Parallel 3D FWI with *simultaneous* shots

Zhilong Fang, Tristan van Leeuwen, Felix J. Herrmann
2013.12.02

SLIM
University of British Columbia

Motivation

1. Challenge for 3D FWI:
   1. Data is extremely large (5D data);
   2. Model is large (3D model);
   3. Solving the wave-equation is expensive.
2. Solutions:
   1. Stochastic optimization;
   2. Simultaneous shots;
   3. Randomly selected shots;
   4. Composite shots.

3D full-waveform inversion

Optimization problem:

$$\min_{\mathbf{m}} \varphi(\mathbf{m}) = \sum_{i=1}^{M} f_i(\mathbf{m}) = \sum_{i=1}^{M} \|F_i(\mathbf{m}) - \mathbf{d}_i\|_2^2 = \sum_{i=1}^{M} \|P_i A_i^{-1} \mathbf{q}_i - \mathbf{d}_i\|_2^2$$

where

$$A_i := \Delta + \mathbf{m}\omega_i^2$$

$\mathbf{m}$ - model (squared slowness);
$\mathbf{d}_i$ - i^{th} observed data;
F_i - i^{th} forward operator;
P_i - i^{th} projection operator;
$\mathbf{q}_i$ - i^{th} source.

3D full-waveform inversion

Optimization method:

- Gradient descent method

  search direction: $\mathbf{s} = -\mathbf{g}$.

- Gauss-Newton method

  search direction: $\mathbf{s} = -[\mathbf{J}^*\mathbf{J}]^{-1}\mathbf{g}$.

- Quasi Newton method (l-bfgs)

  search direction: $\mathbf{s} = -\tilde{\mathbf{H}}^{-1}\mathbf{g}$, $\tilde{\mathbf{H}}^{-1}$ is the low-rank approximation of $\mathbf{H}^{-1}$.

Stochastic optimization

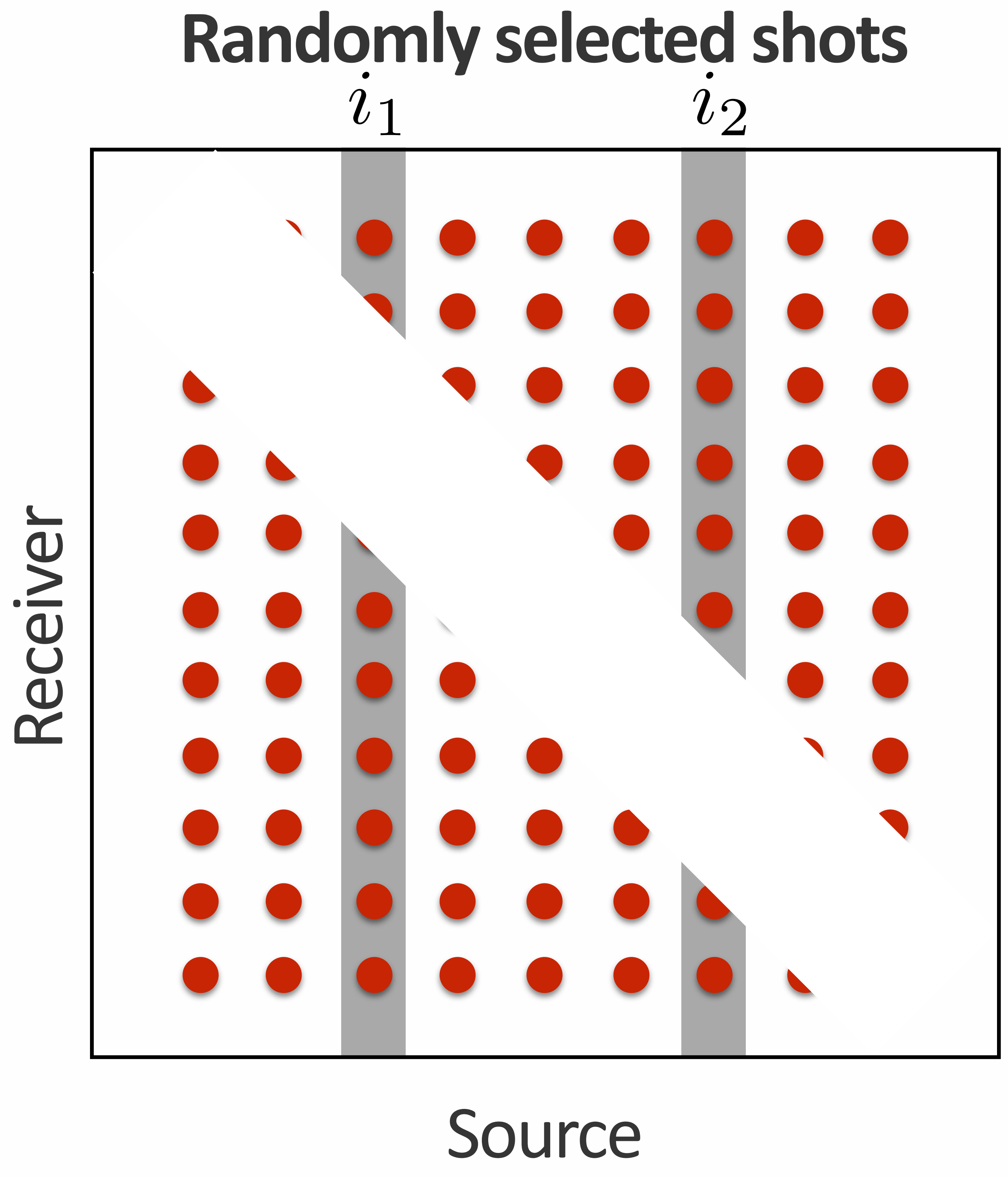
	Deterministic optimization	Stochastic optimization	
Object function	$\min_{\mathbf{m}} \varphi(\mathbf{m}) = \frac{1}{M}\sum_{i=1}^{M} f_i(\mathbf{m})$	$\min_{\mathbf{m}} \overline{\varphi}(\mathbf{m}) = \frac{1}{\lVert\mathcal{I}\rVert}\sum_{i\in\mathcal{I}} f_i(\mathbf{m})$	$\min_{\mathbf{m}} \overline{\varphi}(\mathbf{m}) = \frac{1}{M_s}\sum_{i=1}^{M_s} \tilde{f}_i(\mathbf{m})$
Gradient	$g(\mathbf{m}) = \frac{1}{M}\sum_{i=1}^{M} g_i(\mathbf{m})$	$\overline{g}(\mathbf{m}) = \frac{1}{\lVert\mathcal{I}\rVert}\sum_{i\in\mathcal{I}} g_i(\mathbf{m})$	$\overline{g}(\mathbf{m}) = \frac{1}{M_s}\sum_{i=1}^{M_s} \tilde{g}_i(\mathbf{m})$
Size of batch	M	$\#$ of $\mathcal{I}$	M_s

Michael P. Friedlander, 2012

Tristan van Leeuwen, 2013

Randomly selected shots

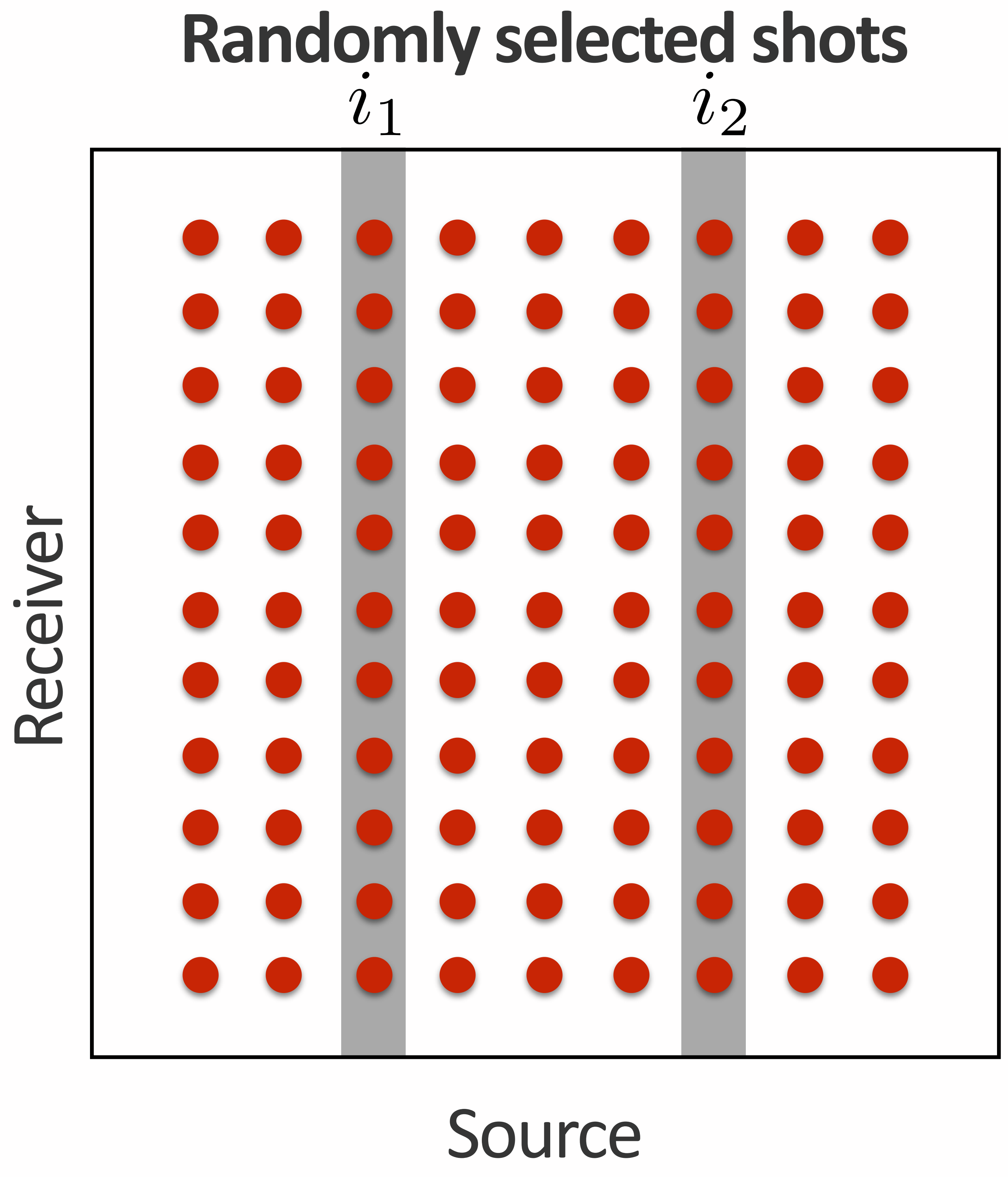
Ocean bottom acquisition:

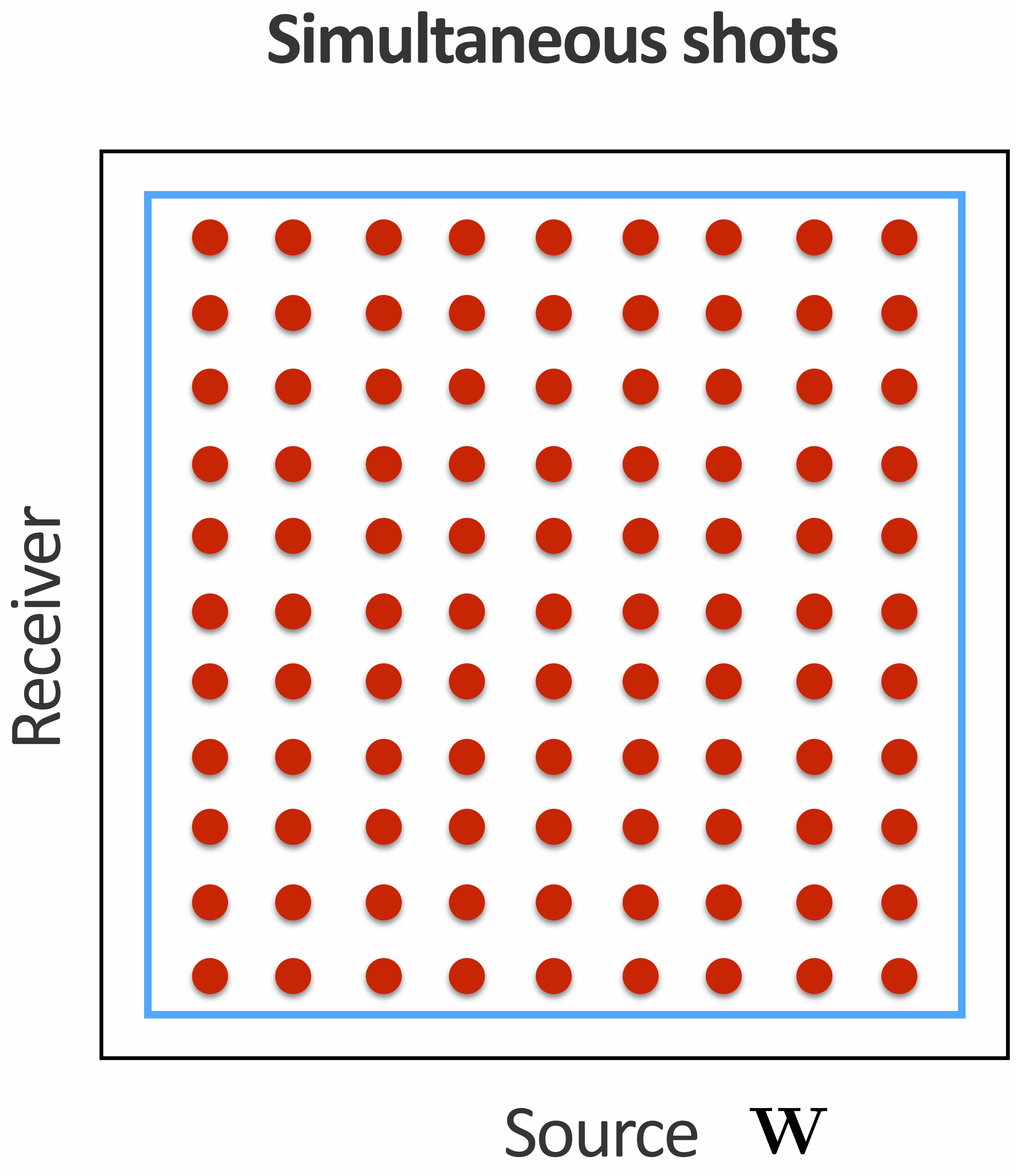


$$\min_{\mathbf{m}} \overline{\varphi}(\mathbf{m}) = \sum_{i \in \mathcal{I}} \|F_i(\mathbf{m}) - \mathbf{d}_i\|_2^2 = \sum_{i \in \mathcal{I}} \|P_i A_i^{-1} \mathbf{q}_i - \mathbf{d}_i\|_2^2$$

Simultaneous shots

Ocean bottom acquisition:





$$\min_{\mathbf{m}} \overline{\varphi}(\mathbf{m}) = \sum_{j=1}^{n_f} \|PA_j^{-1}(\mathbf{m})\mathbf{Q}_j\mathbf{W}_j - \mathbf{D}_j\mathbf{W}_j\|_F^2$$

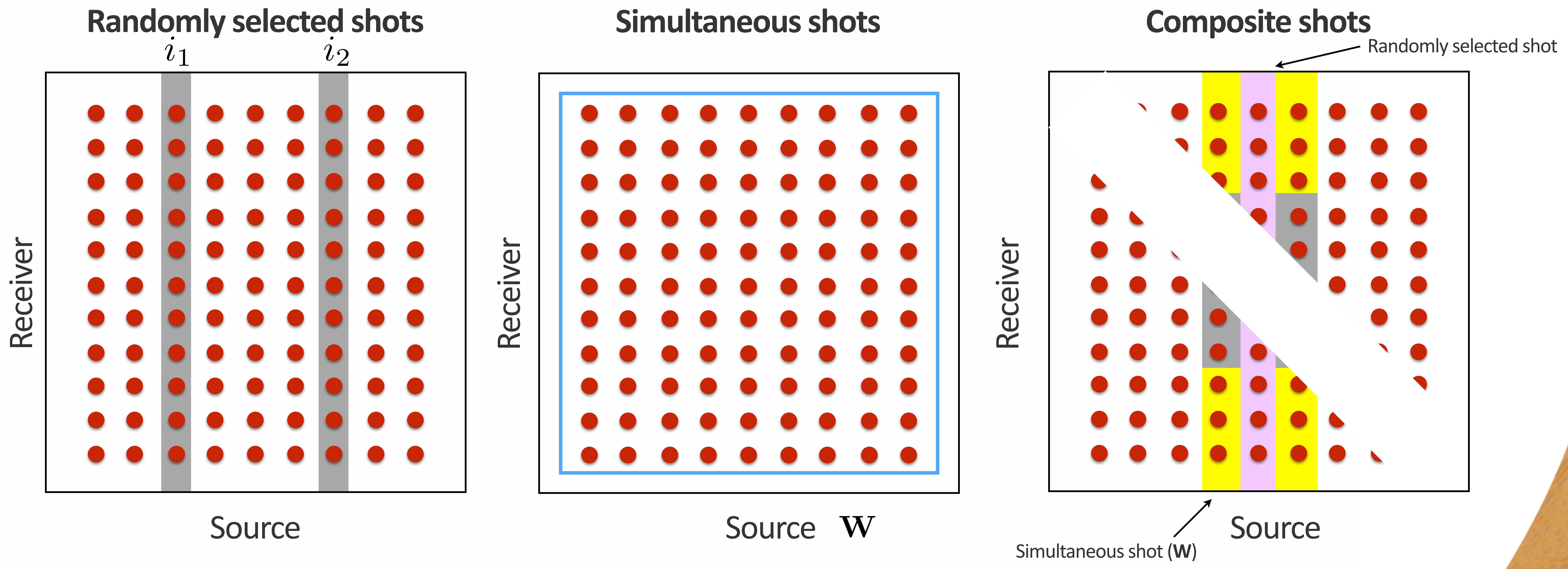
n_f - number of frequencies;
$\|\cdot\|_F$ - Frobenius norm.
$\mathbf{Q}_j = [\mathbf{q}_{j1}, ..., \mathbf{q}_{jn_j}]$;
$\mathbf{D}_j = [\mathbf{d}_{j1}, ..., \mathbf{d}_{jn_j}]$;

Randomly selected shots v.s. *simultaneous* shots

	Advantage	Disadvantage
Randomly selected shots	Can use offset *weighting*	Only use information from *subset* of shots
Simultaneous shots	*All* shots are used	All shots must share *same* receivers

Composite shots

Ocean bottom acquisition:



Composite shots

Optimization problems with composite shots:

$$\min_{\mathbf{m}} \overline{\varphi}(\mathbf{m}) = \sum_{j=1}^{n_f} \sum_{k=1}^{n_{cs}} \|P_k A_j^{-1}(\mathbf{m})\tilde{\mathbf{q}}_k - \mathbf{d}_k\|_2^2$$

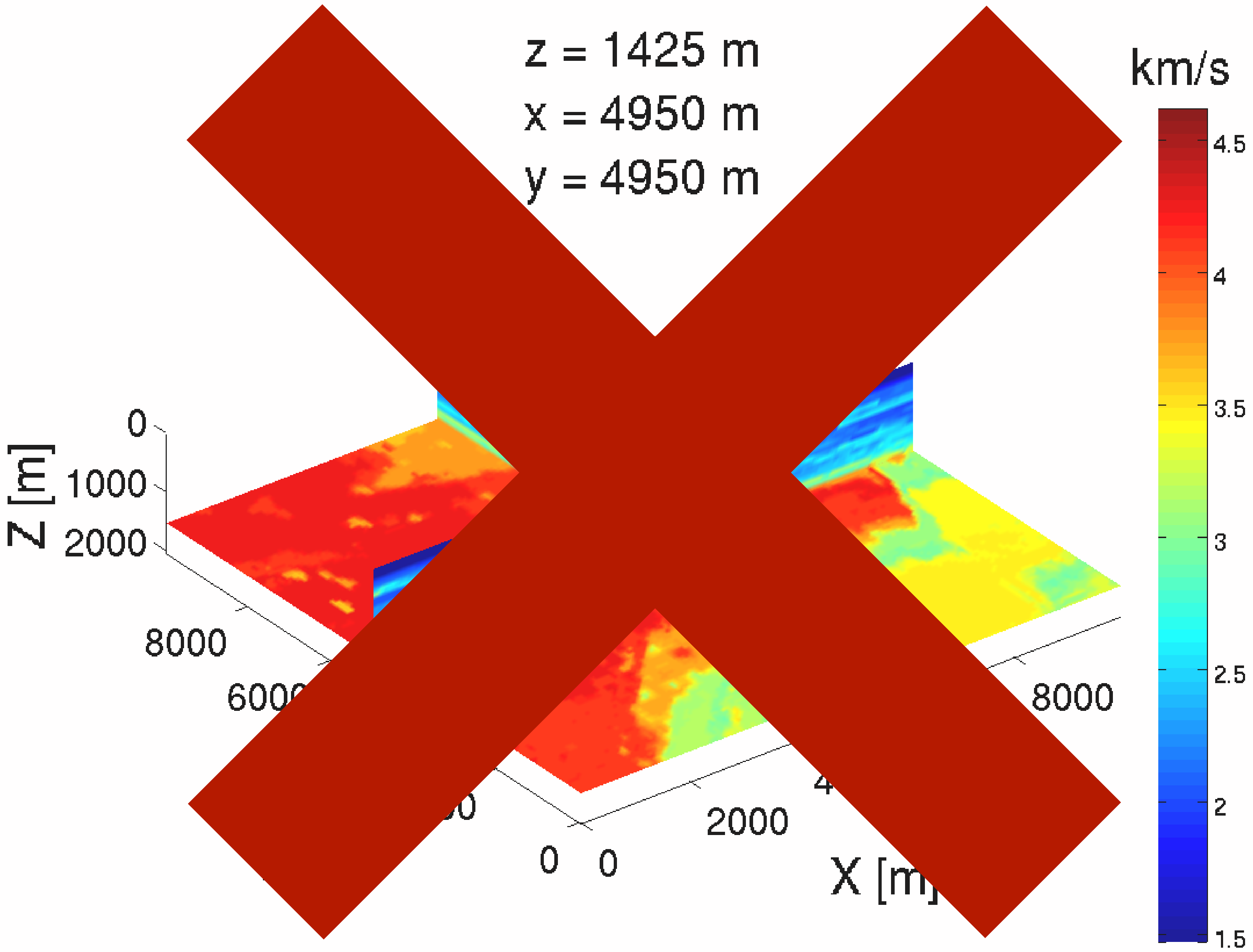
Advantage:

1. Do *not* need *all* shots to share the *same* receivers;

2. Uses *more* information from shots than *randomly* selected shots.

Numerical experiment

BG Model



True Model

Model Infomation:

model size - 28*134*134

2 * 10 *10 km

d - [75 75 75] m

# of shots - 441

# of receivers - 10201

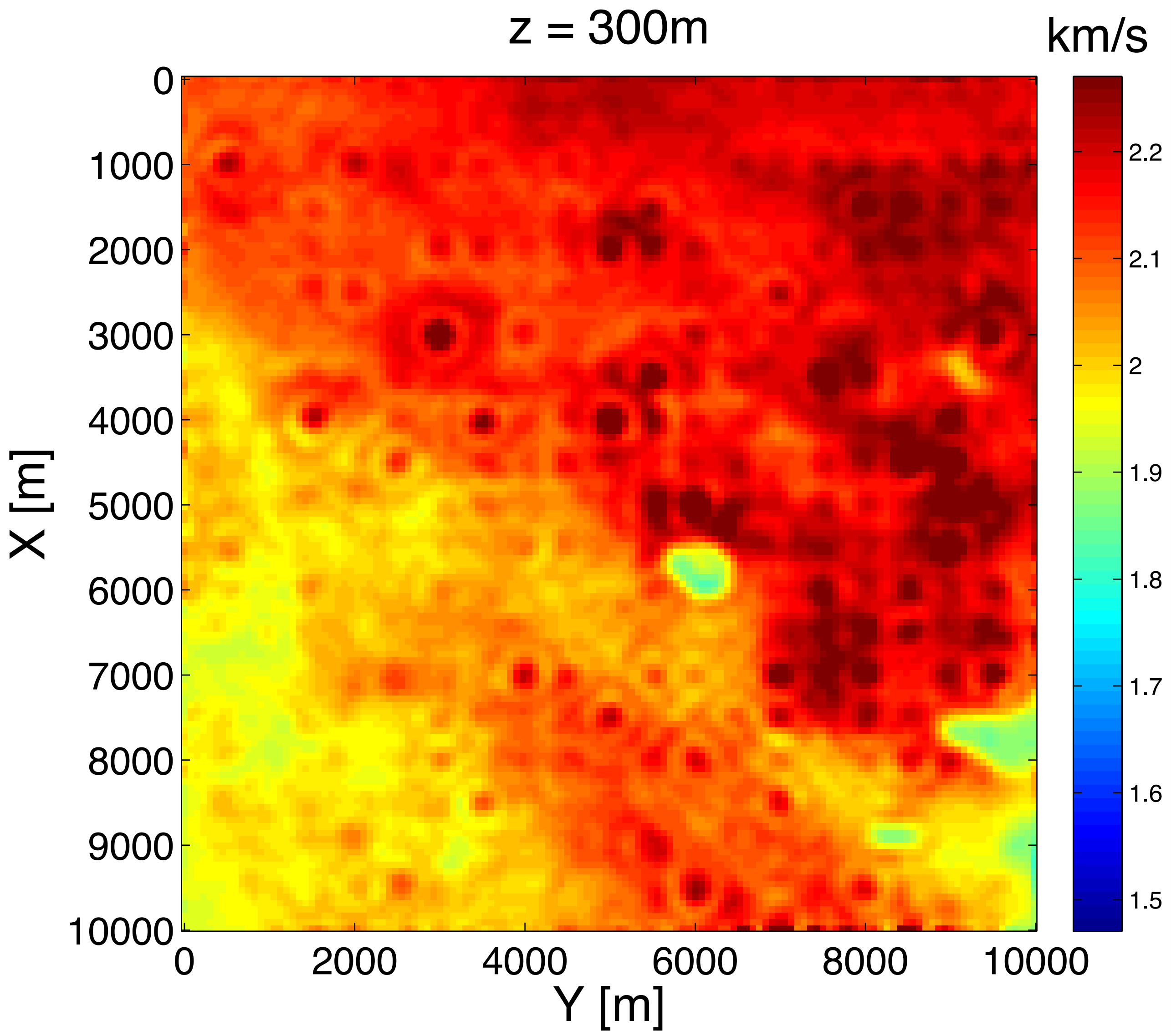
frequency - 3.2 Hz

source depth - 14 m

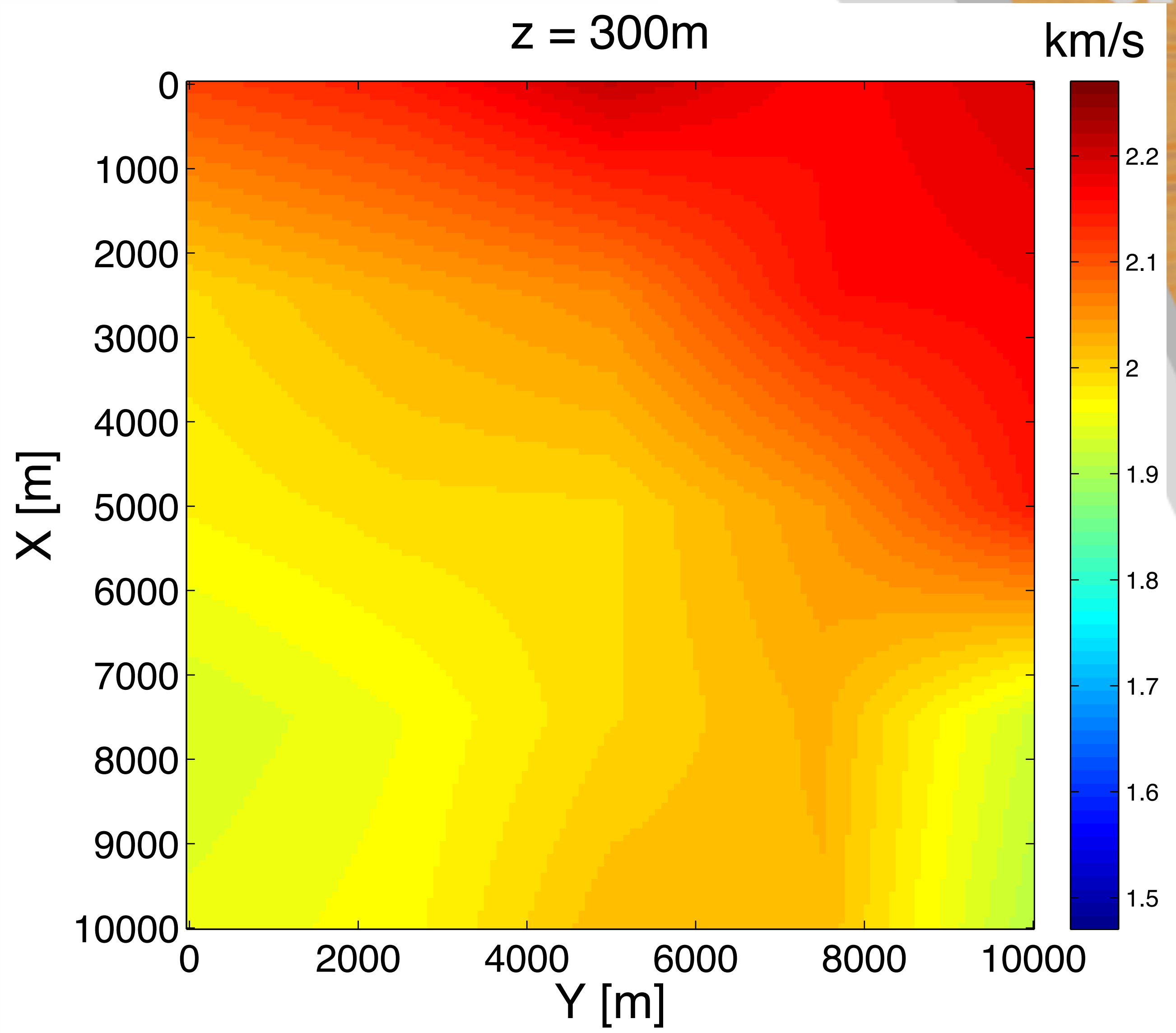
receiver depth - 14 m

l-bfgs - 34 iterations

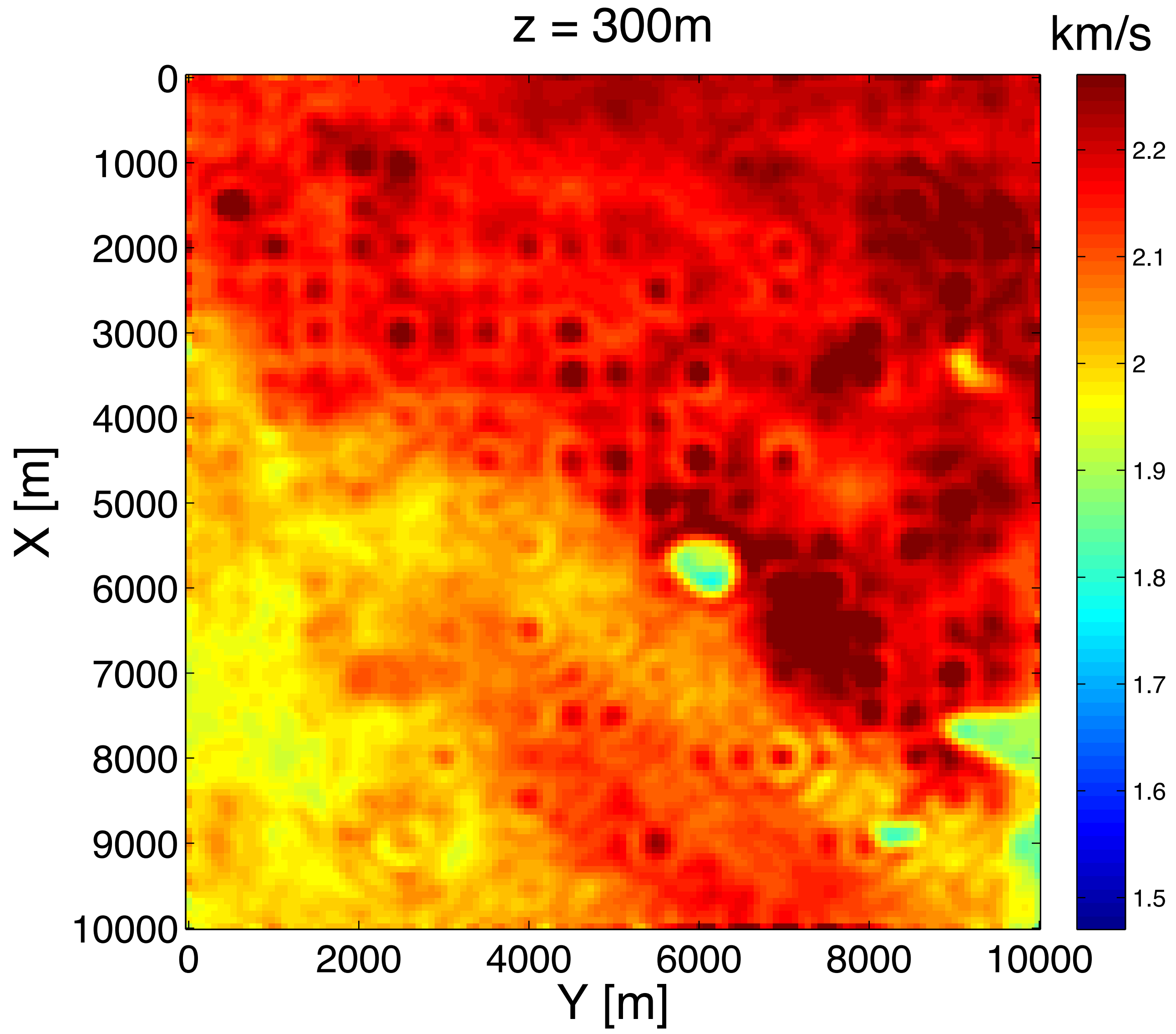
z = 300m



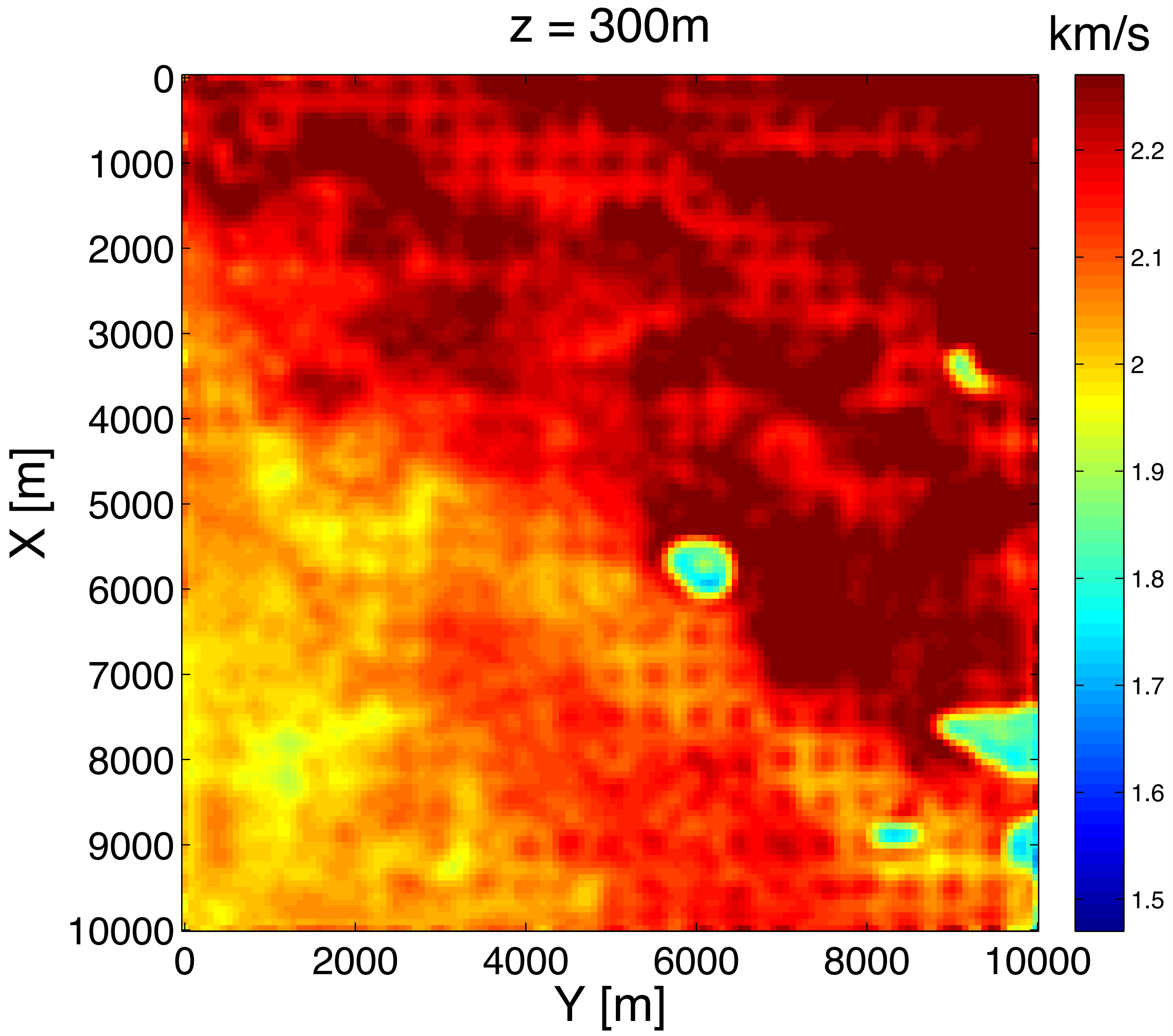
Composite shots



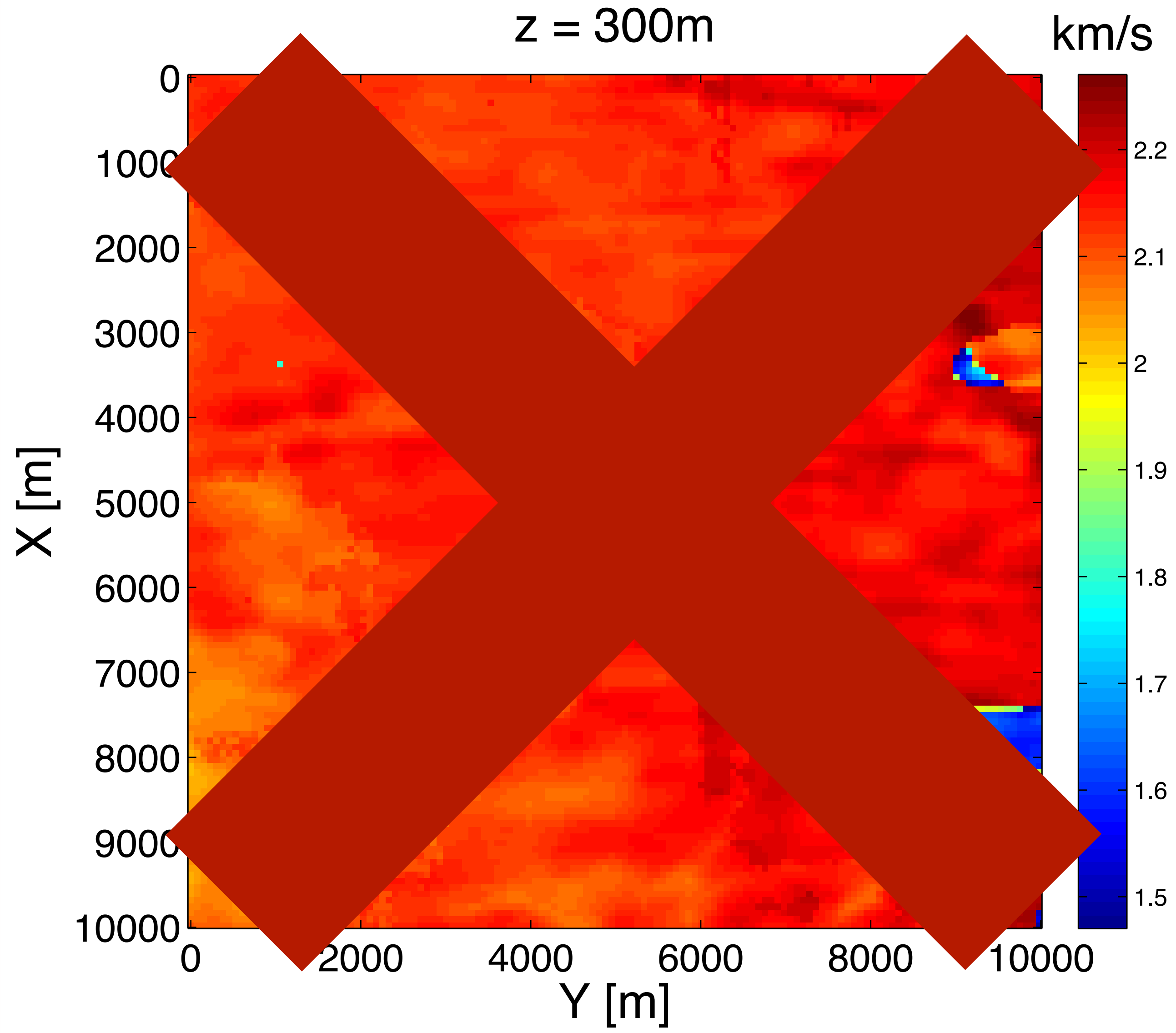
Initial model



Randomly selected shots

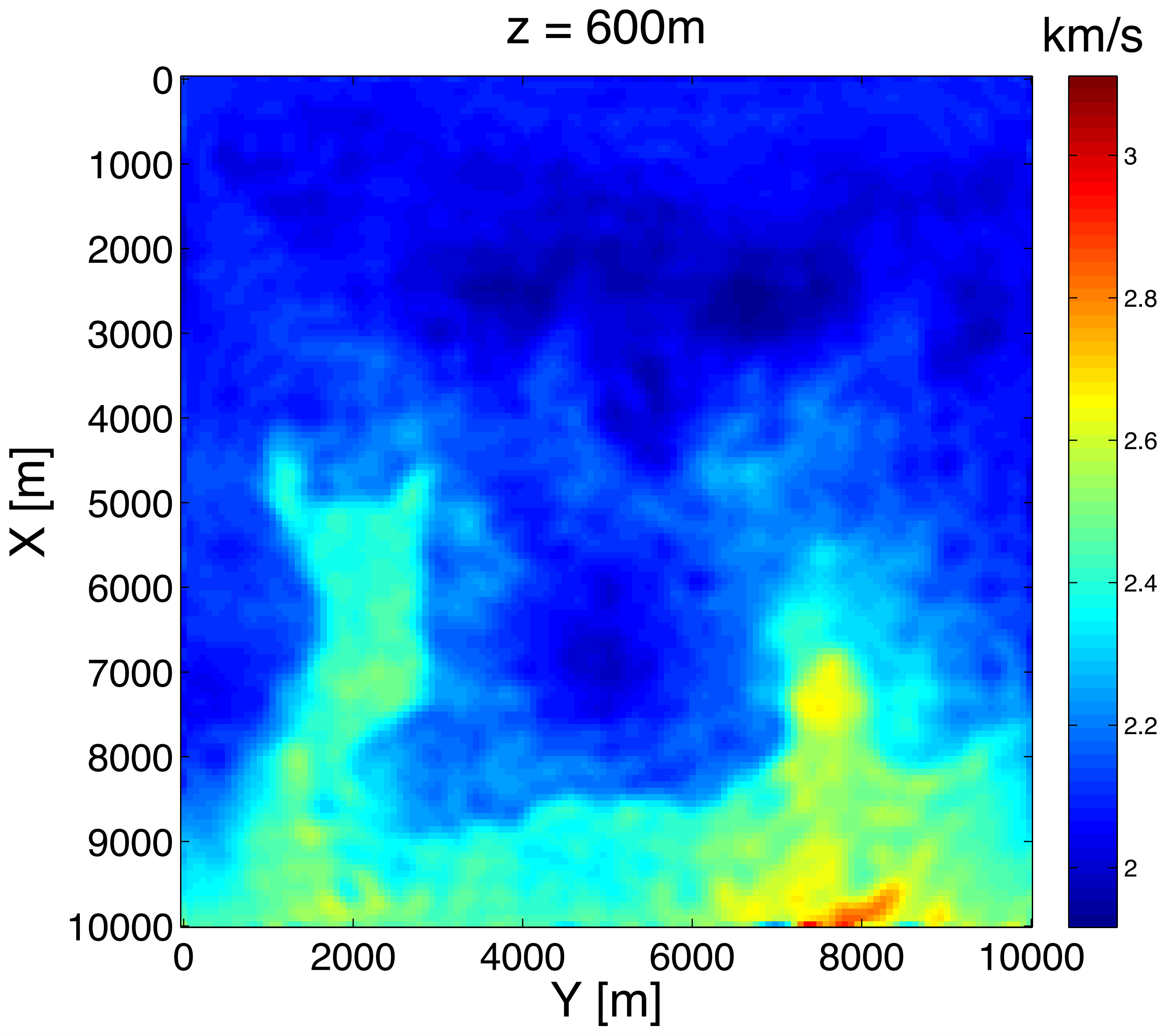


Simultaneous shots

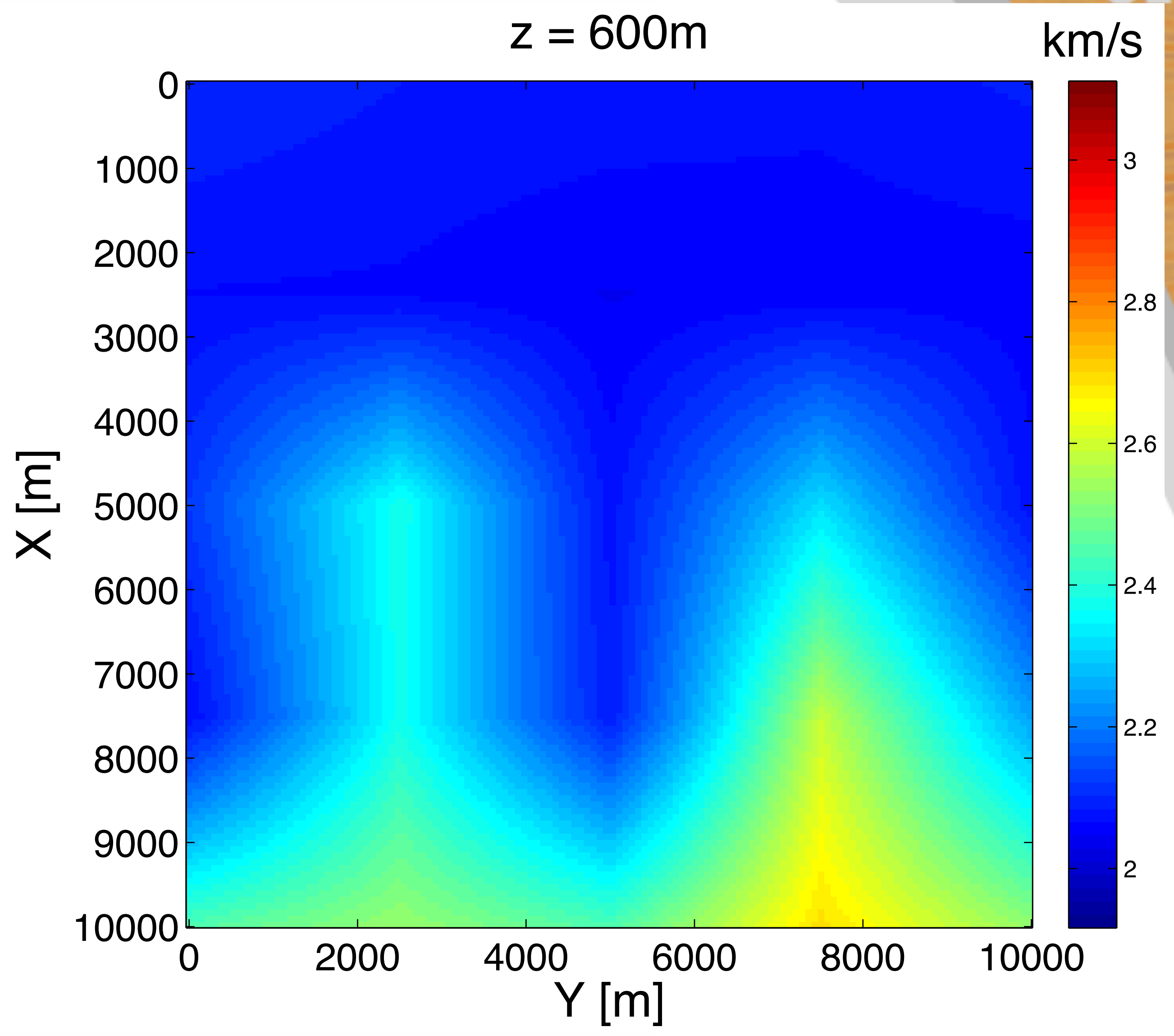


True model

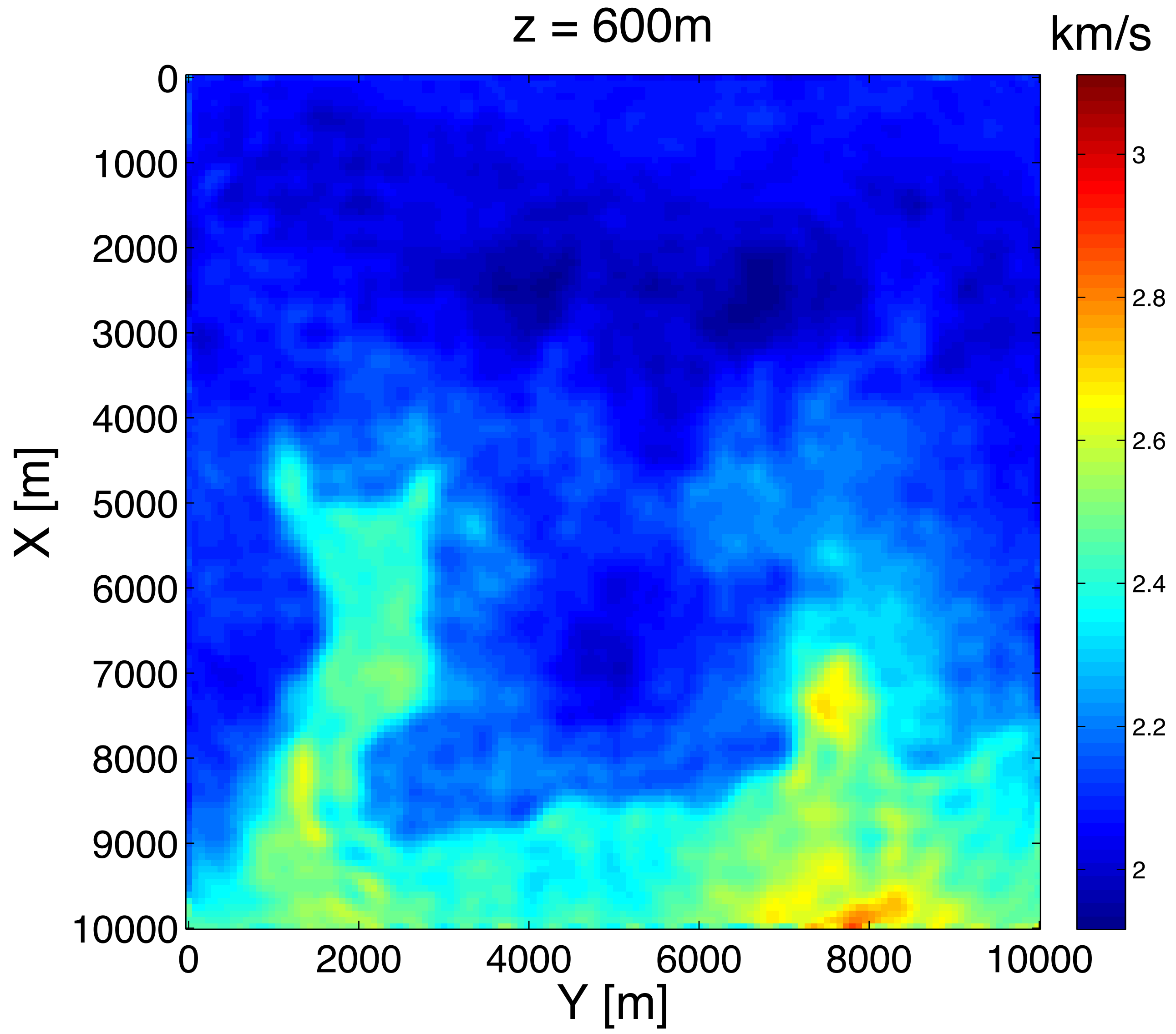
z = 600m



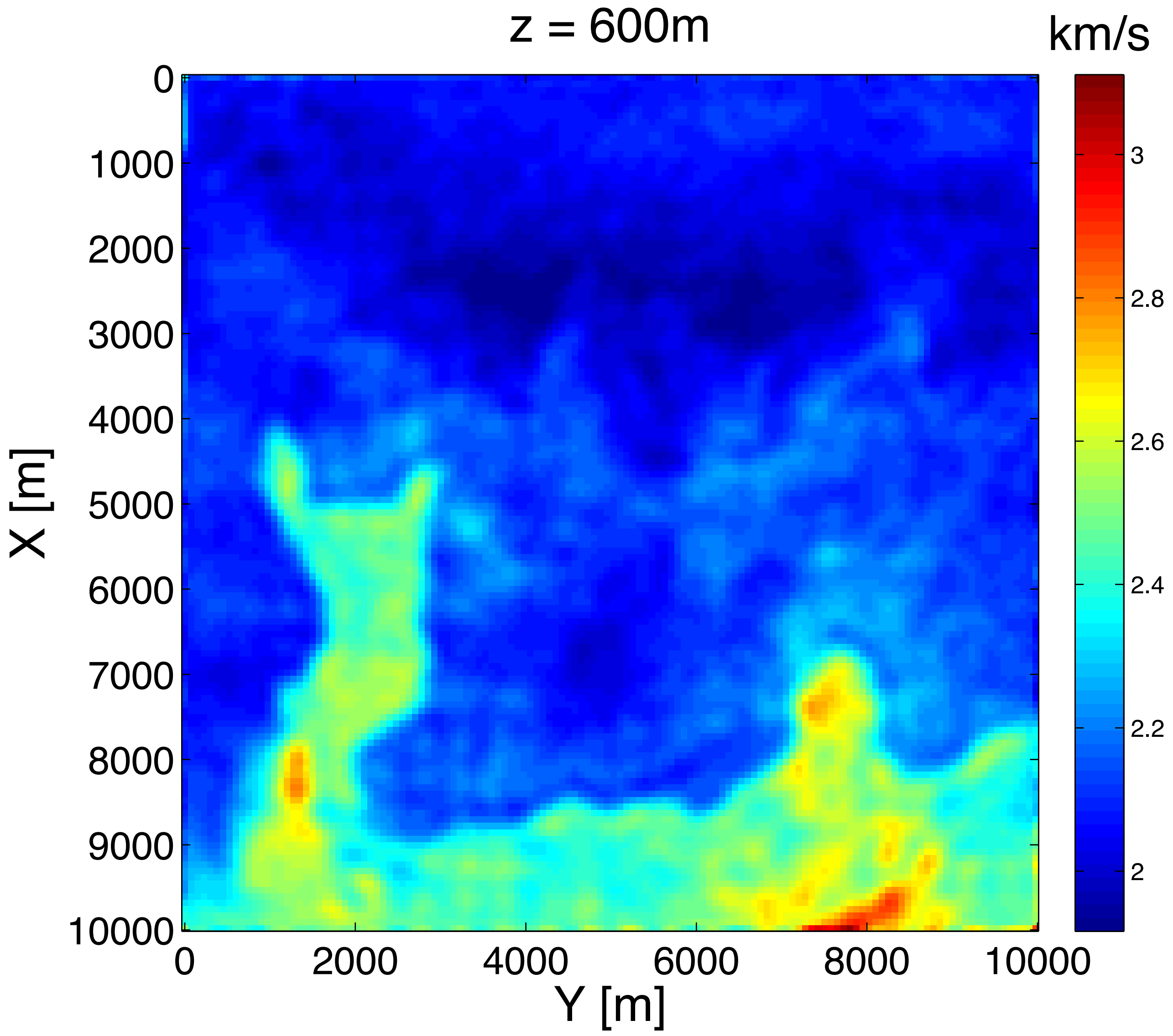
Composite shots



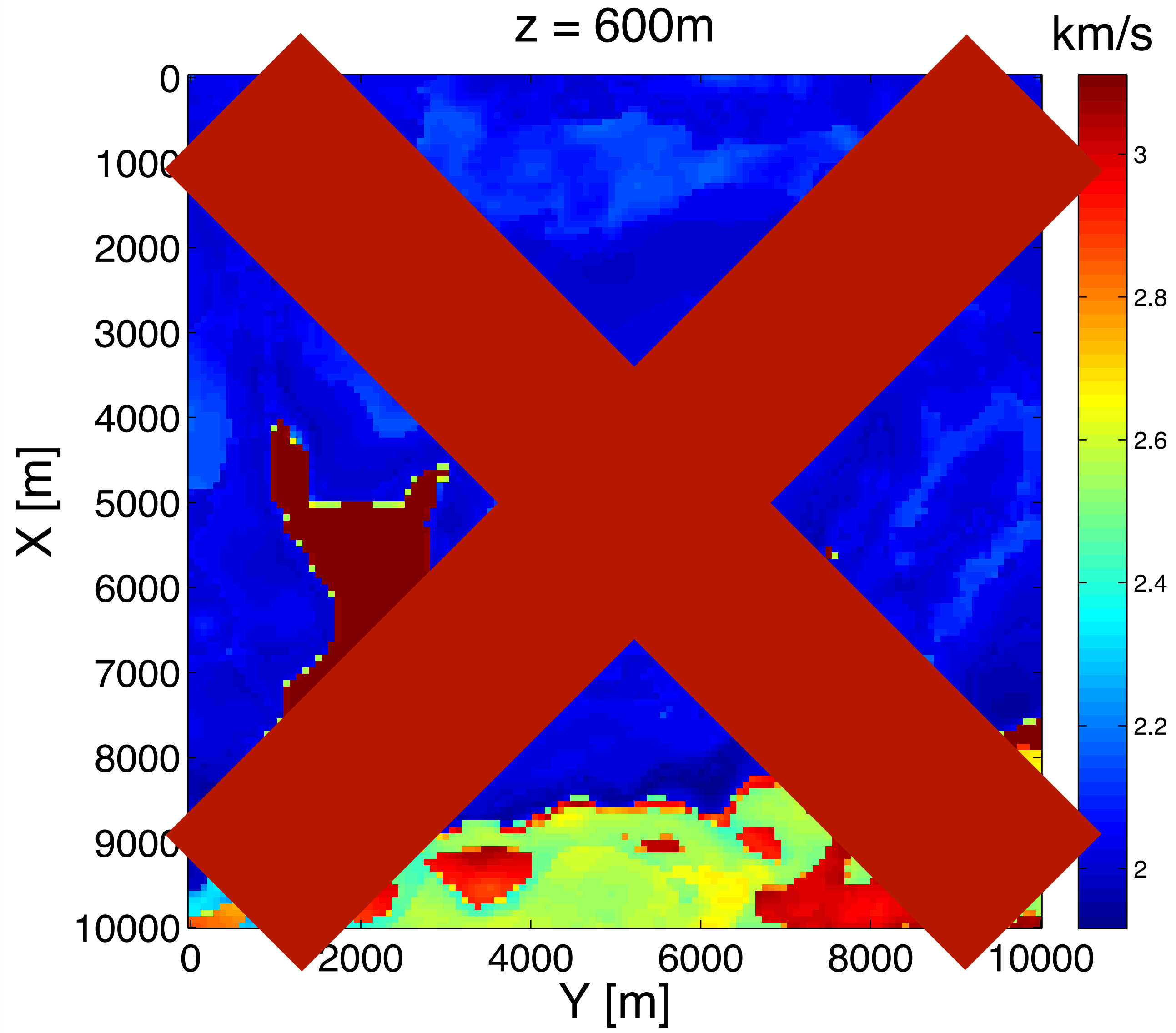
Initial model



Randomly selected shots

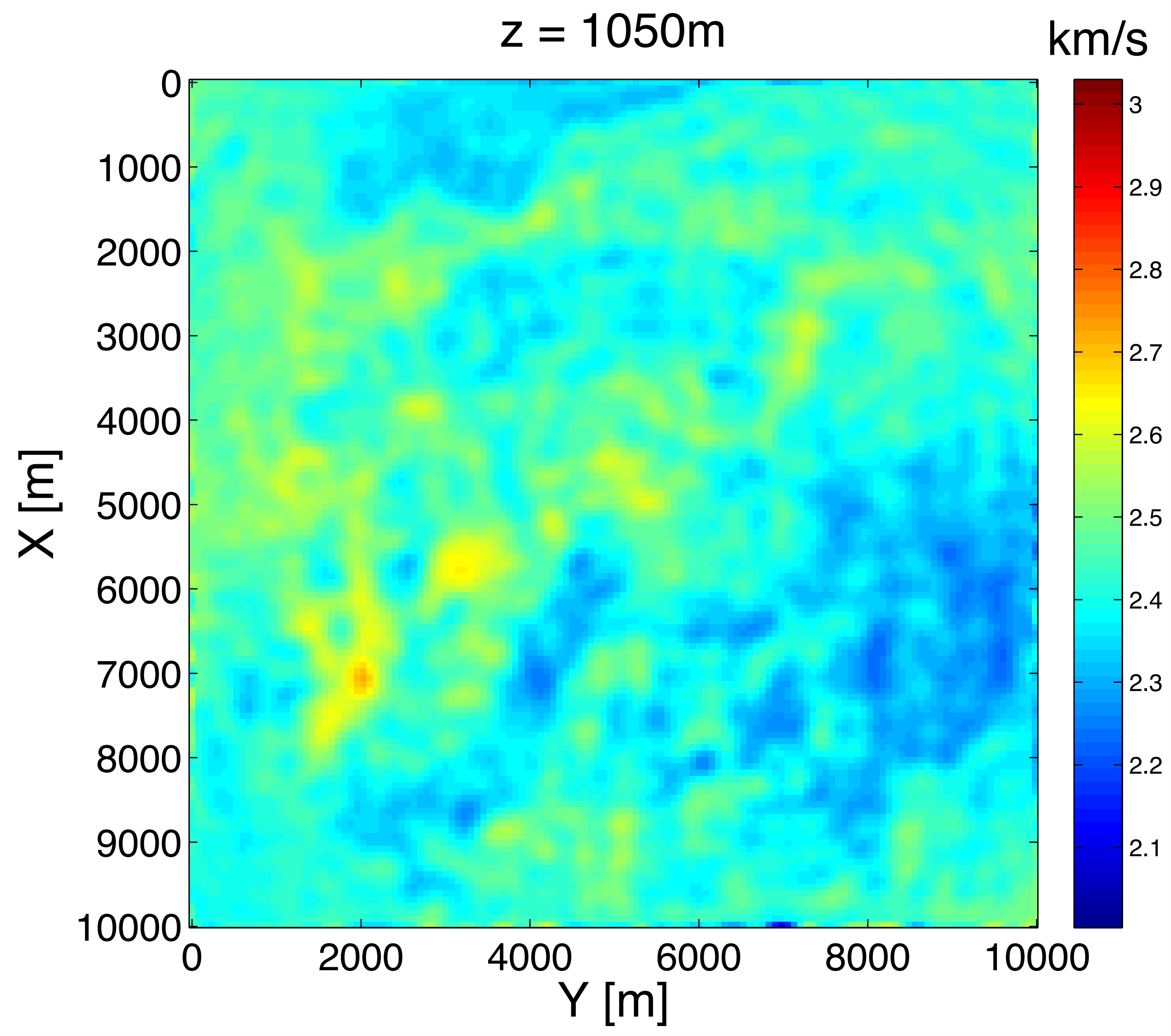


Simultaneous shots

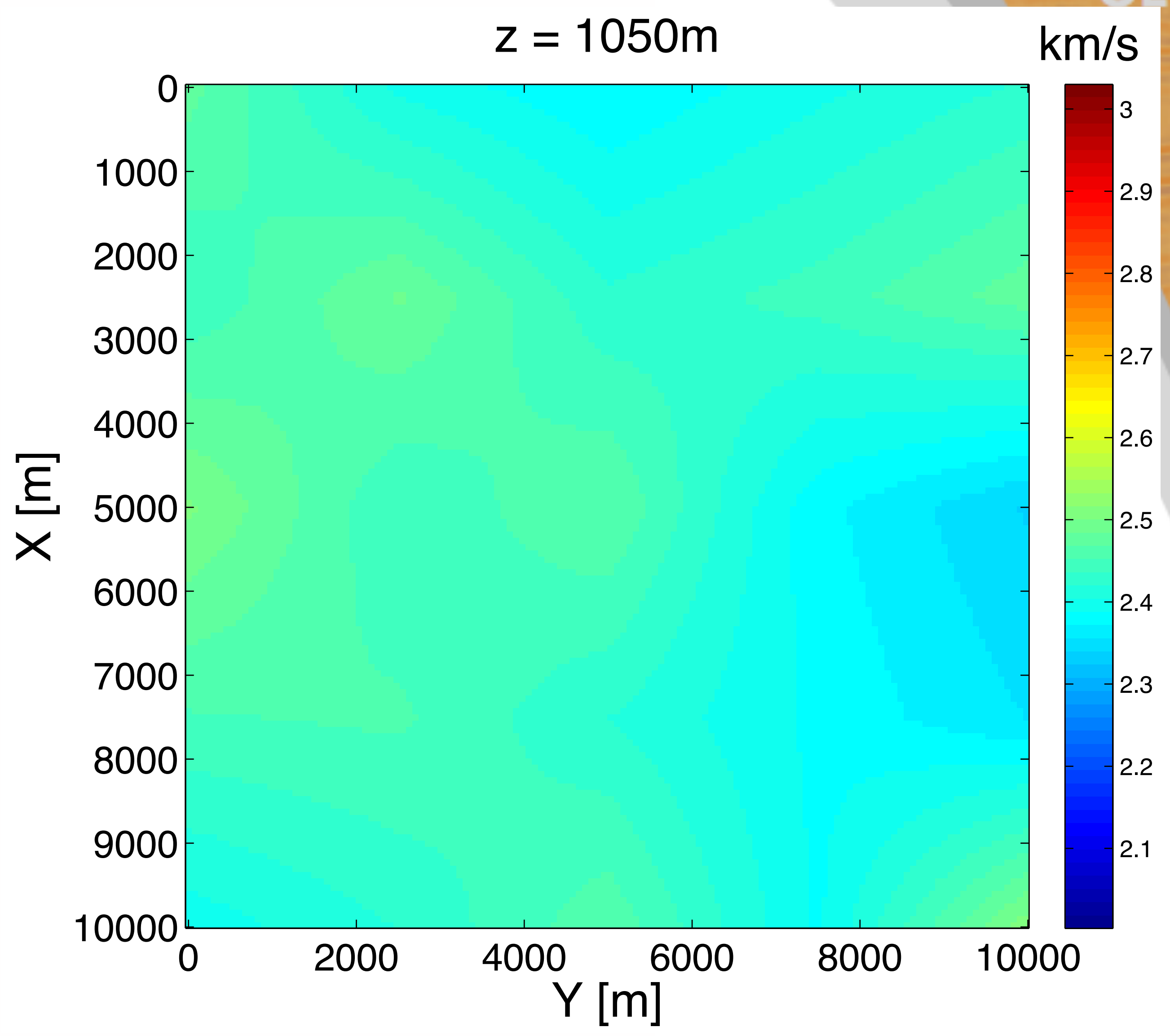


True model

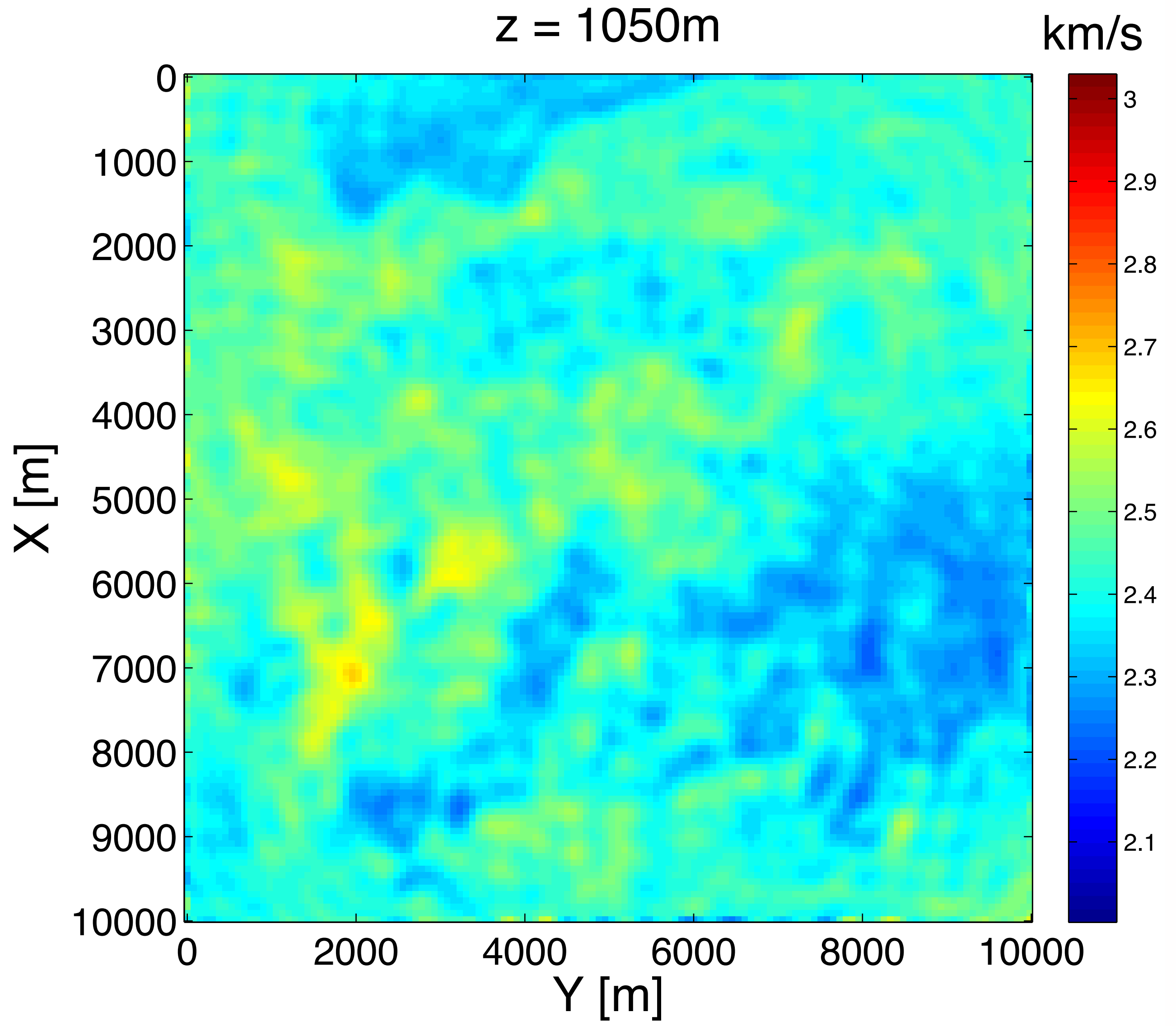
z = 1050m



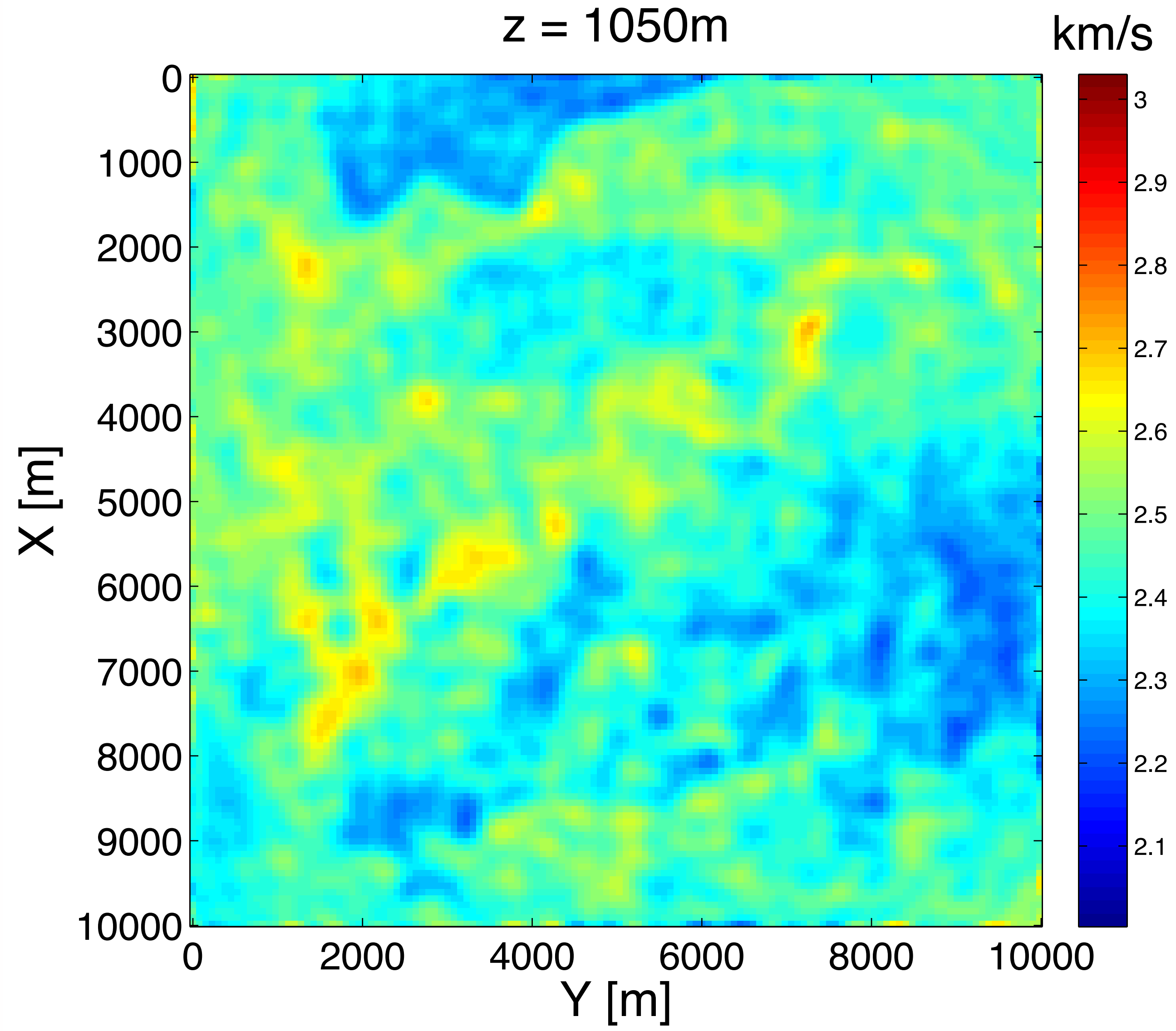
Composite shots



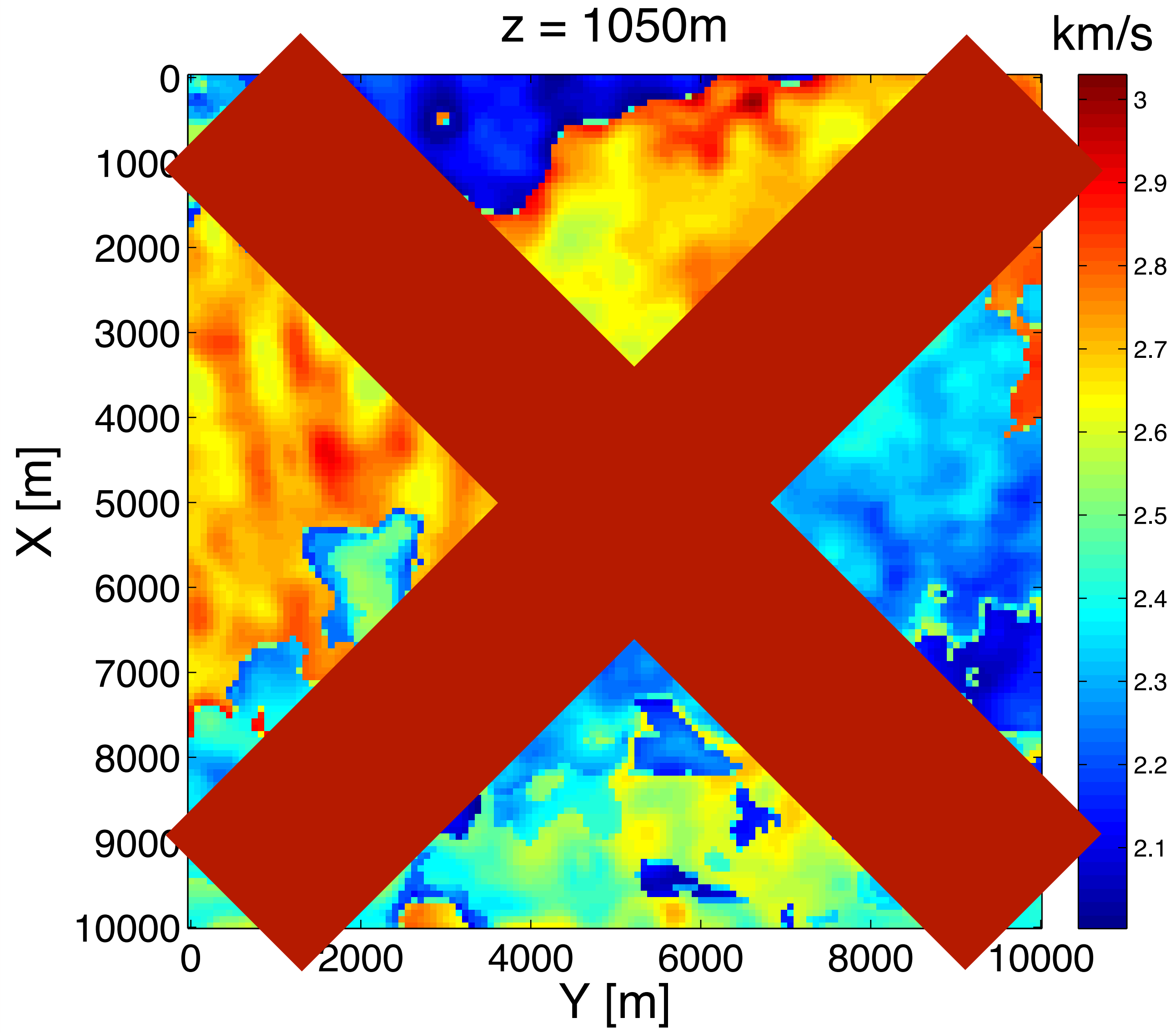
Initial model



Randomly selected shots

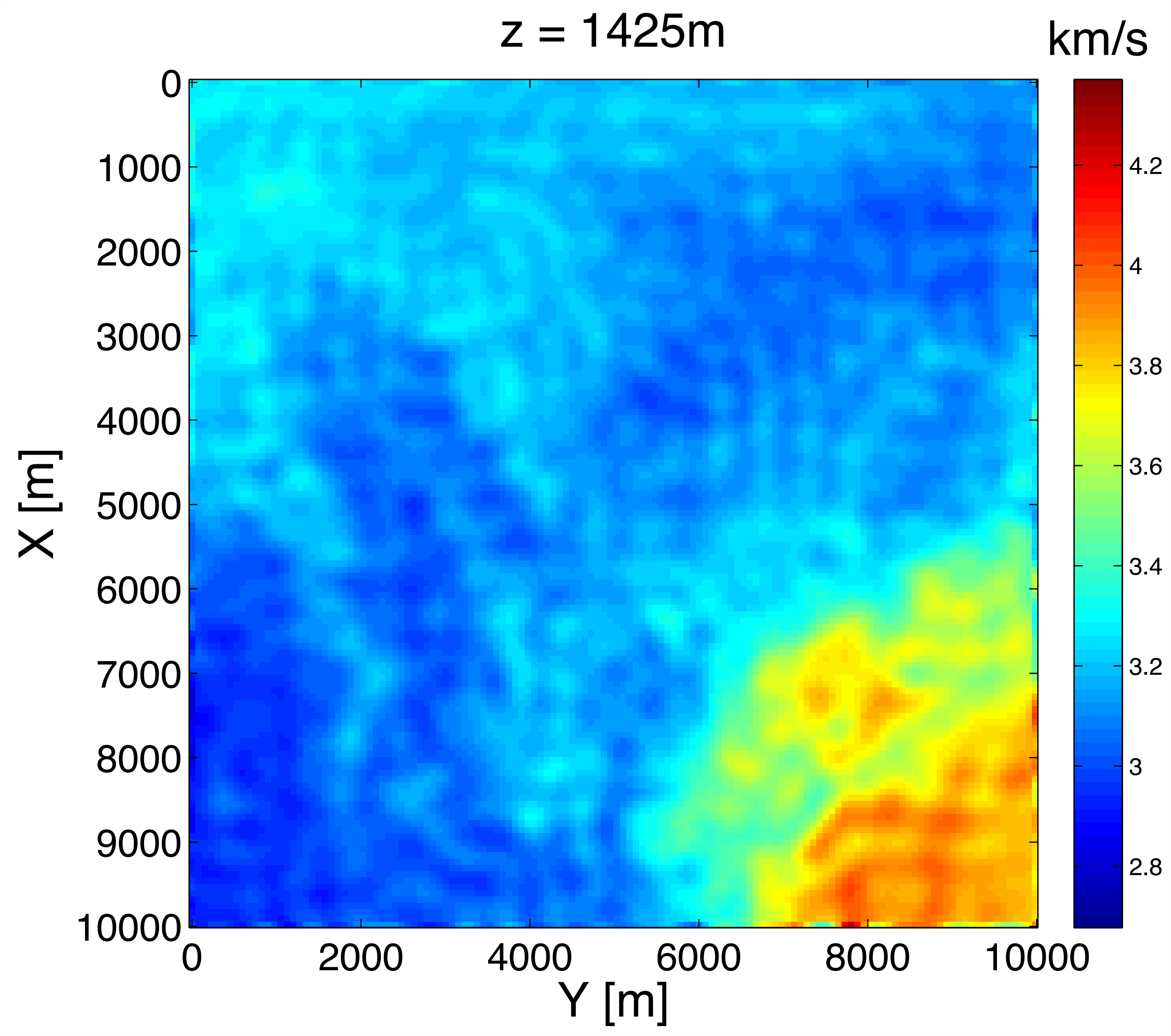


Simultaneous shots

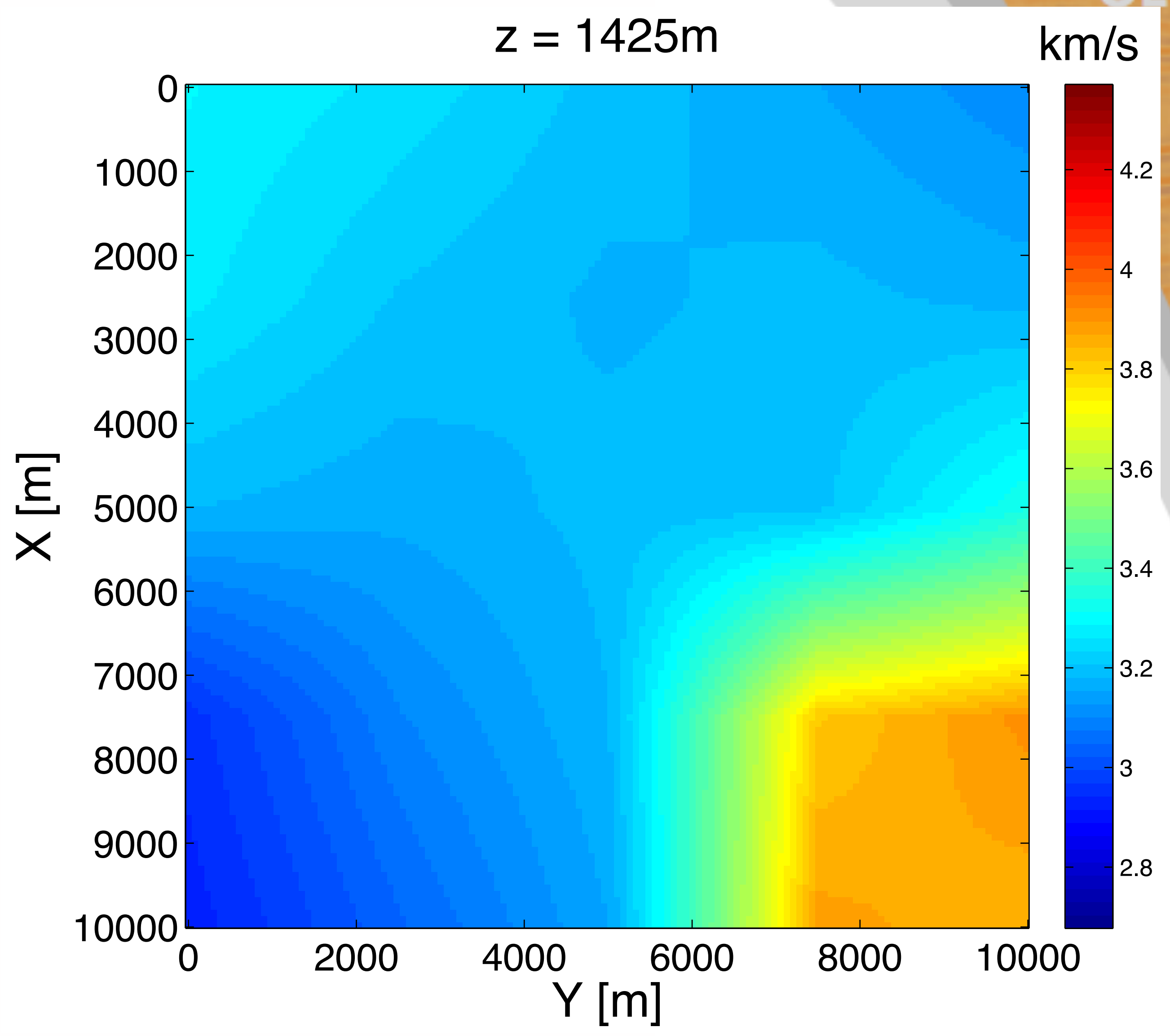


True model

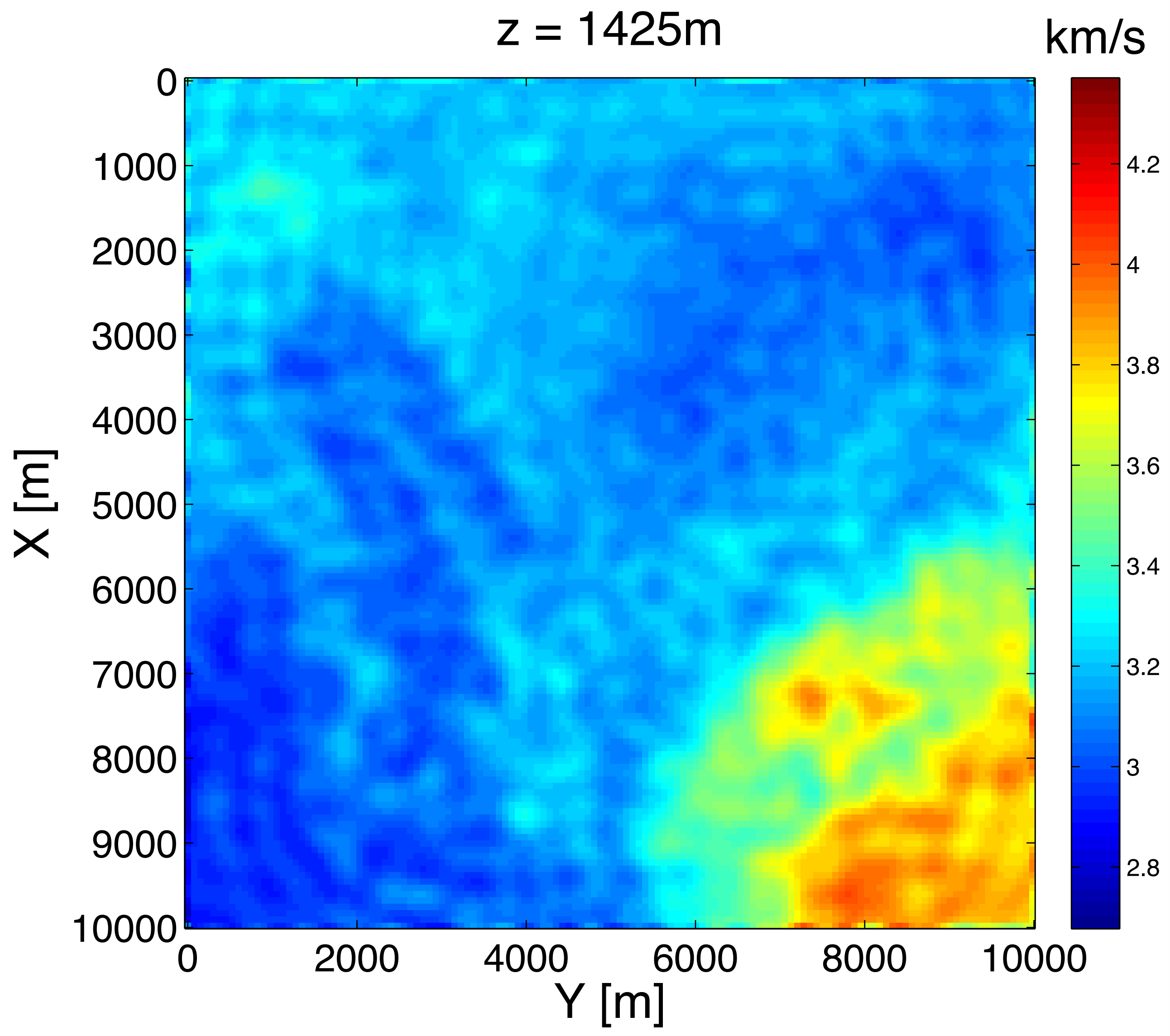
z = 1425m



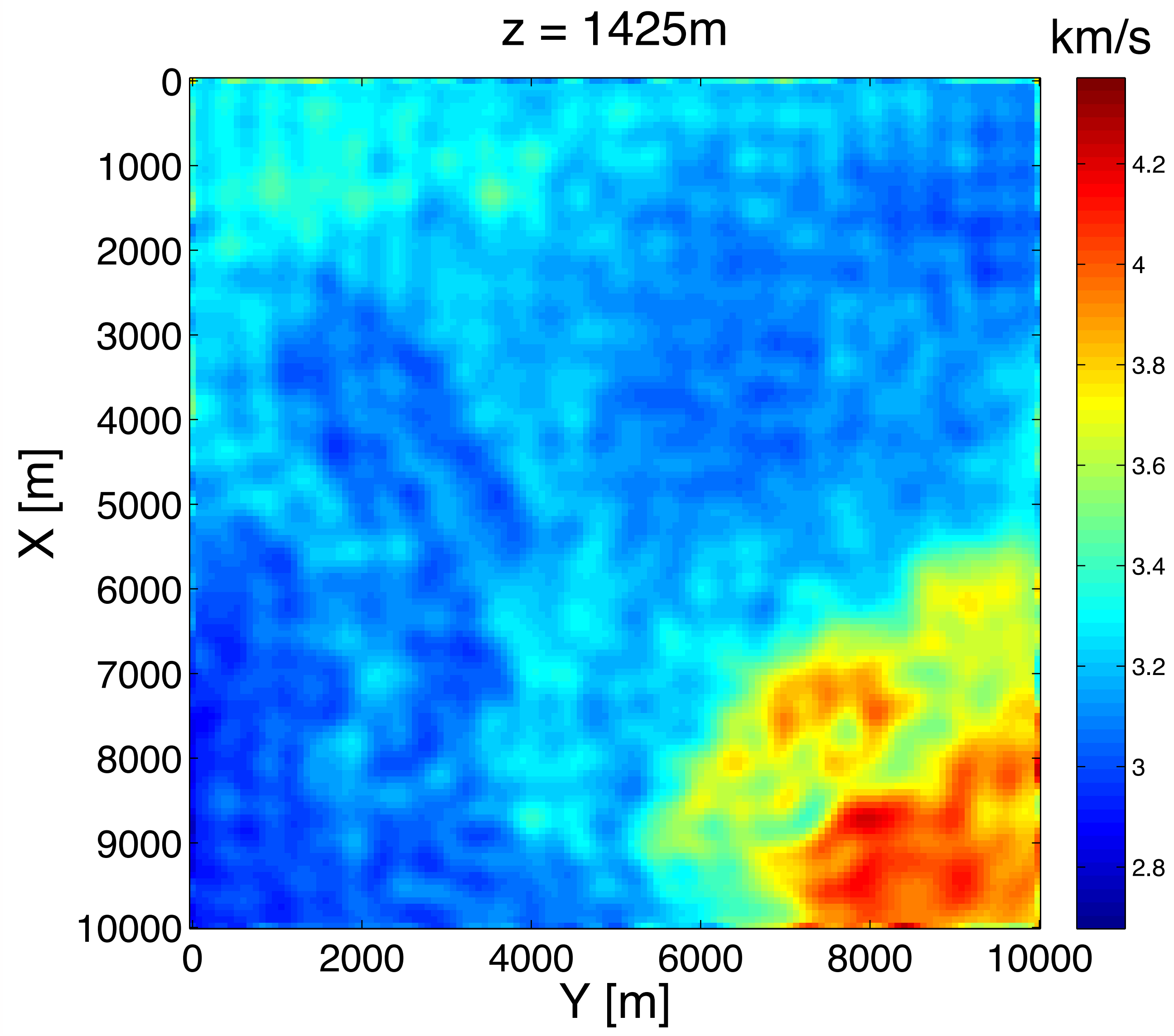
Composite shots



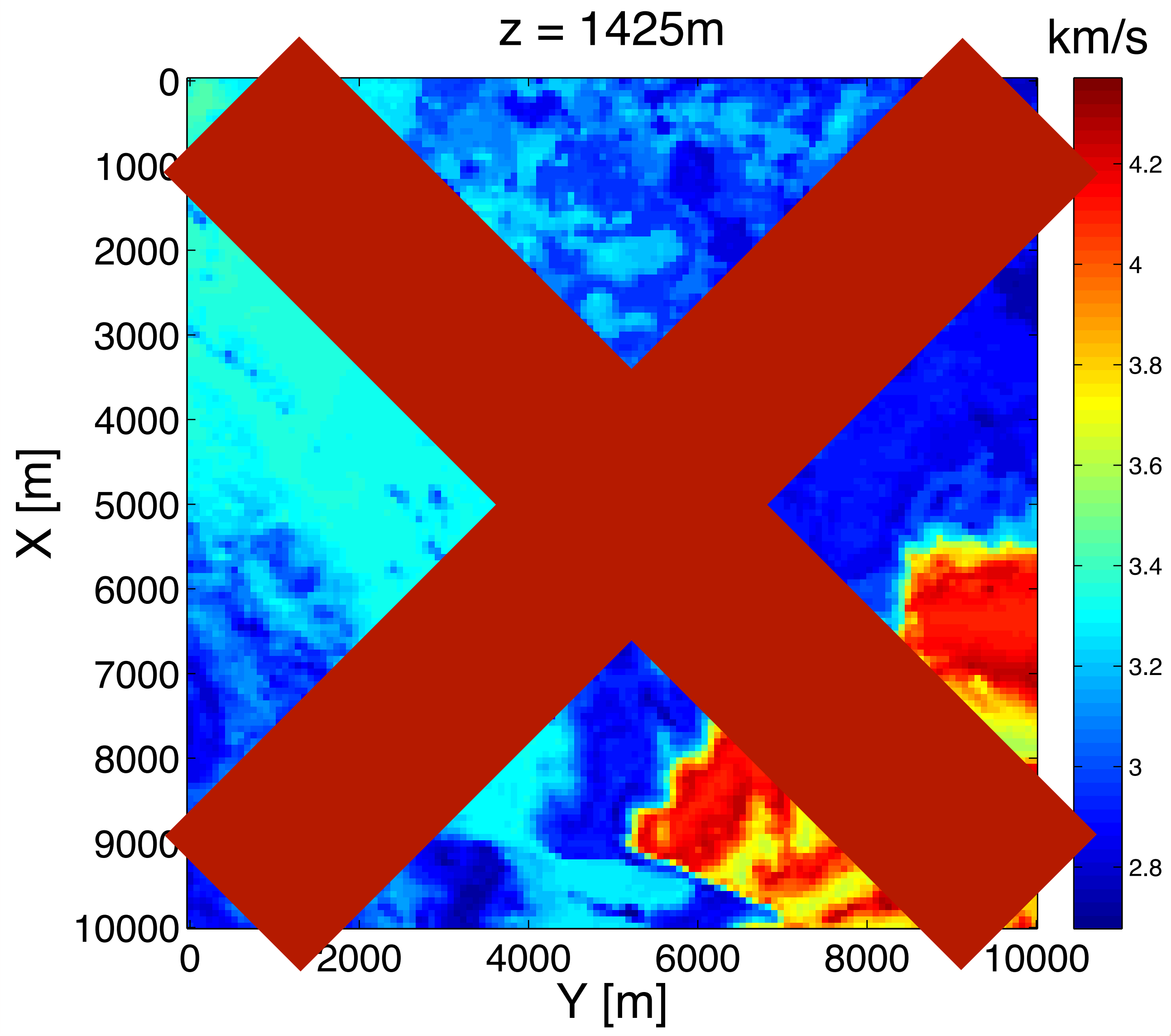
Initial model



Randomly selected shots

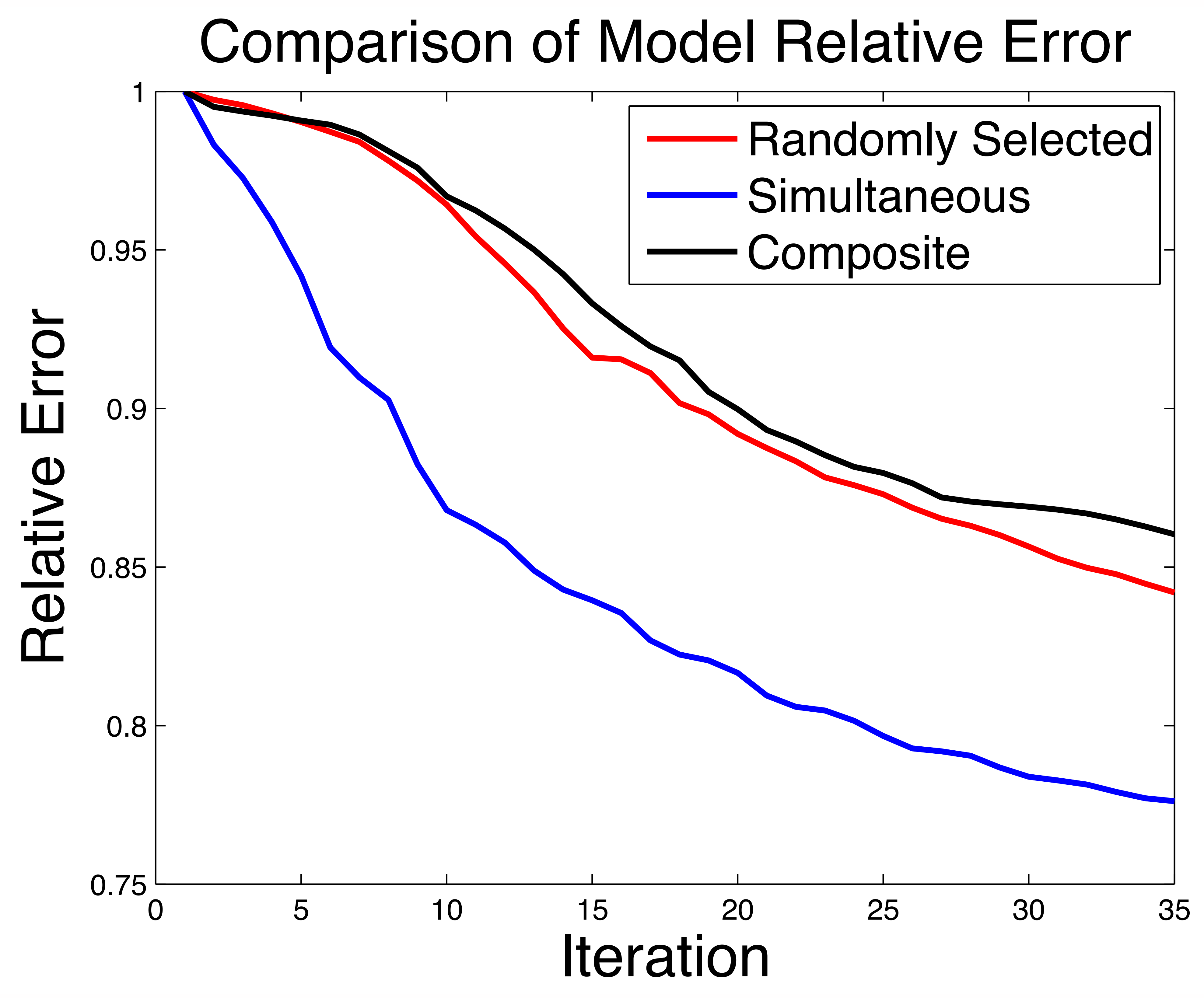


Simultaneous shots

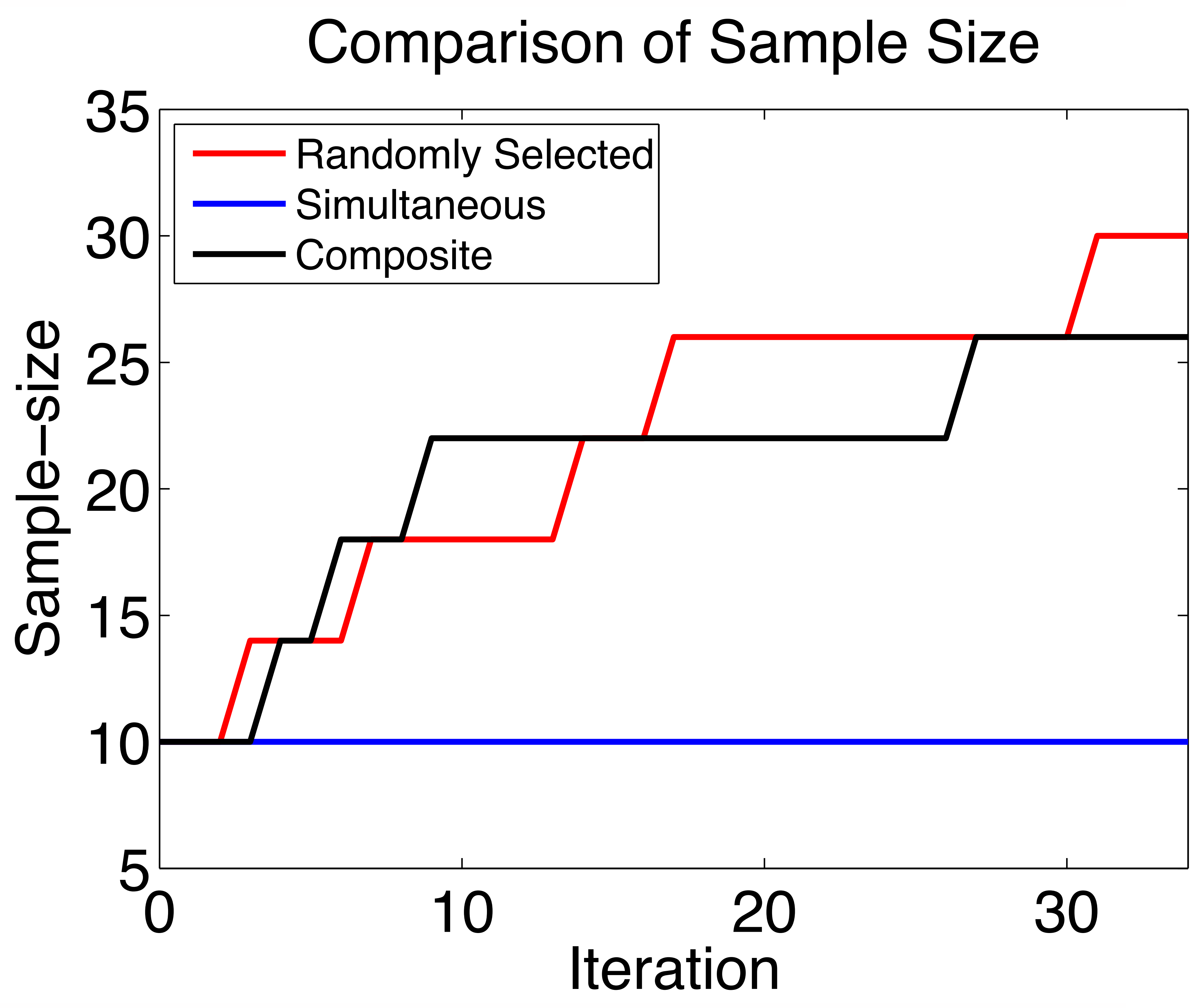


True model

Comparison



Comparison of Model Relative Error
Relative Error
Iteration
1
0.95
0.9
0.85
0.8
0.75
0
5
10
15
20
25
30
35
Randomly Selected
Simultaneous
Composite



Comparison of Sample Size
Sample−size
Iteration
35
30
25
20
15
10
5
0
10
20
30
Randomly Selected
Simultaneous
Composite

Conclusions

1. Applied *stochastic* optimization method where *computational* cost is *reduced* from $\mathcal{O}(n_s)$ to $\mathcal{O}(n_{ss})$.

2. *Simultaneous* shots require *fewer* source experiments (= # PDE solves) to obtain *better* results compared to *other* two methods.

3. We present a *new* sampling method - *composite* shots which
   - can extract more information per *composite* shot
   - allows for *offset* weighting
   - suitable for multi-source vessel marine

Future work

1. Deal with higher frequency data;
2. Study different strategies of composite shots;
3. Adaptive importance sampling;
4. Work with all data (If we have a big machine).

SLIM

Acknowledgements

Thank you for your attention !

https://www.slim.eos.ubc.ca/

SINBAD

NSERC CRSNG

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, Statoil, Total SA, WesternGeco, and Woodside.