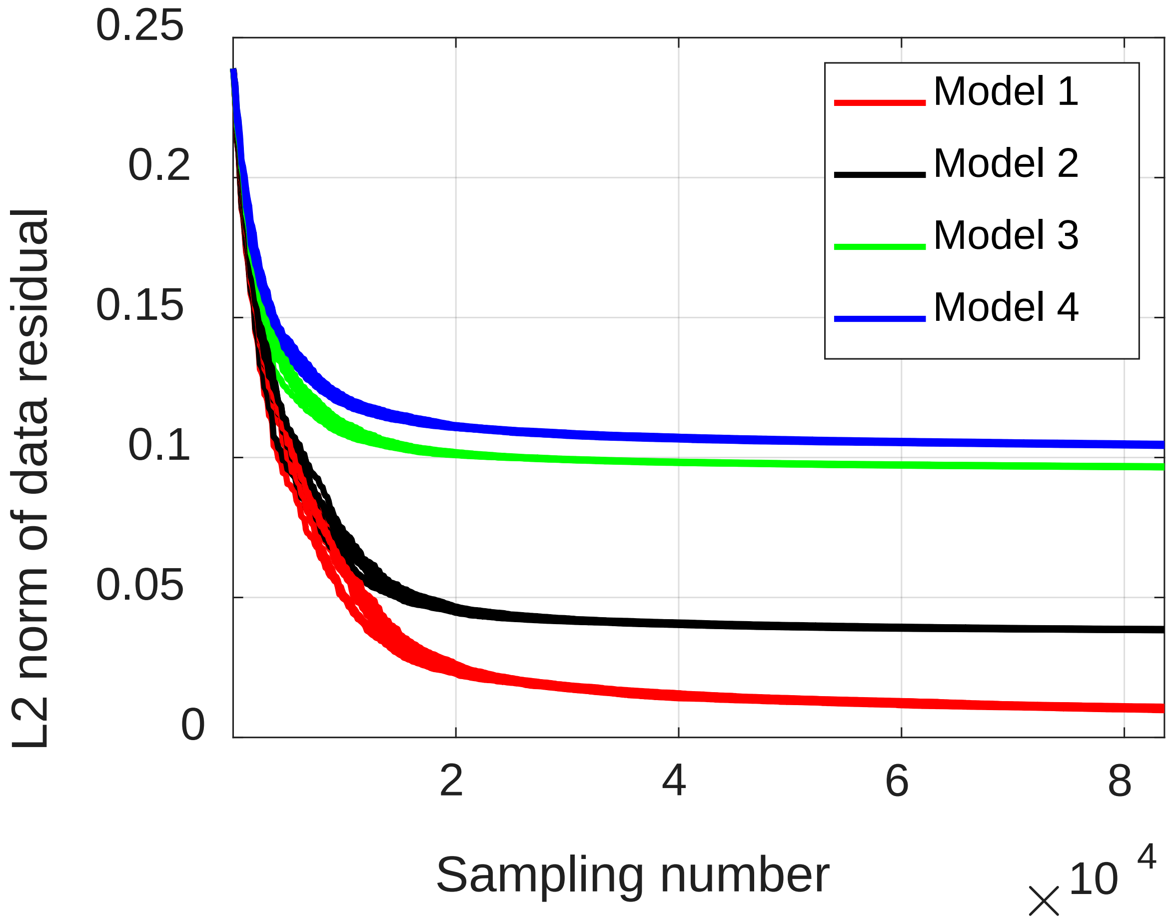


MCMC DDWI

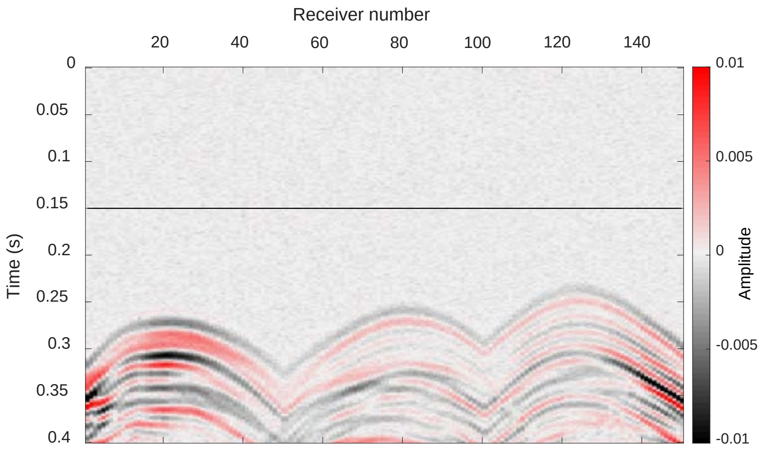


MCMC DDWI + median filter

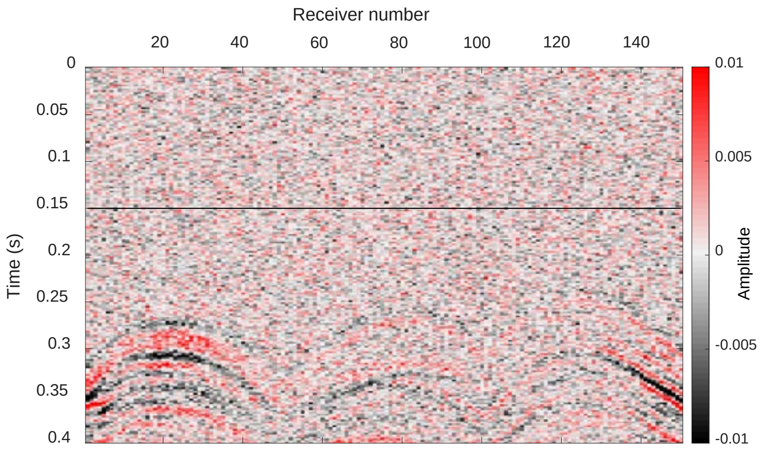
Data misfit


0.25
0.2
0.15
0.1
0.05
0
L2 norm of data residual
2
4
6
8
Sampling number
$\times 10^4$
Model 1
Model 2
Model 3
Model 4

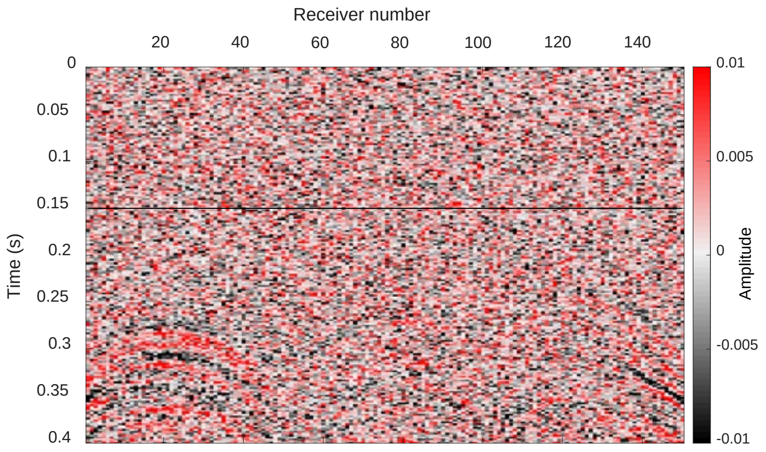
Noisy data tests



Noise 1

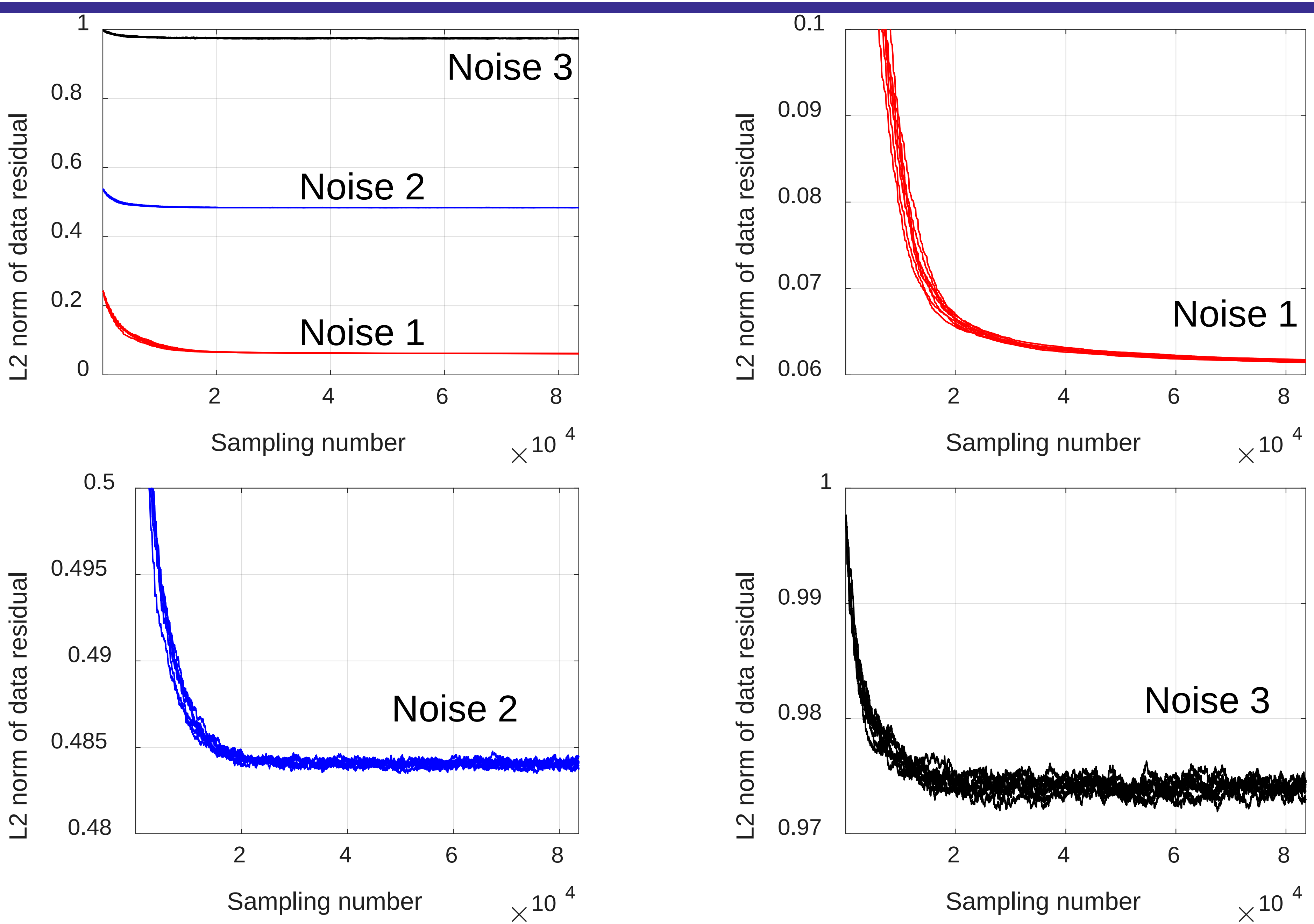


Noise 2



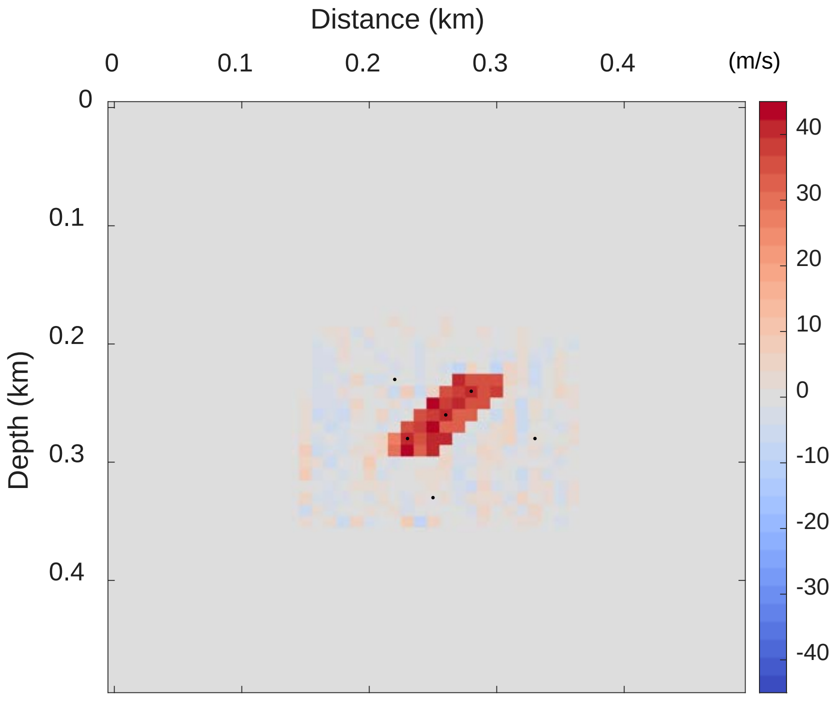
Noise 3

Data misfit

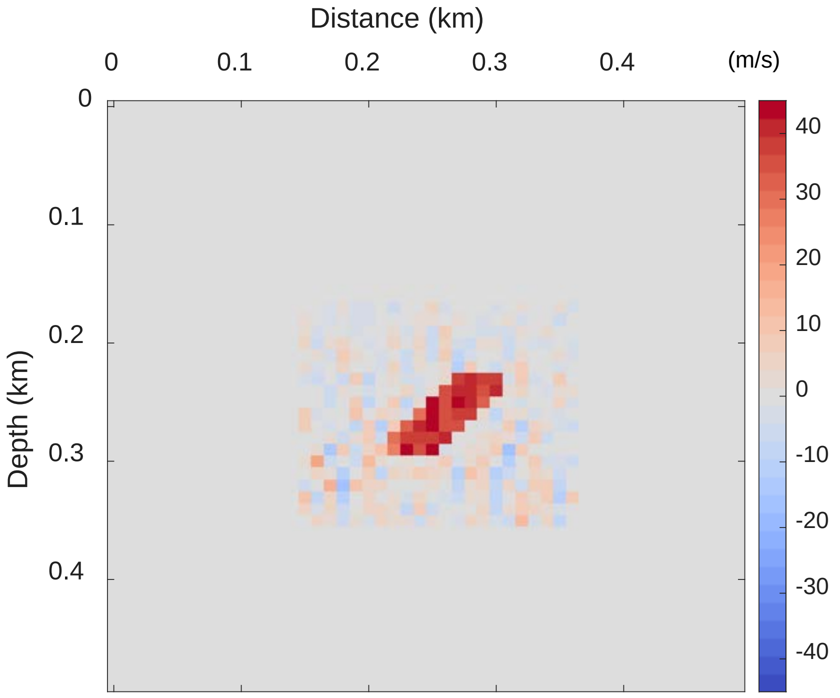


Noise 3
Noise 2
Noise 1
L2 norm of data residual
Sampling number
×10 4
0
0.2
0.4
0.6
0.8
1
2
4
6
8
Noise 1
L2 norm of data residual
Sampling number
×10 4
0.06
0.07
0.08
0.09
0.1
Noise 2
L2 norm of data residual
Sampling number
×10 4
0.48
0.485
0.49
0.495
0.5
Noise 3
L2 norm of data residual
Sampling number
×10 4
0.97
0.98
0.99
1

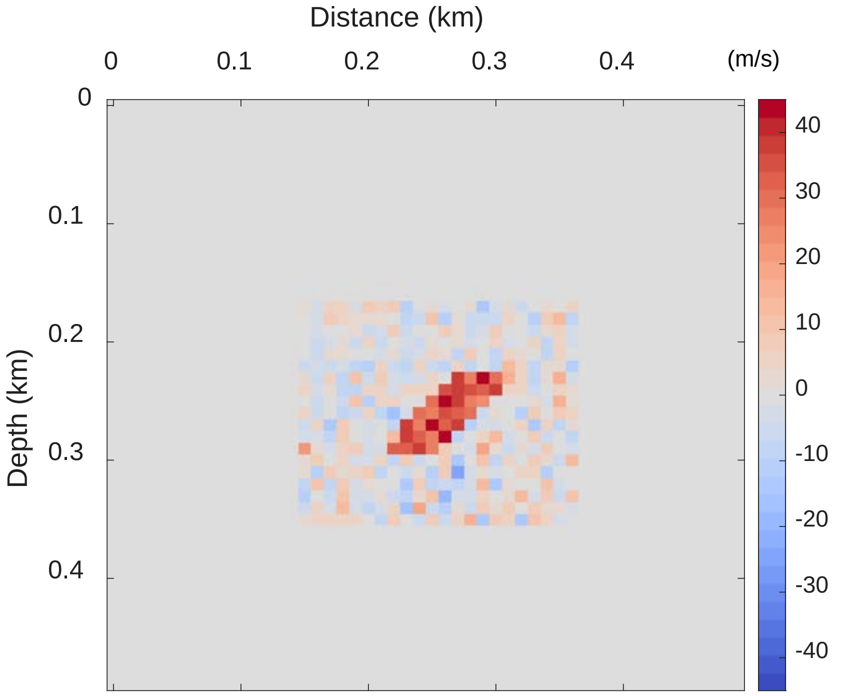
Inverted average models



Noise 1

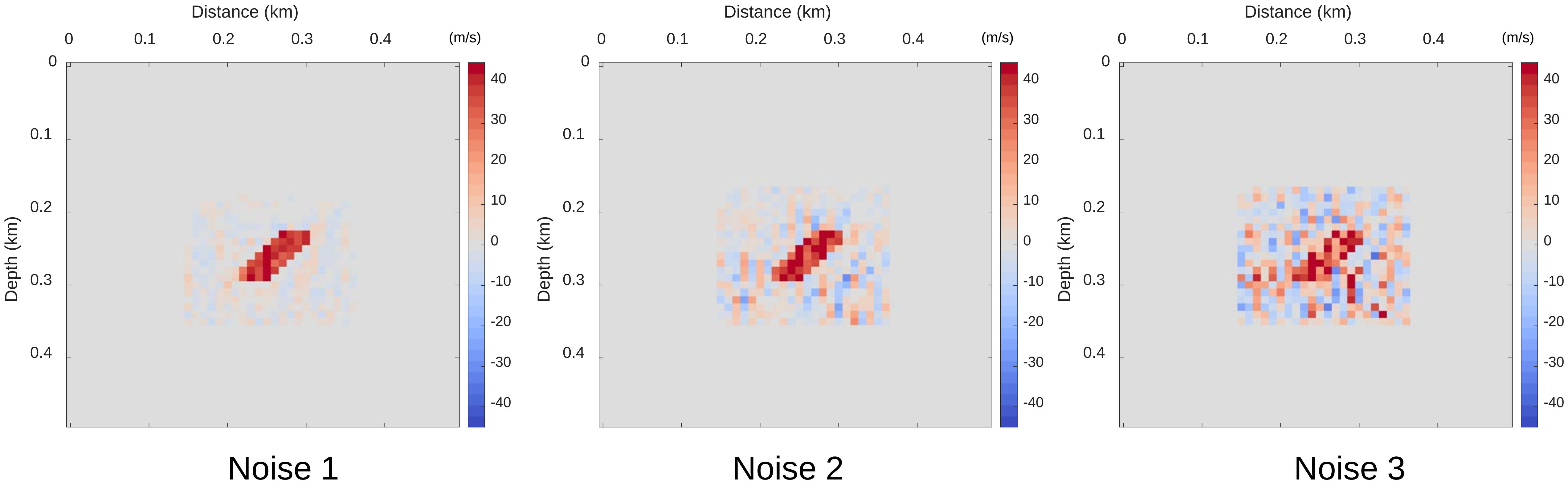


Noise 2

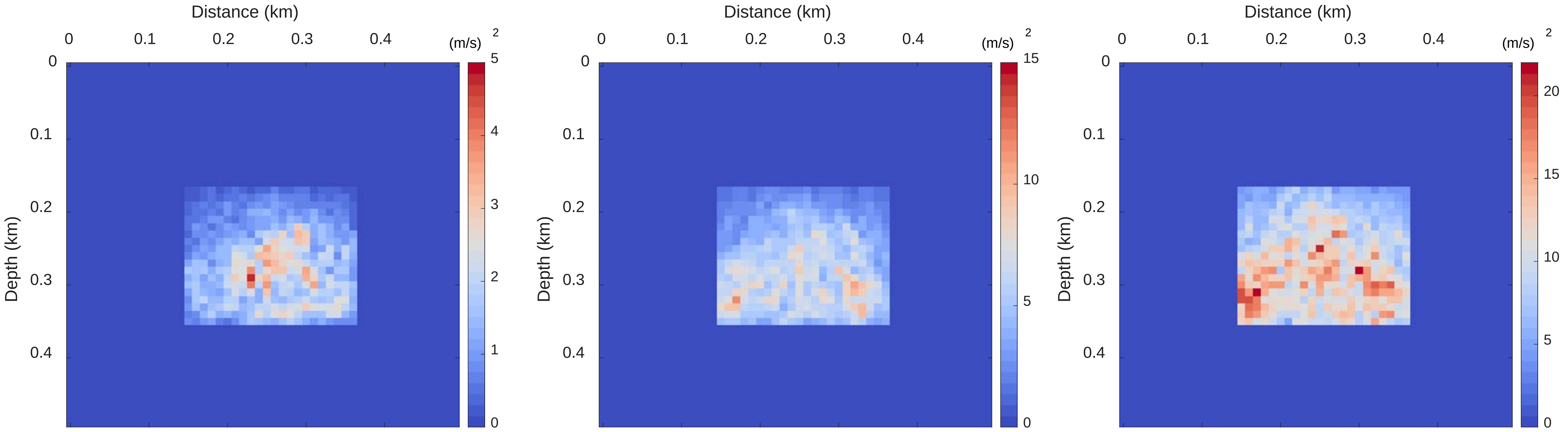


Noise 3

Maximum probability (MAP) models



Model standard deviations

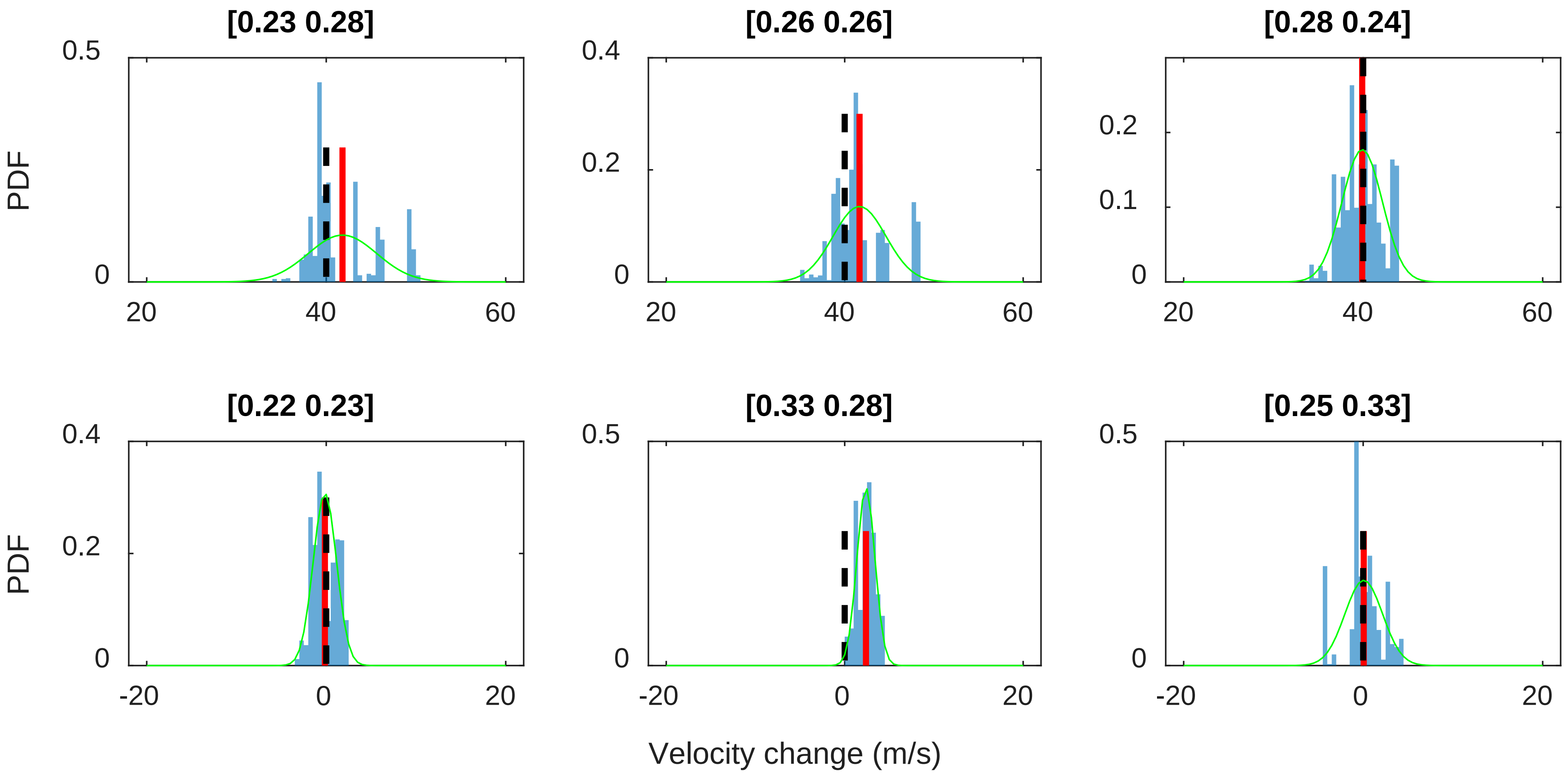


Noise 1

Noise 2

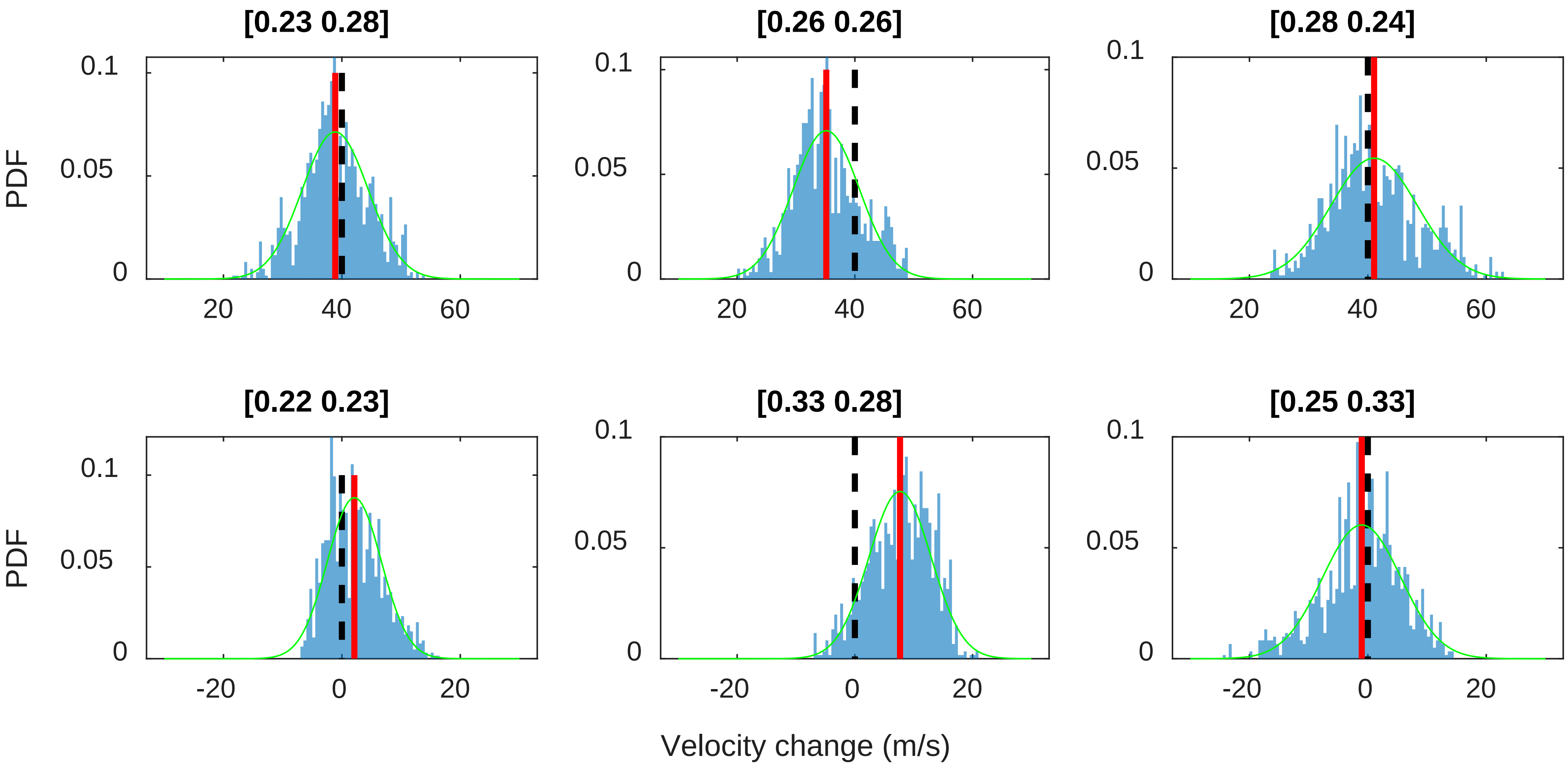
Noise 3

Model distribution



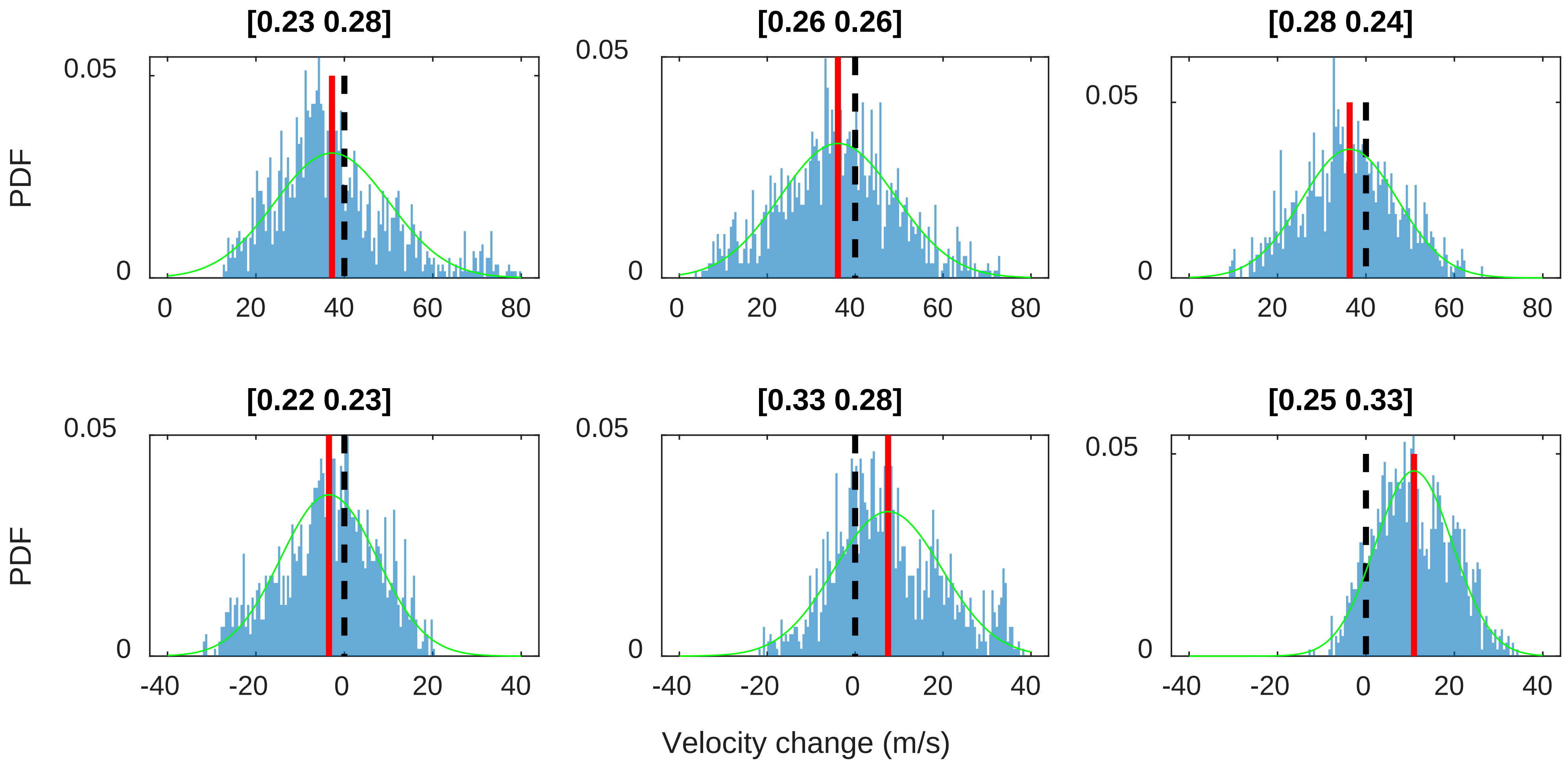
Noise 1

Model distribution



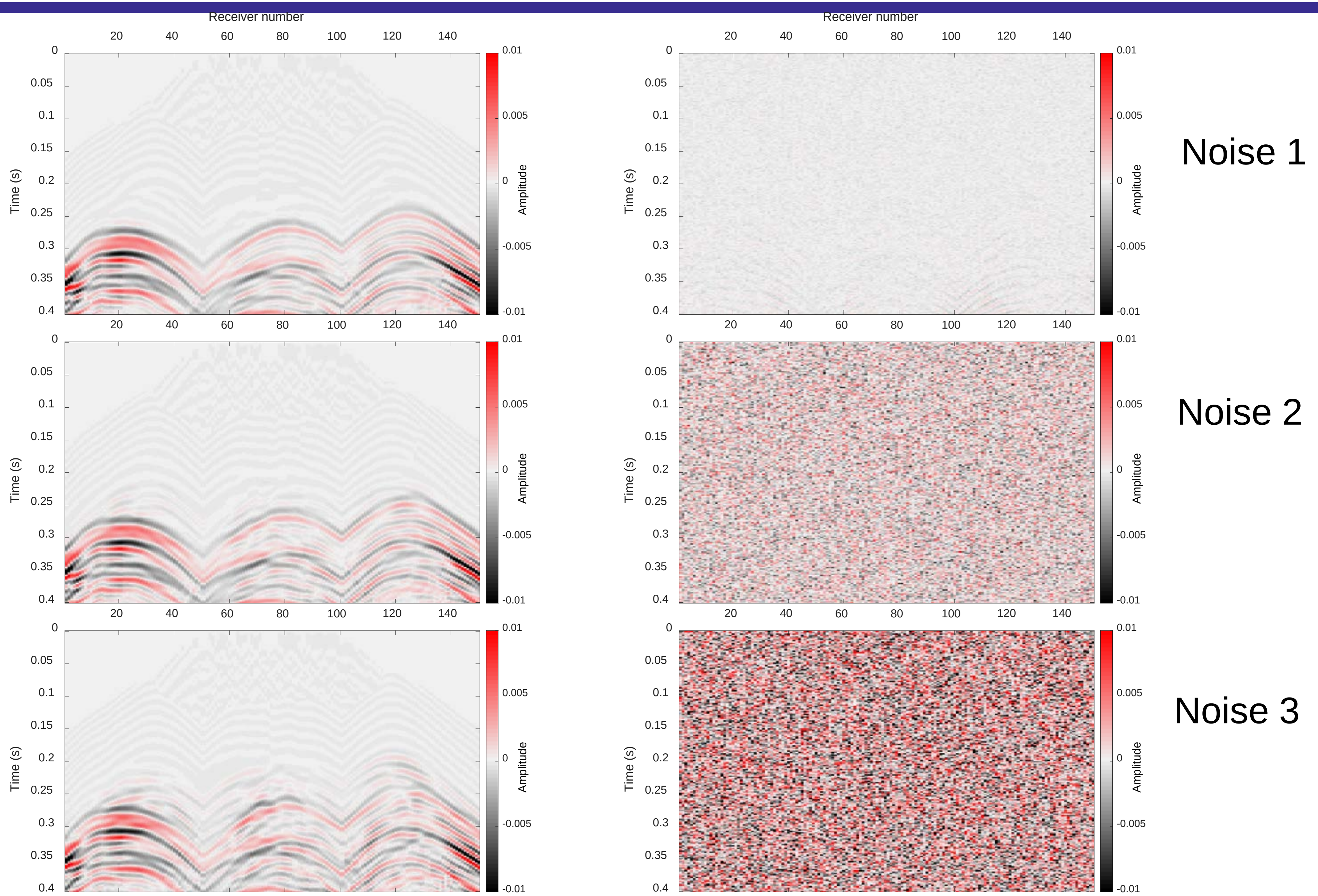
Noise 2

Model distribution

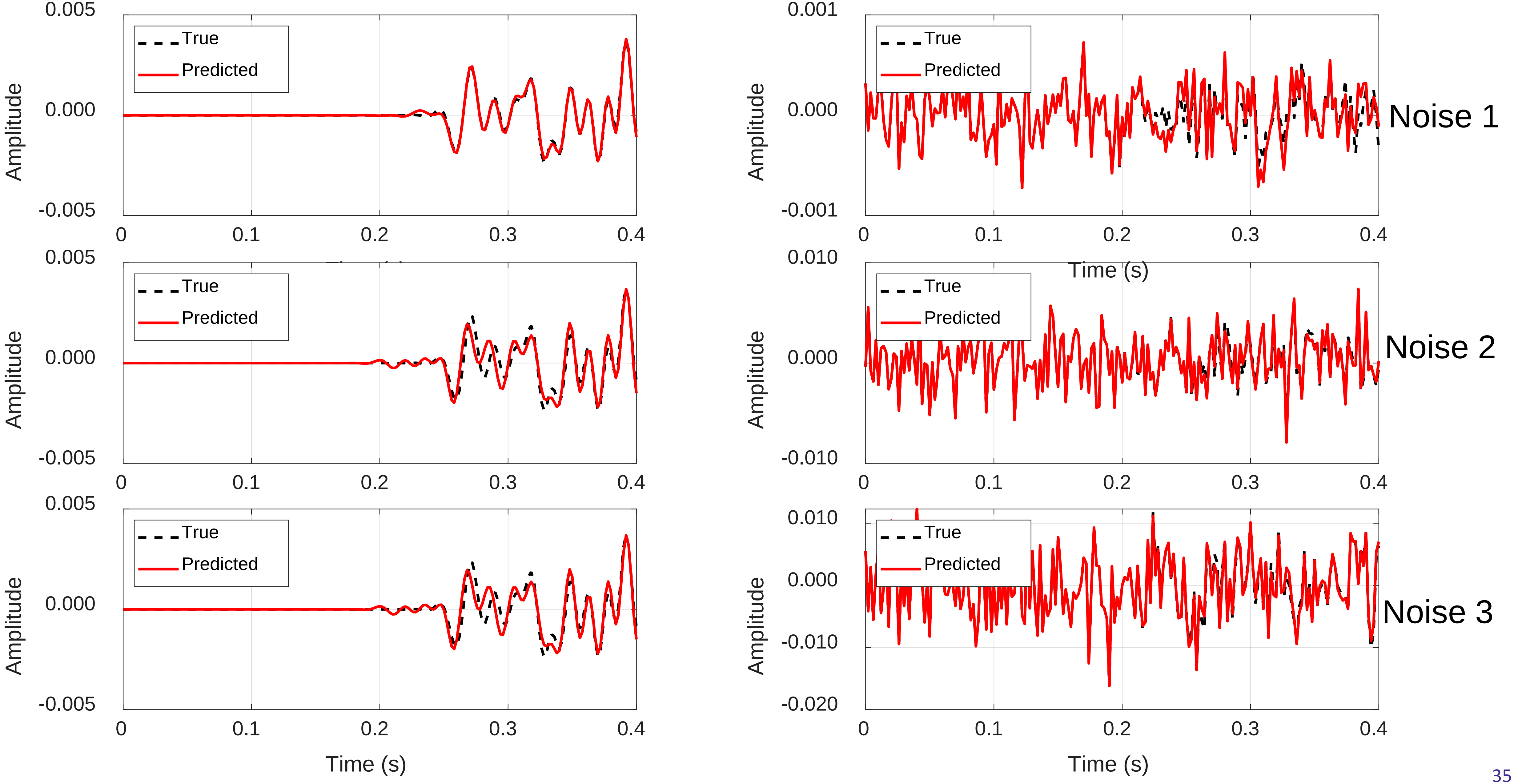


Noise 3

Predicted noise-free data and noise



Predicted noise-free data and noise



Outline

- Introduction
- Deterministic optimization (DO) Full-waveform inversion (FWI)
- Bayesian inference based on Markov chain Monte Carlo (MCMC)
- Inversion Strategies
- Numerical examples
- Conclusions

Conclusions

We have proposed a time-lapse FWI based on an MH MCMC algorithm, and a new method to estimate the data error standard deviation for time-lapse data according to the feature of difference data.

To achieve the MCMC-based time-lapse FWI, we have employed the inversion strategies including DDWI, multisource data, local-updating target-oriented inversion, calculating model covariance with the AM algorithm, and the new data error standard deviation estimation.

Compared the conventional DO DDWI, MCMC DDWI can provide the results with clearer edges of the nonzero time-lapse model change and fewer coherent errors.

Acknowledgements

We thank the sponsors of CREWES for continued support.
NSERC (Natural Science and Engineering Research Council of Canada)
The Canada First Research Excellence Fund

Thank you !