Uncertainty Quantification for Wavefield-Reconstruction Inversion

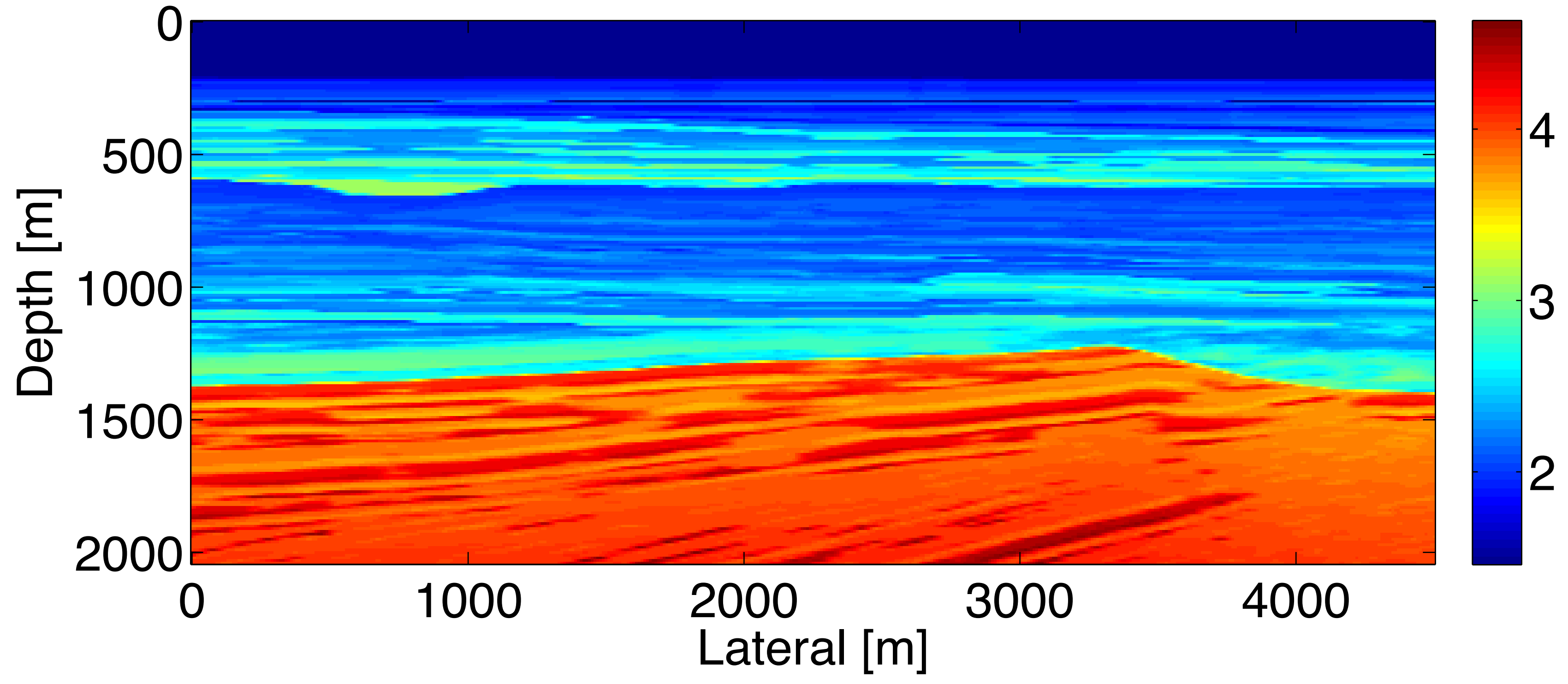
Zhilong Fang, Chia-Ying Lee, Curt Da Silva, Felix J. Herrmann and Rachel Kuske

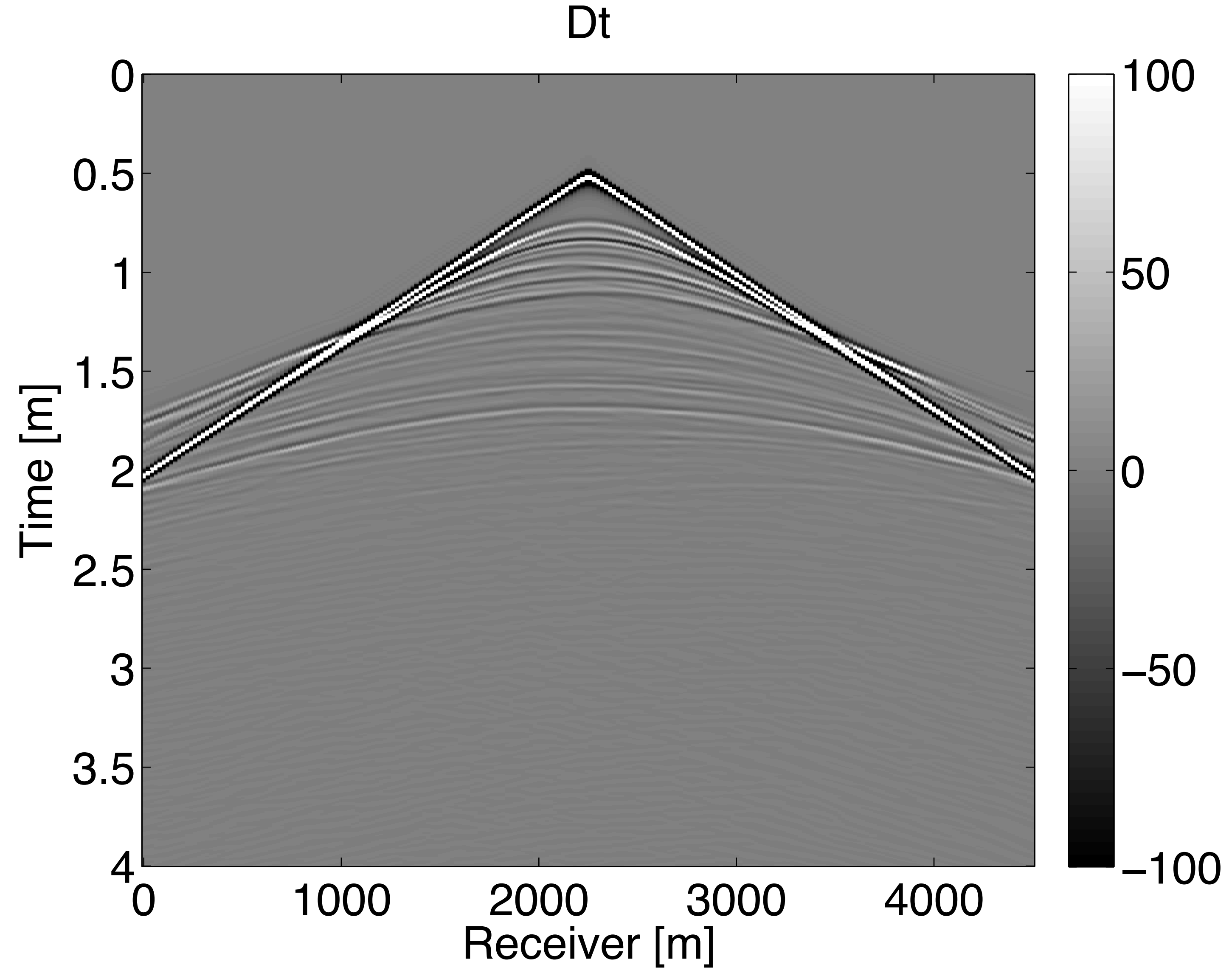


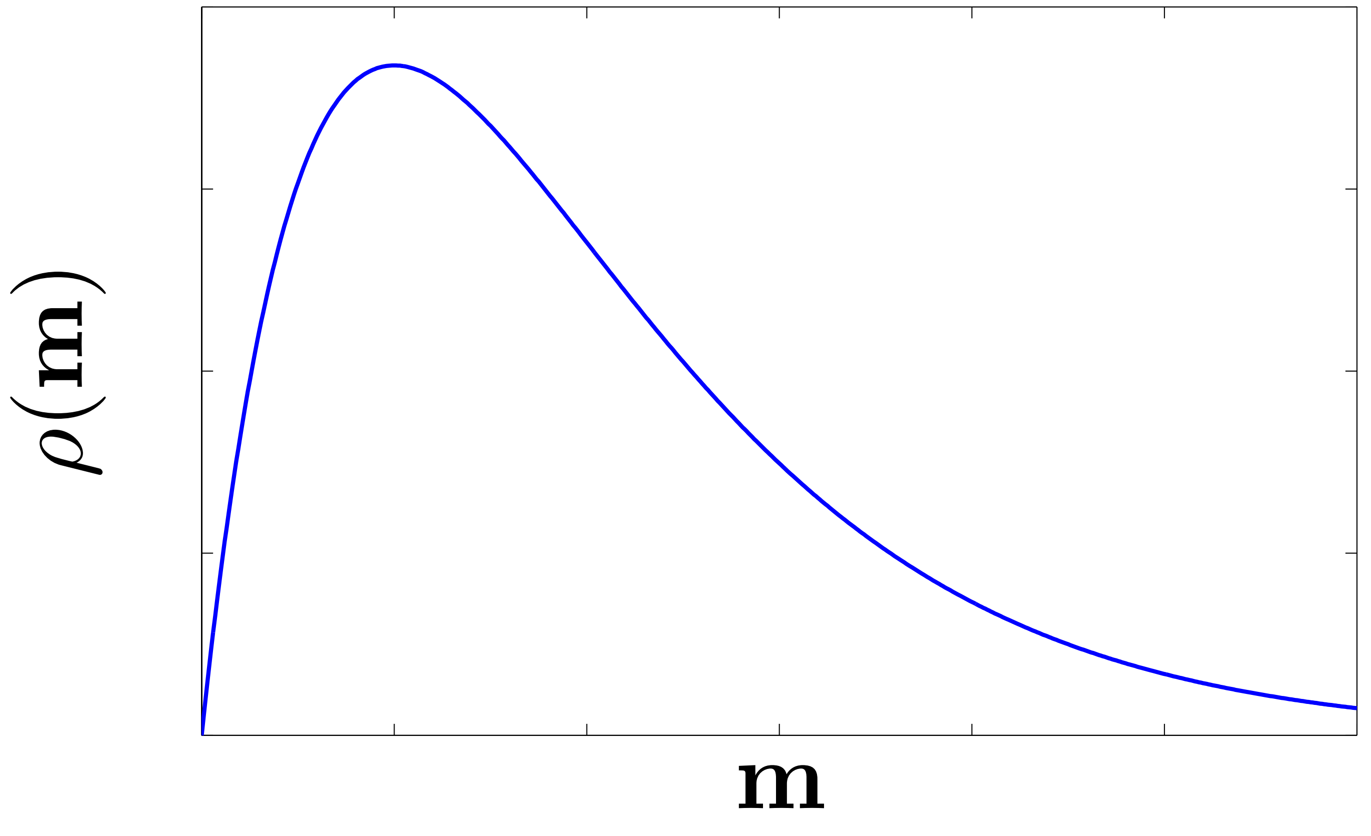


SLIM
University of British Columbia

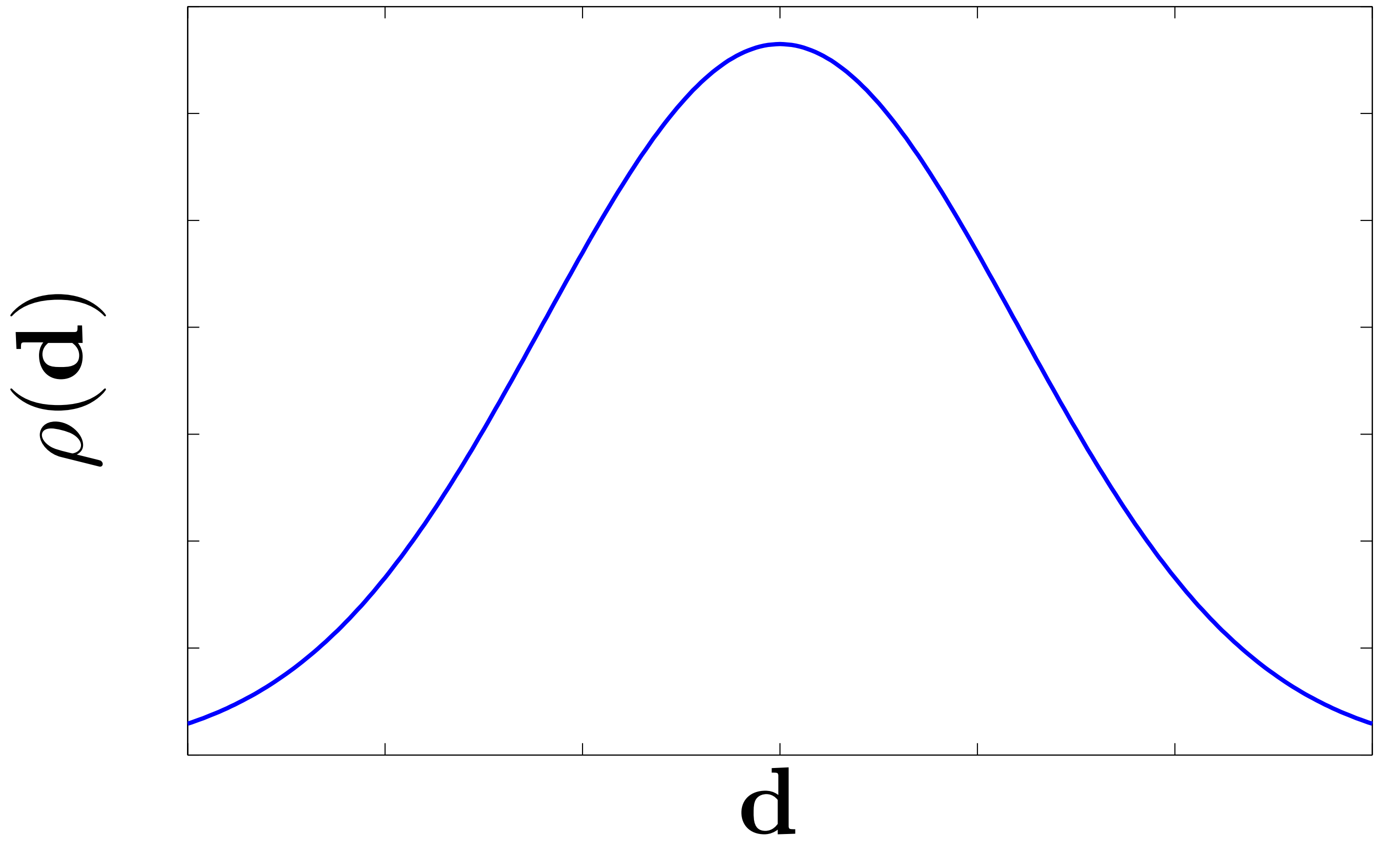
Motivation







Distribution of model



Distribution of data

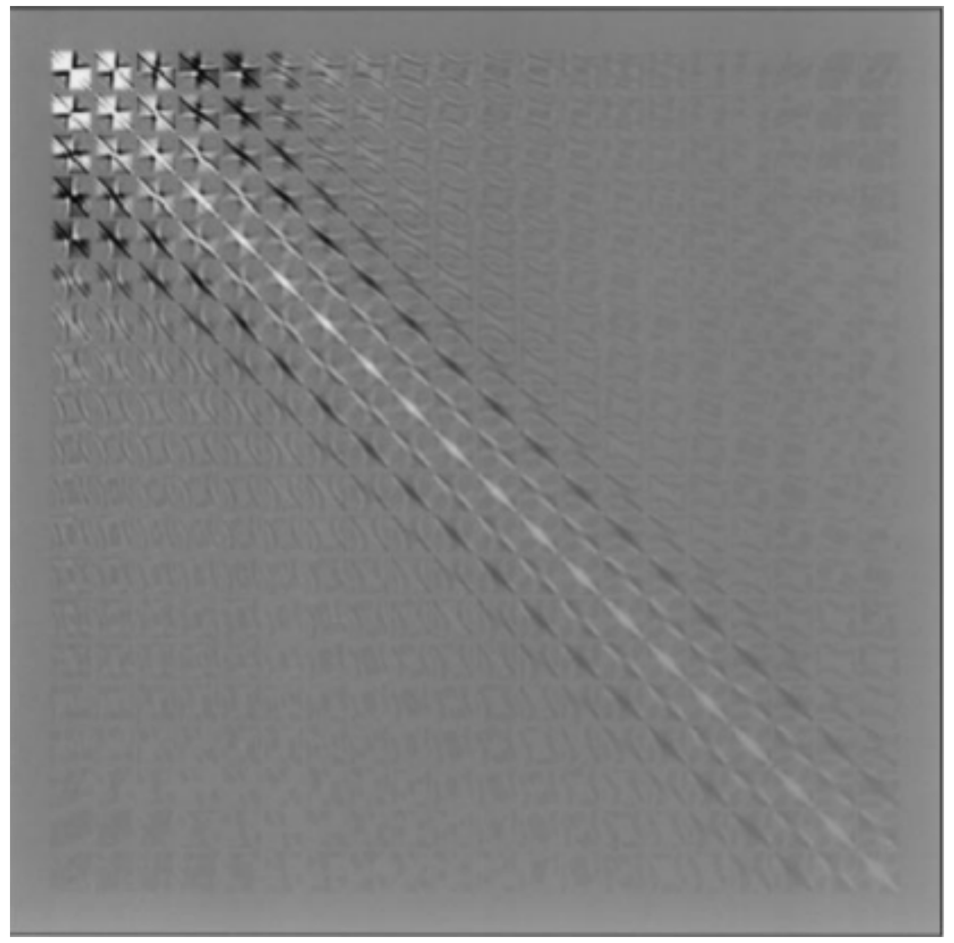
Motivation

FWI	WRI
Strongly nonlinear	Less nonlinear
Dense Gauss-Newton Hessian	Approximately diagonal Hessian

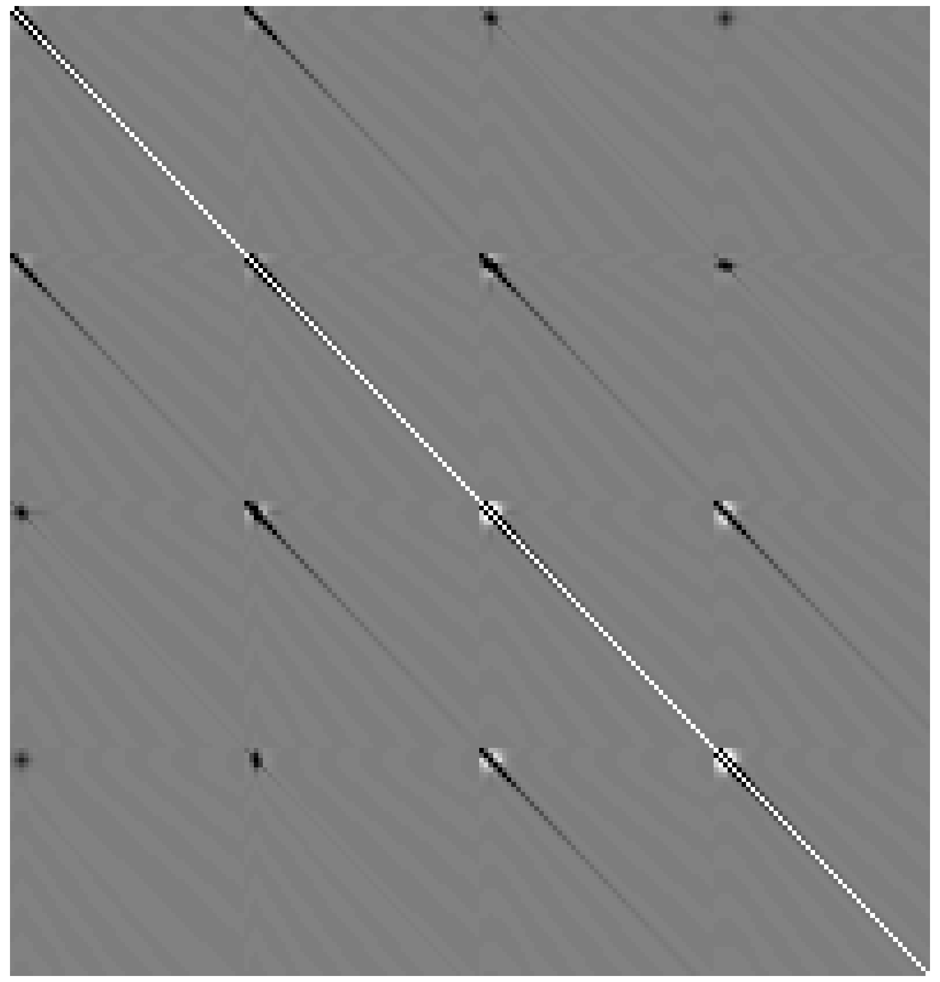
Motivation

[R. Gerhard Pratt, Changsoo Shin and G.J. Hicks, 1998]

[van Leeuwen, T and Herrmann, F J , 2013]

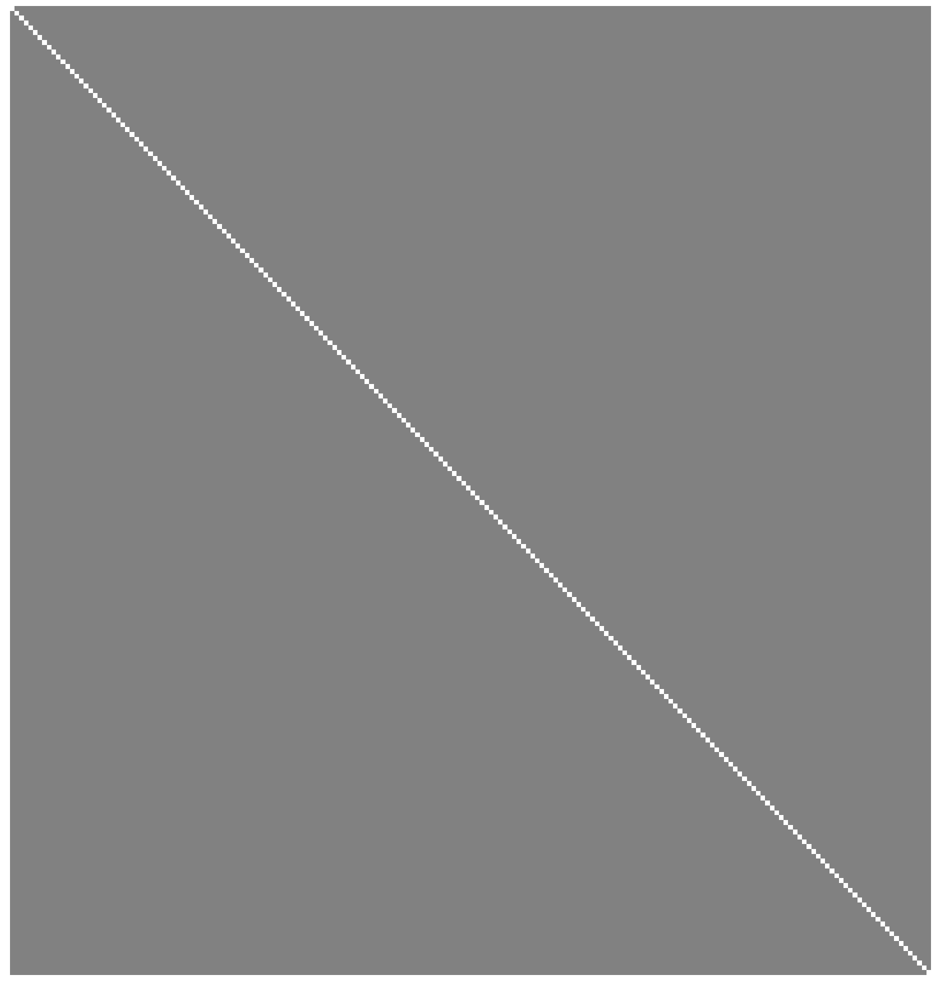


Gauss-Newton Hessian of FWI

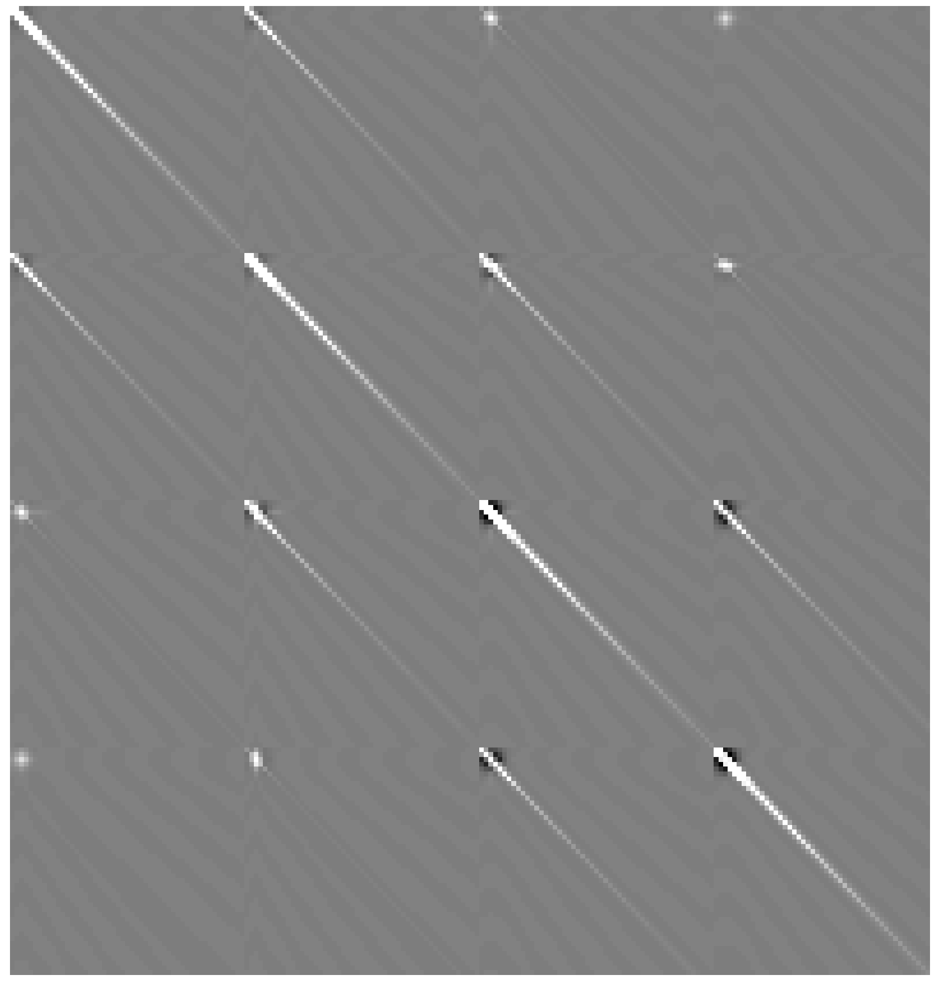


Hessian of WRI

=



−



x 100

Goals

- Set up a reasonable distribution for the model given observed data.
- Derive a practical method to calculate/estimate this distribution.
- Generate different statistical parameters of the model to quantify the uncertainty.

[Tarantola, 1984]

[J. Virieux and S. Operto, 2009]

SLIM

Full-waveform inversion

Original problem:

$$\underset{\mathbf{u},\mathbf{m}}{\text{minimize}} \sum_{k,l} \|\mathbf{P}_k \mathbf{u}_{k,l} - \mathbf{d}_{k,l}\|_2^2$$

$$\text{subject to } \mathbf{A}_{k,l}(\mathbf{m})\mathbf{u}_{k,l} = \mathbf{q}_{k,l},$$

where,

$\mathbf{u}_{k,l}$ – Wavefield of the kth shot at lth frequency

$\mathbf{d}_{k,l}$ – Observed data of the kth shot at lth frequency

$\mathbf{q}_{k,l}$ – Source of the kth shot at lth frequency

$\mathbf{A}_{k,l}$ – Helmholtz of the kth shot at lth frequency

$\mathbf{P}_k$ – Receiver projection operator of the kth shot

$\mathbf{m}$ – Squared-slowness

[van Leeuwen, T and Herrmann, F J , 2013]

[Peters, B, Herrmann, F J and van Leeuwen, T and , 2014]

Wavefield-Reconstruction Inversion (WRI)

Joint optimization problem:

$$\underset{\mathbf{u},\mathbf{m}}{\text{minimize}} \sum_{k,l} \|\mathbf{P}_k \mathbf{u}_{k,l} - \mathbf{d}_{k,l}\|_2^2 + \lambda^2 \|\mathbf{A}_{k,l}(\mathbf{m})\mathbf{u}_{k,l} - \mathbf{q}_{k,l}\|_2^2$$

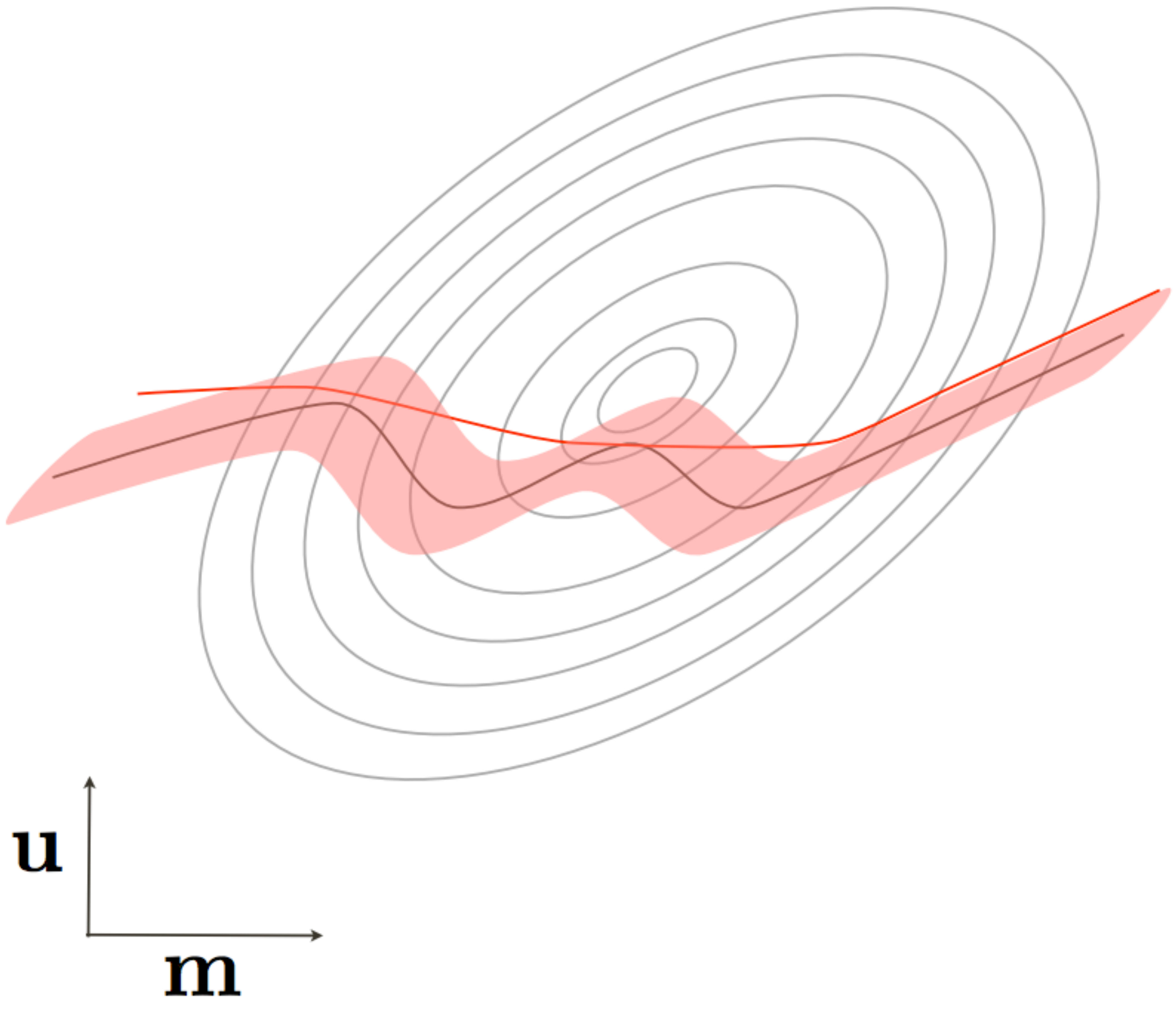
Eliminating $\mathbf{u}$ using variable projection:

$$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \sum_{k,l} \|\mathbf{P}_k \mathbf{u}_{k,l} - \mathbf{d}_{k,l}\|_2^2 + \lambda^2 \|\mathbf{A}_{k,l}(\mathbf{m})\mathbf{u}_{k,l} - \mathbf{q}_{k,l}\|_2^2$$

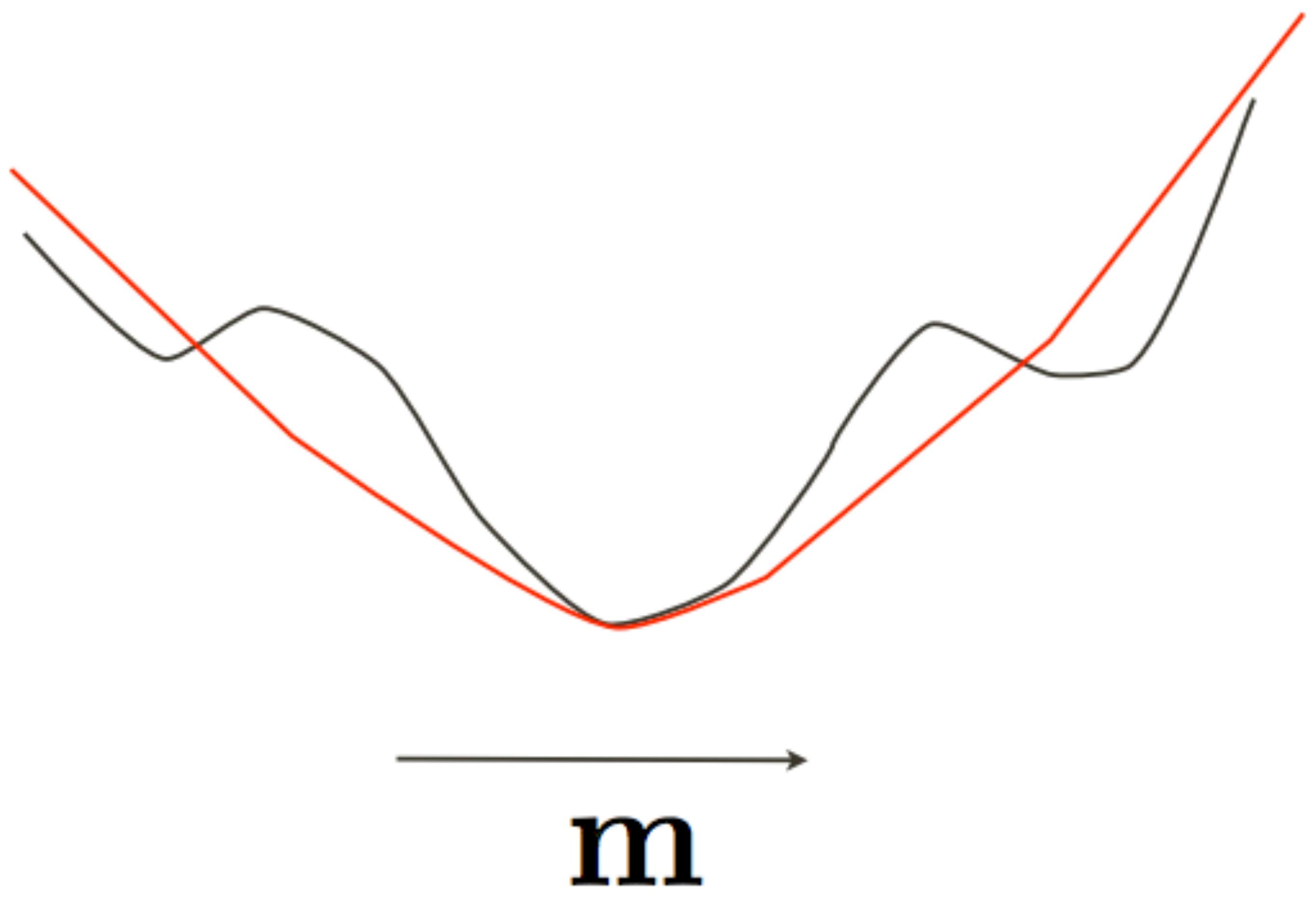
WRI vs FWI

[van Leeuwen, T and Herrmann, F J , 2013]

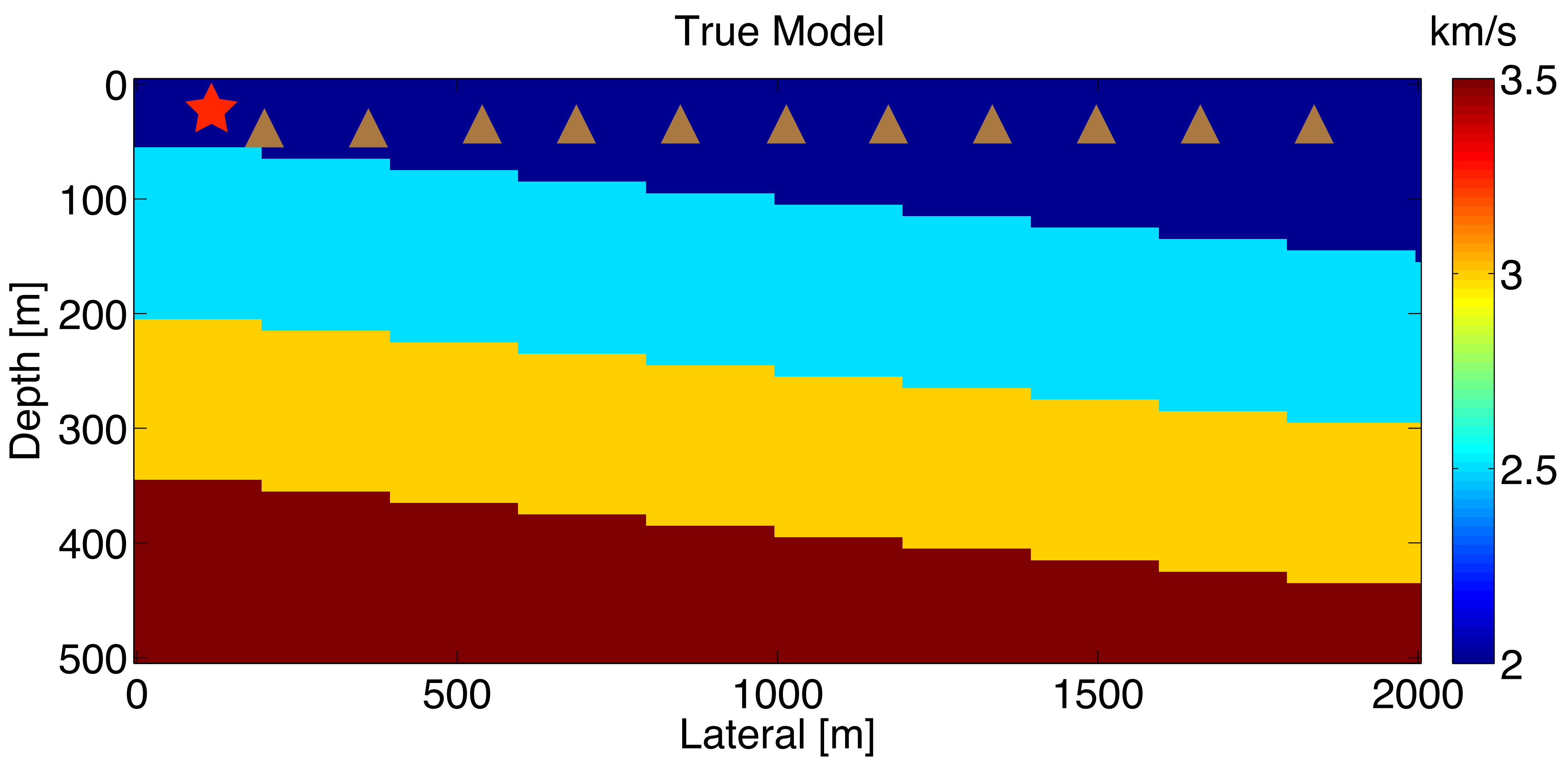
Larger # of degrees of freedom



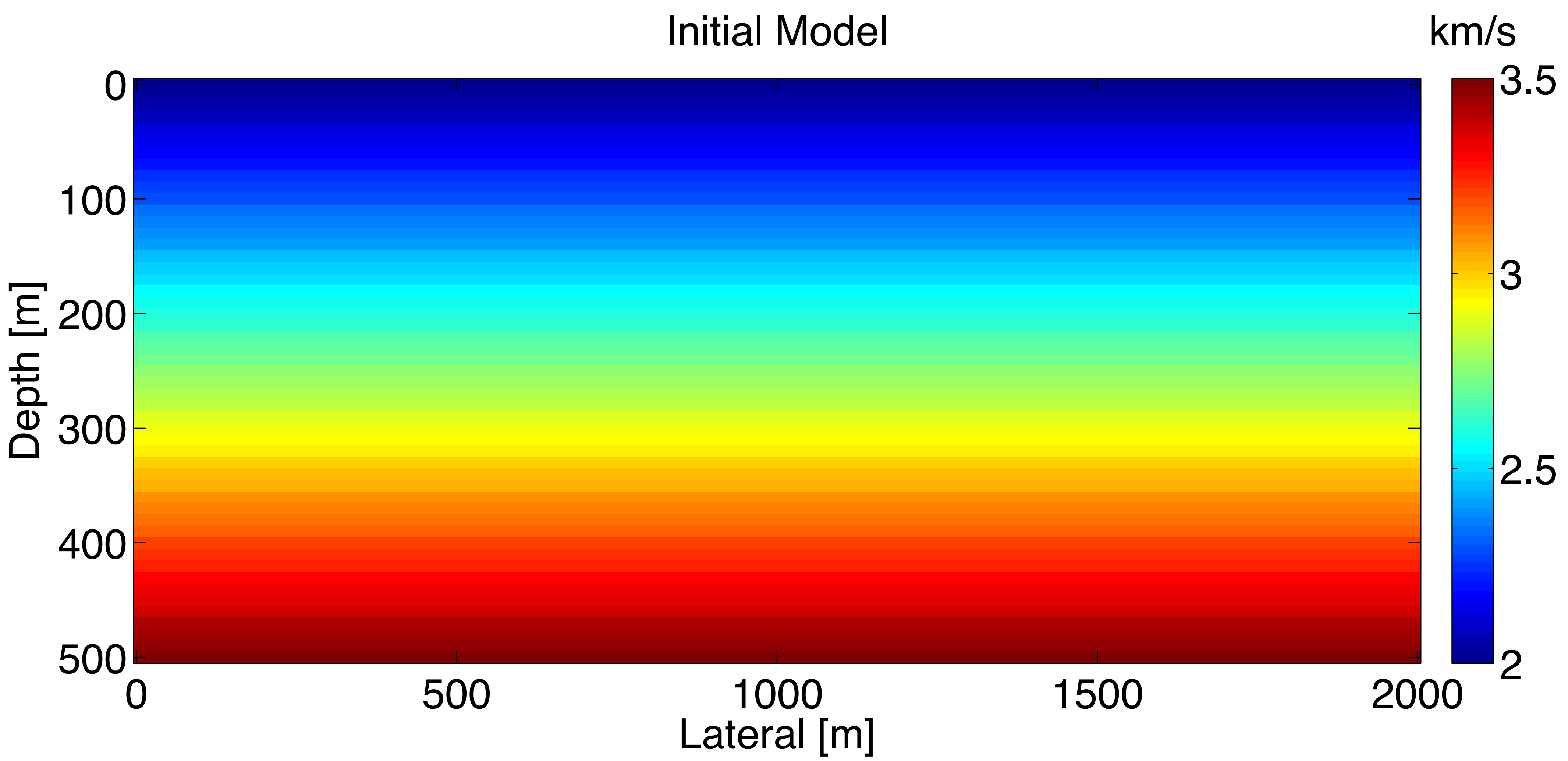
"more convex"



Simple test

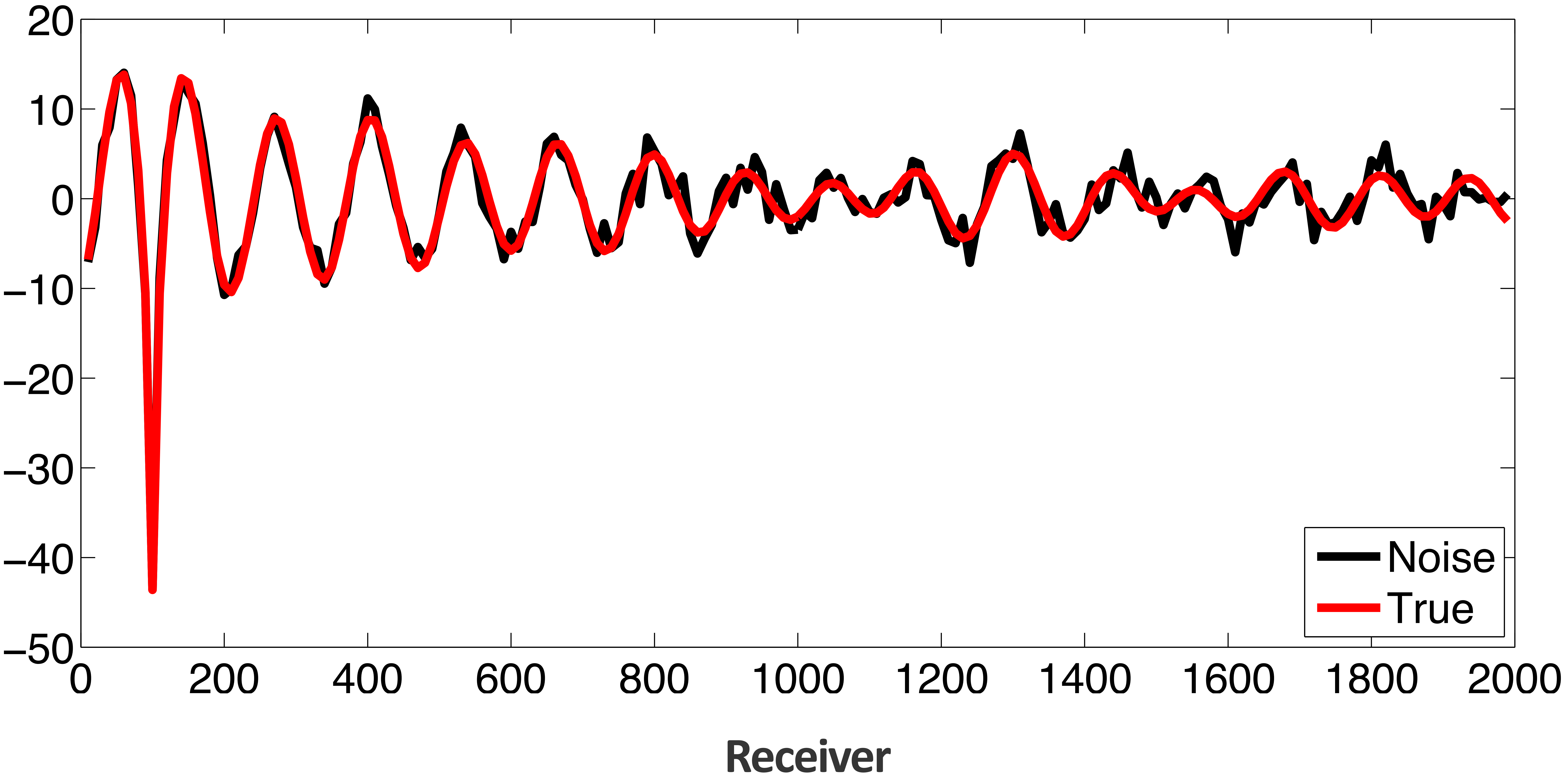


True Model
km/s
3.5
3
2.5
2
0
100
200
300
400
500
Depth [m]
0
500
1000
1500
2000
Lateral [m]



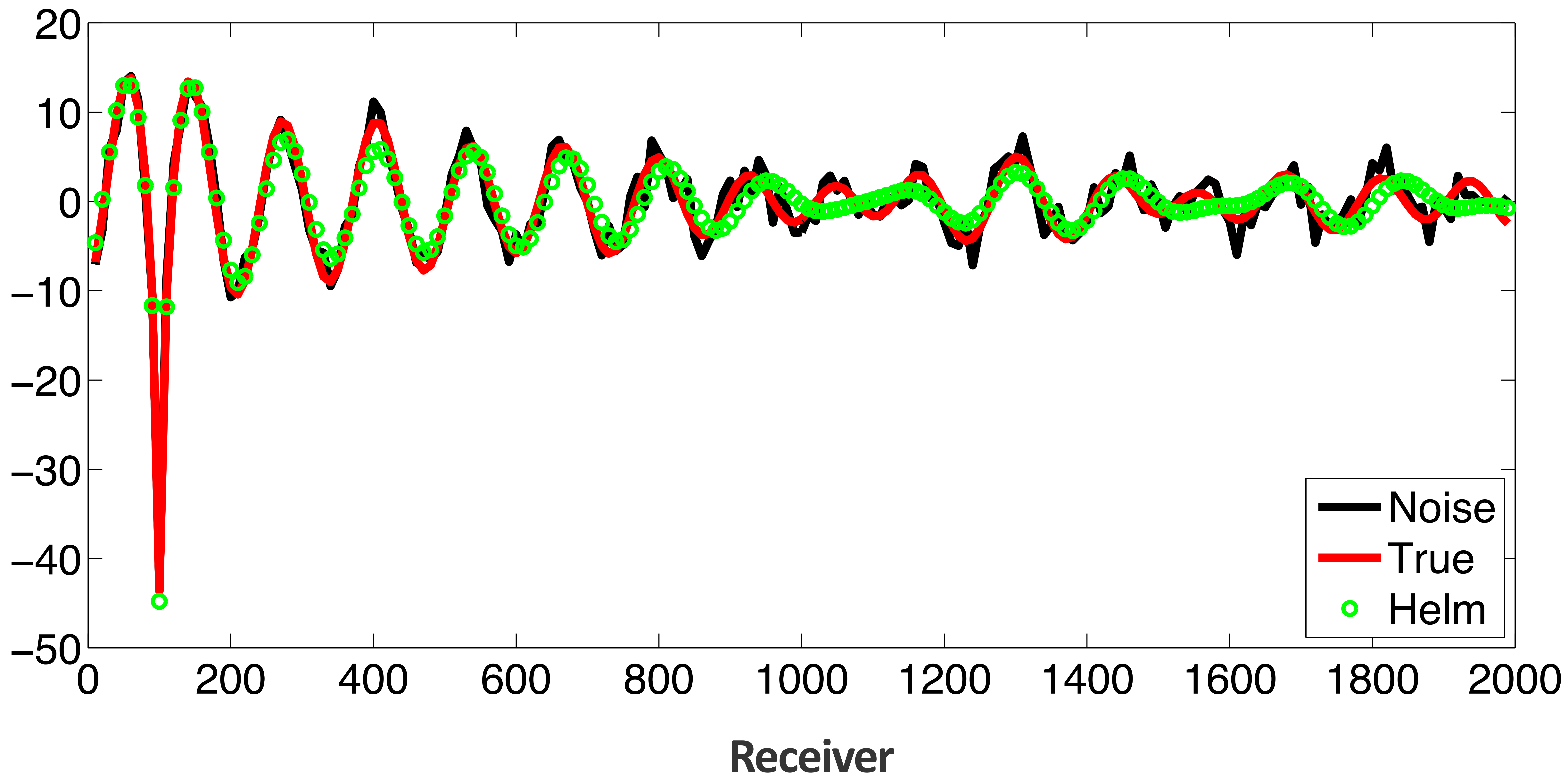
Initial Model
km/s
3.5
3
2.5
2
0
100
200
300
400
500
Depth [m]
0
500
1000
1500
2000
Lateral [m]

Data

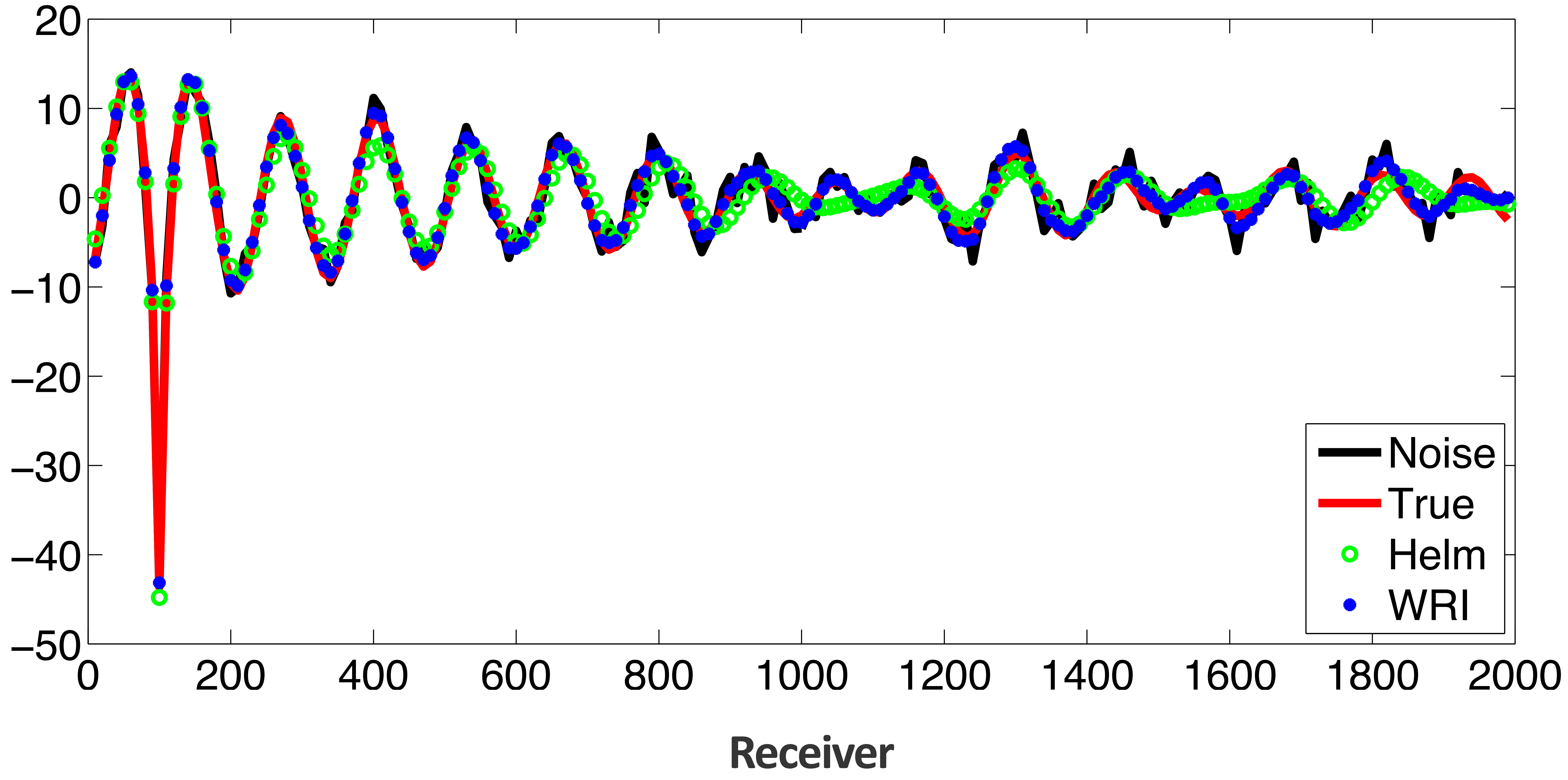


20
10
0
−10
−20
−30
−40
−50
0
200
400
600
800
1000
1200
1400
1600
1800
2000
Receiver
Noise
True

FWI data with initial model



WRI data with initial model



[Tarantola, 1982]

Statistical FWI vs WRI

Full-waveform inversion:

$$\rho_{\text{post}}(\mathbf{m}) \propto \exp\left(\underbrace{-\|\mathbf{P}\mathbf{A}(\mathbf{m})^{-1}\mathbf{q} - \mathbf{d}_{\text{obs}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}}}_{\text{Likelihood}} \underbrace{- \|\mathbf{m} - \mathbf{m}_{\text{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{prior}}}}_{\text{Prior}}\right)$$

Wavefield-reconstruction inversion:

$$\rho_{\text{post}}(\mathbf{m}, \mathbf{u}) \propto \exp(\underbrace{-\|\mathbf{P}\mathbf{u} - \mathbf{d}_{\text{obs}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}} - \lambda^2\|\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{q}\|^2_{\mathbf{\Sigma}^{-1}_{\text{pde}}}}_{\text{Likelihood}}$$

$$\underbrace{-\|\mathbf{m} - \mathbf{m}_{\text{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{prior}}})}_{\text{Prior}}$$

Posterior distribution of WRI

Marginal distribution of **m**:

$$\rho_{\text{post}}(\mathbf{m}) \propto \int \rho_{\text{post}}(\mathbf{m}, \mathbf{u}) \mathrm{d}\mathbf{u}$$

$$= (2\pi)^{N_{\mathbf{u}}/2} |\boldsymbol{\Sigma}_{\mathbf{u}}|^{1/2} \rho_{\text{post}}(\mathbf{m}, \overline{\mathbf{u}}(\mathbf{m}))$$

Here

$$\begin{pmatrix} \lambda \boldsymbol{\Sigma}_{\text{pde}}^{-1/2} \mathbf{A} \\ \boldsymbol{\Sigma}_{\text{noise}}^{-1/2} \mathbf{P} \end{pmatrix} \overline{\mathbf{u}}(\mathbf{m}) = \begin{pmatrix} \lambda \boldsymbol{\Sigma}_{\text{pde}}^{-1/2} \mathbf{q} \\ \boldsymbol{\Sigma}_{\text{noise}}^{-1/2} \mathbf{d}_{\text{obs}} \end{pmatrix}$$

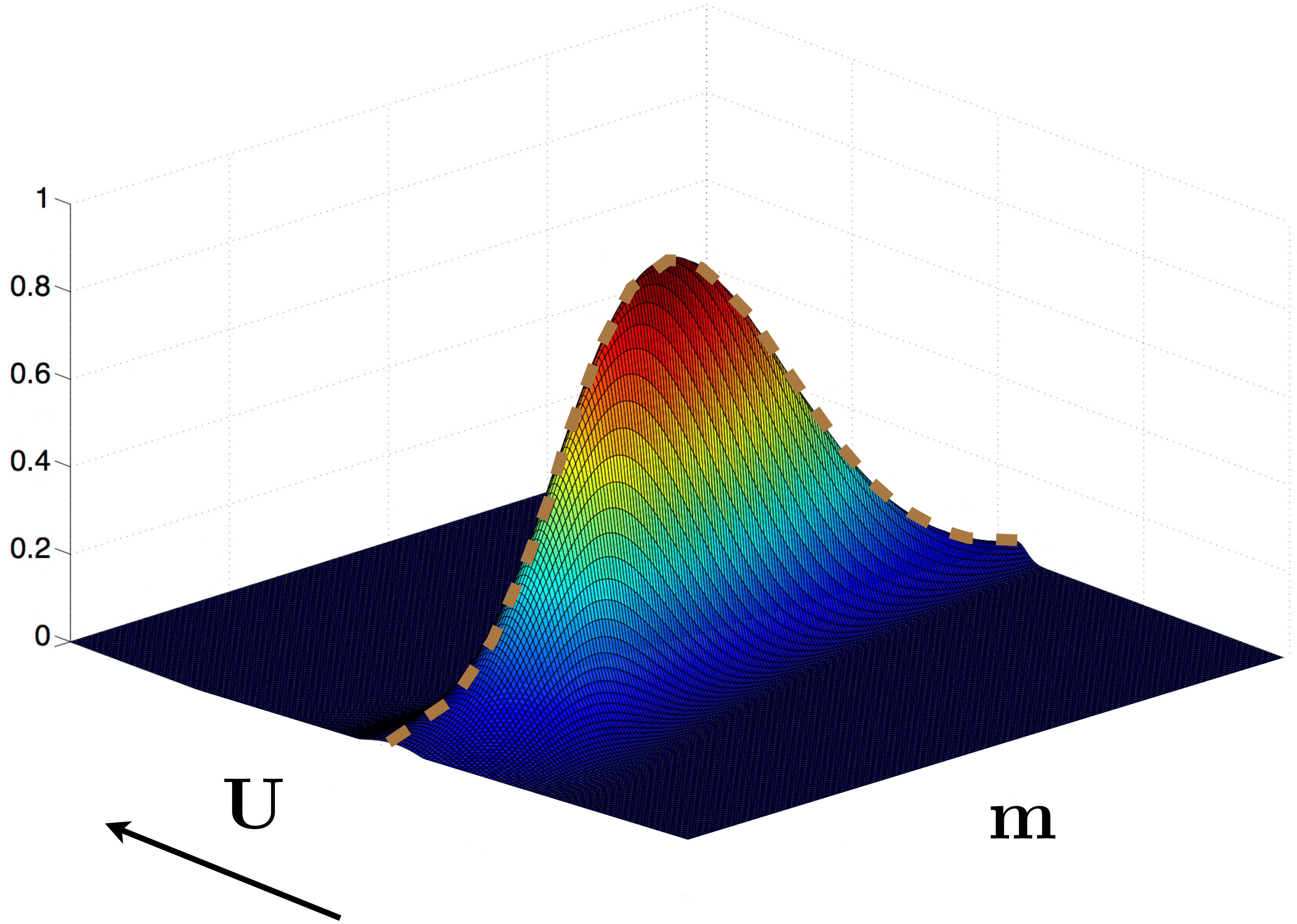
$$|\boldsymbol{\Sigma}_{\mathbf{u}}| = \det((\lambda^2 \mathbf{A}^T \boldsymbol{\Sigma}_{\text{pde}}^{-1} \mathbf{A} + \mathbf{P}^T \boldsymbol{\Sigma}_{\text{noise}}^{-1} \mathbf{P})^{-1})$$ **Huge computational cost !!!**

Posterior distribution of WRI

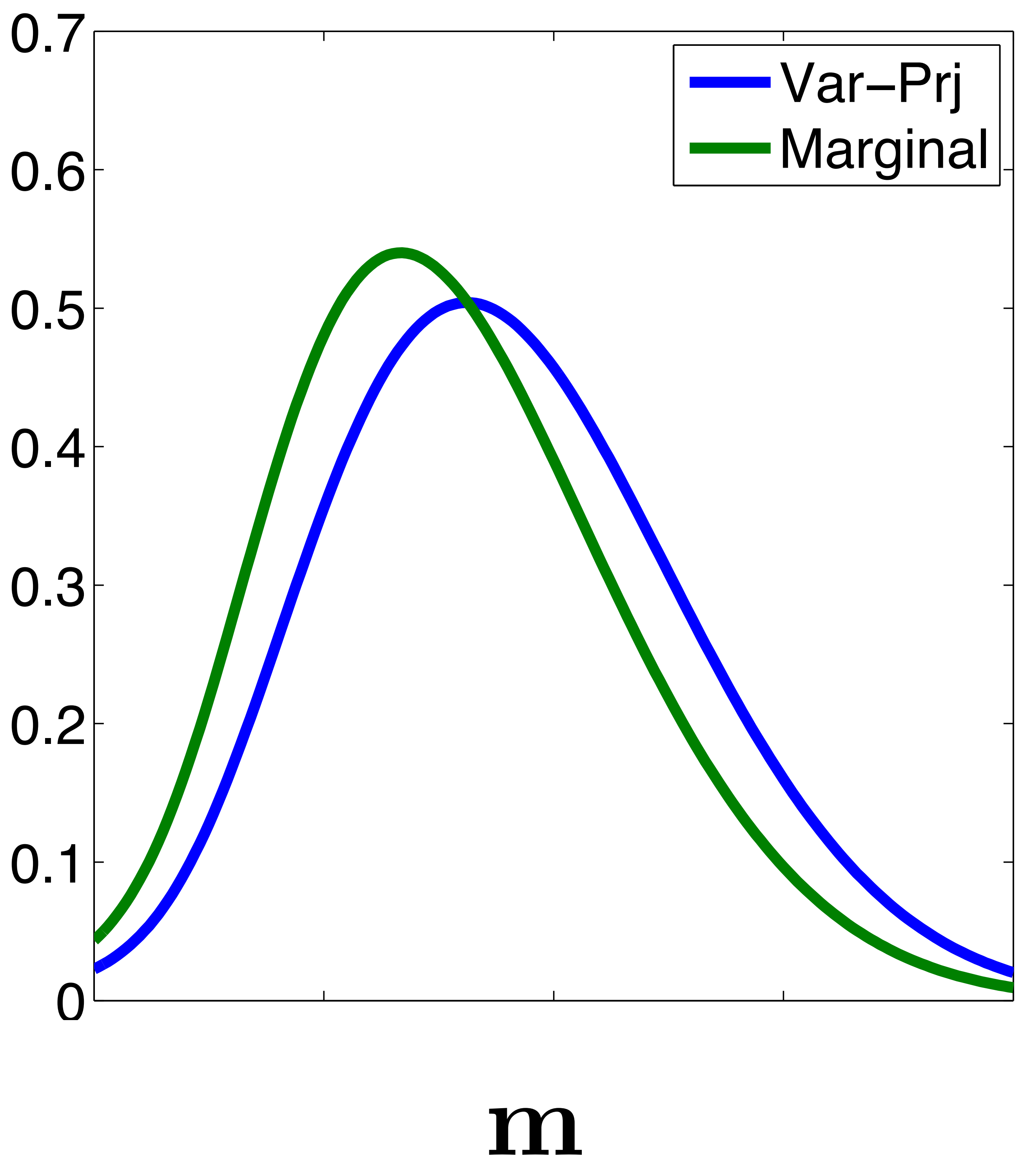
Approximate the marginal distribution:

$$
\begin{aligned}
\rho_{\mathrm{post}}(\mathbf{m}) &\propto \int \rho_{\mathrm{post}}(\mathbf{m}, \mathbf{u}) \mathrm{d}\mathbf{u} \\
&= (2\pi)^{N_{\mathbf{u}}/2} |\boldsymbol{\Sigma}_{\mathbf{u}}|^{1/2} \rho_{\mathrm{post}}(\mathbf{m}, \overline{\mathbf{u}}(\mathbf{m})) \\
&\approx \mathbf{C} \rho_{\mathrm{post}}(\mathbf{m}, \overline{\mathbf{u}}(\mathbf{m})) \\
&\propto \rho_{\mathrm{post}}(\mathbf{m}, \overline{\mathbf{u}}(\mathbf{m}))
\end{aligned}
$$

Marginal distribution vs Approximate distribution



Joint distribution



Marginal distribution
vs Approximate distribution

Quantify the uncertainty

Goal : Quantify the uncertainty based on the posterior distribution $\rho_{\text{post}}(\mathbf{m})$

Solution:

- Integrate the posterior distribution

Huge computational cost!!!

Quantify the uncertainty

Goal : Quantify the uncertainty based on the posterior distribution $\rho_{\mathrm{post}}(\mathbf{m})$

Solution:

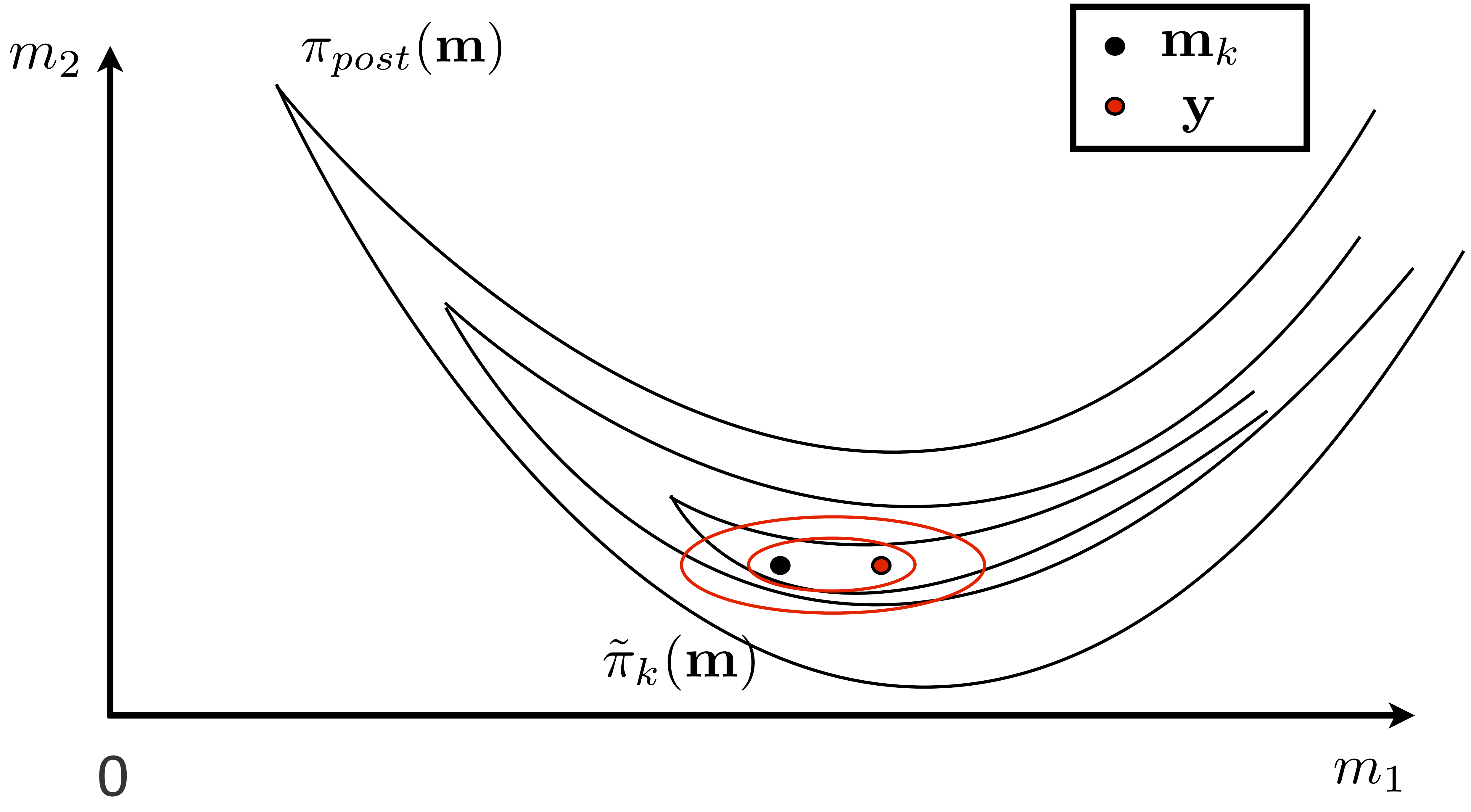
- Integrate the posterior distribution
- McMC method to sample the posterior distribution

[James Martin *et al*, 2012]

SLIM

McMC method

Newton Type McMC: $\tilde{\pi}_k(\mathbf{m}) \backsim \mathcal{N}(\mathbf{m}_k - \mathbf{H}_k^{-1}\mathbf{g}_k, \mathbf{H}_k^{-1})$



Computational cost:

1) Low rank approximation of the Hessian.

2) Number of PDE solvers ~ Number of samples.

Quantify the uncertainty

Goal : Quantify the uncertainty based on the posterior distribution $\rho_{\text{post}}(\mathbf{m})$

Solution:

- Integrate the posterior distribution

- McMC method to sample the posterior distribution
  - Advantage: the true uncertainty can be quantified
  - Disadvantage: Huge computational cost

Quantify the uncertainty

Goal : Quantify the uncertainty based on the posterior distribution $\rho_{\text{post}}(\mathbf{m})$

Solution:

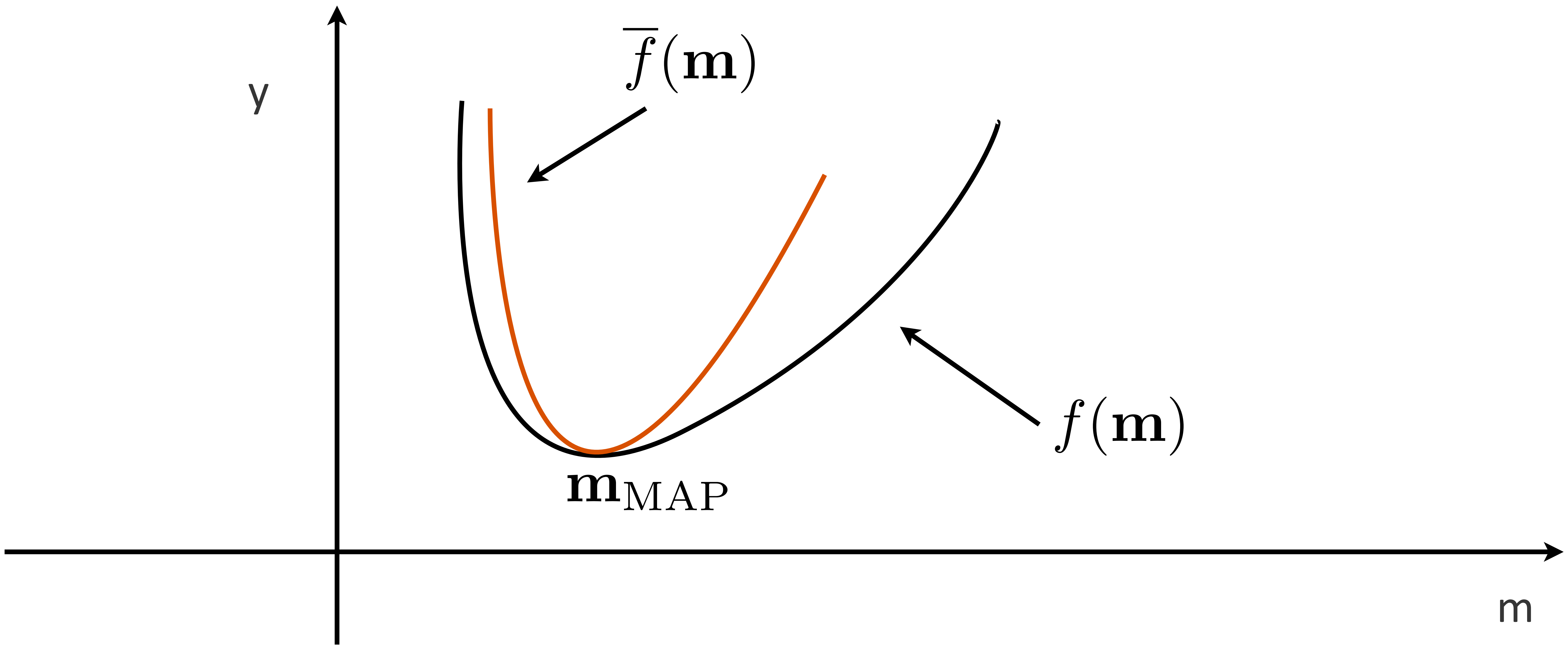
- Integrate the posterior distribution

- McMC method to sample the posterior distribution
  - Advantage: the true uncertainty can be quantified
  - Disadvantage: Huge computational cost

- Use an approximate distribution to quantify the uncertainty

Quadratic approximation

$$f(\mathbf{m}) \approx f(\mathbf{m}_{\mathrm{MAP}}) + \mathbf{g}^T(\mathbf{m} - \mathbf{m}_{\mathrm{MAP}}) + \frac{1}{2}(\mathbf{m} - \mathbf{m}_{\mathrm{MAP}})^T \mathbf{H}(\mathbf{m} - \mathbf{m}_{\mathrm{MAP}}) := \overline{f}(\mathbf{m})$$



Hessian of FWI

[R. Gerhard Pratt, Changsoo Shin and G.J. Hicks, 1998]

Full Hessian of FWI:

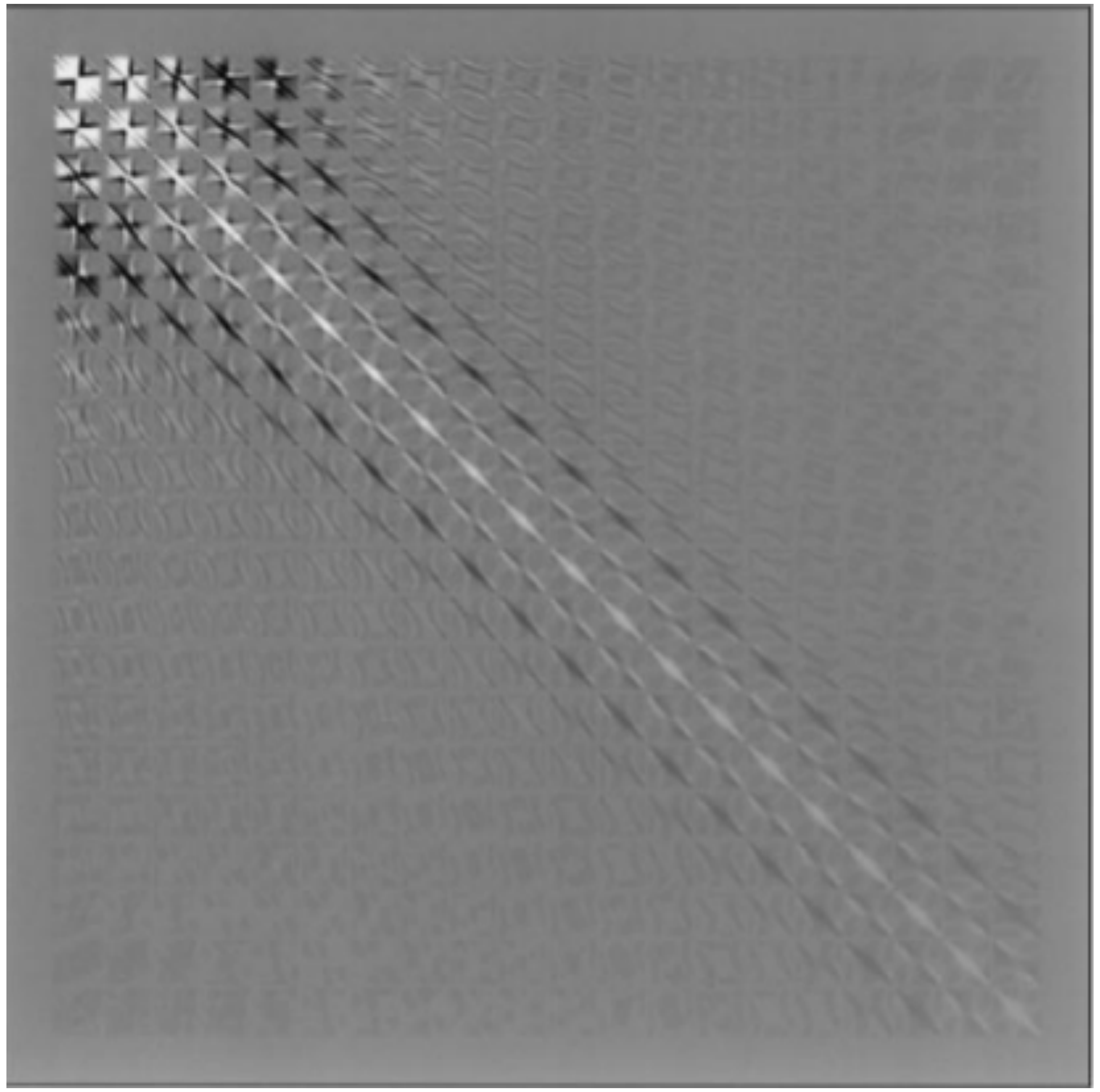
$$\mathbf{H} = \mathbf{H_{GN}} + \mathbf{H_2}$$

Gauss-Newton Hessian of FWI:

$$\mathbf{H_{GN}} = \mathbf{G^T A^{-T} P^T P A^{-1} G}$$

where

$$\mathbf{G} = \frac{\partial \mathbf{A(m)u}}{\partial \mathbf{m}}$$ sparse



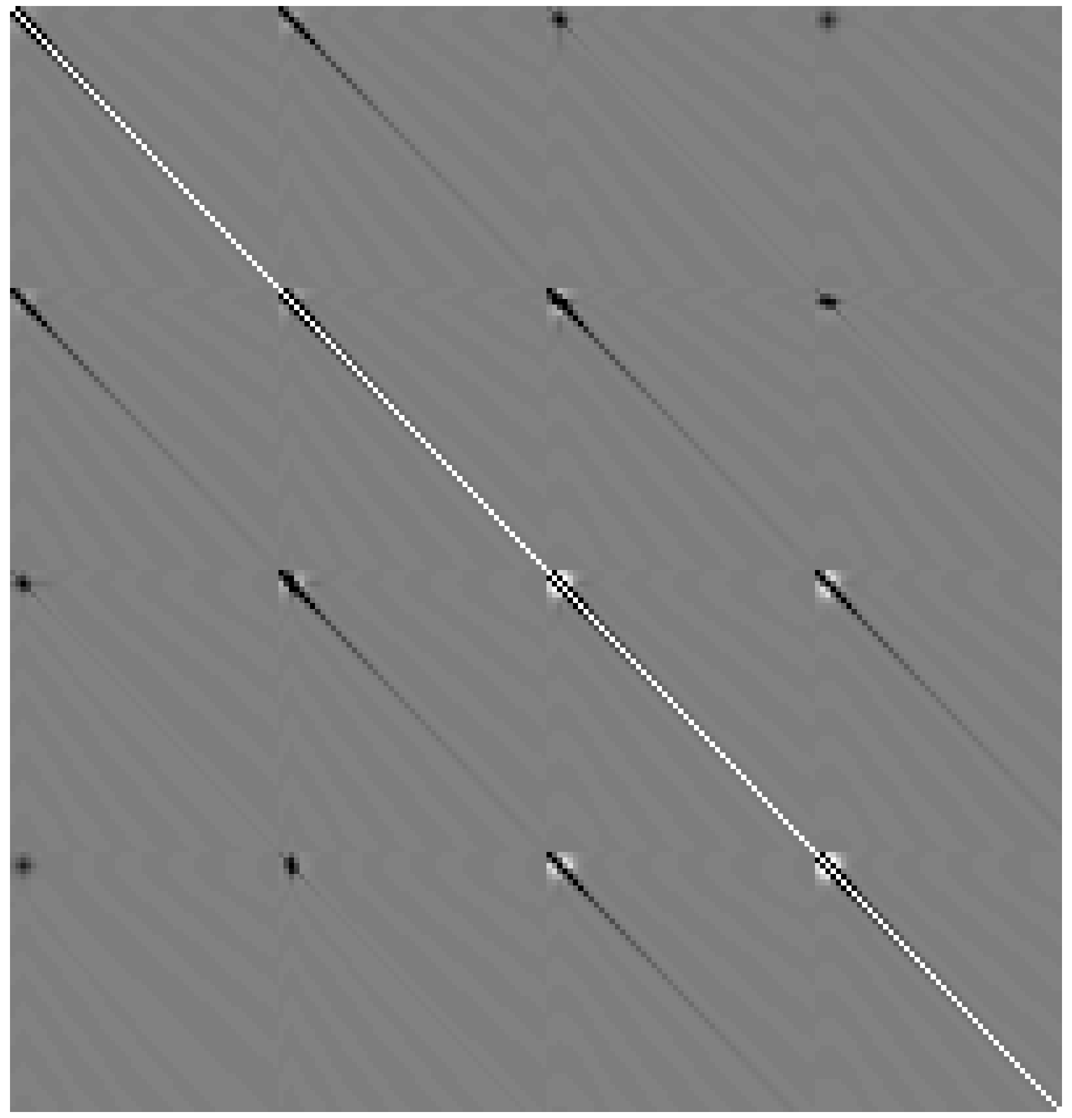
Gauss-Newton Hessian of FWI

Hessian of WRI

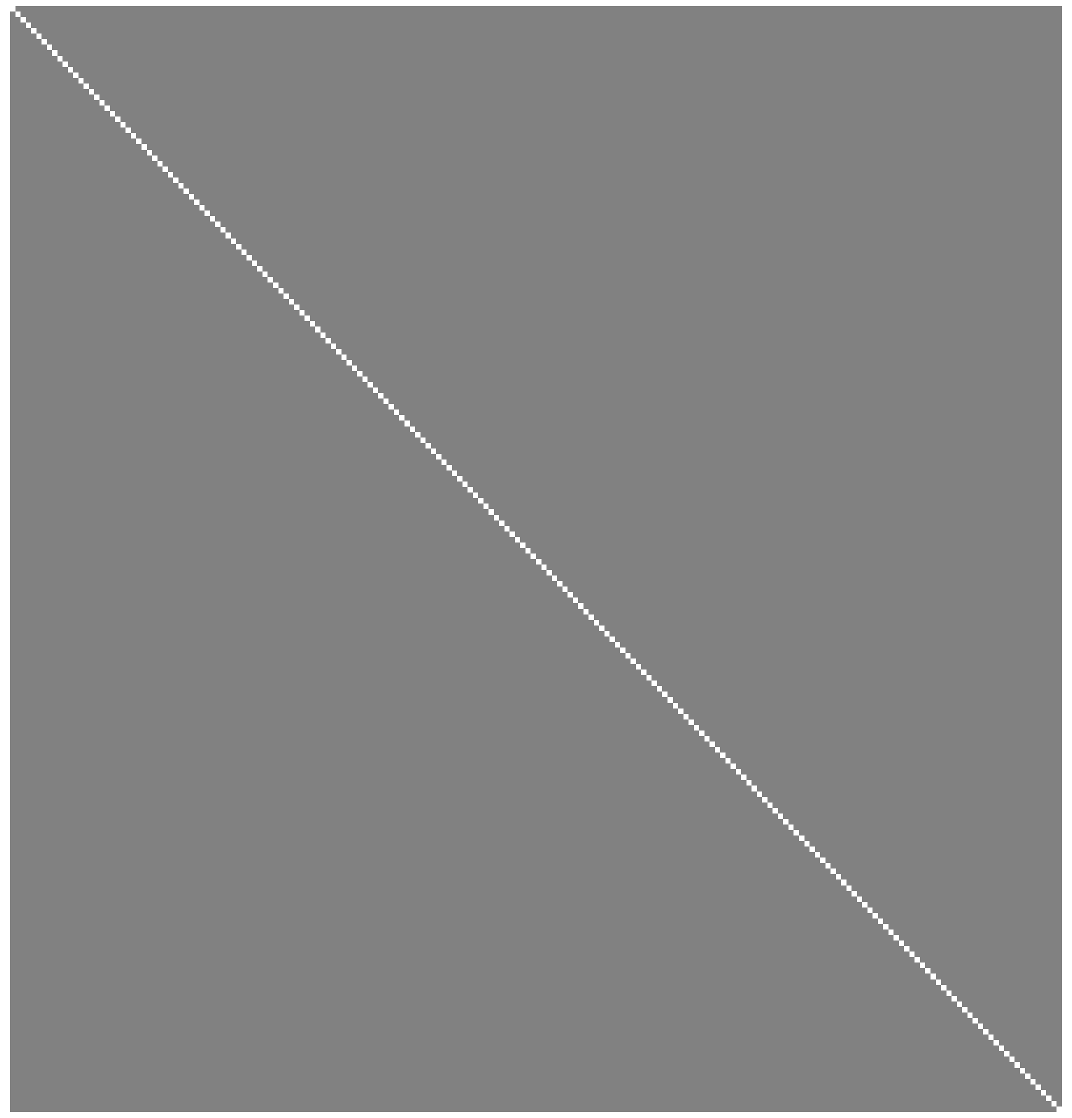
[van Leeuwen, T and Herrmann, F J , 2013]

Hessian of WRI:

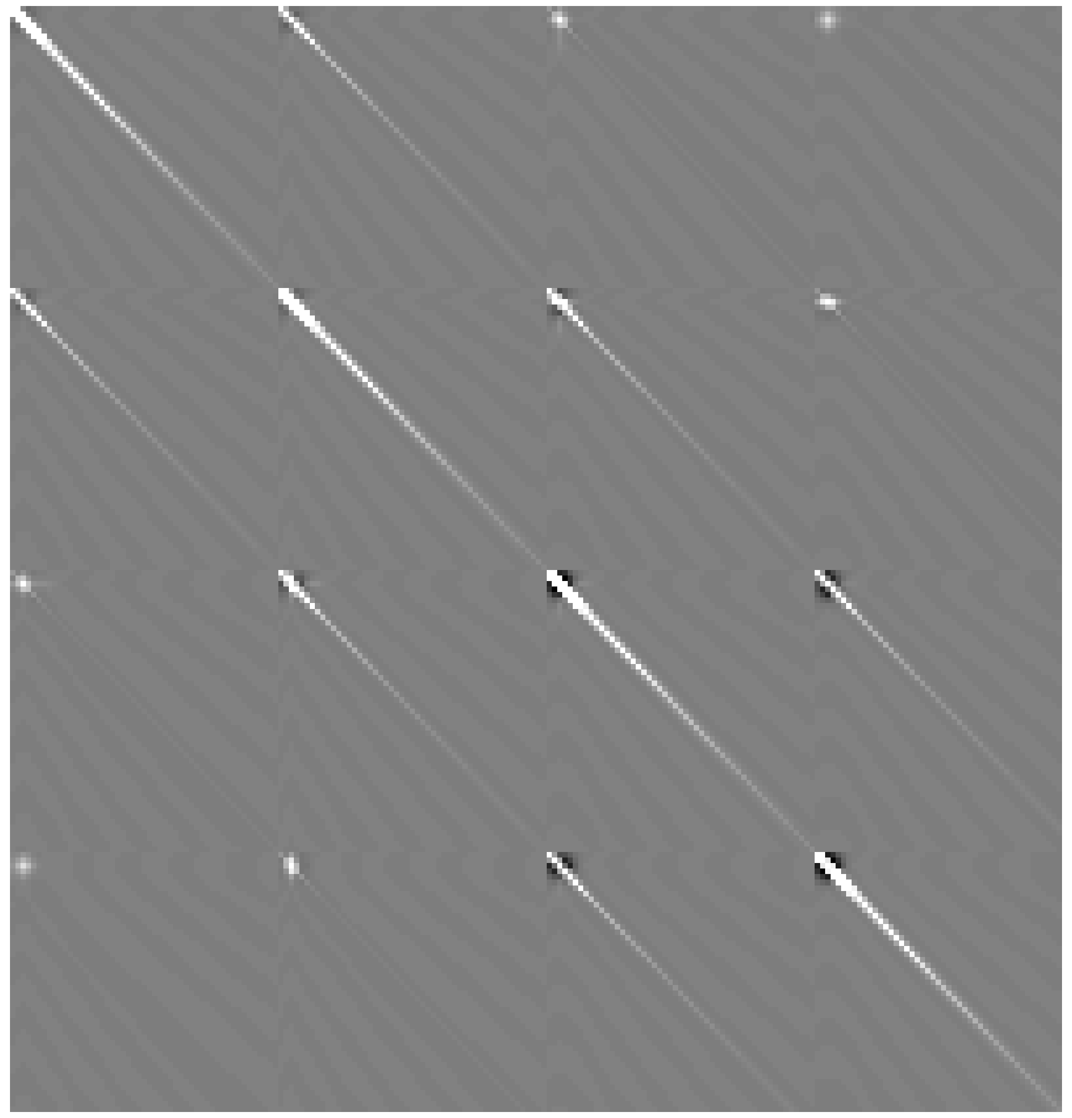
$$\mathbf{H} = \lambda^2(\mathbf{G^TG} - \mathbf{G^T}(\mathbf{I} + \lambda^{-2}\mathbf{A^{-T}P^TPA^{-1}})^{-1}\mathbf{G})$$



=



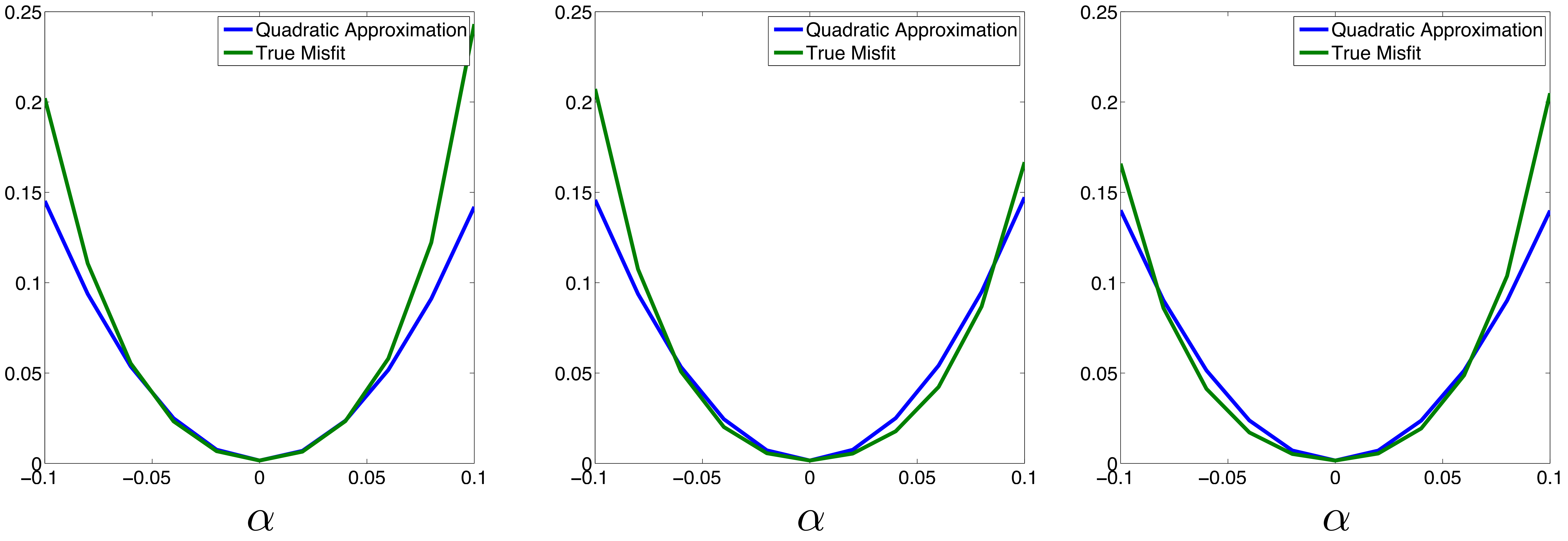
−



x 100

Quadratic approximation

$$f(\mathbf{m}_{MAP} + \alpha \mathrm{d}\mathbf{m})$$



dm – Different random directions

Quantify the uncertainty

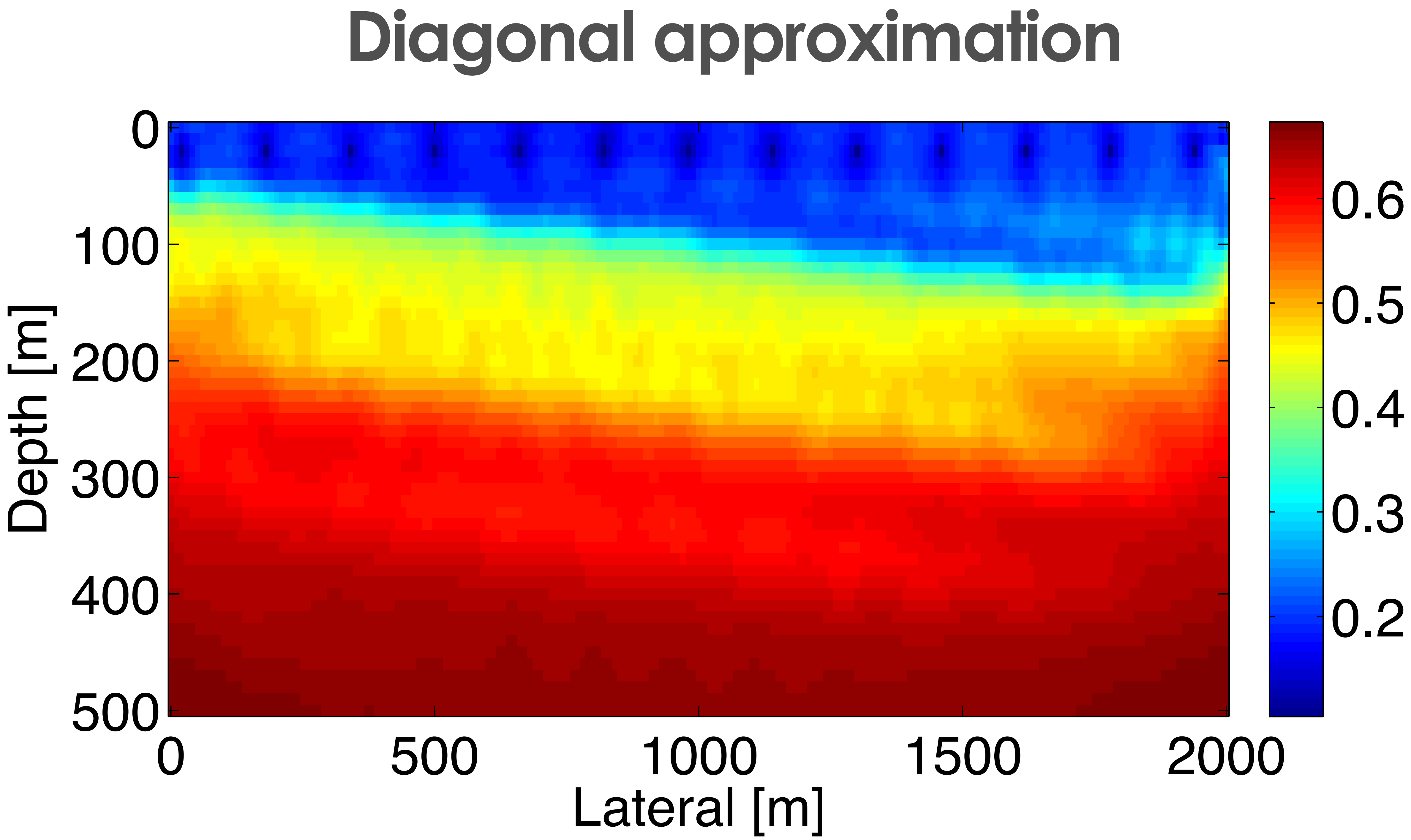
Goal : Quantify the uncertainty based on the posterior distribution $\rho_{\text{post}}(\mathbf{m})$

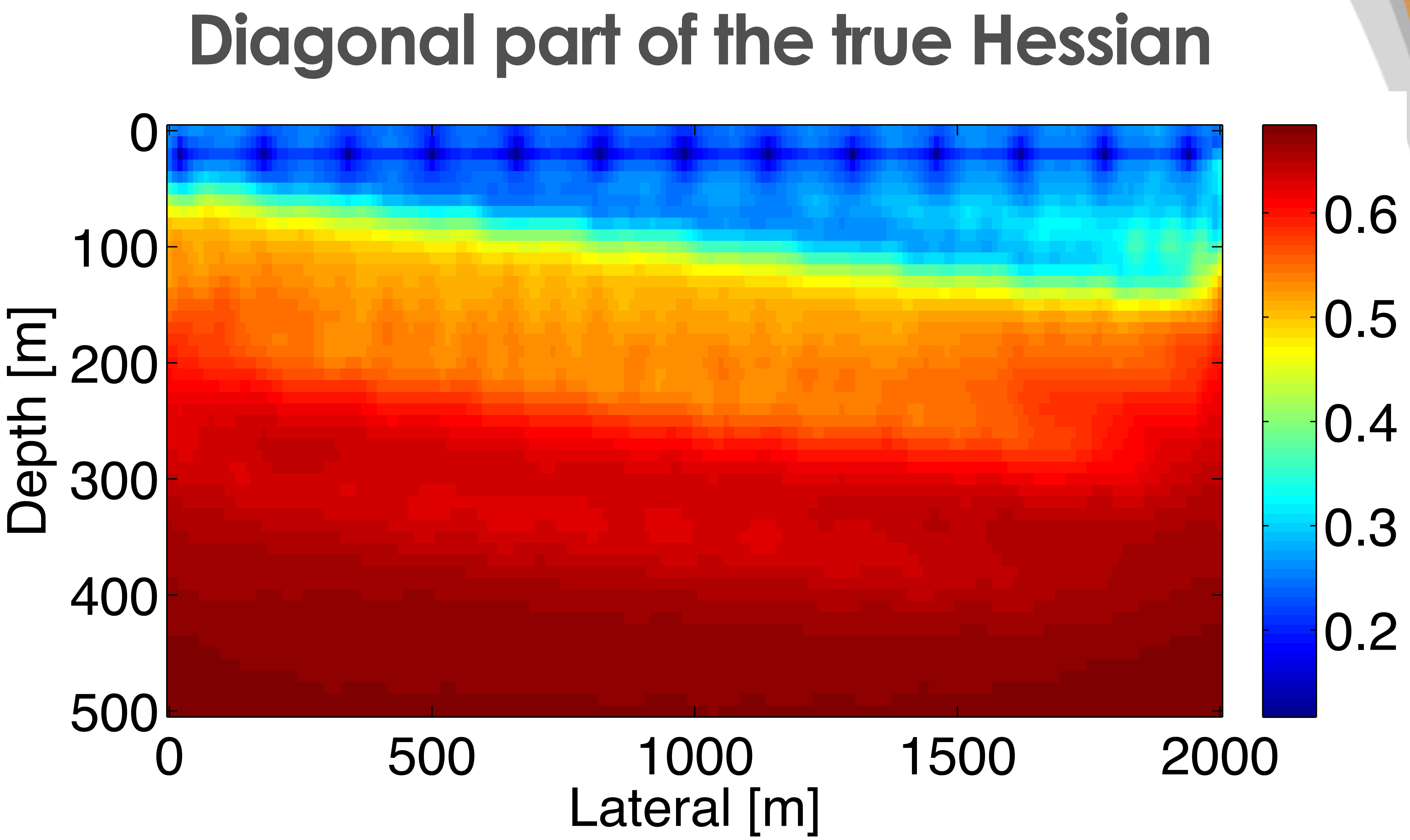
Solution:

- Use an approximate distribution to quantify the uncertainty.
  - Quantify the uncertainty by estimating the diagonal part of the inverse of the Hessian.

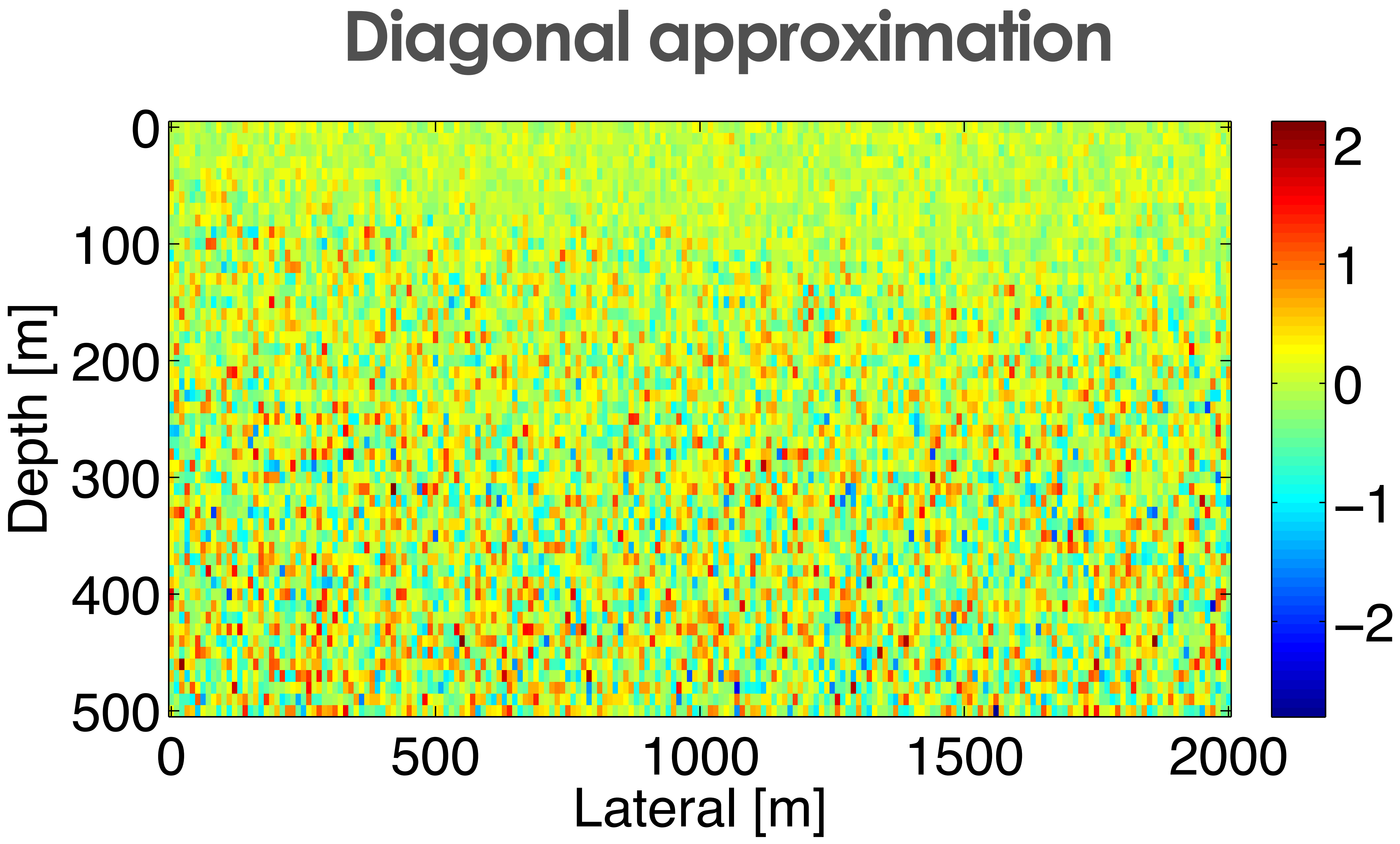
Sparse, no additional PDE solves are required!

Diagonal approximation vs true Hessian

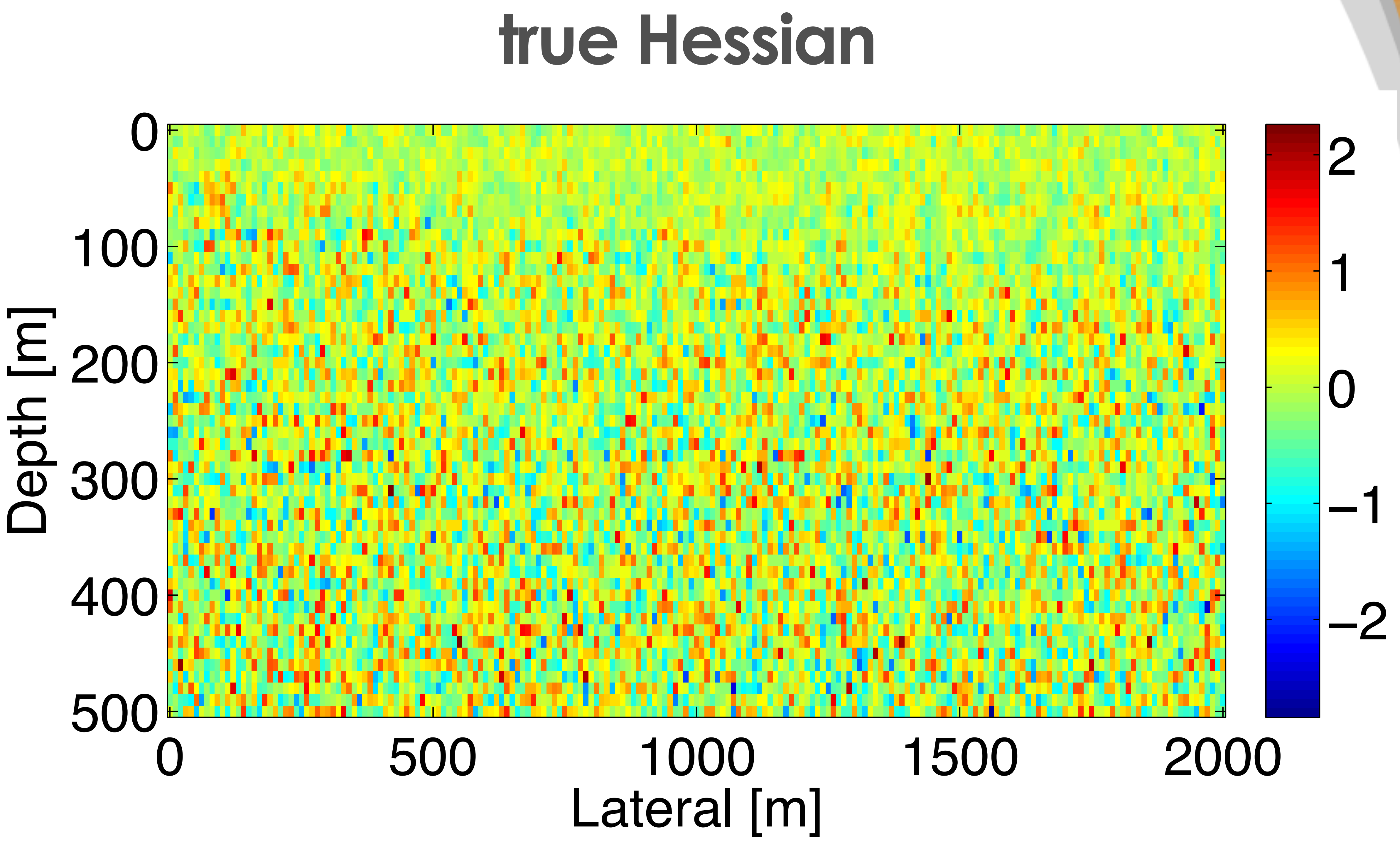




Diagonal approximation vs true Hessian
– random realizations



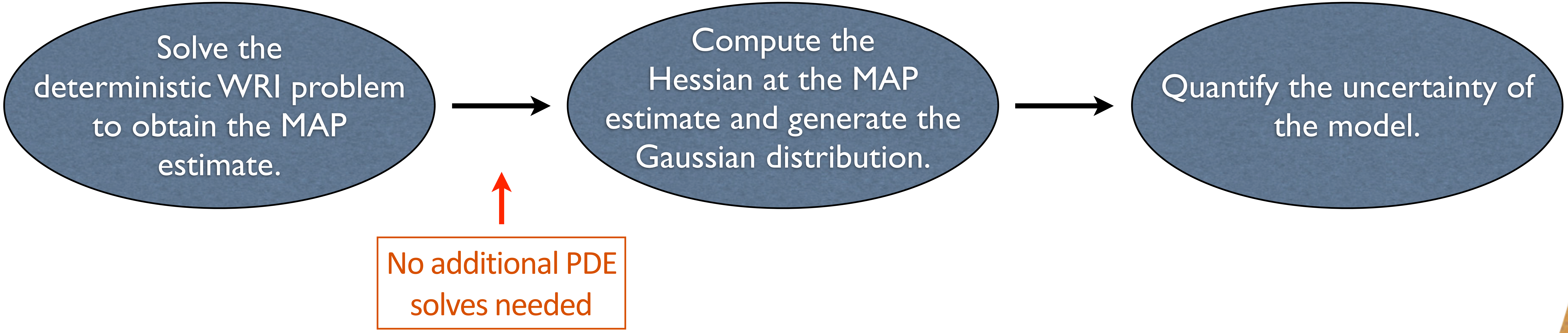
$\mathbf{H}_a^{-1/2}\mathbf{r}$



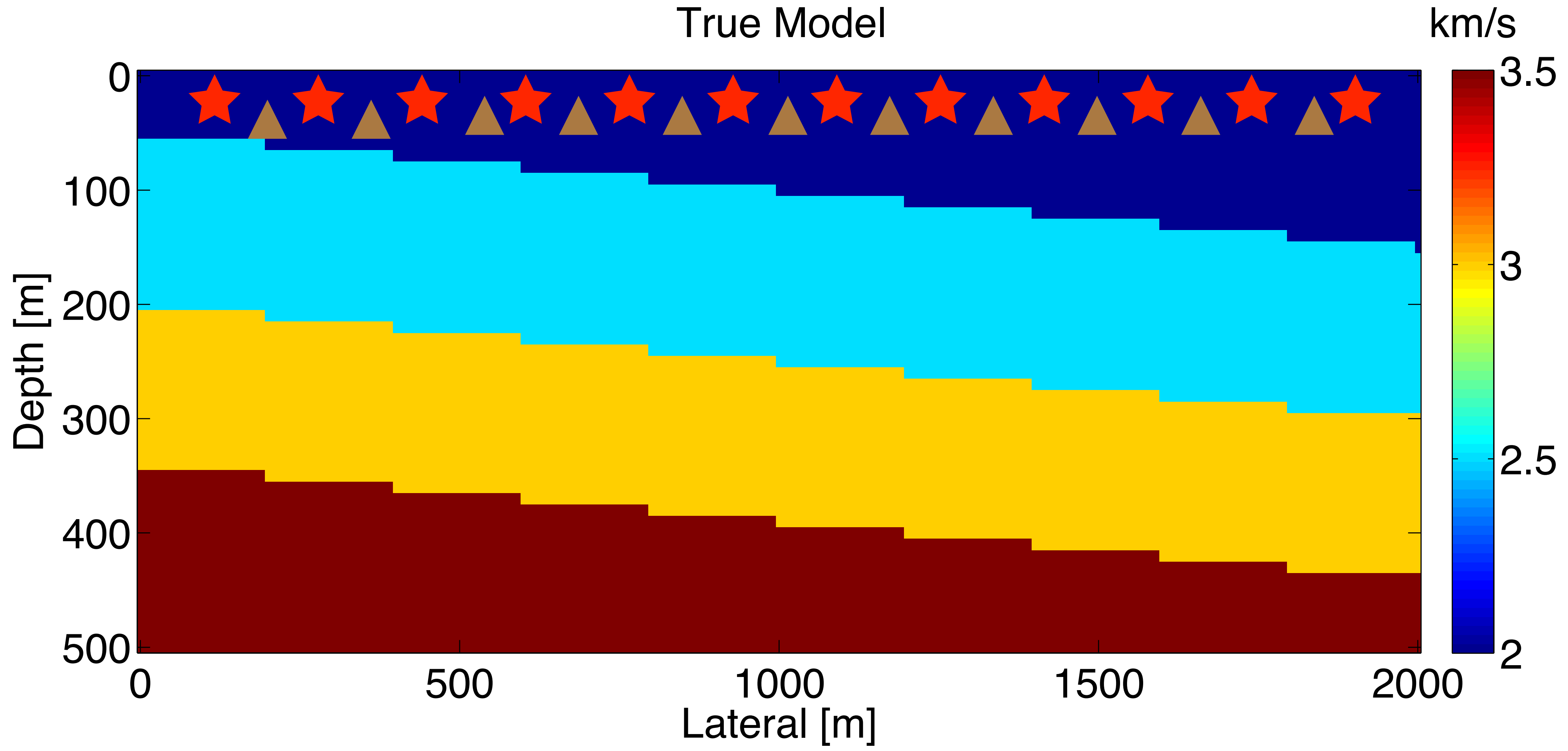
$\mathbf{H}_t^{-1/2}\mathbf{r}$

$$\mathbf{r} \sim \mathcal{N}(0, \mathbf{I})$$

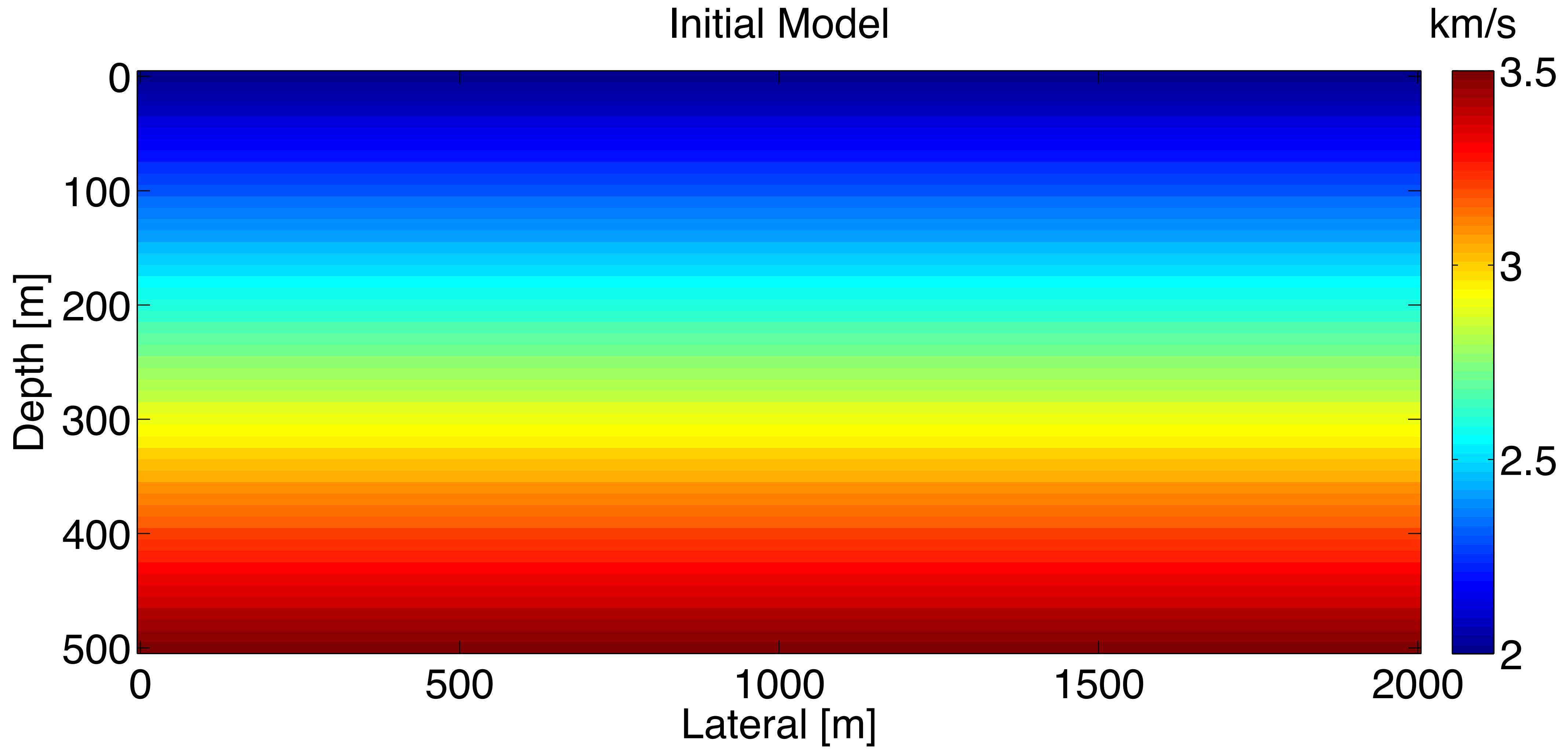
Workflow
– uncertainty quantification



Numerical experiment



(a)

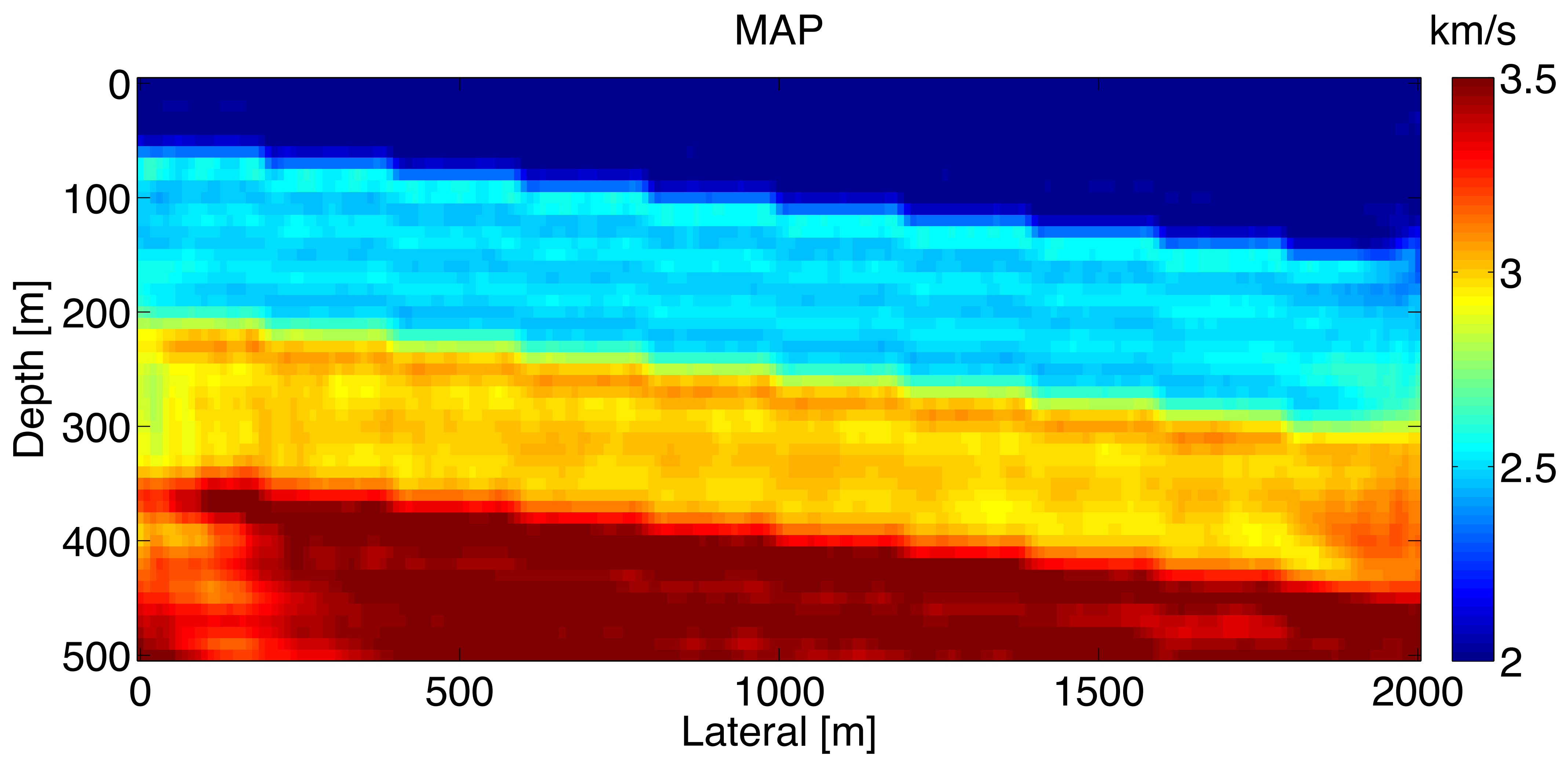


(b)

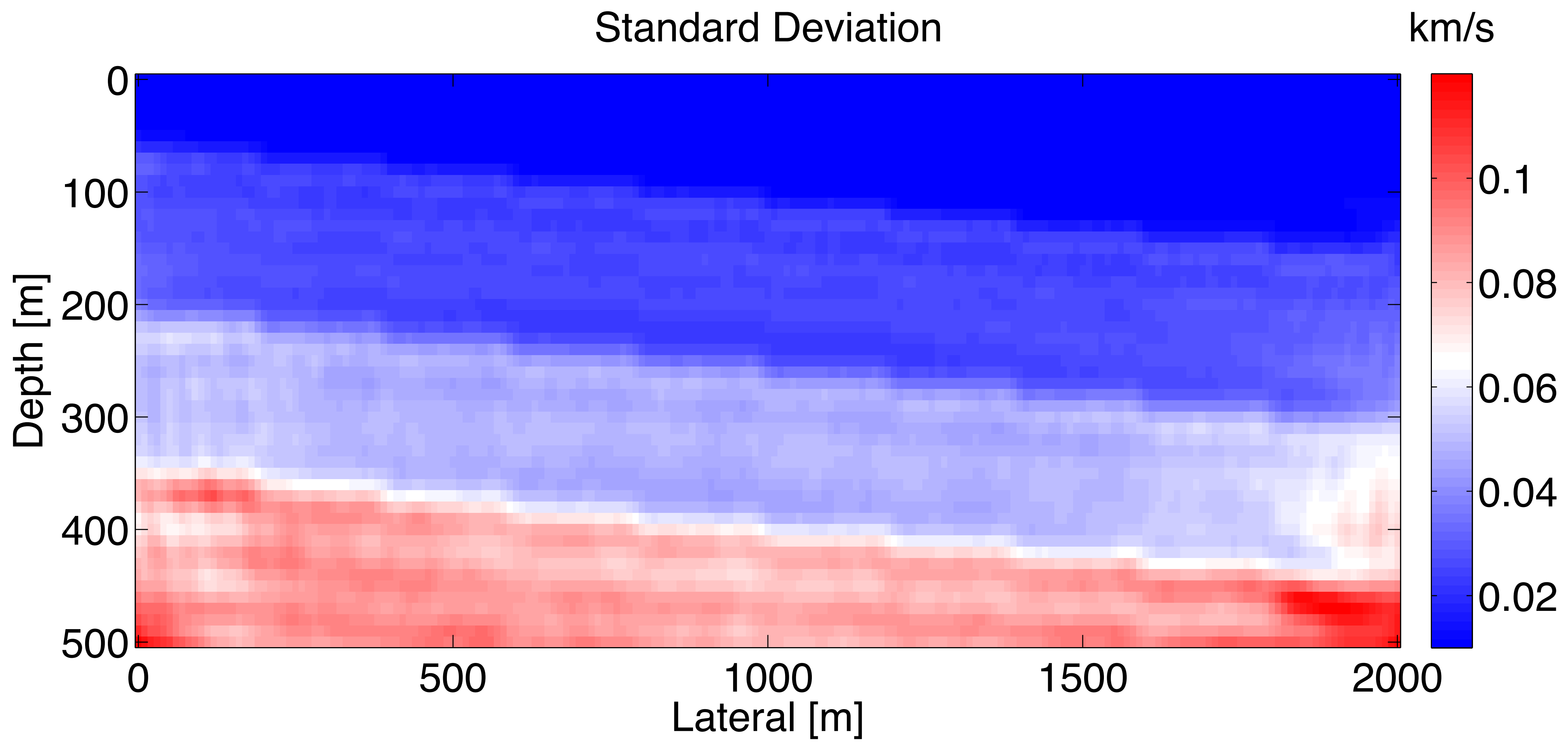
Model size: 500m x 2000m
Source spacing: 80m
Receiver spacing: 20m
Fixed spread 2km
Frequency : 10-30 Hz

Standard deviation of data noise: 0.5
Standard deviation of pde: 0.5
lambda: 1

Simple model

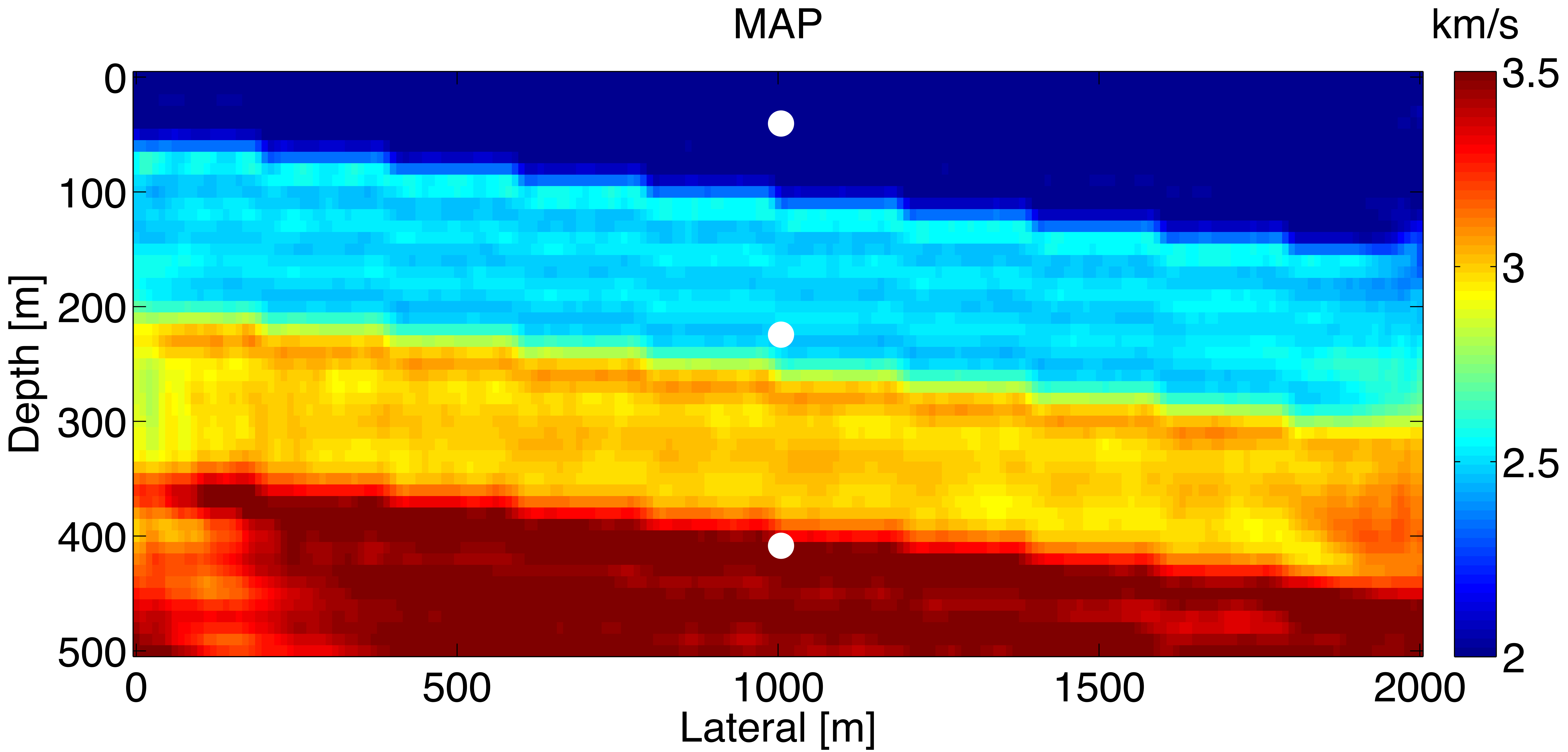


a) Maximum a posteriori estimate



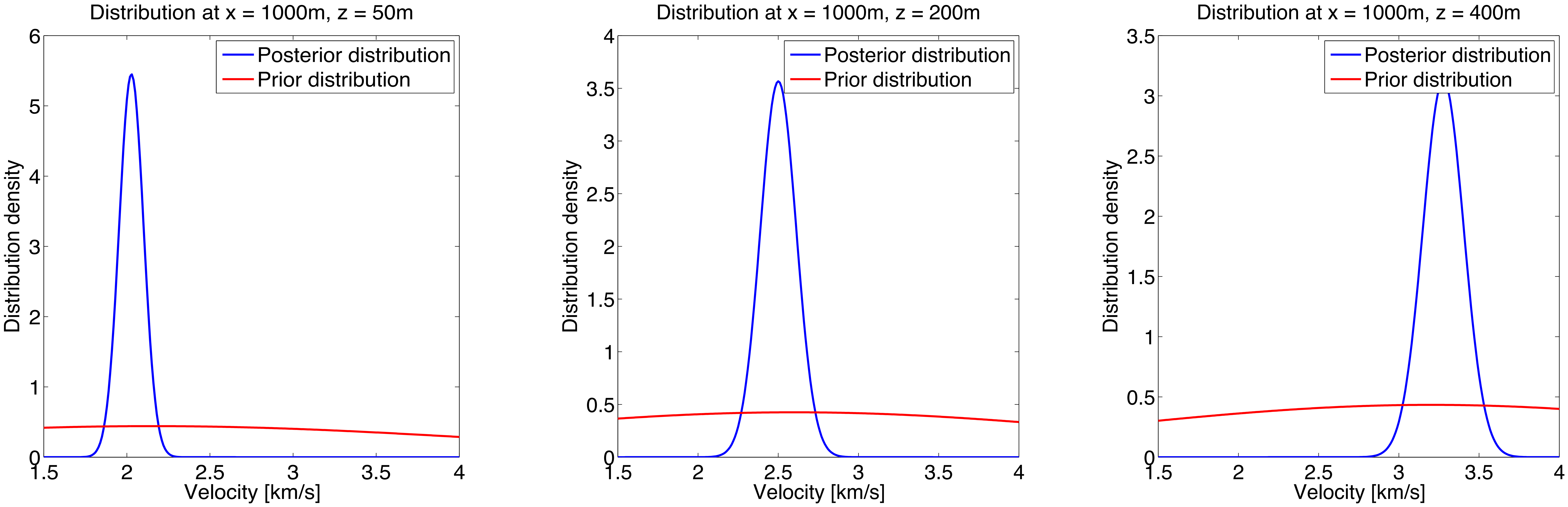
b) The standard deviation

Posterior distribution



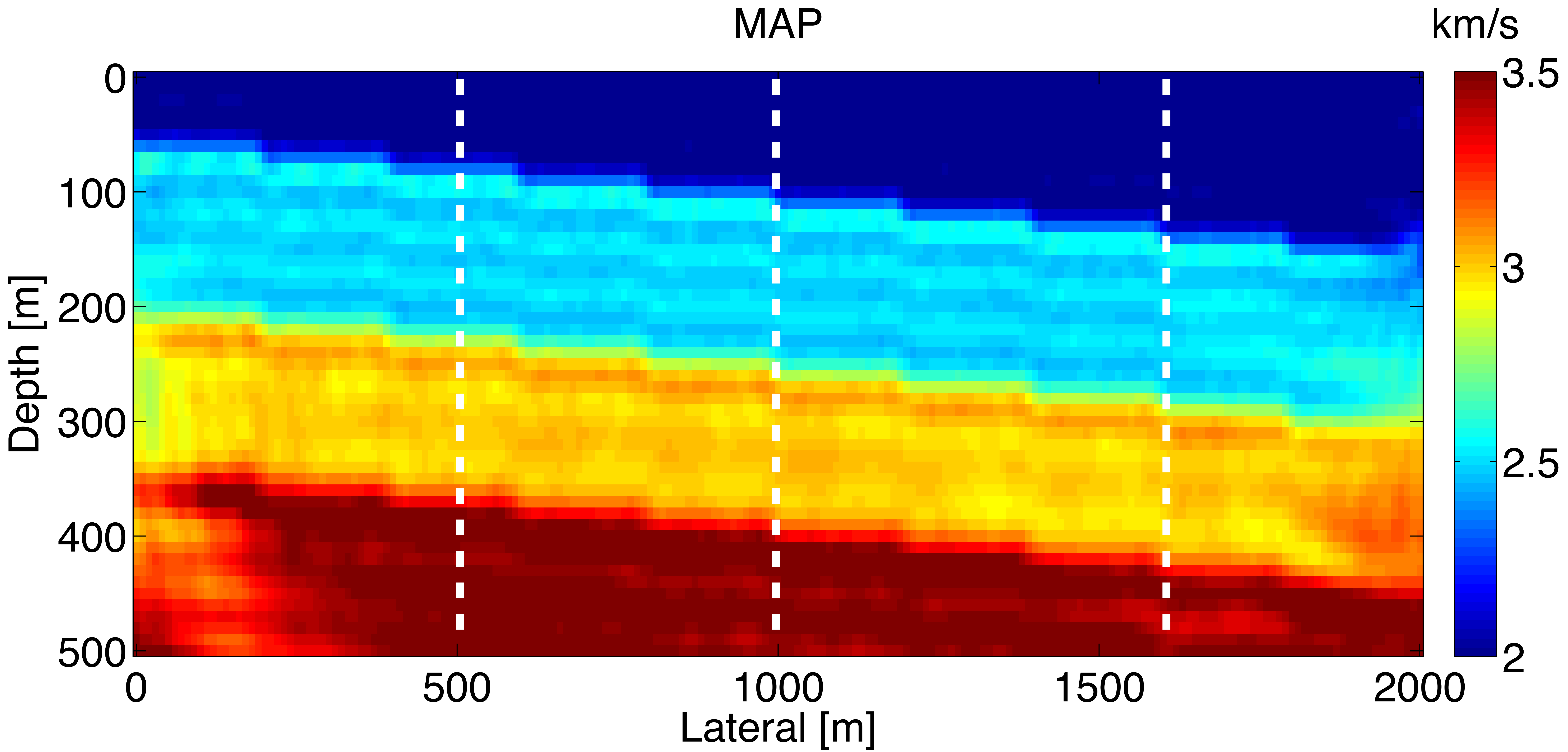
Maximum a posterior

Posterior distribution



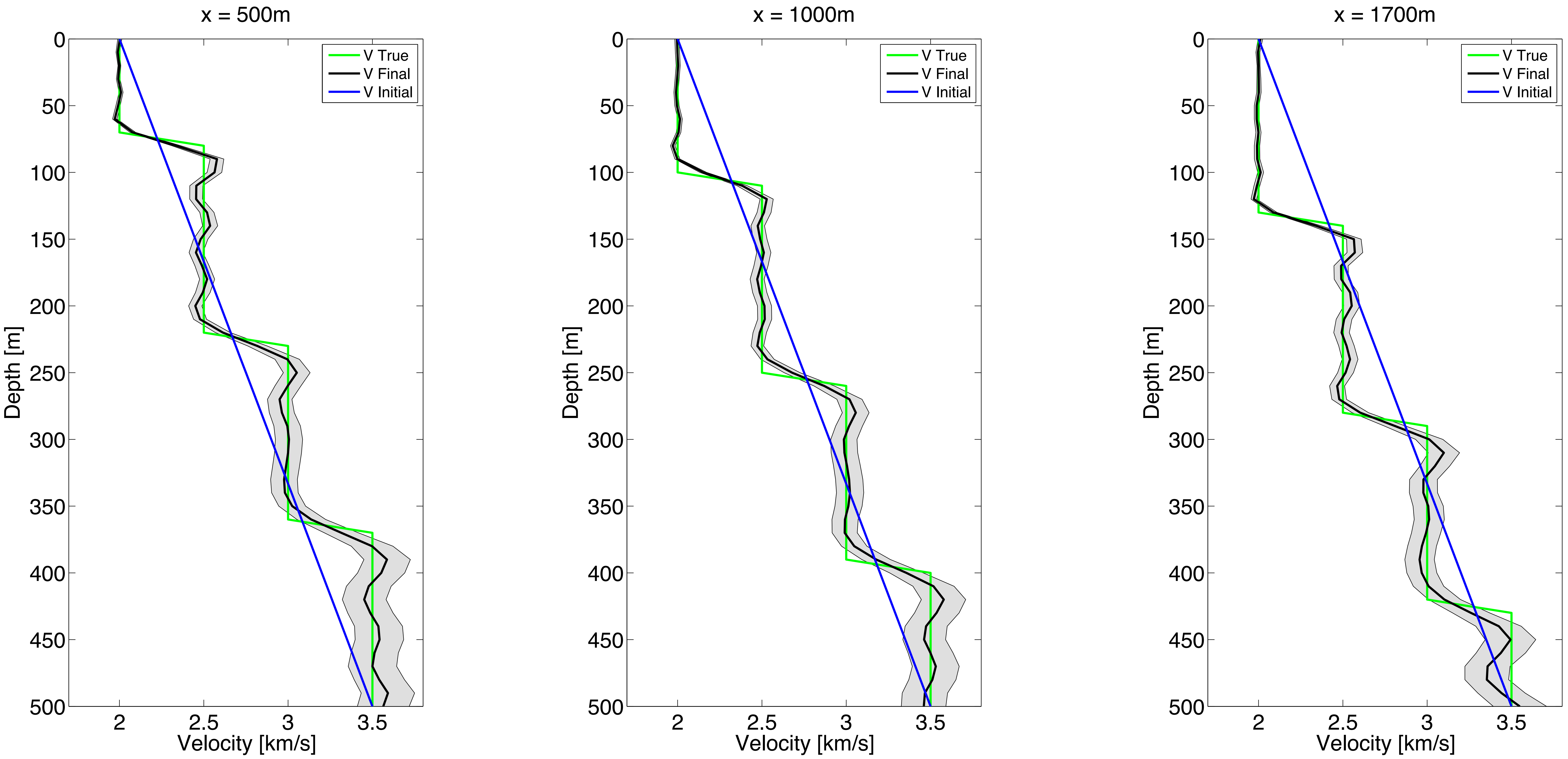
(a) (b) (c)

Confidence intervals



Maximum a posterior

Confidence intervals



(a) (b) (c)

Confidence intervals

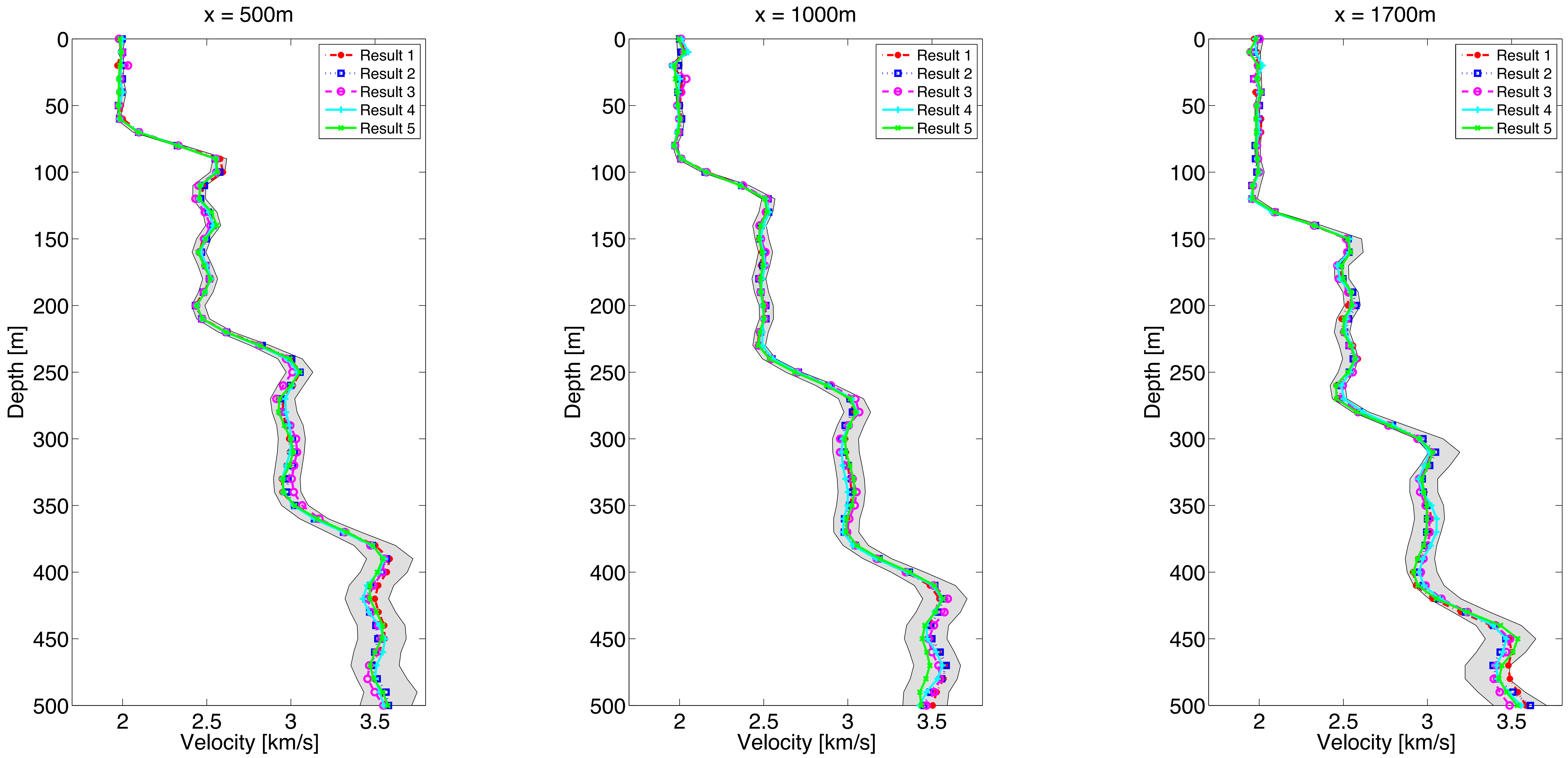
Five random realizations of data:

$$\mathbf{d}_i = \mathbf{F}(\mathbf{m}_{\mathrm{t}}) + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$$

Inversion results corresponds to these data:

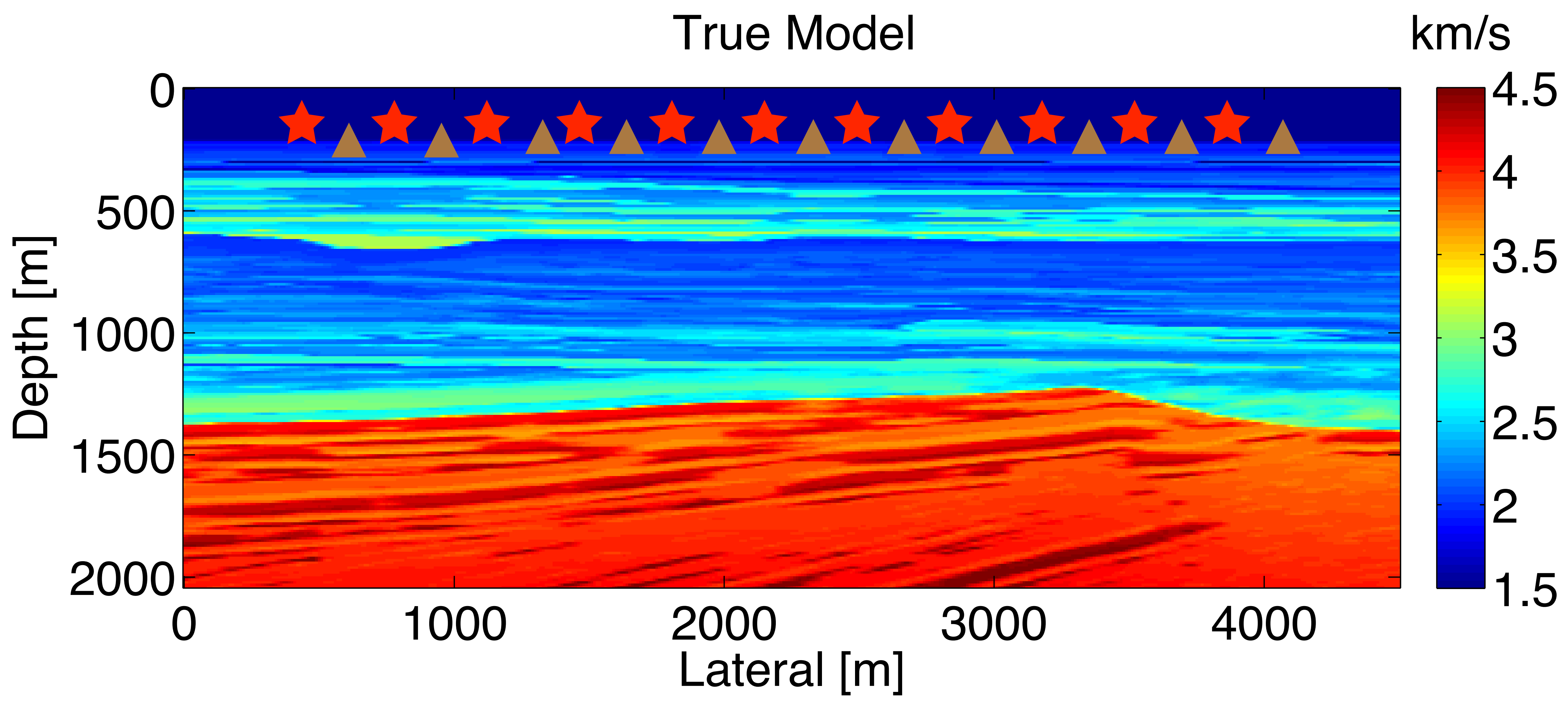
$$\mathbf{d}_i \rightarrow \mathbf{m}_{\mathrm{i}}$$

Confidence intervals

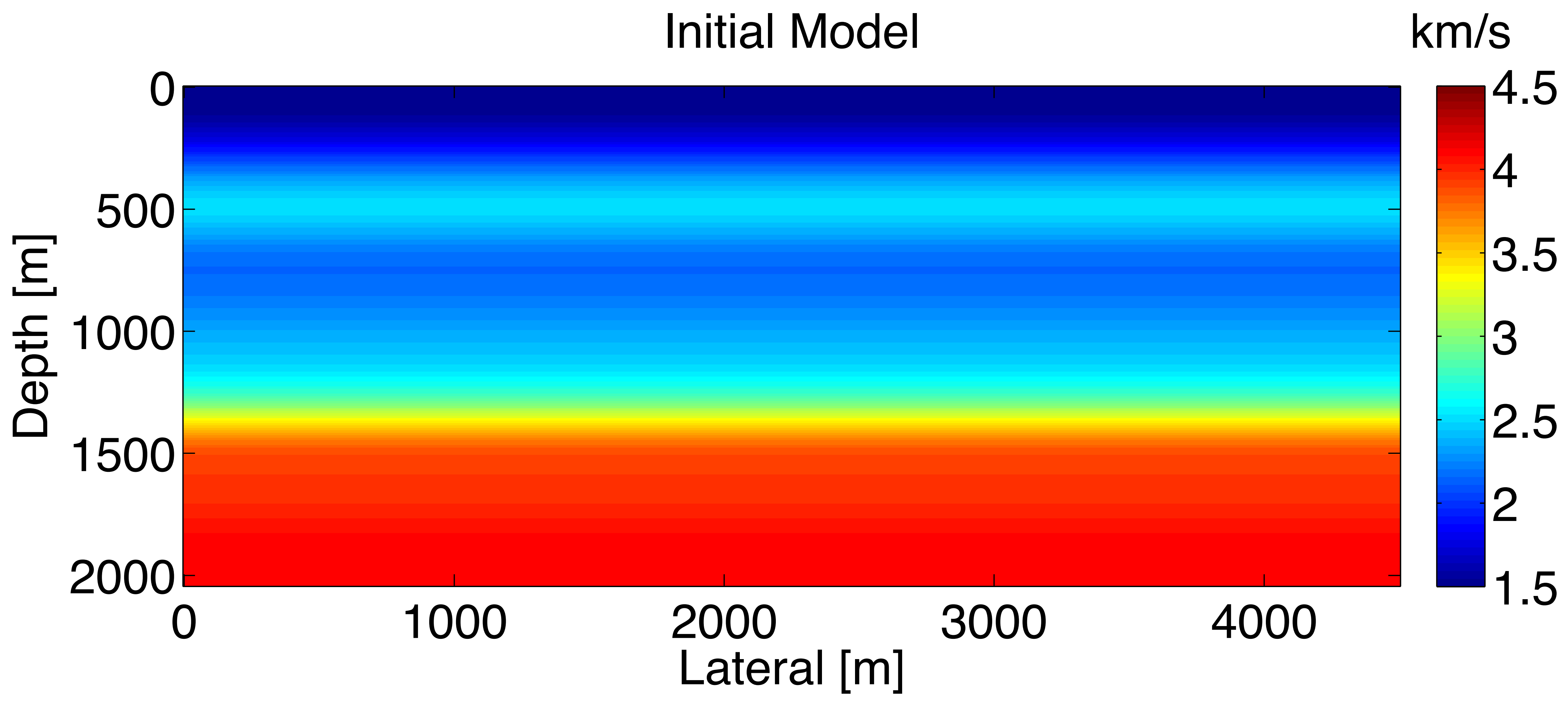


(a) (b) (c)

BG model



(a)

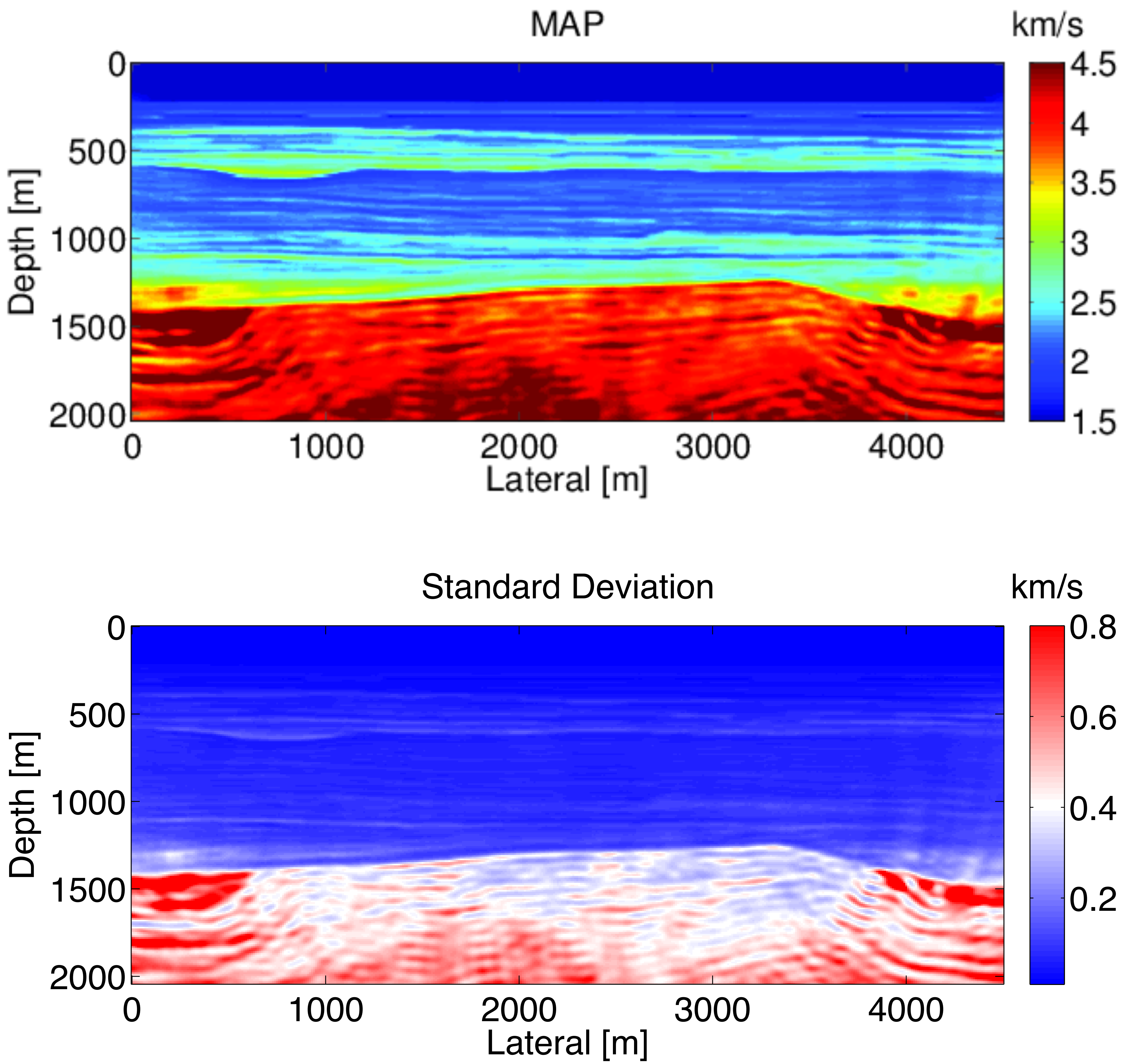


(b)

Model size: 2000m x 4500m
Source spacing: 50m
Receiver spacing: 10m
Fixed spread 4.5km
Frequency : 5~31 Hz

Standard deviation of data noise: 0.5
Standard deviation of pde: 0.5
lambda: 1

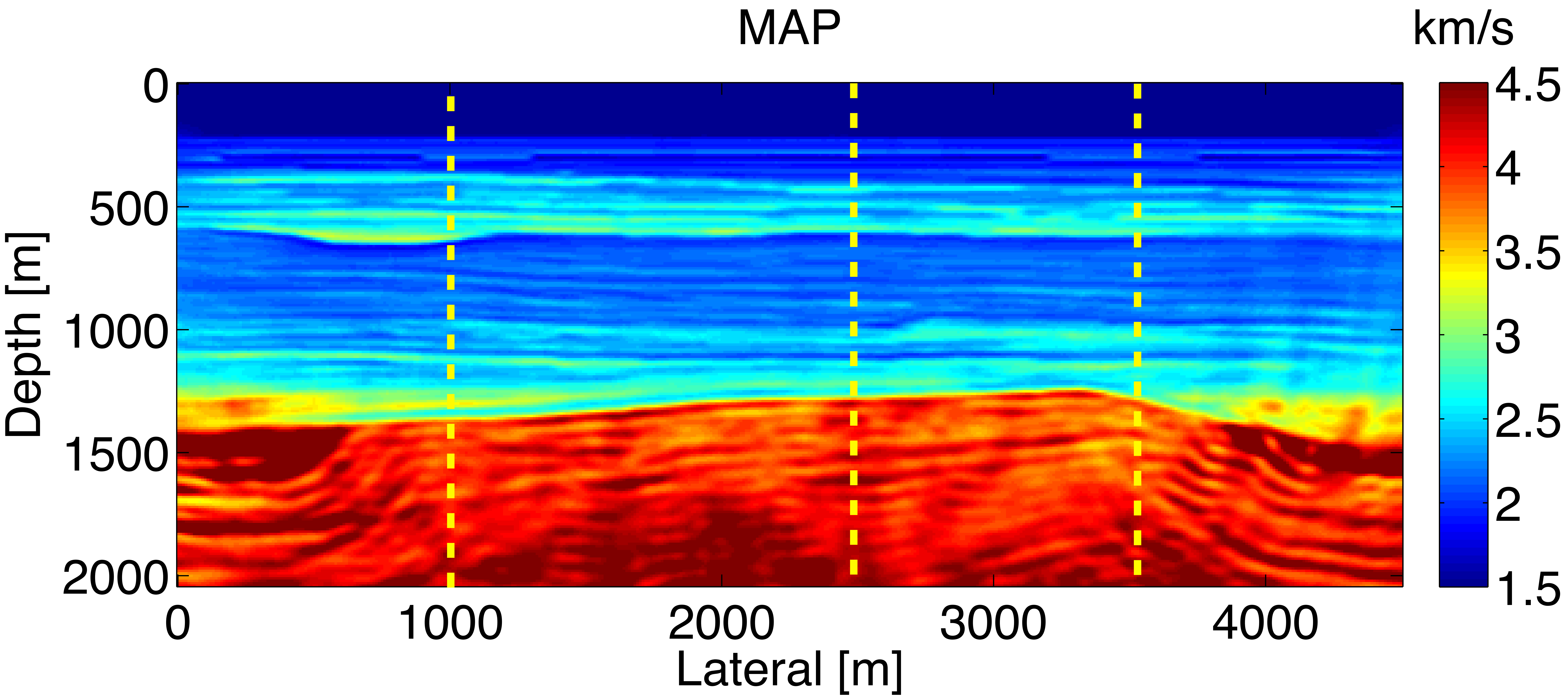
BG model



a) Maximum a posterior estimate

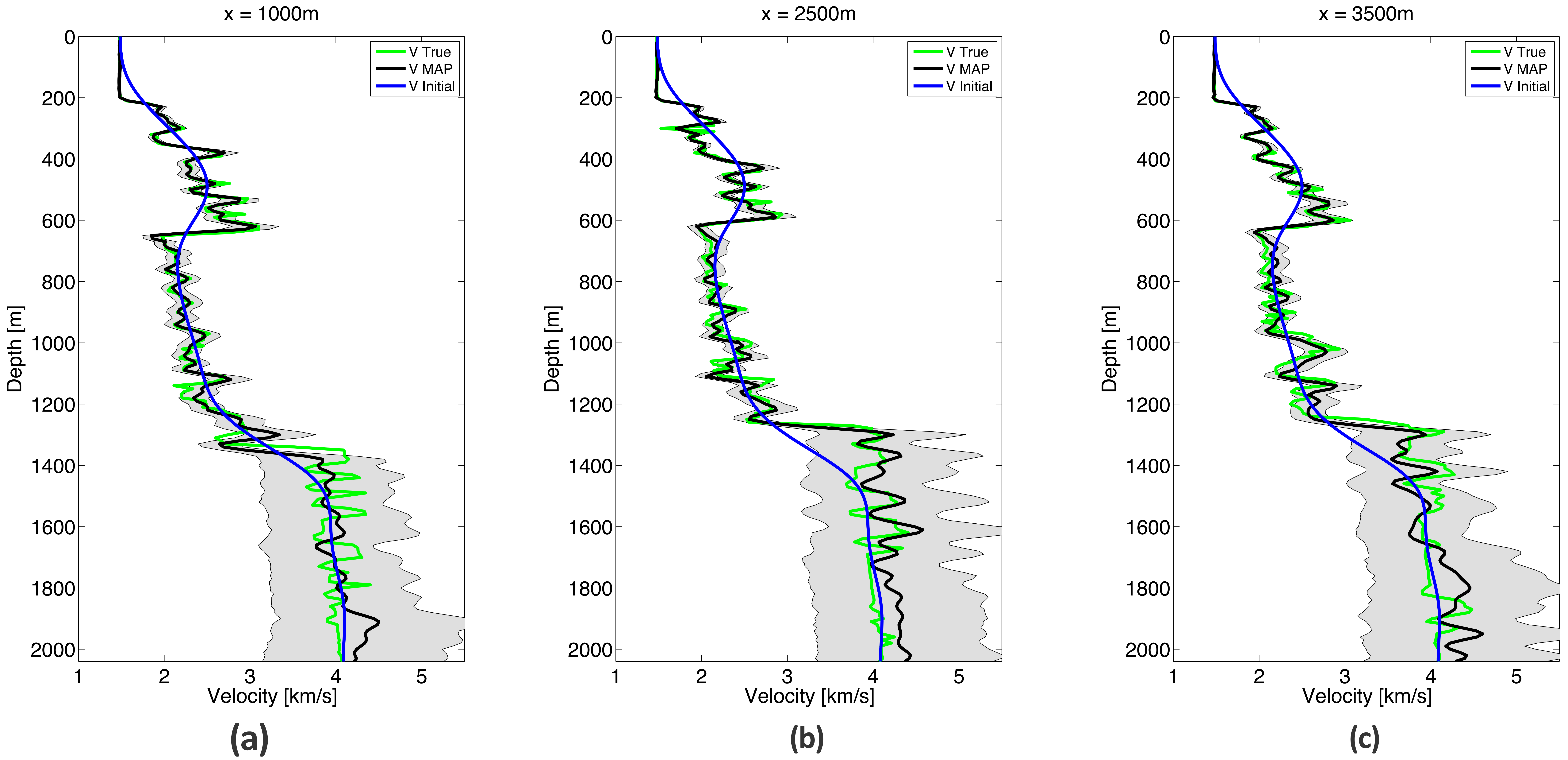
b) The standard deviation

Confidence intervals

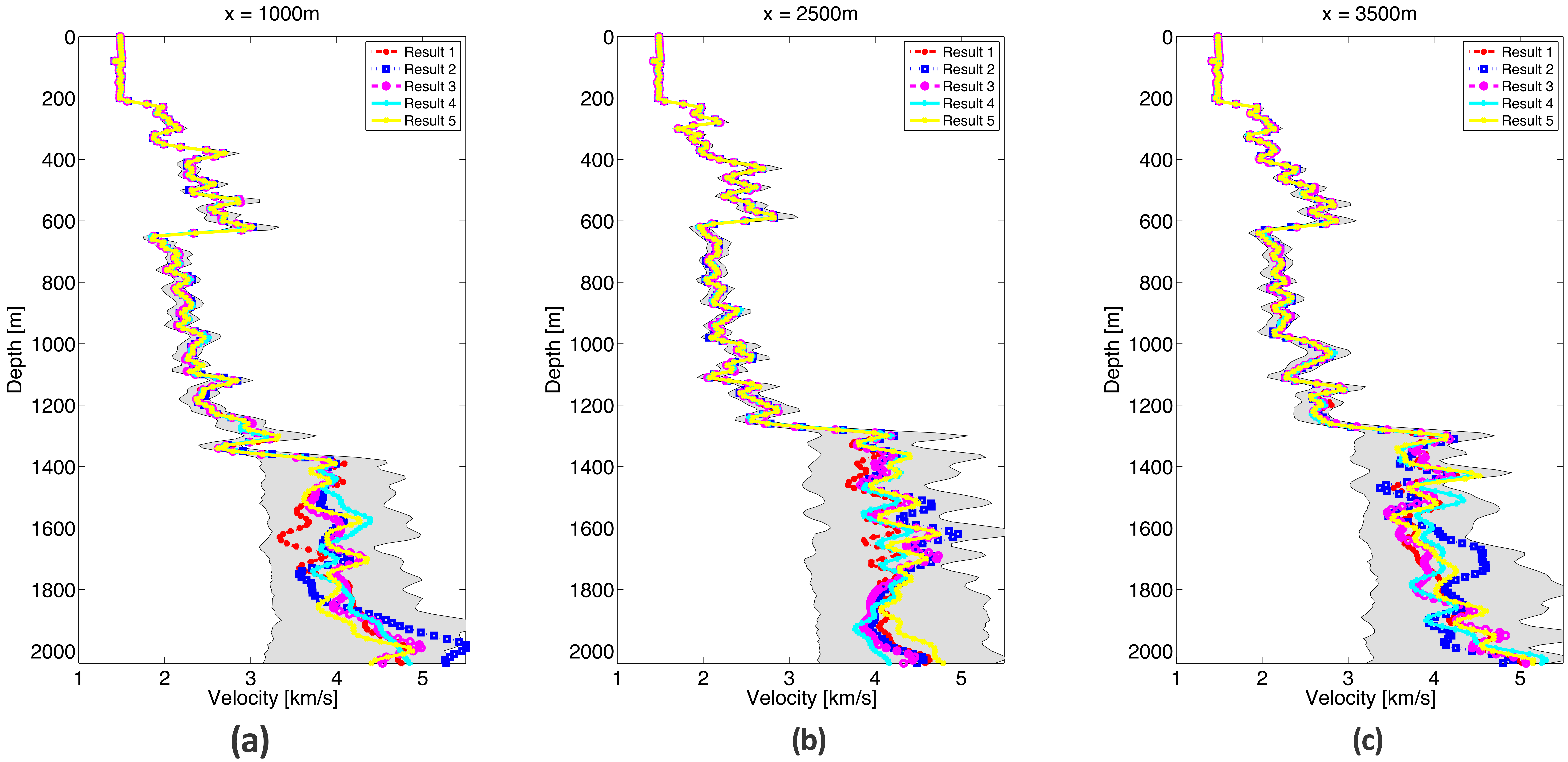


Maximum a posterior

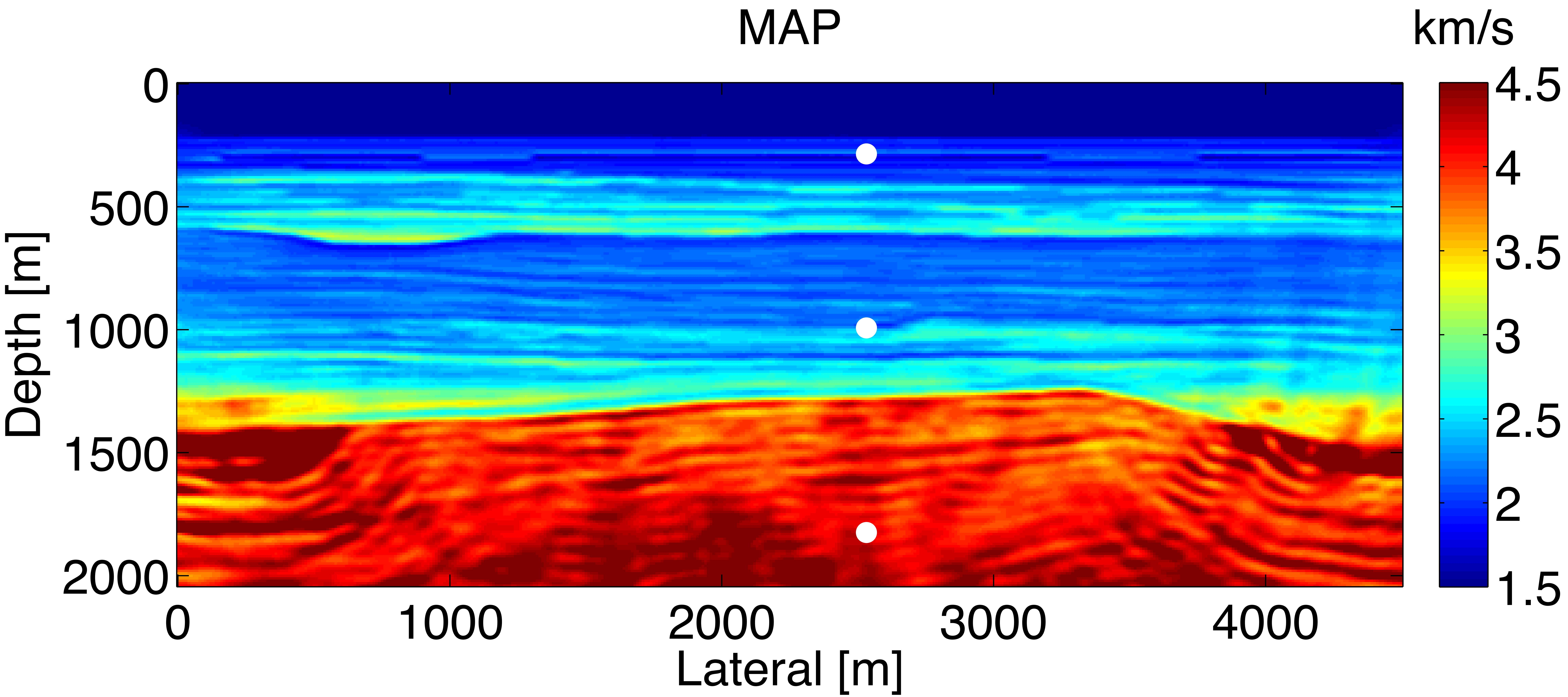
Confidence intervals



Confidence intervals

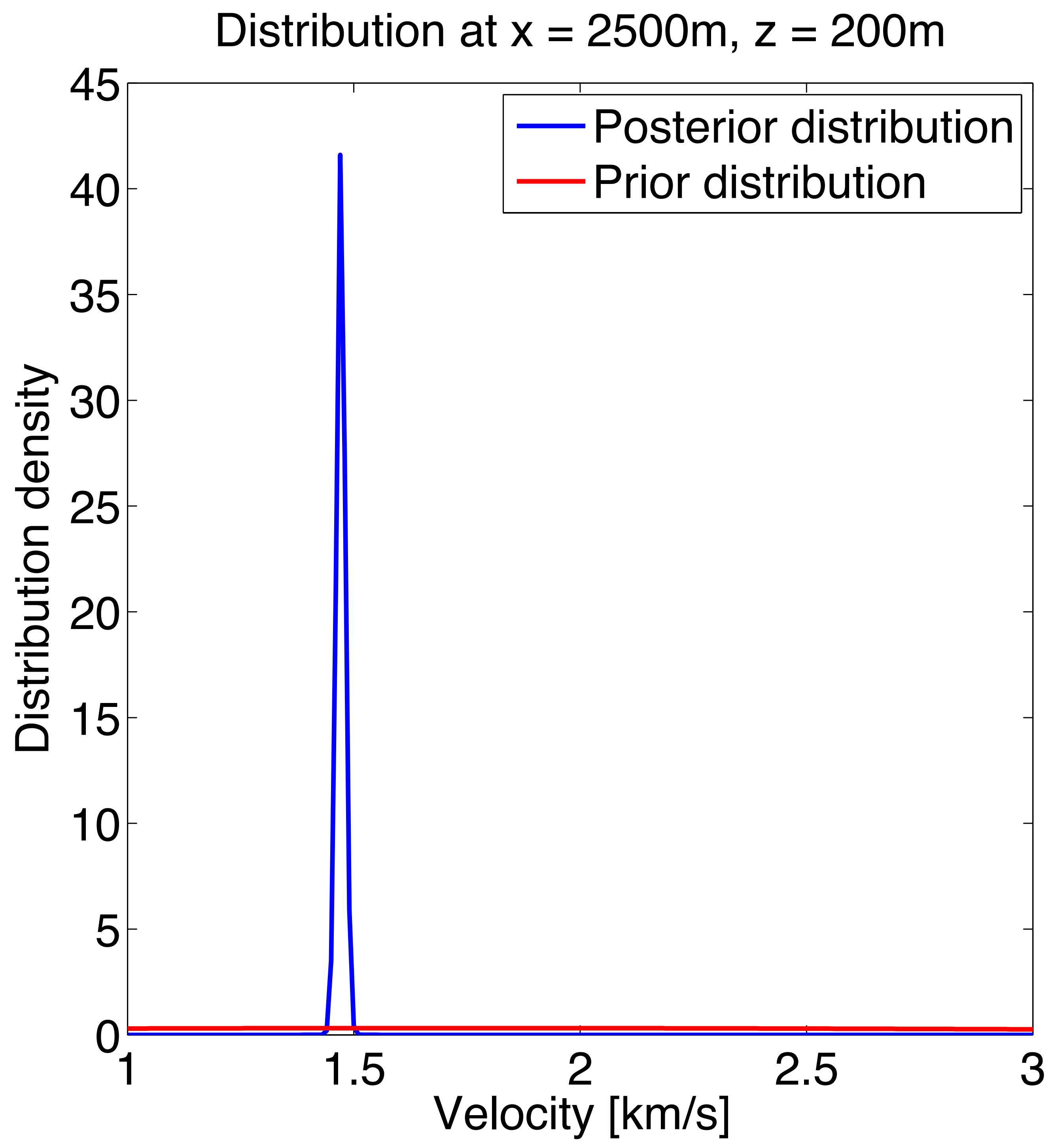


Posterior distribution

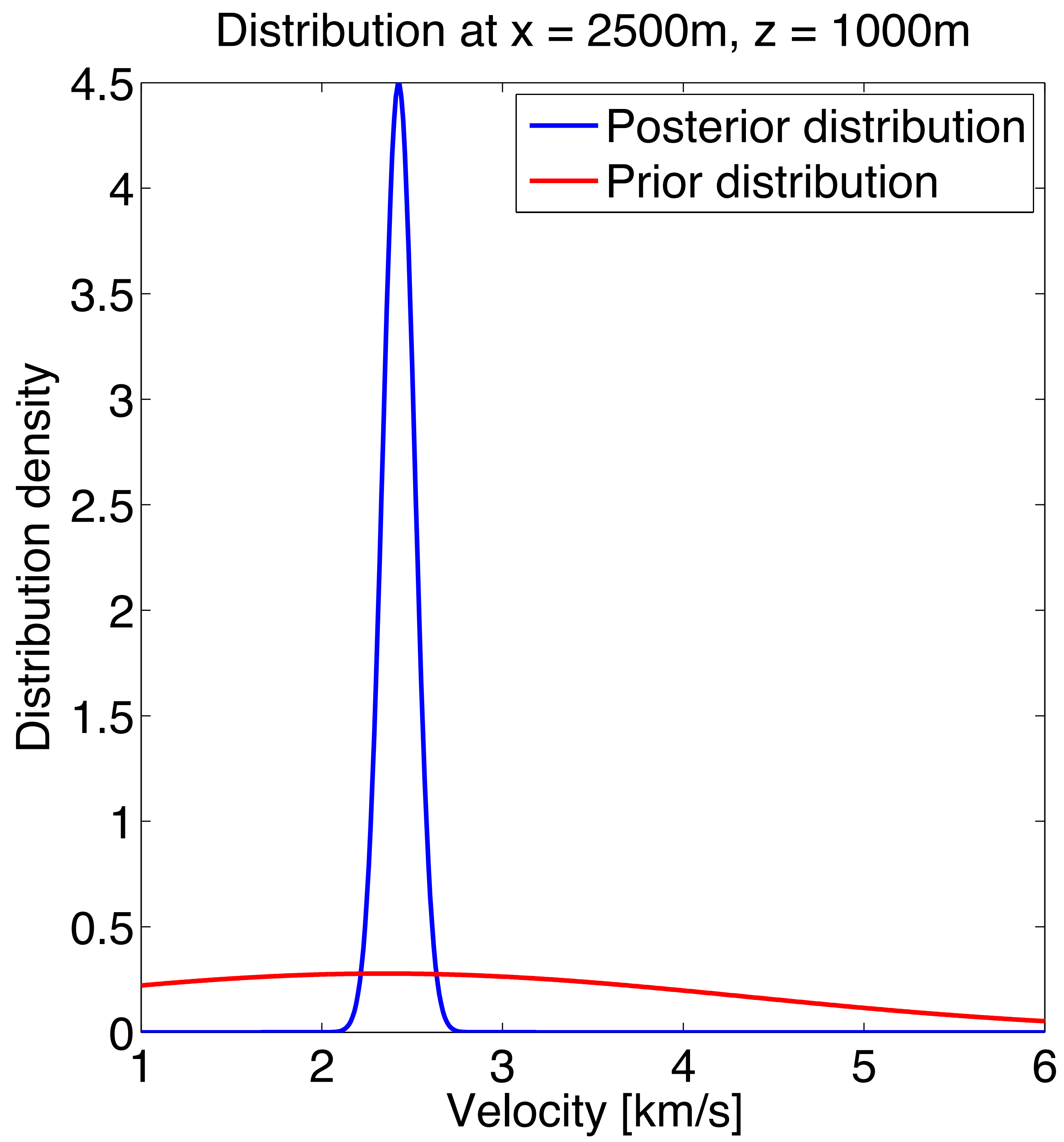


Maximum a posterior

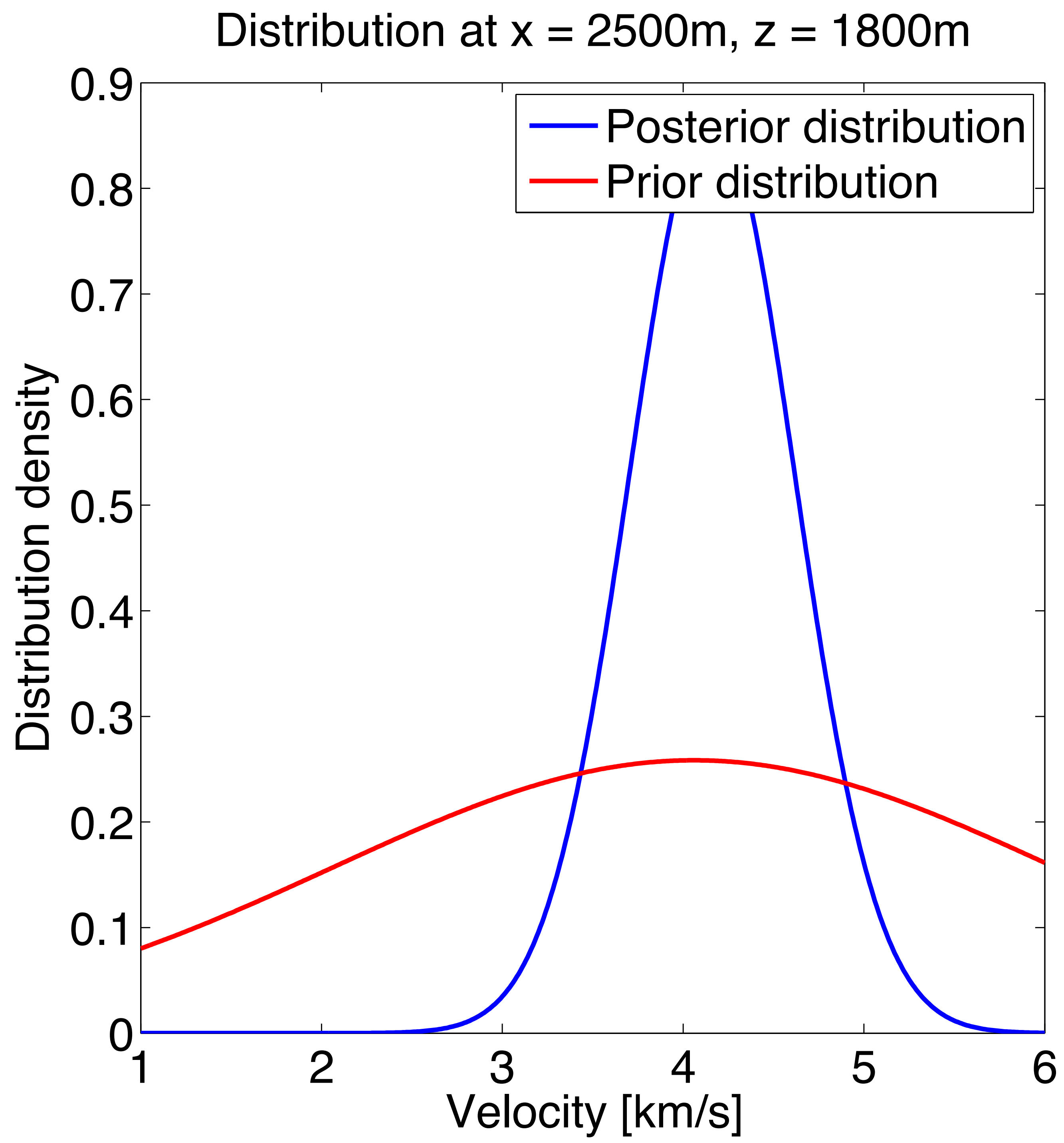
Posterior distribution vs Prior distribution



(a)



(b)



(c)

Summary

- We can approximate the marginal distribution in a computationally feasible manner
    - No extra PDE solves are required

- The approximate diagonal Hessian is a good approximation of the true Hessian to quantify the uncertainty.

- The results of inverting noisy data still lie in the confidence intervals
    - This gives us confidence in our confidence intervals

Acknowledgements

Thank you for your attention !!

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada via the Collaborative Research and Development Grant DNOISEII (375142--08). This research was carried out as part of the SINBAD II project which is supported by the following organizations: BG Group, BGP, CGG, Chevron, ConocoPhillips, DownUnder GeoSolutions, Hess, Petrobras, PGS, Schlumberger, Sub Salt Solutions and Woodside