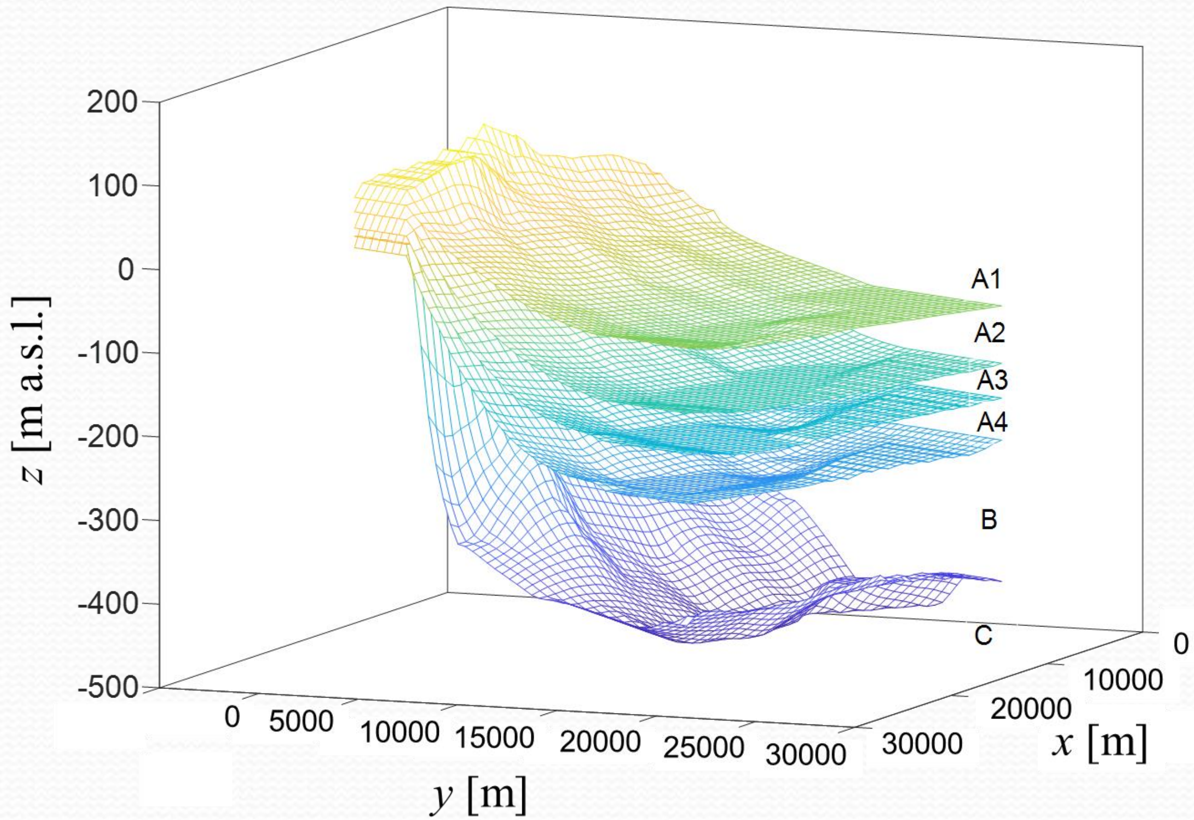
Geostatistical characterization and groundwater-flow model calibration in the Bologna aquifer system

Martina Siena, Monica Riva, Alberto Guadagnini

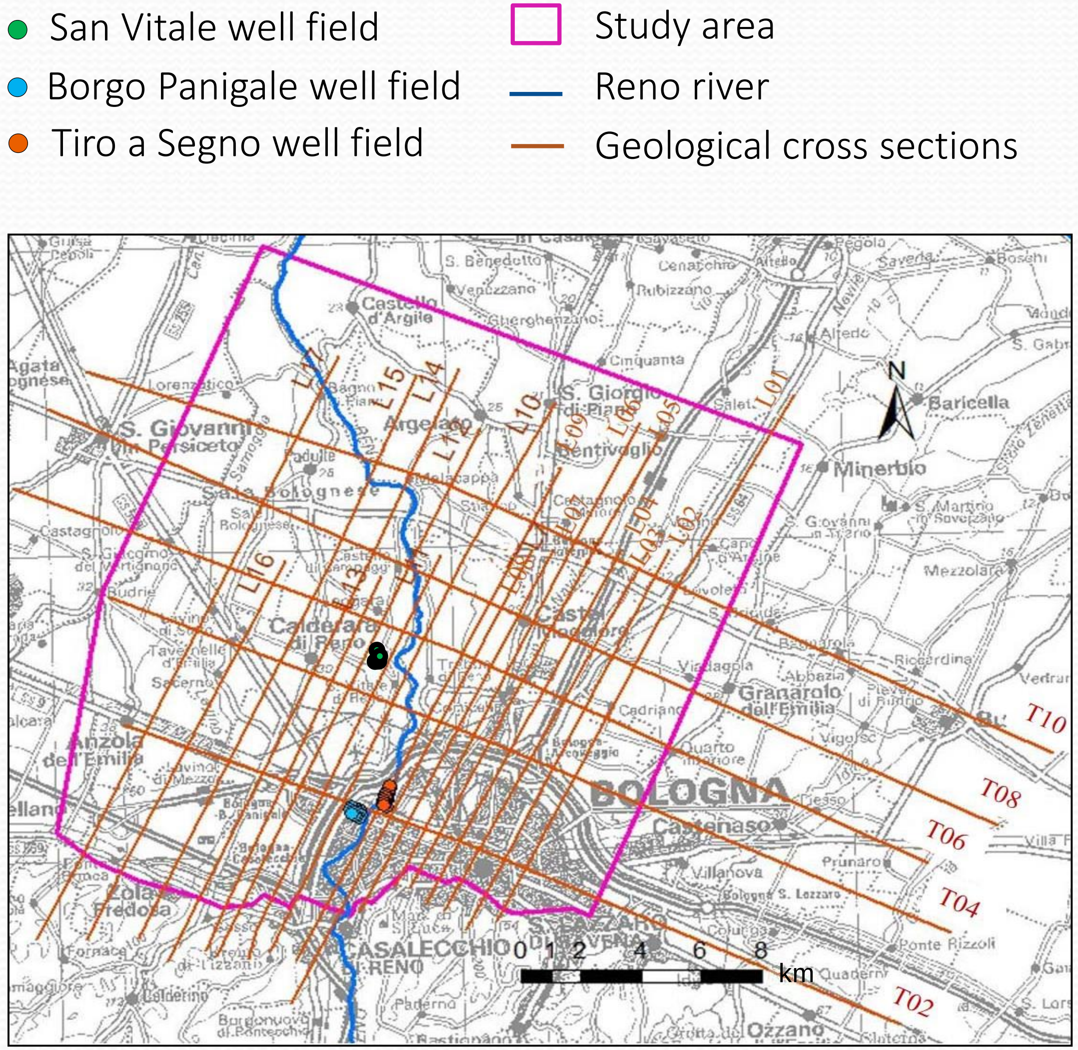
WE-NEED final conference and international workshop

Bologna site: investigated domain

- Medium alluvial Po Plain
- Surface area ≈ 460 km^2
- Three main groundwater bodies (A-B-C)
- Groundwater resources intensively exploited for urban supply



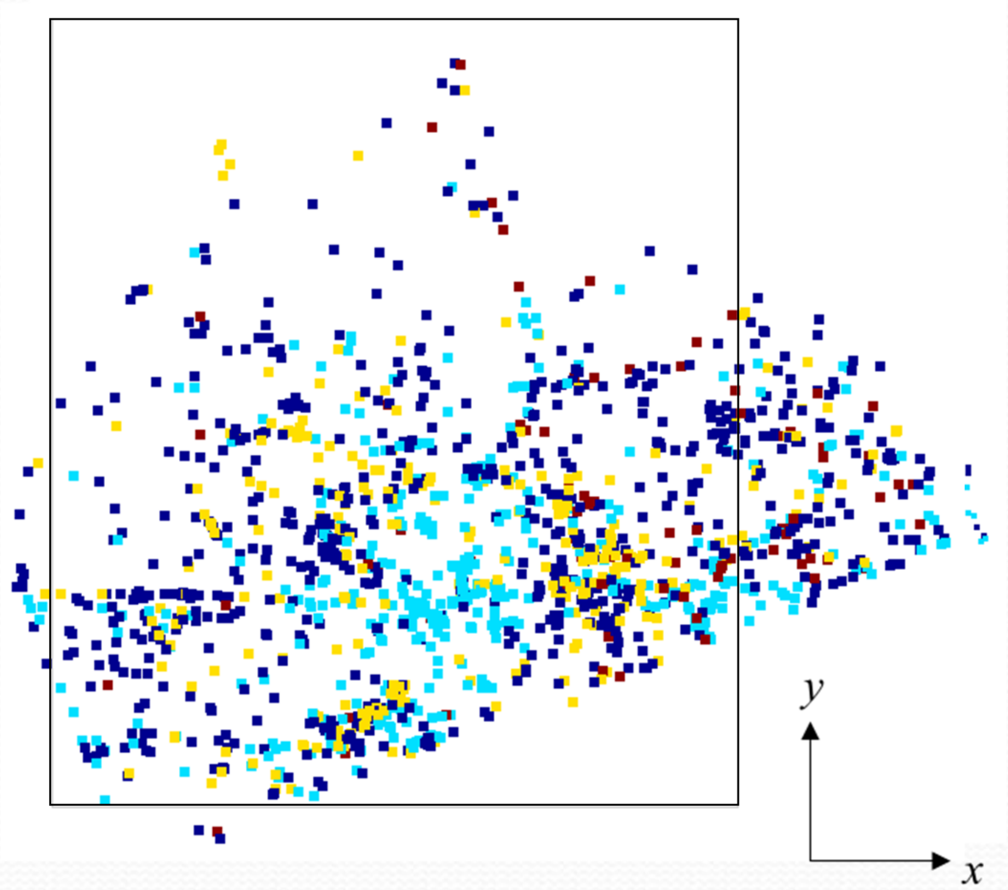
Bologna site: investigated domain

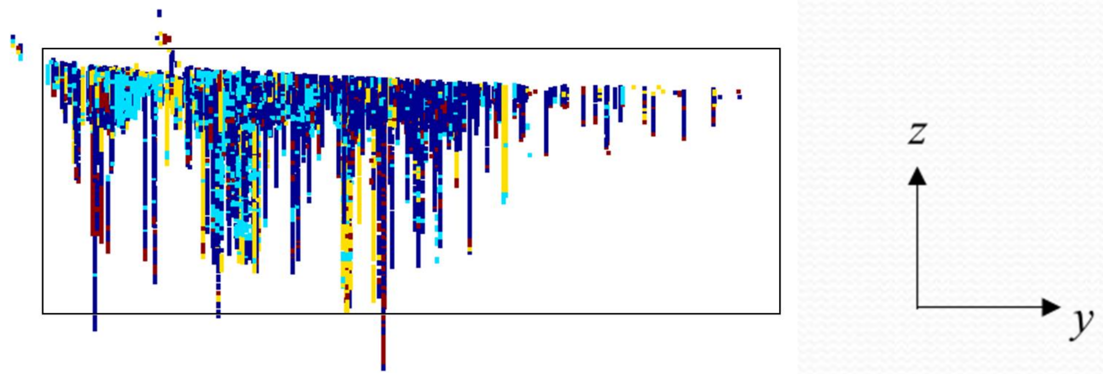


- Medium alluvial Po Plain
- Surface area ≈ 460 km^2
- Three main groundwater bodies (A-B-C)
- Groundwater resources intensively exploited for urban supply

Lithological data

Lithological data from 1303 boreholes allowed to identify 4 main categories:





- Clay 52% — Least-conductive facies
- Gravel 28% — Most-conductive facies
- Silt 13%
- Sand 7%

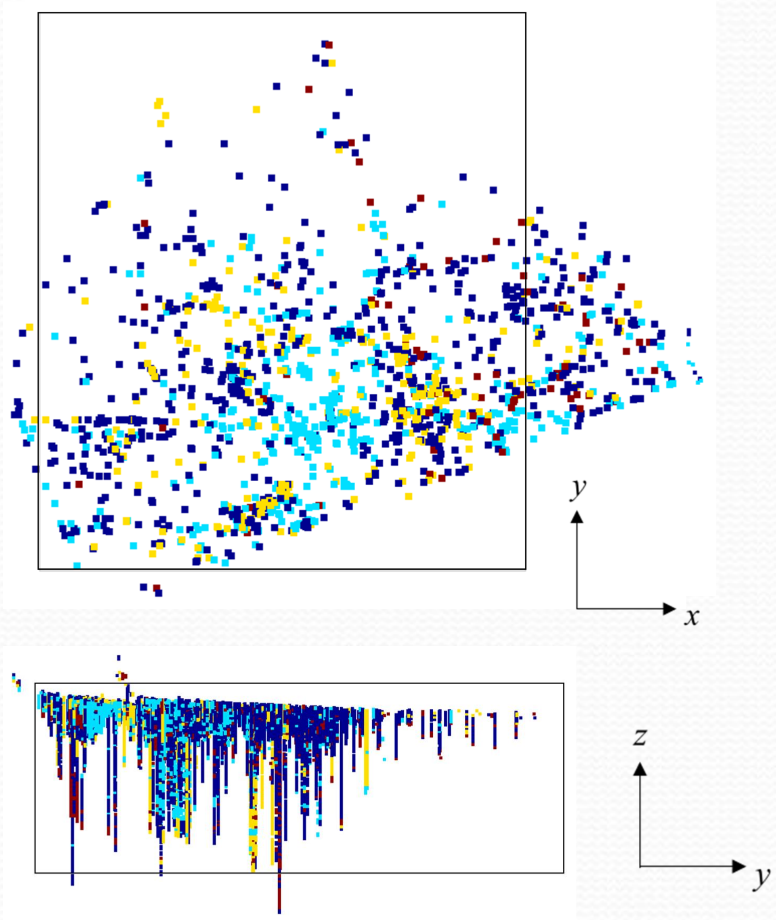
Outline of the work

- Apply 2 geostatistical reconstruction methods to describe the architecture of the aquifer system, on the basis of lithological data
- Evaluate possible impacts of the reconstruction method on the connectivity of lithological facies
- Develop a 3D groundwater flow model
- Evaluate possible impacts of the reconstruction method on model outputs

Reconstruction methods

Sequential Indicator Simulation (SISIM)

Transitional Probability Simulation (T-PROGS)



Reconstruction methods: SISIM

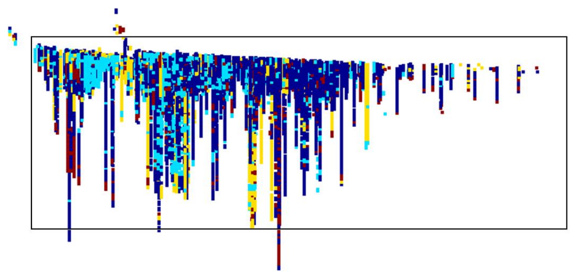
Sequential Indicator Simulation (SISIM)

1. Discrete variable

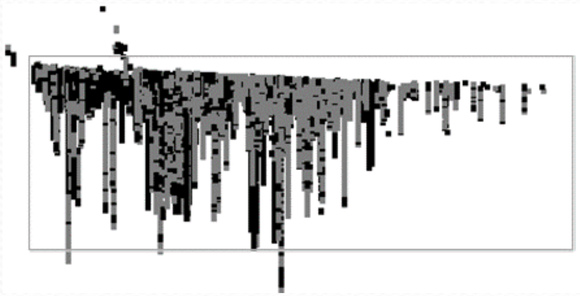
$$Z(x) = 1, 2, 3, 4$$

Indicator function

$$I(x;k) = \begin{cases} 1 \text{ if } Z(x) = k \\ 0 \text{ otherwise} \end{cases}$$



$$Z(x) = \begin{cases} 1 & \text{Clay} \\ 2 & \text{Gravel} \\ 3 & \text{Silt} \\ 4 & \text{Sand} \end{cases}$$



$$I(x,1) = \begin{cases} 1\ if\ Z(x) = 1 \\ 0\ elsewhere \end{cases}$$

Reconstruction methods: SISIM

<u>Sequential Indicator Simulation (SISIM)</u>

1. Discrete variable

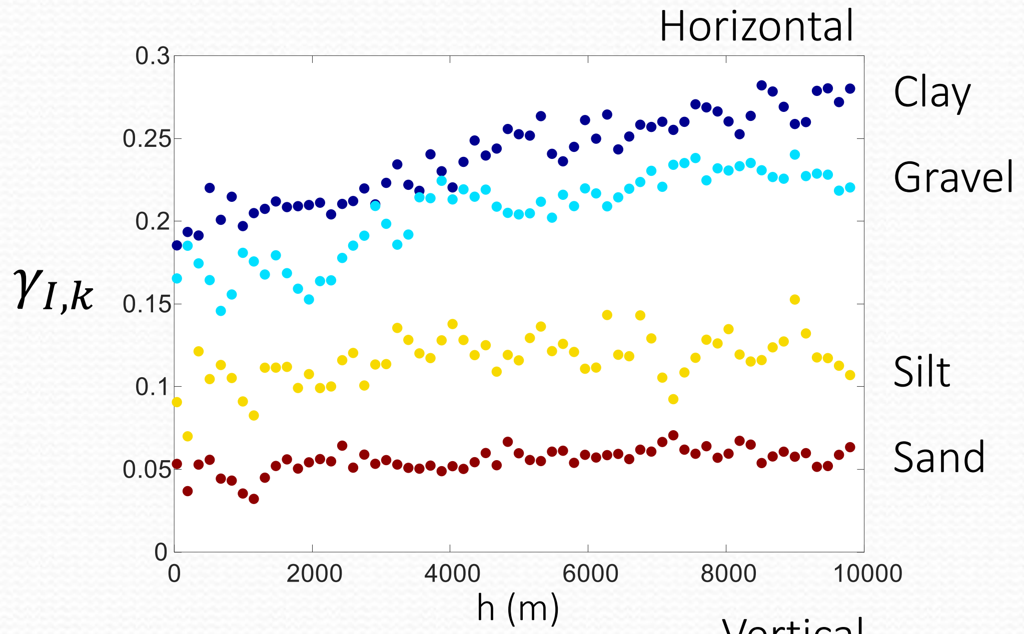
$$Z(x) = 1, 2, 3, 4$$

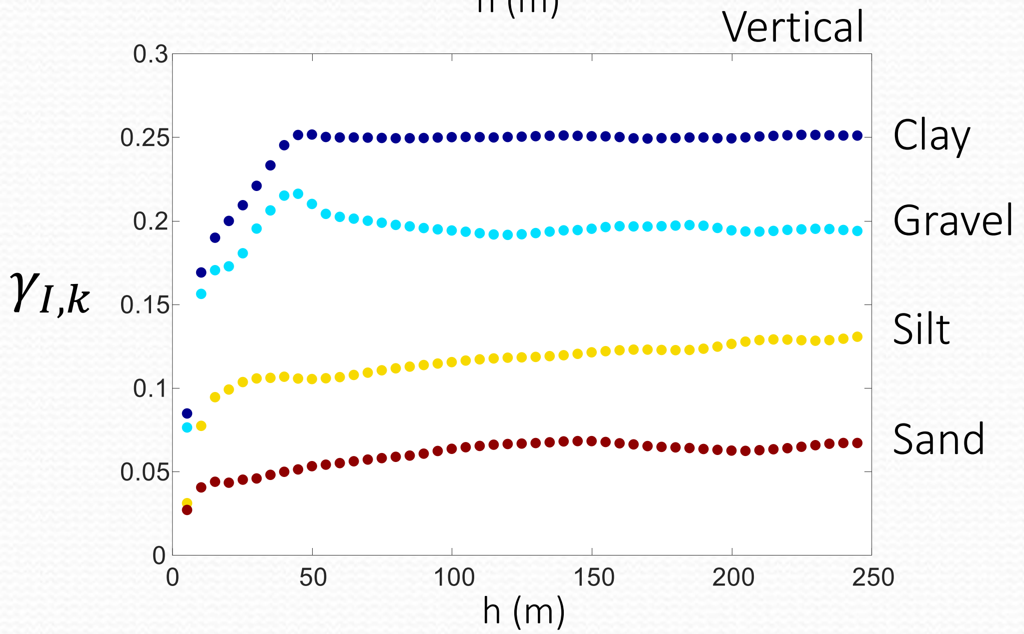
Indicator function

$$I(x;k) = \begin{cases} 1 \text{ if } Z(x) = k \\ 0 \text{ otherwise} \end{cases}$$

2. Sample (directional) variograms based on data

$$\gamma_{I,k}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [I(x;k) - I(x+h;k)]^2$$





Reconstruction methods: SISIM

Sequential Indicator Simulation (SISIM)

1. Discrete variable

$$Z(x) = 1, 2, 3, 4$$

Indicator function

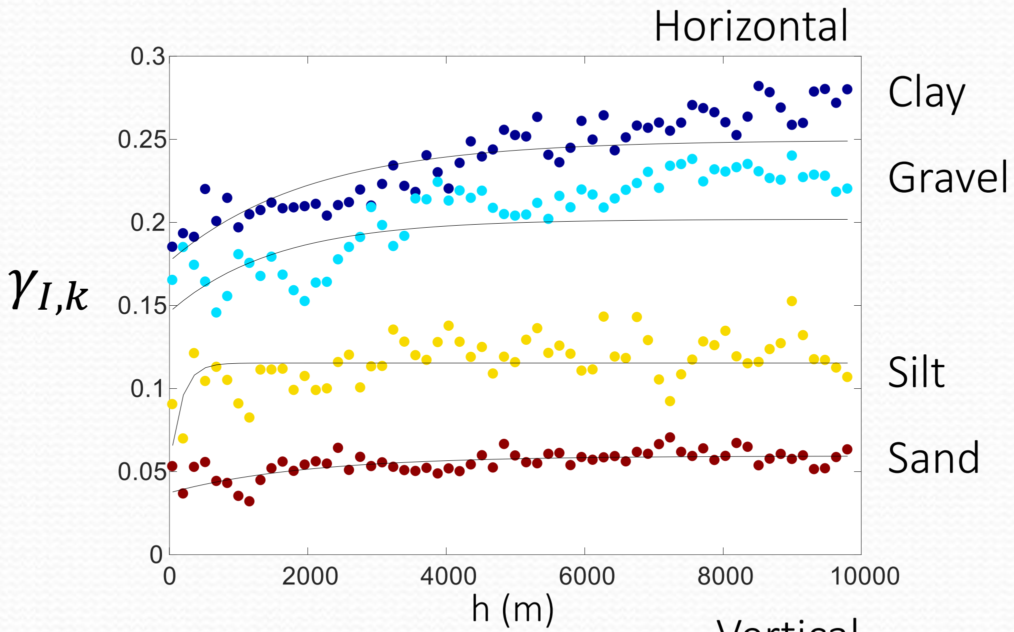
$$I(x;k) = \begin{cases} 1 \text{ if } Z(x) = k \\ 0 \text{ otherwise} \end{cases}$$

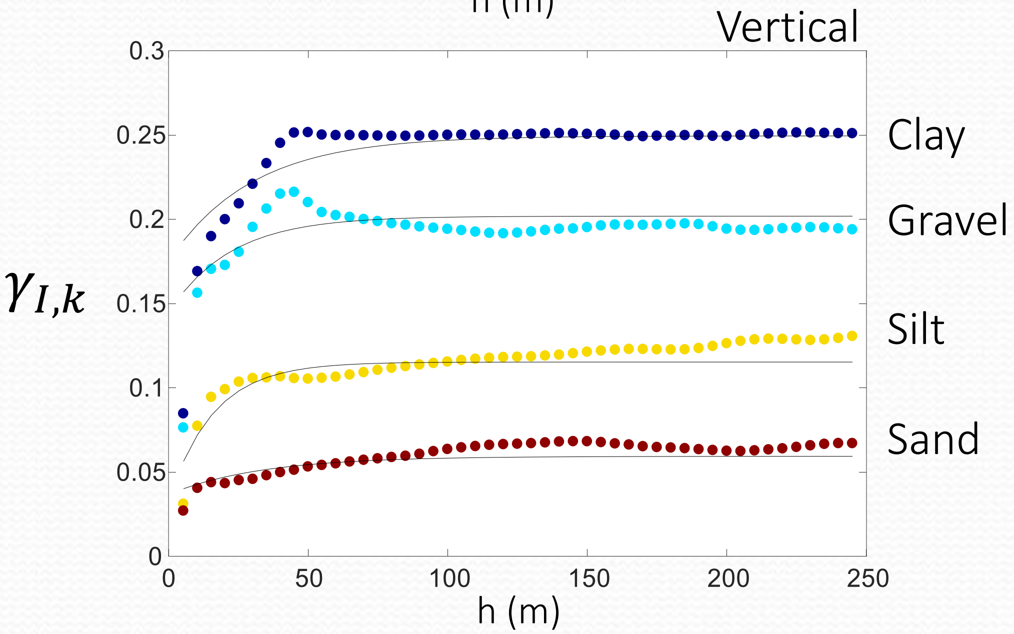
2. Sample (directional) variograms based on data

$$\gamma_{I,k}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [I(x;k) - I(x+h;k)]^2$$

3. Best-fitting variogram model for each k

$$\gamma(h) = c_0 + \sigma^2 [1 - \exp(h/a)]$$



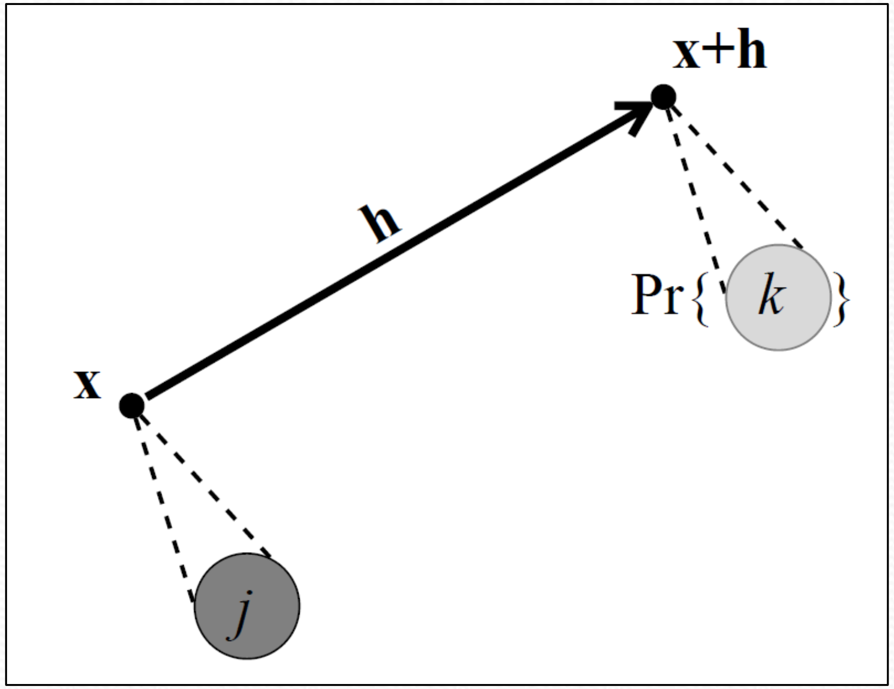


Reconstruction methods: T-PROGS

Transitional Probability Simulation (T-PROGS)

1. Transition probability:

$$t_{jk}(h) = \Pr\left\{Z(x+h) = k \,\middle|\, Z(x) = j\right\}$$



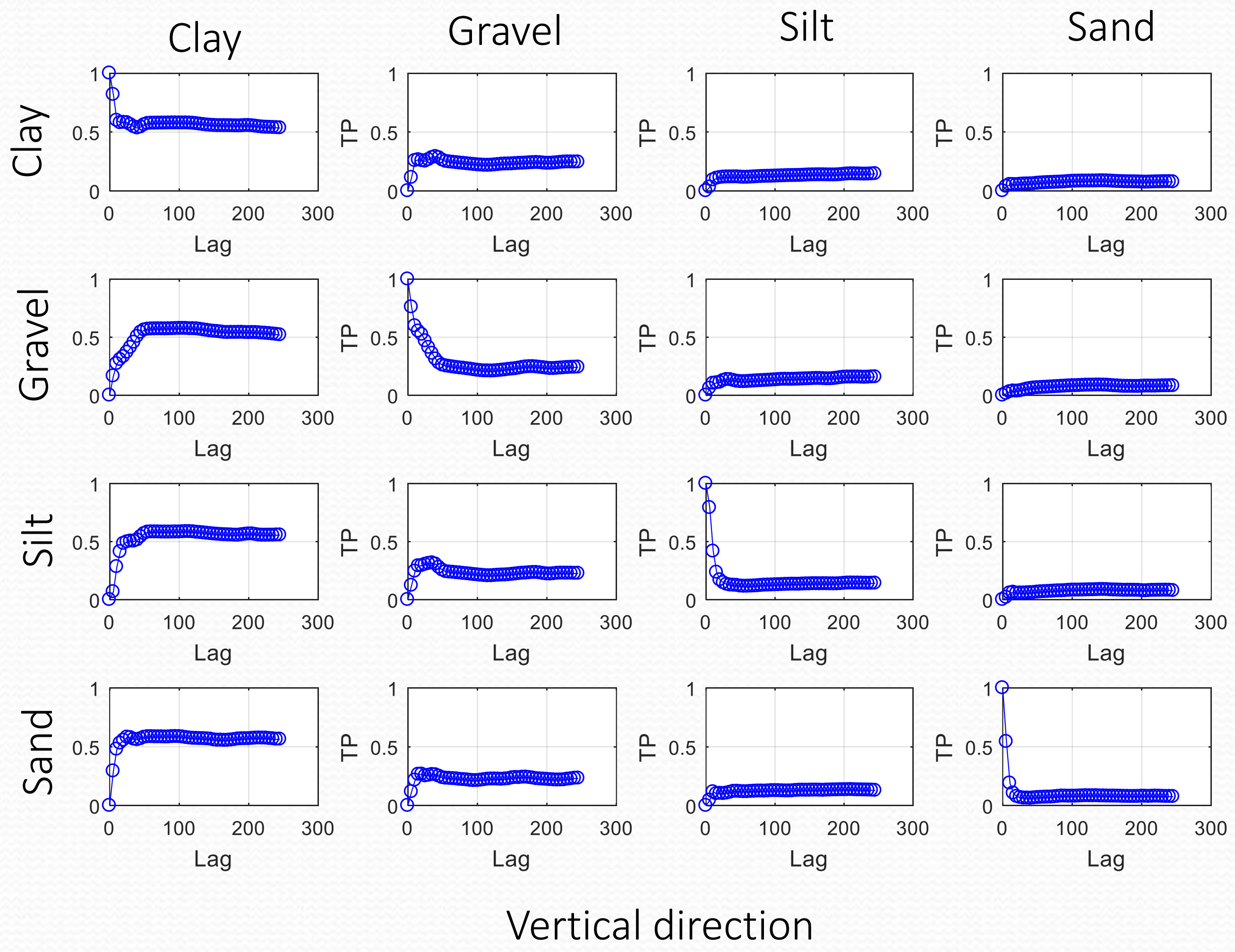
Reconstruction methods: T-PROGS

Transitional Probability Simulation (T-PROGS)

1. Transition probability:

$$t_{jk}(h) = \Pr\{Z(x+h) = k \,|\, Z(x) = j\}$$

2. Sample (directional) transiograms based on data



Reconstruction methods: T-PROGS

Transitional Probability Simulation (T-PROGS)

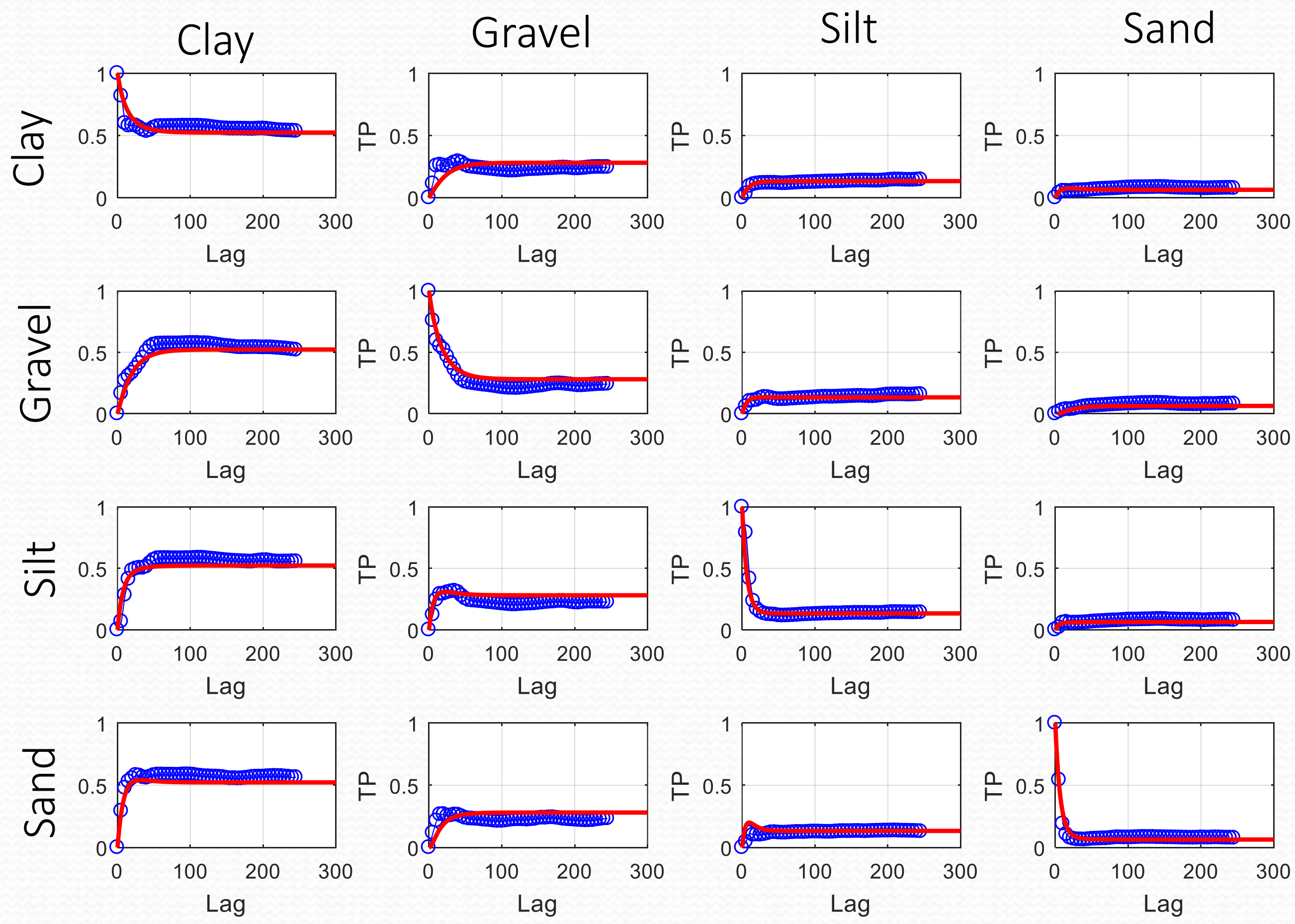
1. Transition probability:

$$t_{jk}(h) = \Pr\{Z(x+h) = k \mid Z(x) = j\}$$

2. Sample (directional) transiograms based on data

3. Best-fitting Markov-chain model

$$\mathbf{T}(h) = \exp(\mathbf{R}h)$$

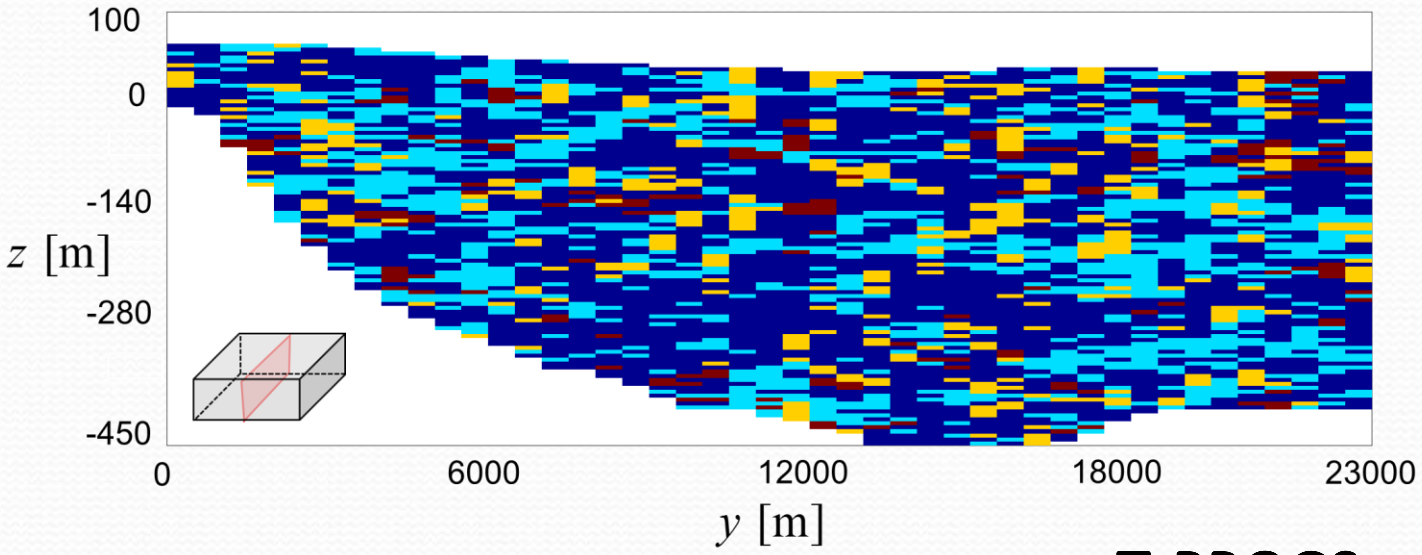


Vertical direction

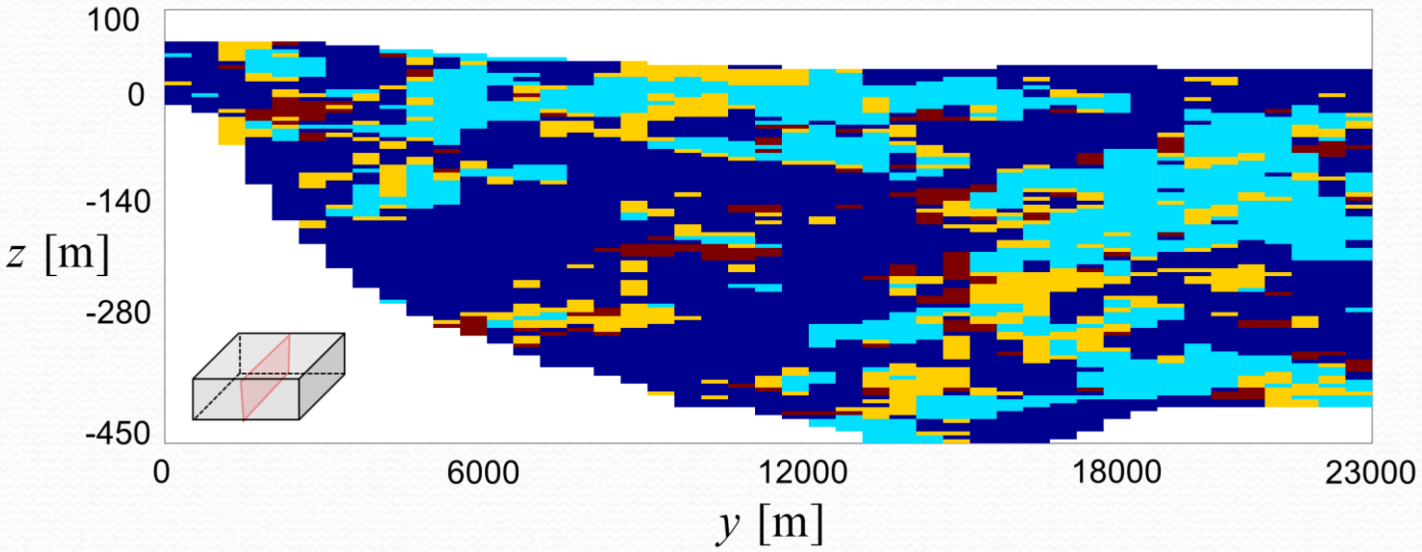
Reconstruction methods: results

Facies proportions	Dataset	Generated fields	SISIM	T-PROGS
Clay	0.523	mean	0.543	0.523
		std	5×10^{-3}	2×10^{-5}
Gravel	0.281	mean	0.267	0.281
		std	6×10^{-3}	2×10^{-5}
Silt	0.133	mean	0.118	0.133
		std	1×10^{-3}	1×10^{-5}
Sand	0.063	mean	0.071	0.063
		std	5×10^{-3}	2×10^{-5}

SISIM



T-PROGS

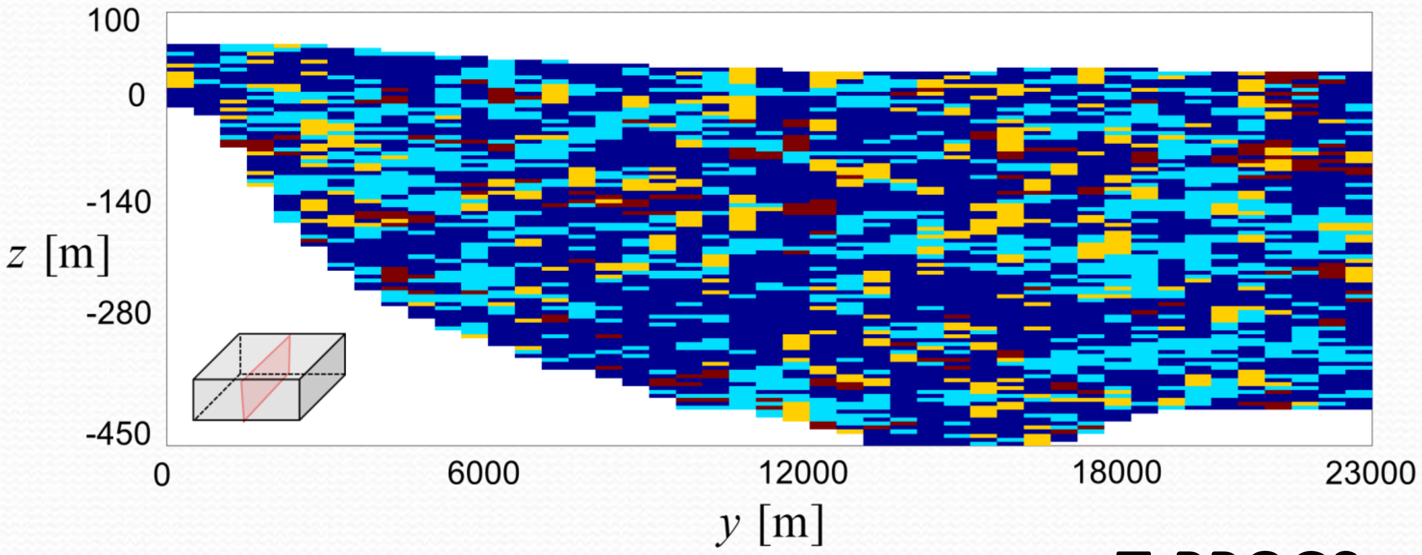


Reconstruction methods: results

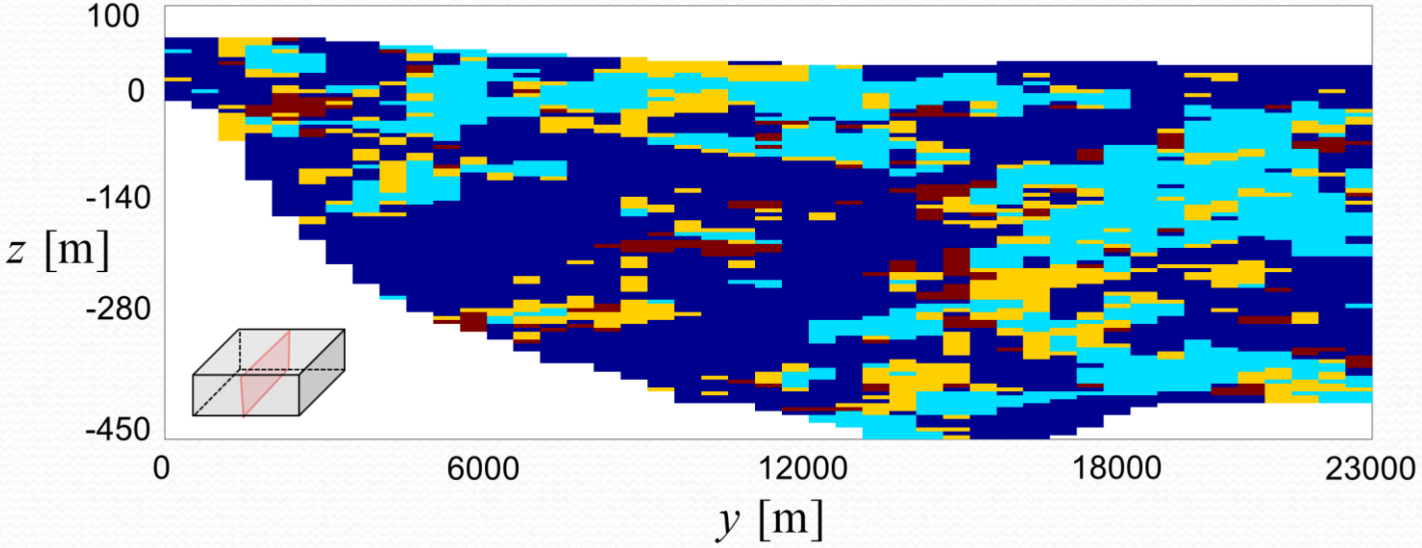
Facies proportions		Dataset	Generated fields	SISIM	T-PROGS
Clay		0.523	mean	0.543	0.523
			std	5×10^{-3}	2×10^{-5}
Gravel		0.281	mean	0.267	0.281
			std	6×10^{-3}	2×10^{-5}
Silt		0.133	mean	0.118	0.133
			std	1×10^{-3}	1×10^{-5}
Sand		0.063	mean	0.071	0.063
			std	5×10^{-3}	2×10^{-5}

SISIM simulations show more fragmented facies distributions compared to T-PROGS

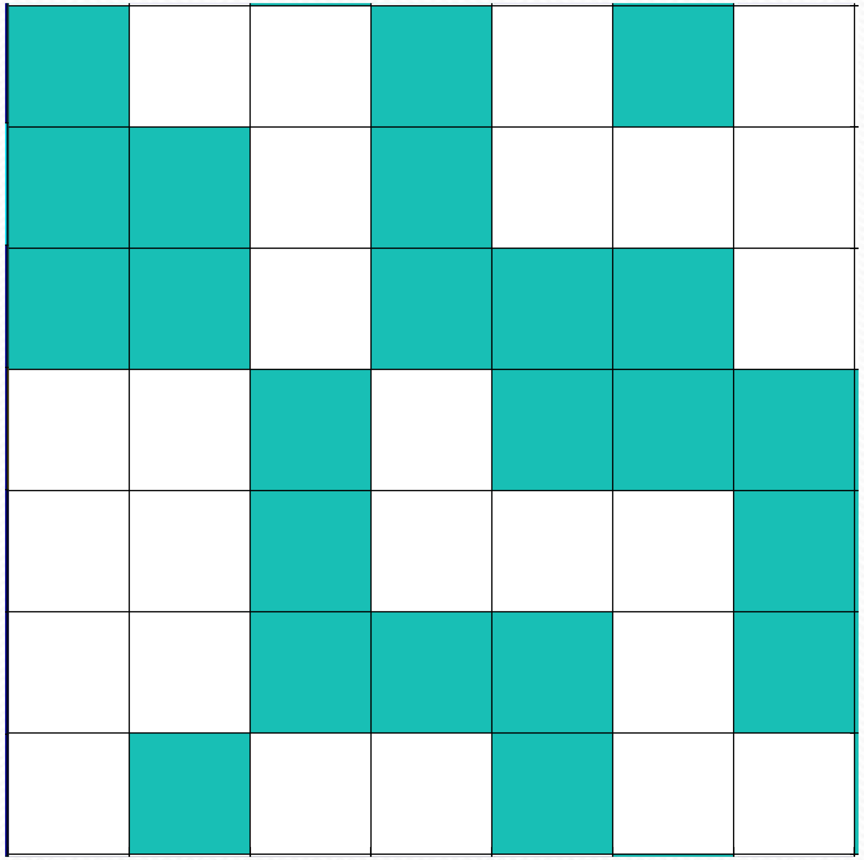
SISIM



T-PROGS



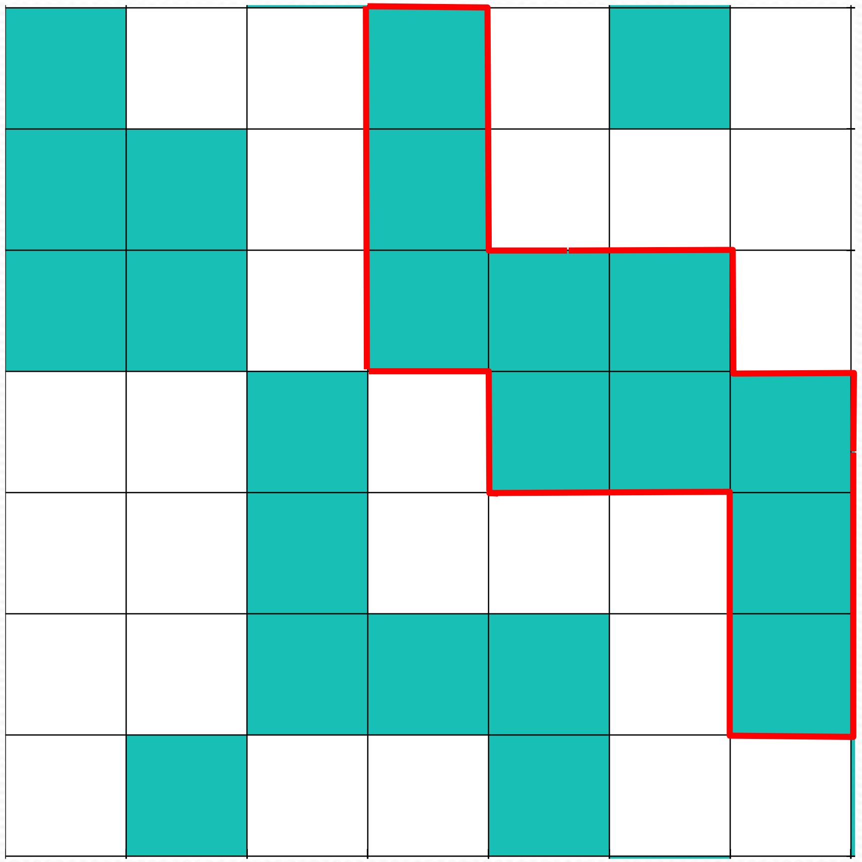
Analysis of facies connectivity



Connectivity indices

Number of clusters: $N_C = 5$

Analysis of facies connectivity

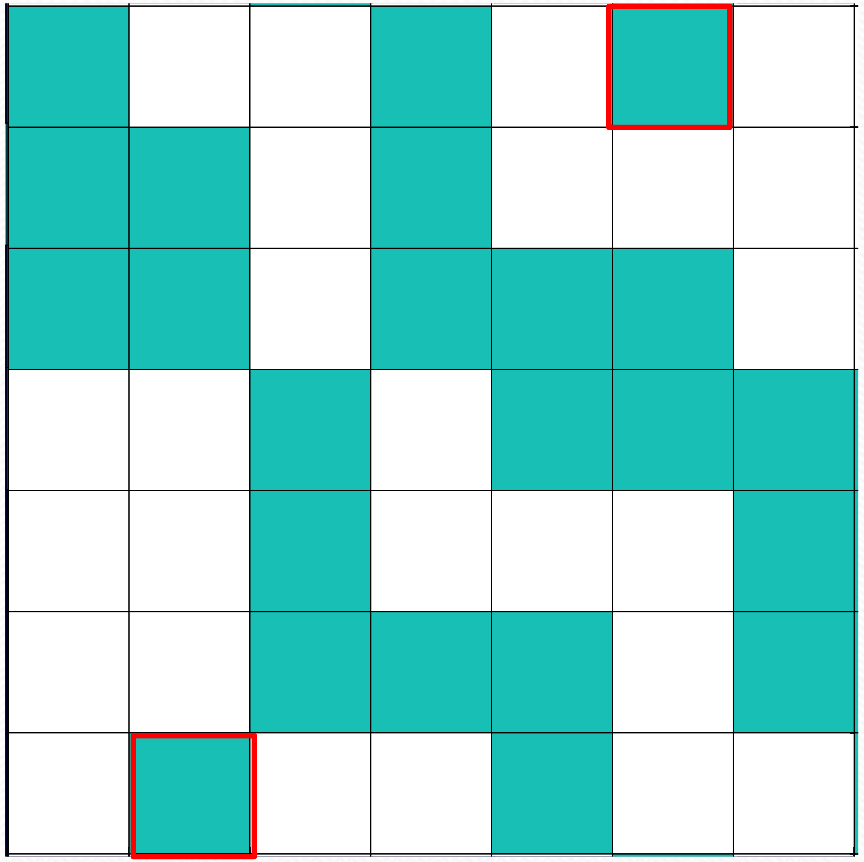


Connectivity indices

Number of clusters: $N_C = 5$

Max cluster size: $C_{max} = 10$

Analysis of facies connectivity



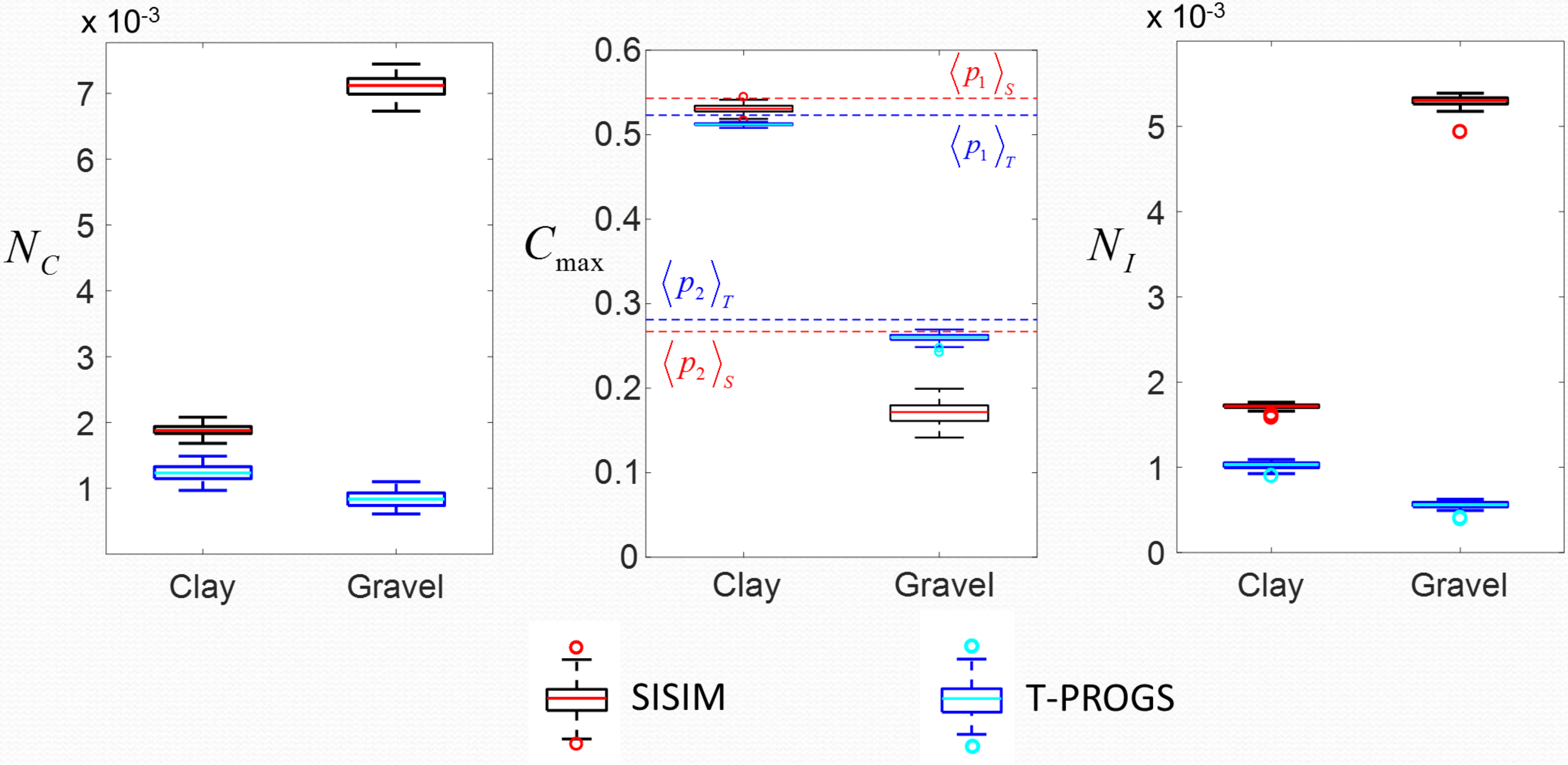
Connectivity indices

Number of clusters: $N_C = 5$

Max cluster size: $C_{max} = 10$

Number of isolated cells: $N_I = 2$

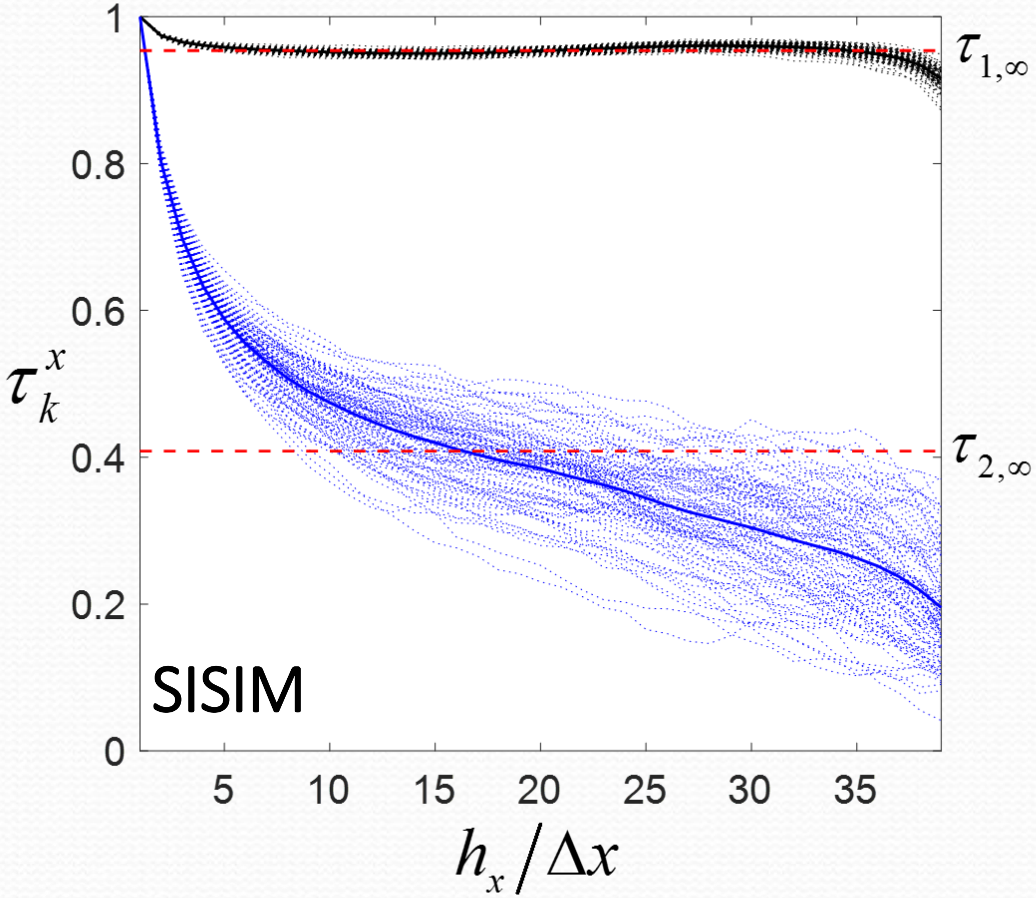
Analysis of facies connectivity

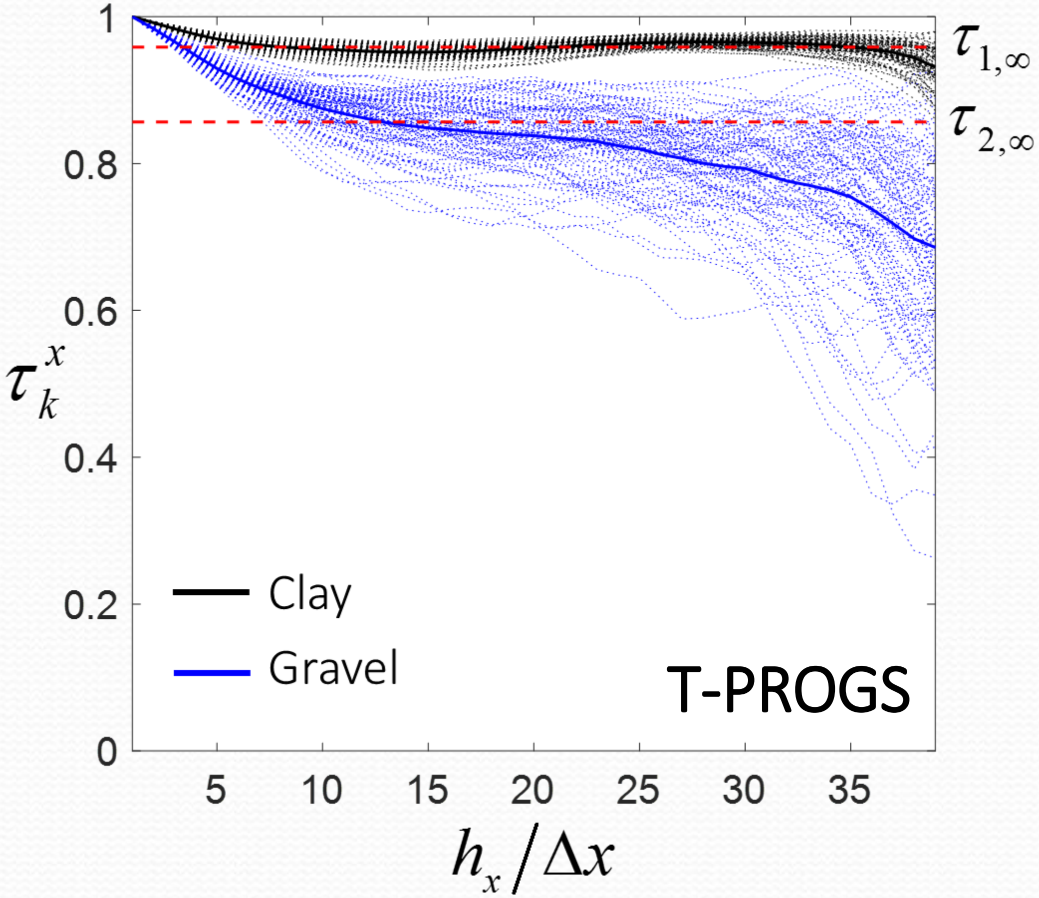


Analysis of facies connectivity

Connectivity function:
i = 1, …, 4 and j = x, y, z

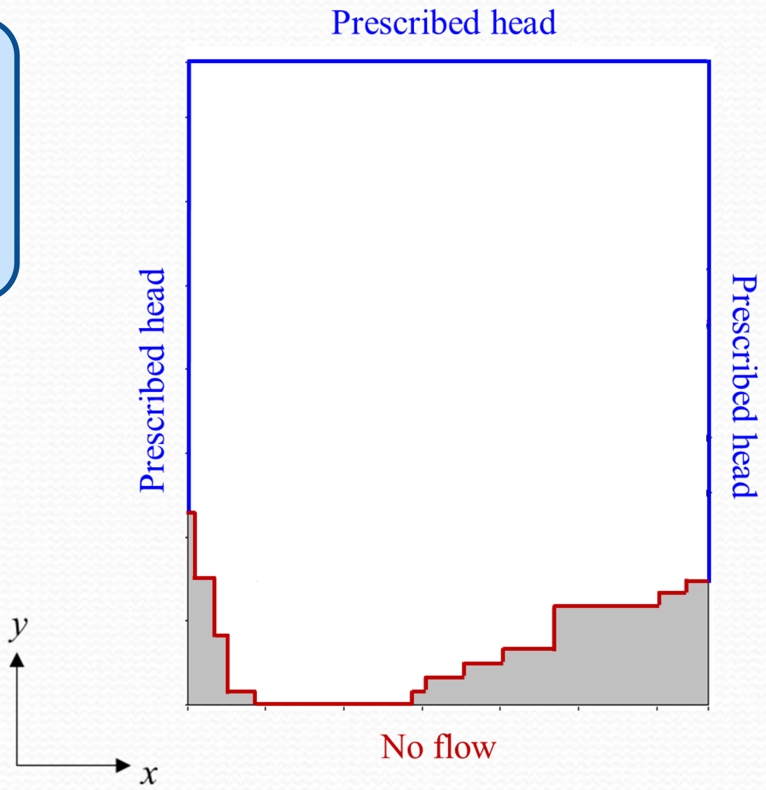
$$\tau_i^j(h) = \frac{N\left(\mathbf{x}_A \leftrightarrow \mathbf{x}_B \middle| \mathbf{x}_A \in \Omega_i, \mathbf{x}_B \in \Omega_i, \mathbf{x}_A - \mathbf{x}_B = h\mathbf{e}_j\right)}{N\left(\mathbf{x}_A \in \Omega_i, \mathbf{x}_B \in \Omega_i, \mathbf{x}_A - \mathbf{x}_B = h\mathbf{e}_j\right)}$$

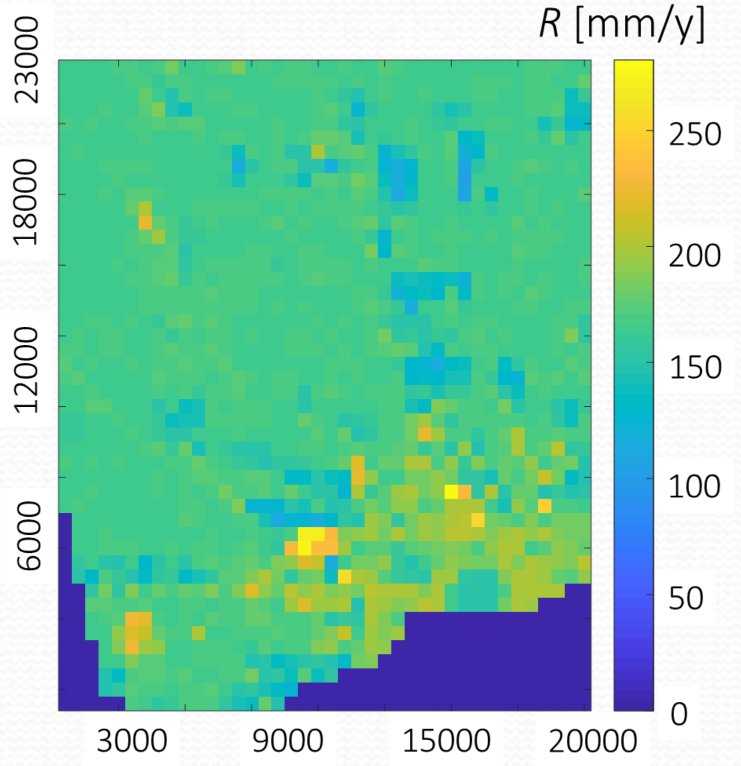


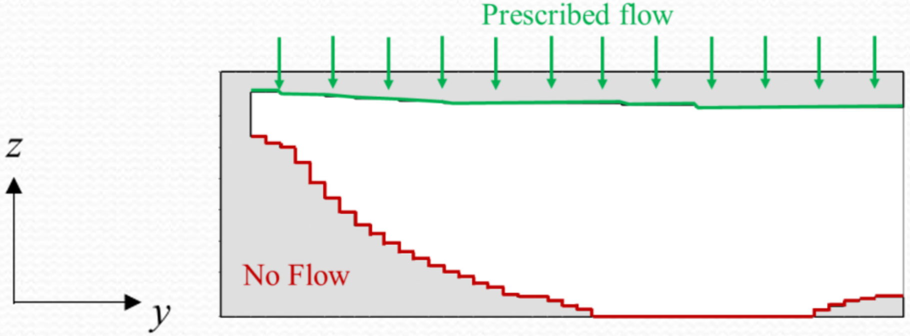


Groundwater flow model

Steady-state three-dimensional groundwater flow model (MODFLOW-2005).

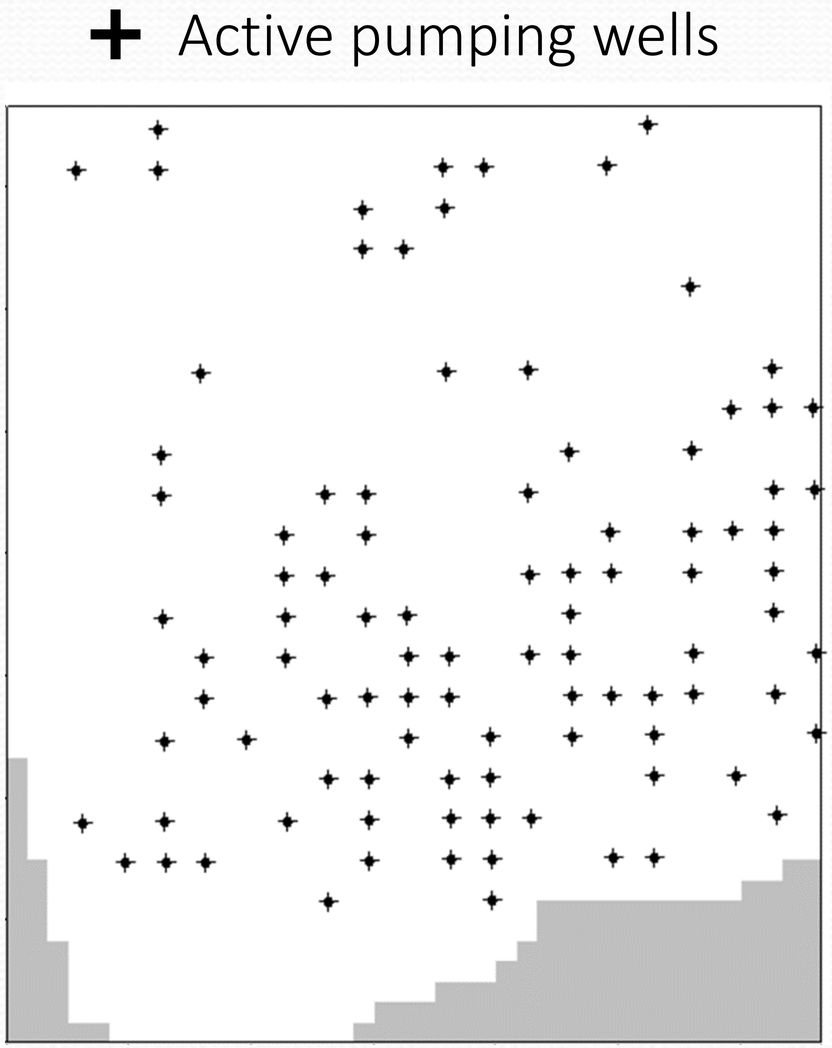




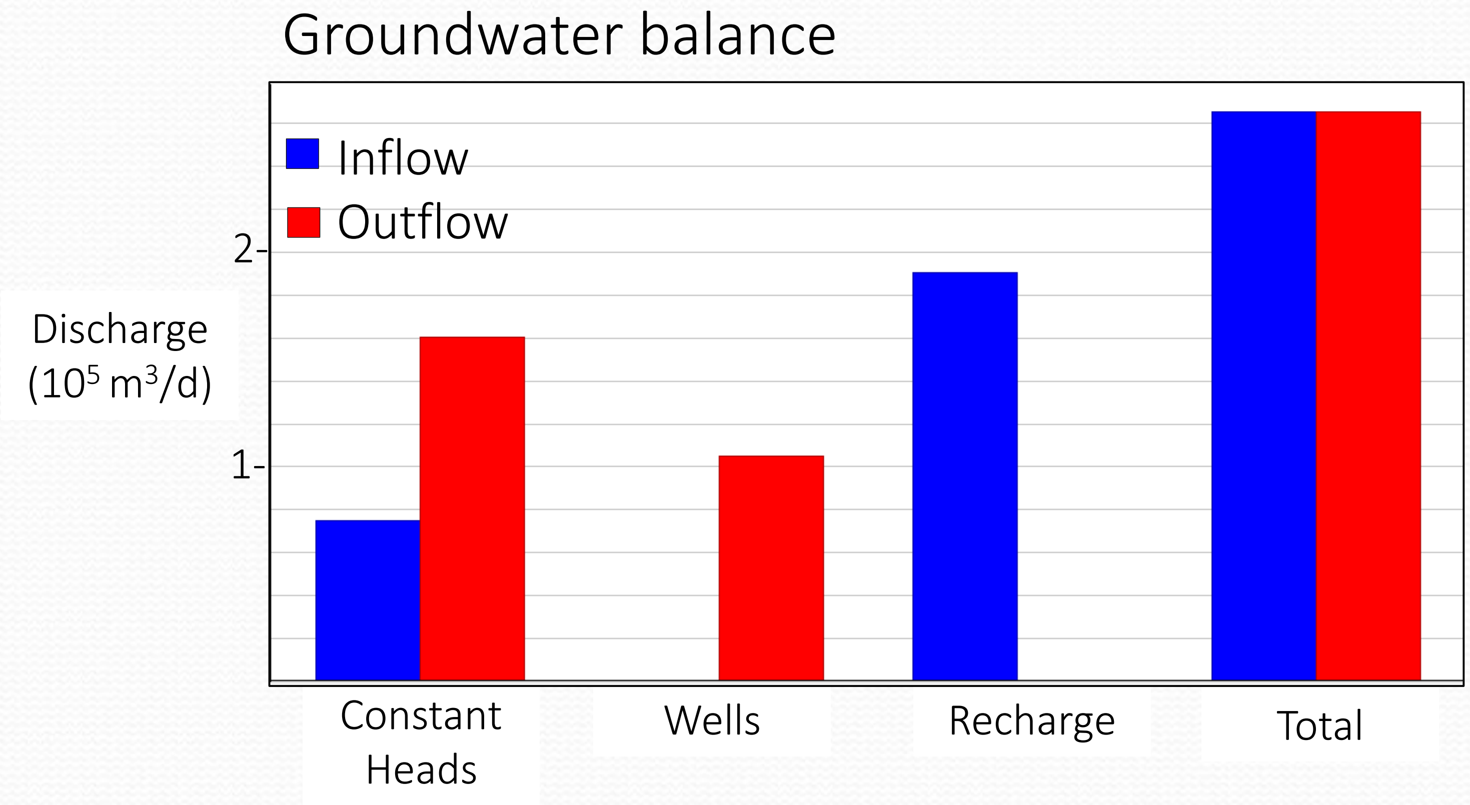


Natural recharge: land-use dependent (18% of the rainfall)

Groundwater flow model

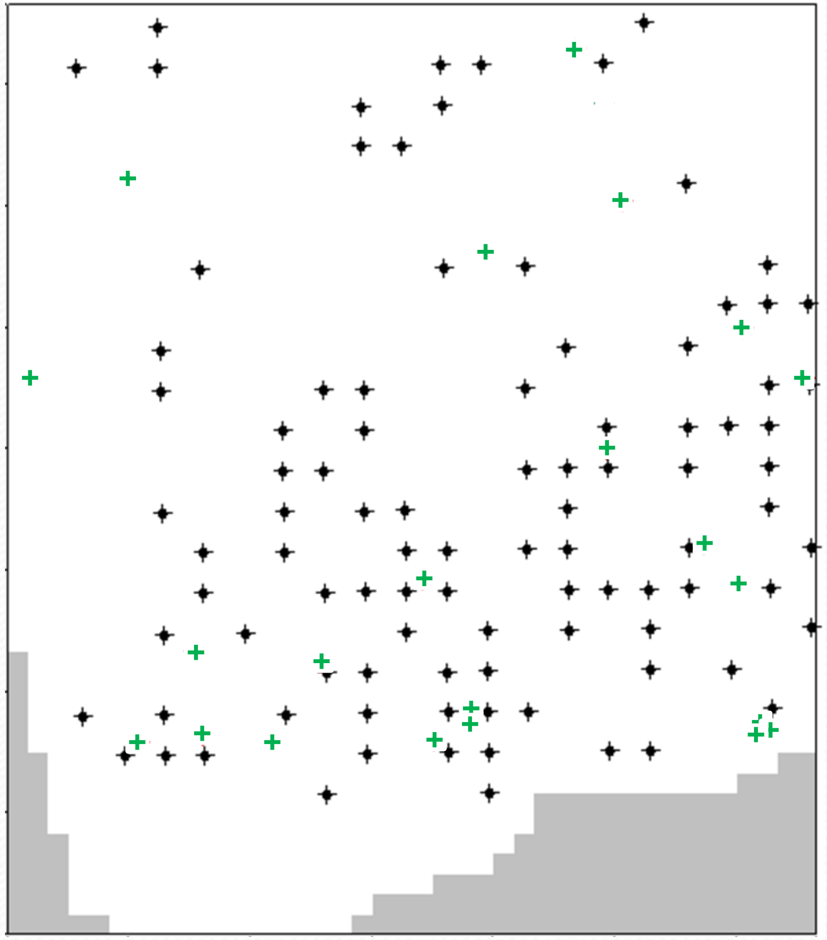


Groundwater withdrawal data of 2011: for civil (78%), industrial (18%) and agricultural (4%) purposes.



Groundwater flow model

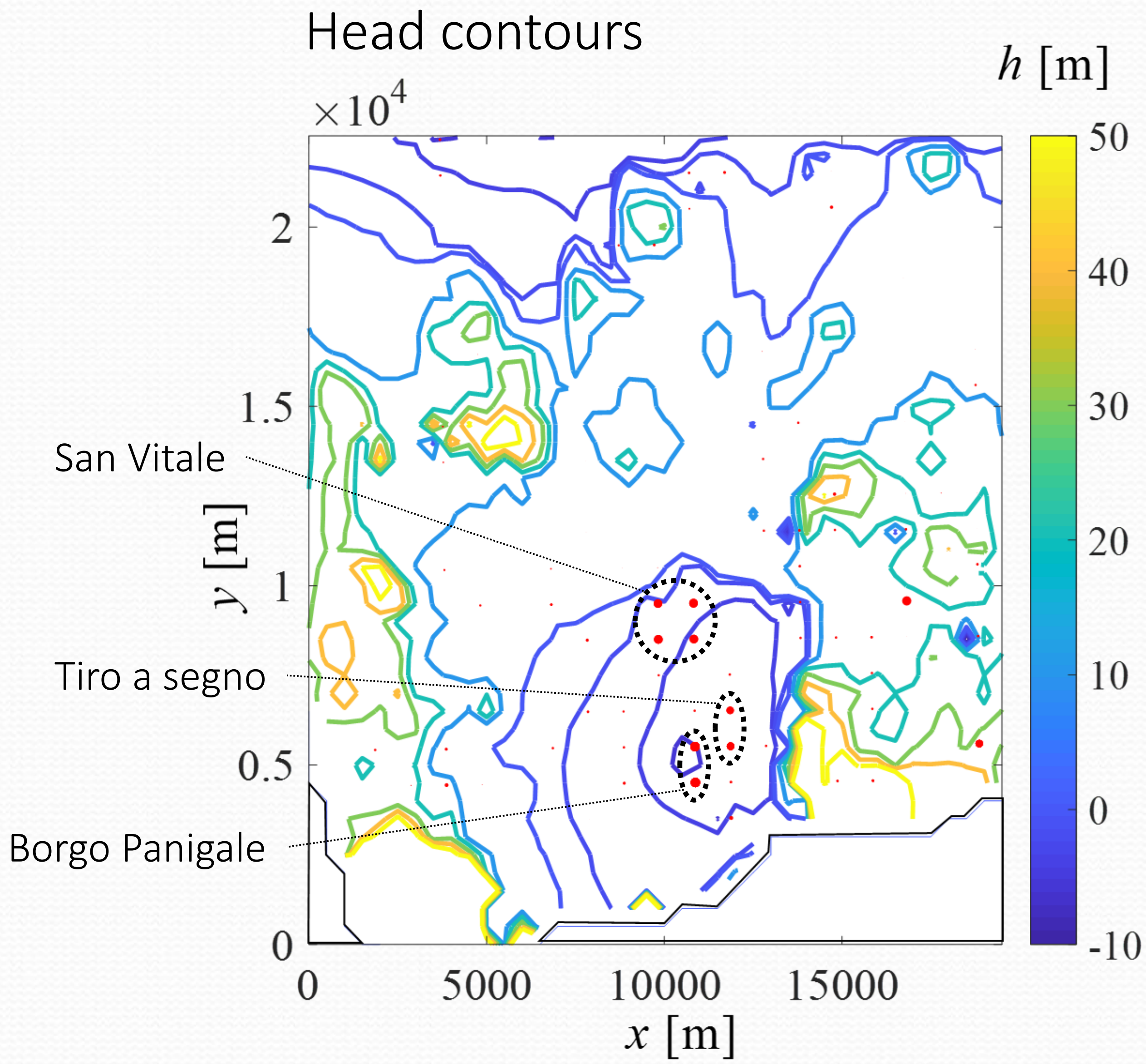
+ Head observations

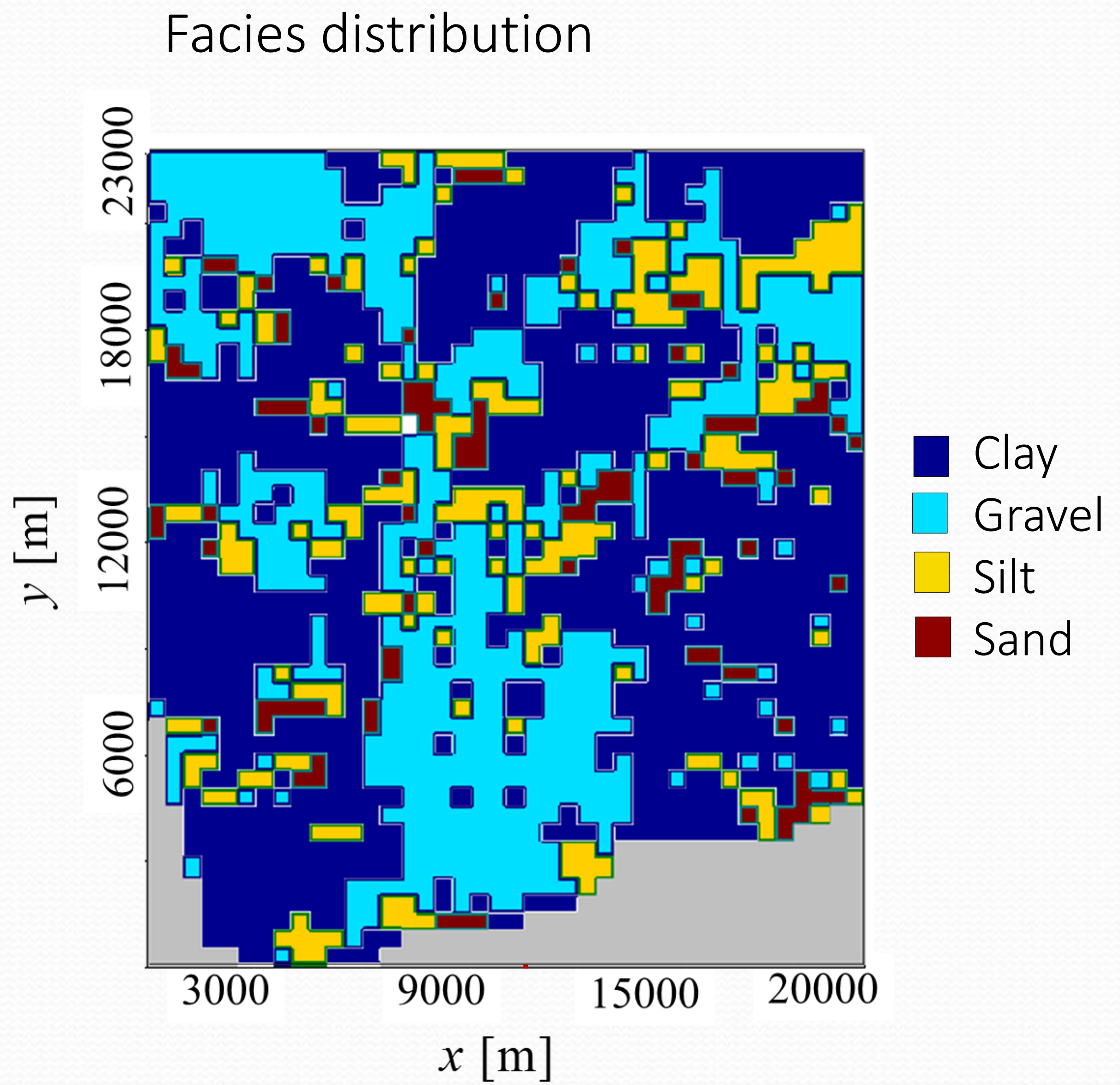


Hydraulic head data of 2011 used for the calibration of 2 hydraulic conductivity values

Parameter		Type	Value range
K_{CLAY}	[m/s]	*Adjustable*	$10^{-9} \div 10^{-6}$
K_{GRAVEL}	[m/s]	*Adjustable*	$10^{-4} \div 10^{-2}$
K_{SILT}	[m/s]	*Fixed*	10^{-6}
K_{SAND}	[m/s]	*Fixed*	10^{-5}

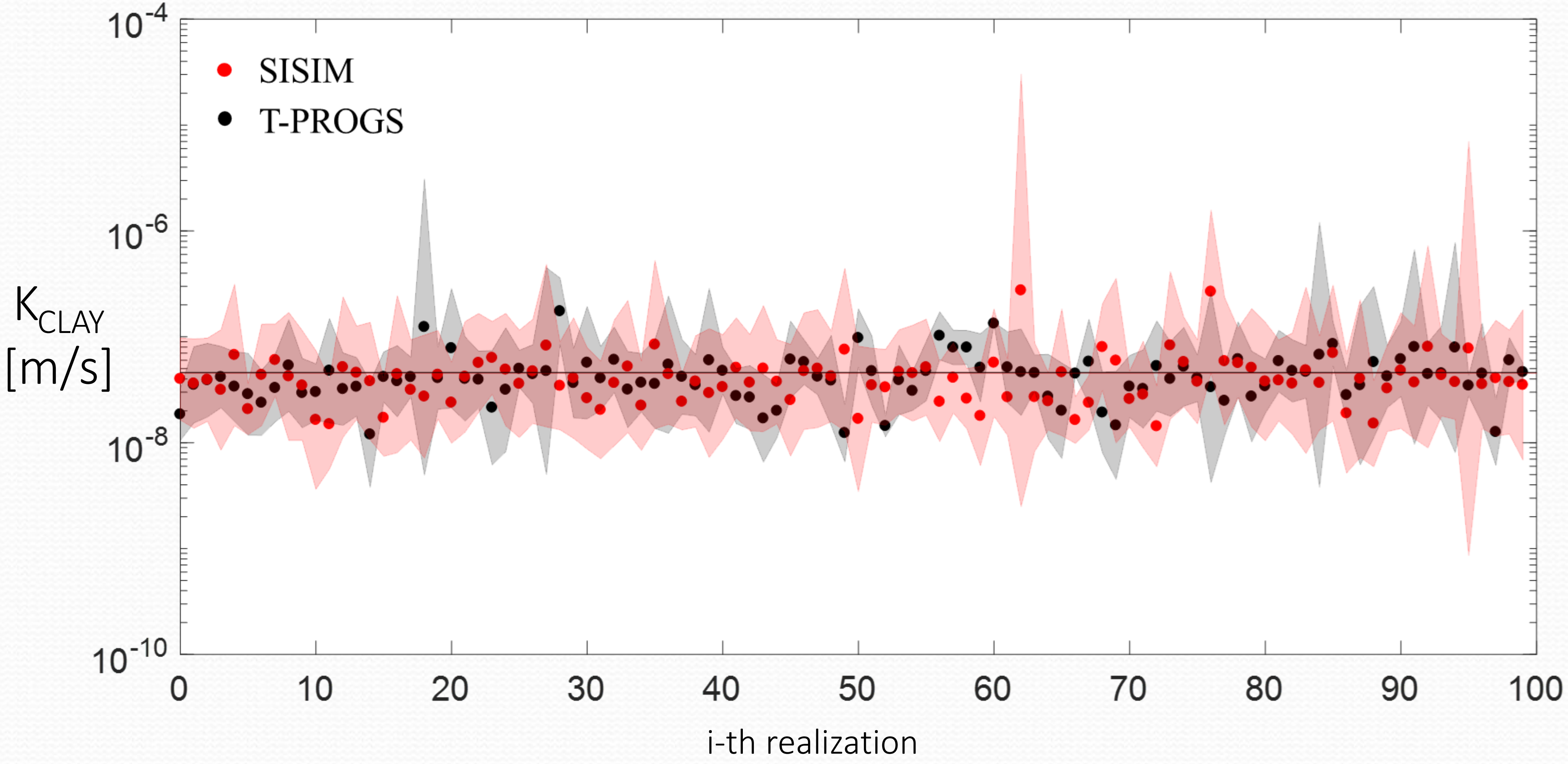
Groundwater flow model outputs





Model calibration results

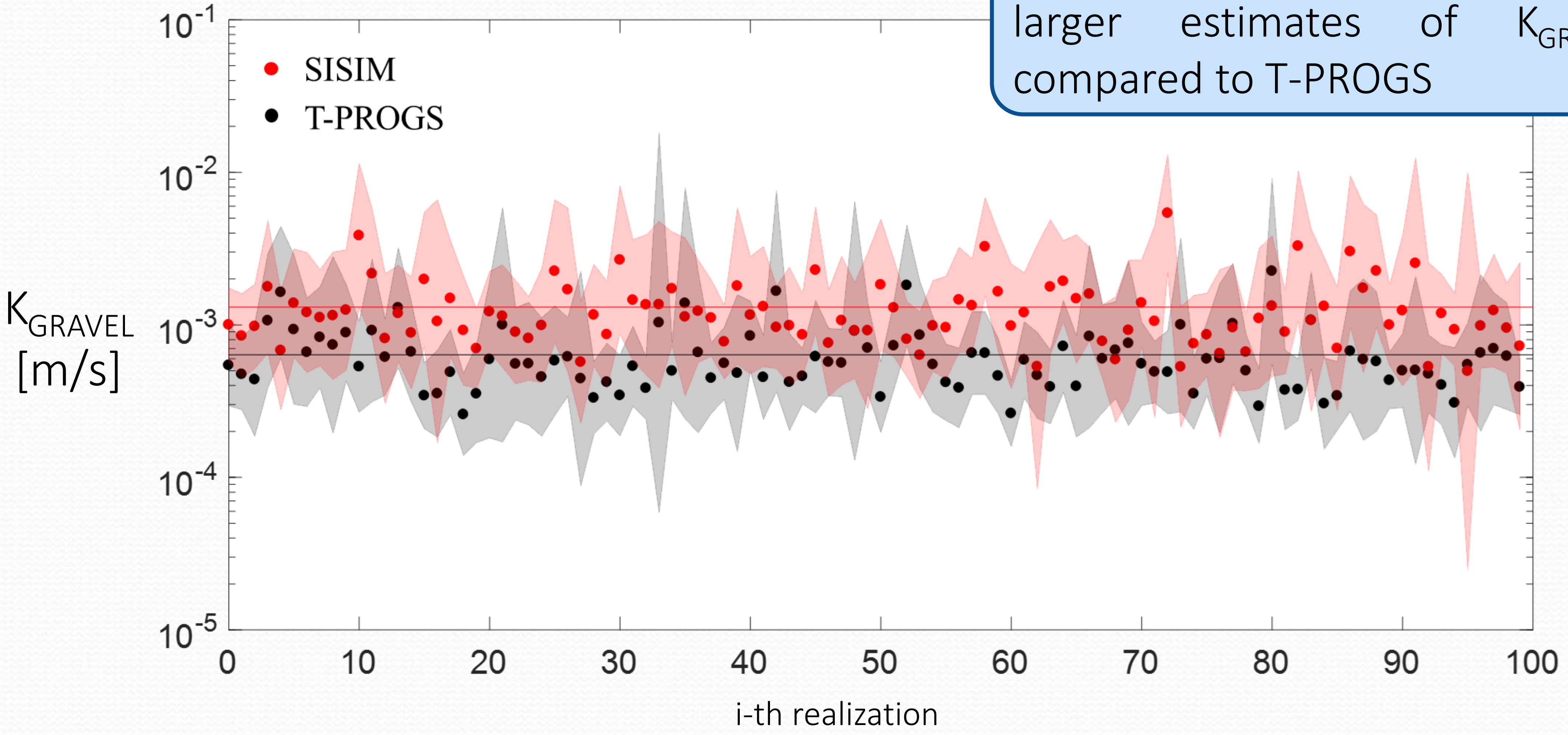
Maximum Likelihood approach: minimization of *NLL*



Model calibration results

Maximum Likelihood approach: minim

In SISIM, lower facies connectivity is compensated by larger estimates of K_{GRAVEL} compared to T-PROGS



Maximum-Likelihood Bayesian Model Averaging

Maximum Likelihood approach: minimization of *NLL*

⇩

Diverse realizations = Competing models of the aquifer

⇩

1. Rank realizations based on model identification criteria:

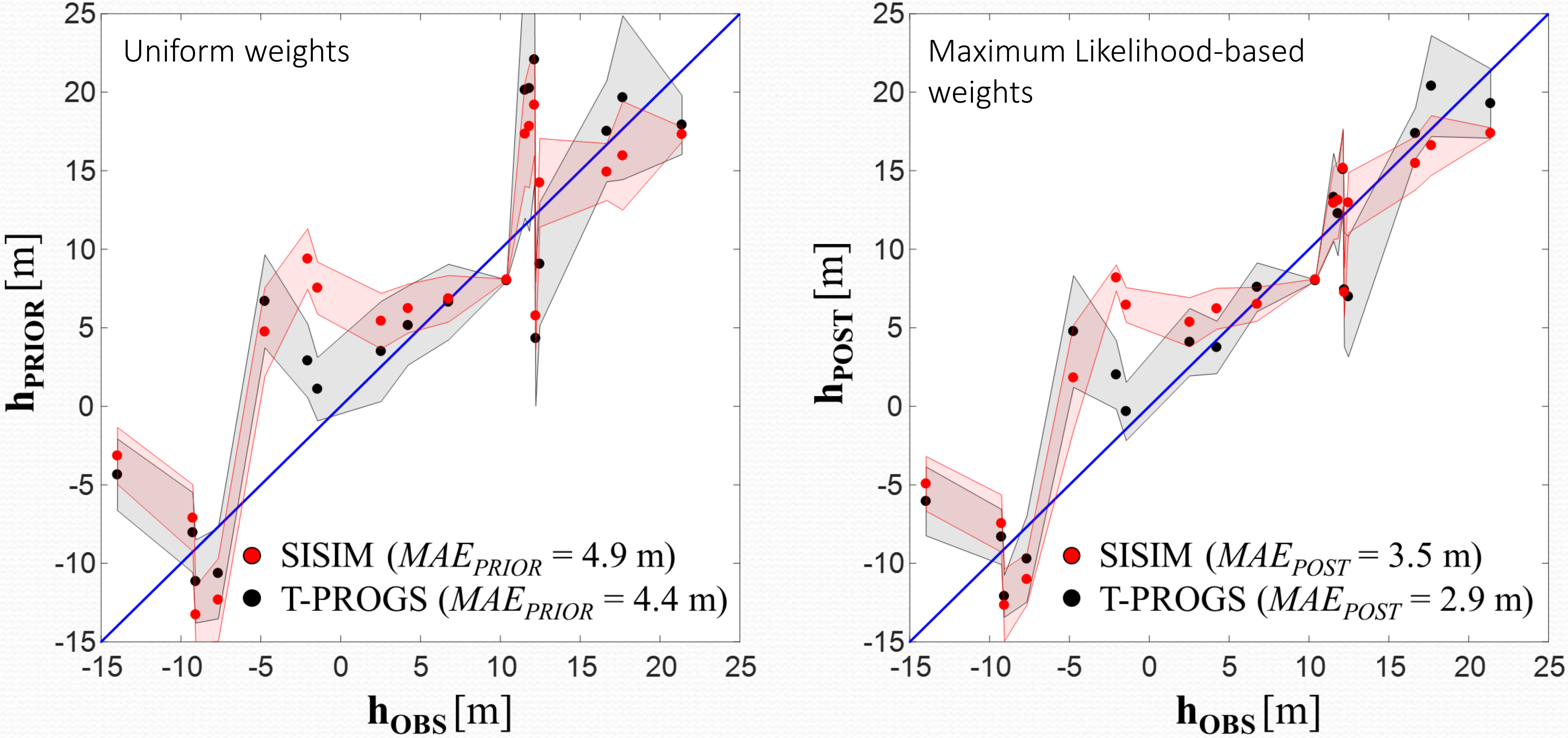
$$\mathrm{KIC} = \mathrm{NLL} - N_P \ln(2\pi) - \ln|\mathbf{Q}|$$

2. Assign to hydraulic head model predictions a ML-based weight:

$$\mathbf{h}_{\mathbf{POST}} = \sum_{i=1}^{n} \mathbf{h} \cdot p\left(M_i \middle| \mathbf{h}_{\mathbf{OBS}}\right) \quad \text{where} \quad p\left(M_i \middle| \mathbf{h}_{\mathbf{OBS}}\right) = \frac{\exp\left(-\frac{1}{2}\left(\mathrm{KIC}_i - \mathrm{KIC}_{\min}\right)\right) p(M_i)}{\sum_{\ell=1}^{n}\left[\exp\left(-\frac{1}{2}\left(\mathrm{KIC}_\ell - \mathrm{KIC}_{\min}\right)\right) p(M_\ell)\right]}$$

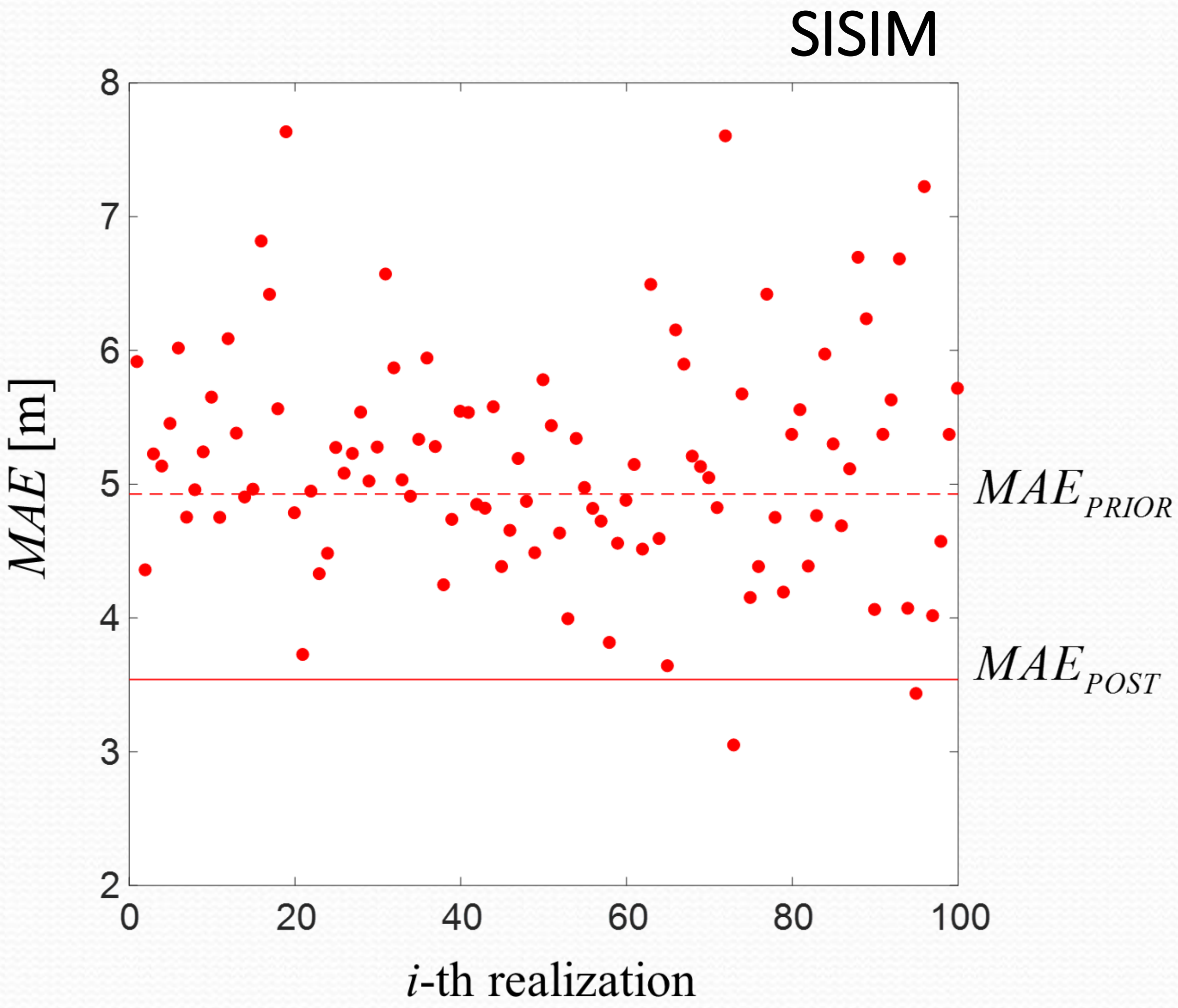
Maximum-Likelihood Bayesian Model Averaging

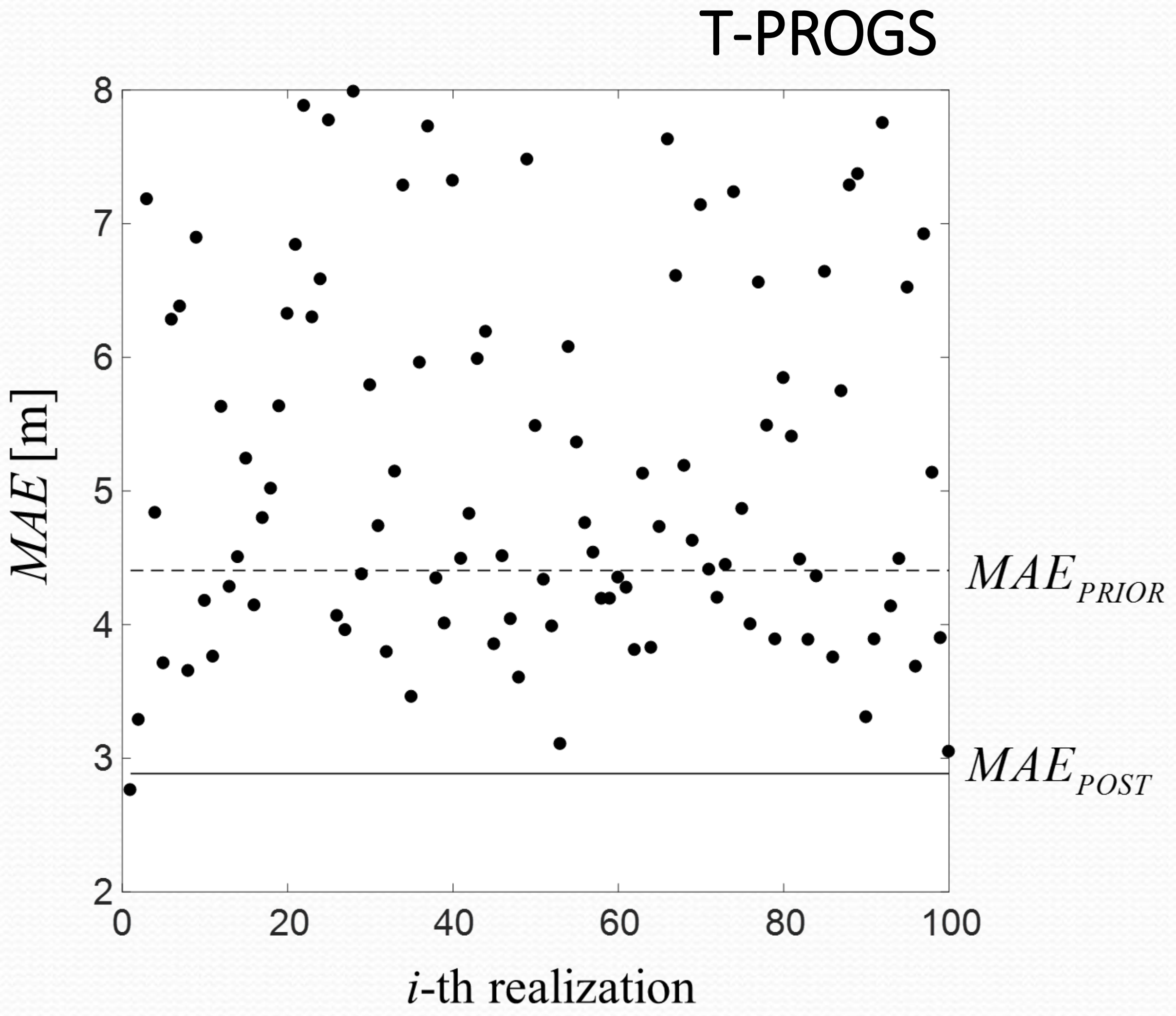
Model-averaged heads vs observed heads:



Maximum-Likelihood Bayesian Model Averaging

Averaged model performance:





Concluding remarks

Clay has similar degree of connectivity in the two ensembles and similar results in terms of calibrated conductivities

To compensate for the lower degree of connectivity, gravel hydraulic conductivity estimates obtained in the SISIM realizations are generally larger than their T-PROGS counterparts.

The best individual model (i.e., the realization minimizing KIC) as well as the average model obtained via MLBMA in the T-PROGS set are more skillful than their counterparts obtained for the SISIM set.