Inverse Modeling with the aid of Surrogate Models

Dongxiao Zhang, Qinzhuo Liao, Haibin Chang
College of Engineering
Peking (Beijing) University
dxz@pku.edu.cn

The 2017 EnKF Workshop

Inverse modeling

- **Estimate parameters from physical models and observations**

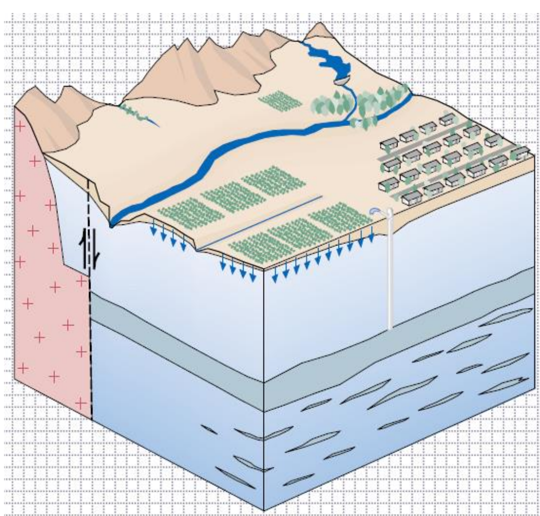
input parameter: θ ← model: f — output observation: s

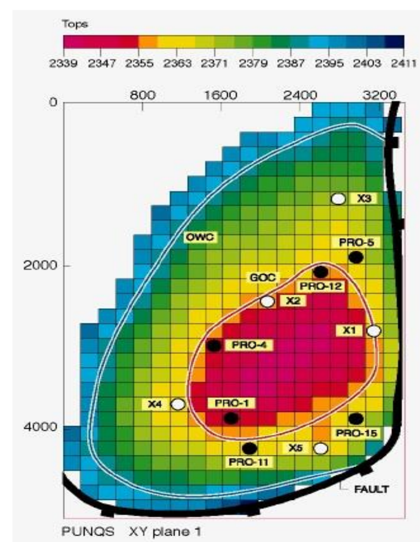
$$s = f(\theta) + e$$

Input
- conductivity
- porosity
- boundary condition
- source & sink

Model
- groundwater supply
- contaminant control
- oil and gas production
- CO_2 sequestration

Output
- hydraulic head
- velocity/flux
- phase saturation
- solute concentration





Stochastic approach

- **Bayesian inference**

$$s = f(\theta) + e$$

$$p(\theta \mid s) = \frac{p(\theta)\, p(s \mid \theta)}{p(s)}$$

$p(\theta)$: prior density

$p(\theta \mid s)$: posterior density

$p(s \mid \theta)$: likelihood function

$p(s)$: normalization factor

- **Markov chain Monte Carlo**
  - Use Monte Carlo simulations to construct a Markov chain
  - Computationally expensive: repeated evaluations of the forward model
- **Surrogate model**
  - Can generate a large number of samples at low cost
  - Posterior error depends on forward solution error

Forward Stochastic Formulation

- SPDE:

$$L(u;\boldsymbol{\xi},x)=g(\boldsymbol{\xi},x),\quad \boldsymbol{\xi}\in P, x\in D,$$

$$\text{where } \boldsymbol{\xi}=(\xi_1,\xi_2,\ldots,\xi_N)^T$$

which has a finite (random) dimensionality.

- Weak form solution:

$$\int_P L(\hat{u};\theta,x)w(\theta)p(\theta)d\theta=\int_P g(\theta,x)w(\theta)p(\theta)d\theta$$

where

$$\hat{u}(\theta,x)\in V,\ \text{where } V-\text{trial function space}$$

$$w(\theta)\in W,\ \text{where } W-\text{test (weighting) function space}$$

$$p(\theta)-\ \text{probability density function of } \boldsymbol{\xi}(\theta)$$

Forward Stochastic Methods

- **Galerkin polynomial chaos expansion (PCE)** [e.g., Ghanem and Spanos, 1991]:

$$V = span\{\Psi_i(\xi)\}_{i=1}^{M}, \quad W = span\{\Psi_i(\xi)\}_{i=1}^{M}$$

where $\{\Psi_i(\xi)\}_{i=1}^{M}$ – orthogonal polynomials

- **Probabilistic collocation method (PCM)** [*Tatang et al.,* 1997; *Sarma et al*., 2005; *Li and Zhang*, 2007, 2009]:

$$V = span\{\Psi_i(\xi)\}_{i=1}^{M}, \quad W = span\{\delta(\xi - \boldsymbol{\theta}_i)\}_{i=1}^{M}$$

- **Stochastic collocation method (SCM)** [Mathelin et al., 2005; Xiu and Hesthaven, 2005; Chang and Zhang, 2009]:

$$V = span\{L_i(\boldsymbol{\theta})\}_{i=1}^{M}, \quad W = span\{\delta(\boldsymbol{\theta} - \boldsymbol{\theta}_i)\}_{i=1}^{M}$$

where $\{L_i(\boldsymbol{\theta})\}_{i=1}^{M}$ – lagrange interpolation basis

Key Components for Stochastic Methods

- **Random dimensionality** of underlying stochastic fields
  - How to effectively approximate the input random fields with finite dimensions
  - Karhunen-Loeve and other expansions may be used
- **Trial function space**
  - How to approximate the dependent random fields
  - Perturbation series, polynomial chaos expansion, or Lagrange interpolation basis
- **Test (weighting) function space**
  - How to evaluate the integration in random space?
  - Intrusive or non-intrusive schemes?

Stochastic collocation method （SCM）

- **Based on polynomial interpolations in random space**

$\Theta_N = \{\xi_i\}_{i=1}^{M}$ is a set of M nodes in N-dimensional random space

Xiu and Hesthaven (2005)

$$\mathrm{I}(f)(\xi) = \sum_{i=1}^{M} f(\xi_i) L_i(\xi)$$

$$L_i(\xi) = \prod_{\substack{j=1 \\ j\neq i}}^{M} \frac{\xi - \xi_j}{\xi_i - \xi_j}, \quad L_i(\xi_j) = \delta_{ij}, 1 \le i, j \le M$$

- **Collocation points: Smolyak sparse grid algorithm**

Xiu and Hesthaven (2005)

$$\mathrm{A}(q,N) = \sum_{q-N+1 \le |i| \le q} (-1)^{q-|i|} \binom{N-1}{q-|i|} \cdot \left(U^{i_1} \otimes \ldots \otimes U^{i_N}\right)$$

U is univariate interpolation

- **Converge fast in case of smooth functions**

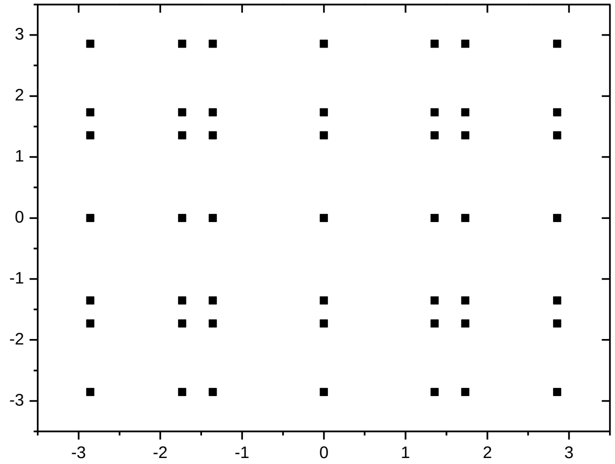
Chang & Zhang (2009)
Lin & Tartakovsky (2009)

Choices of Collocation Points

- **Tensor product of one-dimensional nodal sets**

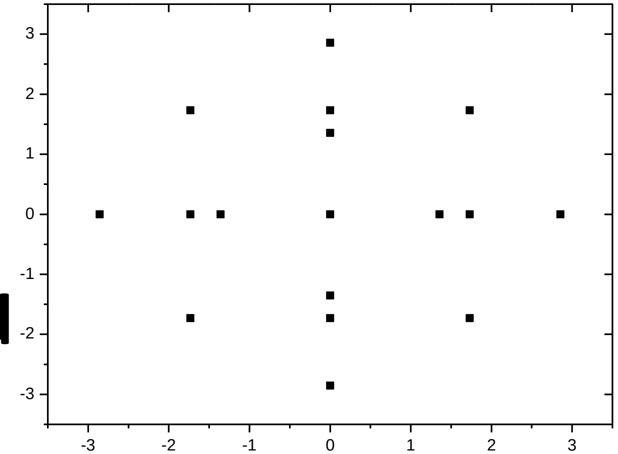
  Each dimension: m knots

  N dimension: $M = m^N$



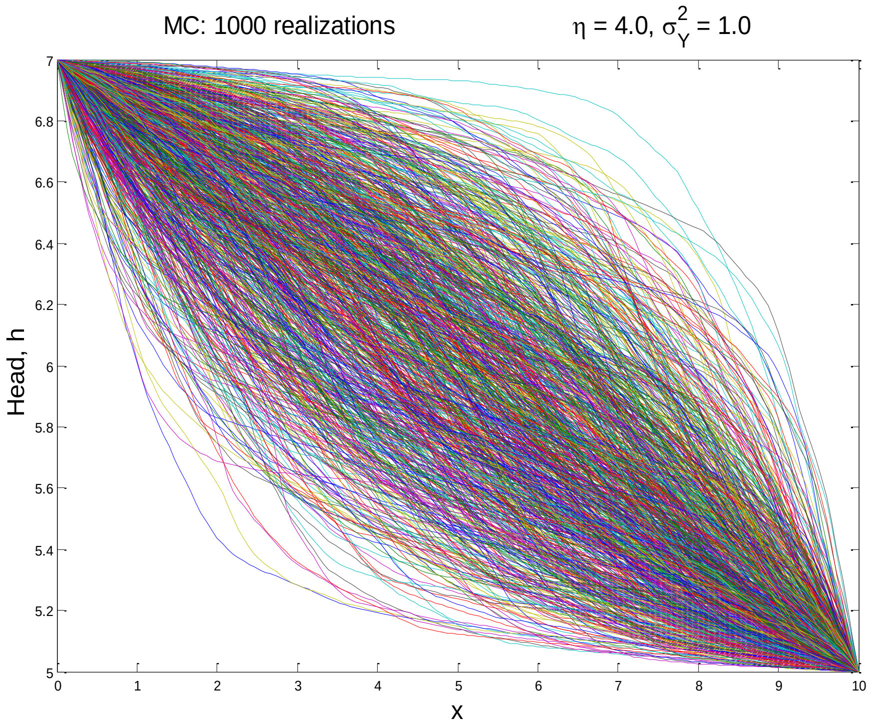
- **Smolyak sparse grid** (level: *k*=*q*-*N*)

  For N>1, preserving interpolation property of N=1 with a small number of knots



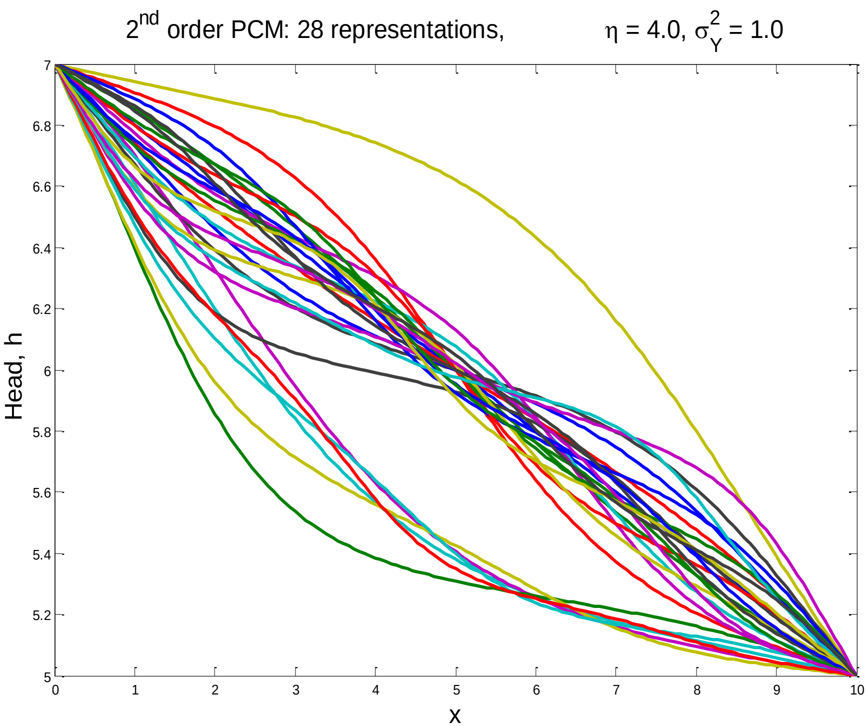
- **Tensor product vs. level-2 sparse grid**
  - *N*=2, 49 knots vs. 17 (shown right)
  - *N*=6, 117,649 knots vs. 97

MCS vs. PCM/SCM



MCS:

- Random sampling of (**realizations**)
- Equal weights for h_j (**realizations**)

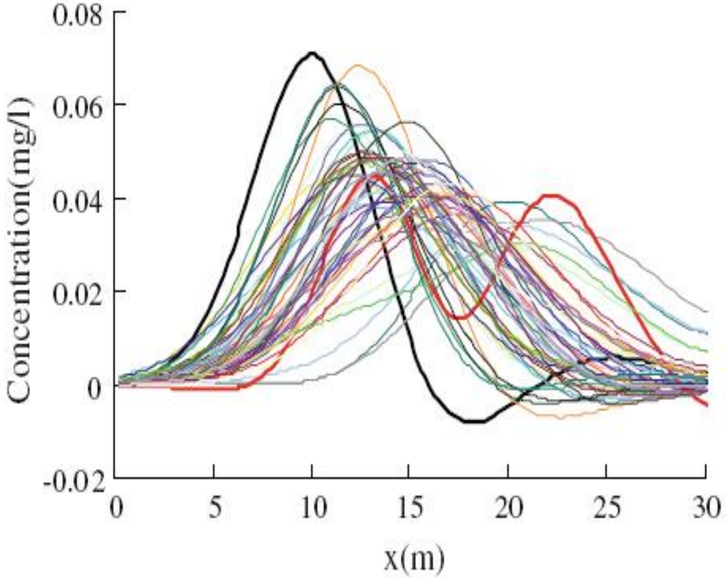


PCM/SCM:

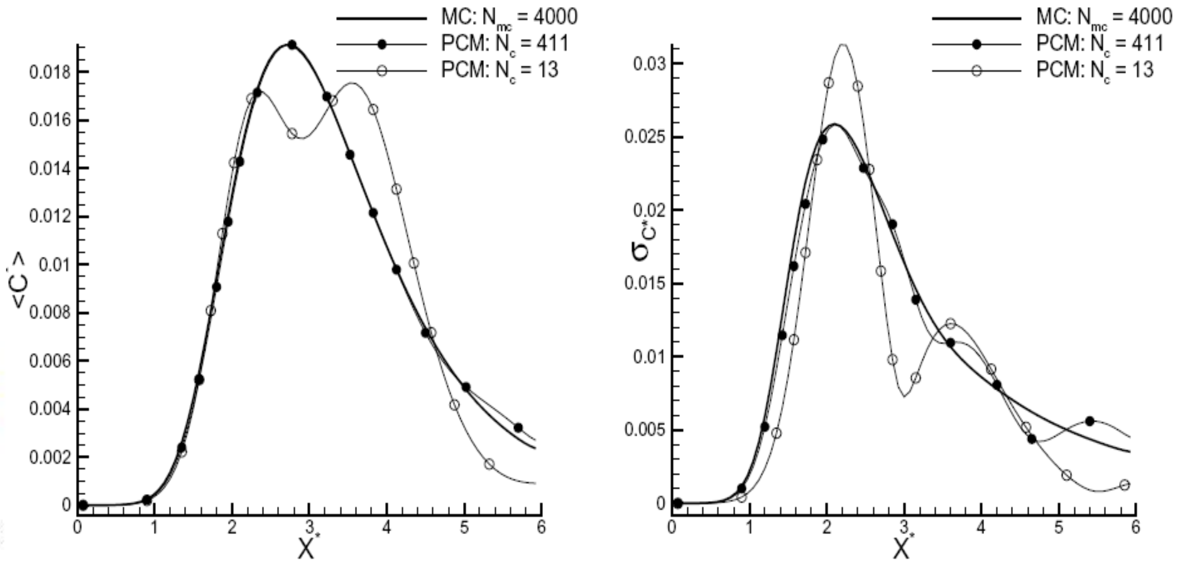
- Structured sampling (**collocation points**)
- Non-equal weights for h_j (**representations**)

Stochastic collocation method

- **Inaccurate results**
  - Non-physical realizations/Gibbs oscillation
  - Inaccurate statistical moments and probability density functions



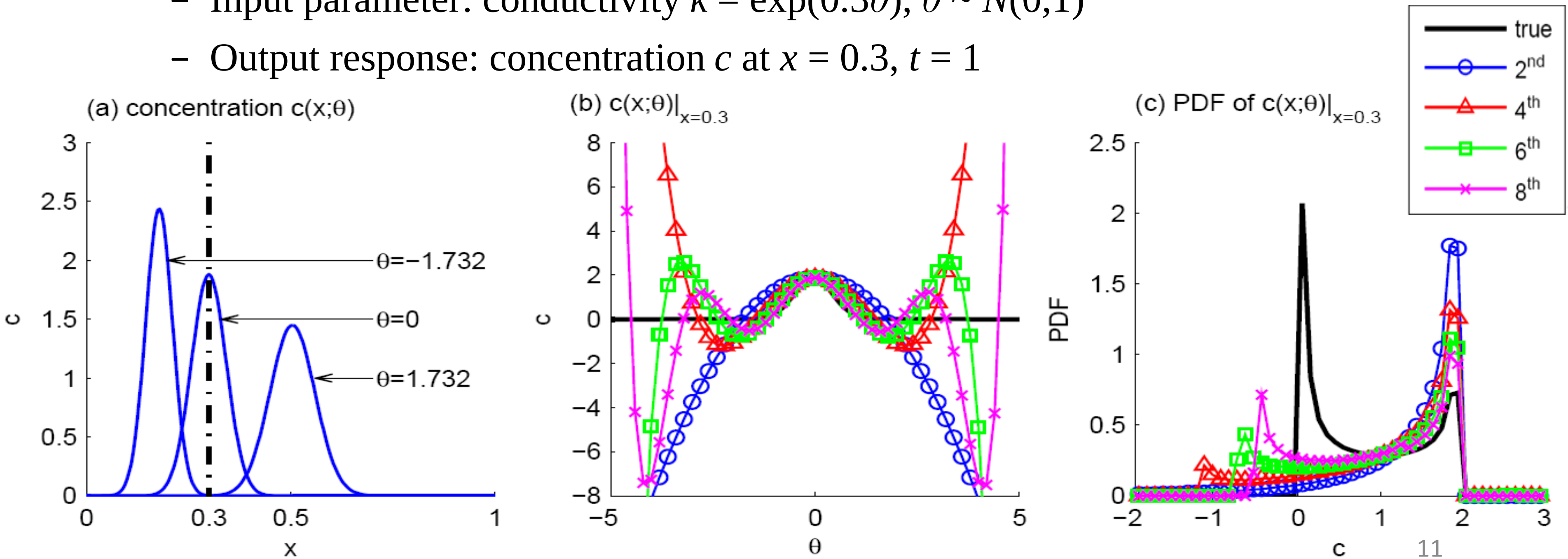
Zhang et al. (2010)



Lin & Tartakovsky (2009)

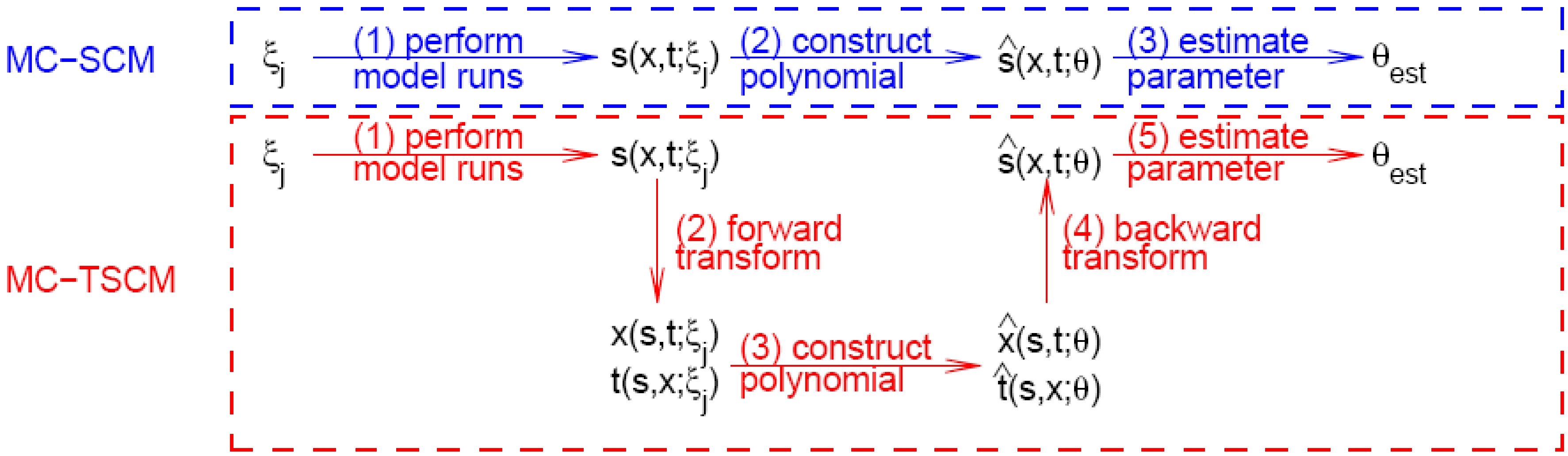
Stochastic collocation method

- **Inaccurate results**
  - When: advection dominated (Pe = 100) → low regularity
  - Why: physical space → random space
- **Illustration**
  - Unit mass instantaneously released at $x = 0$, $t = 0$
  - Input parameter: conductivity $k = \exp(0.3\theta)$, $\theta \sim N(0,1)$
  - Output response: concentration c at $x = 0.3$, $t = 1$



Transformed stochastic collocation method

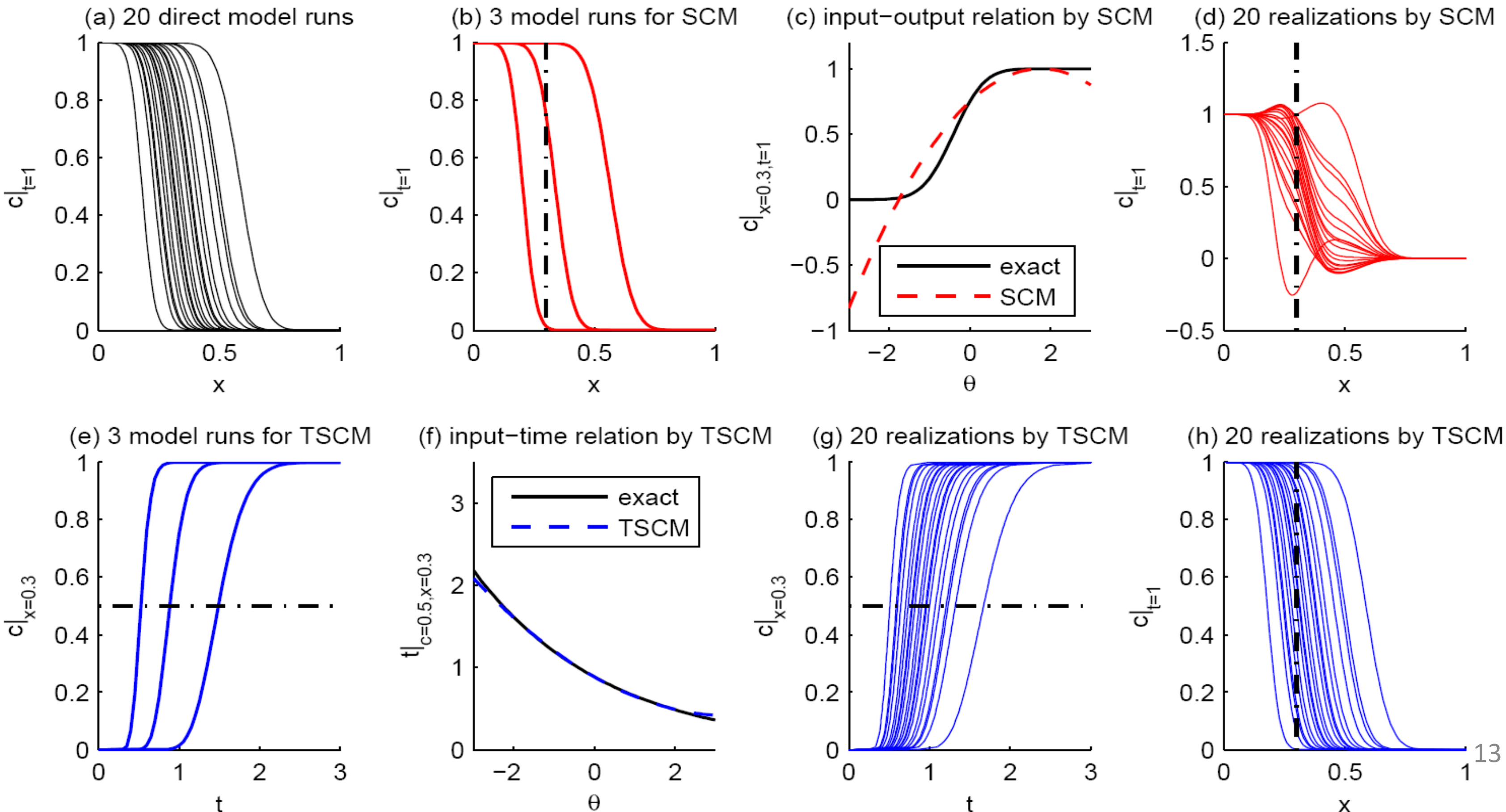
- **Stochastic collocation method (SCM)**
  - Approximate s as a function of θ at fixed x and t $\quad s(\mathbf{x}, t; \theta)$
- **Transformed stochastic collocation method (TSCM)**
  - Approximate x as a function of θ for a given s at fixed t $\quad \mathbf{x}(s, t; \theta)$
  - Approximate t as a function of θ for a given s at fixed x $\quad t(\mathbf{x}, s; \theta)$



Liao & Zhang (WRR, 2016)

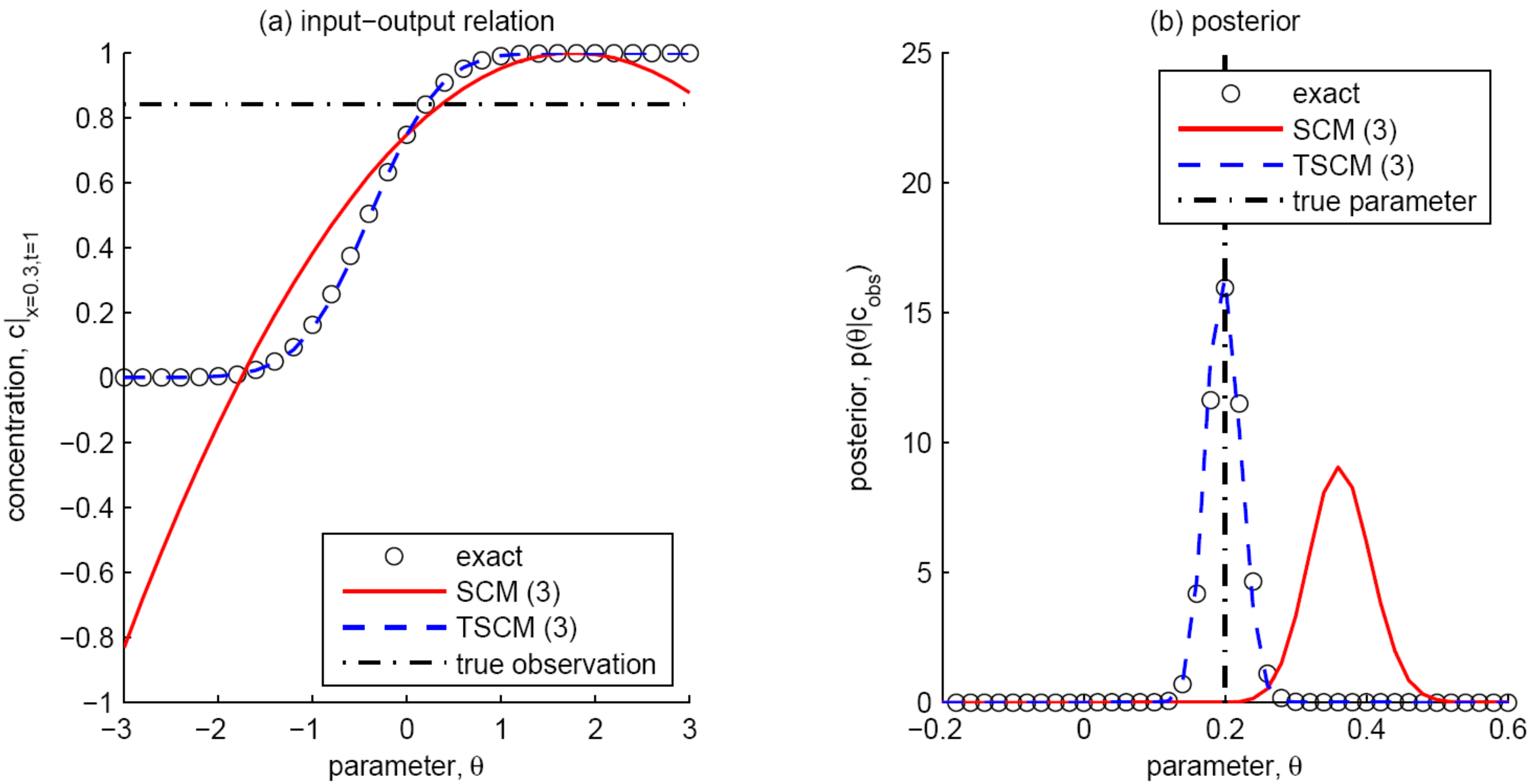
1D example

- **Continuous injection**
  - Input parameter: conductivity $k = \exp(0.3\theta)$, $\theta \sim N(0,1)$,
  - Output response: concentration c at $x = 0.3$, $t = 1$



1D example

- **Forward solution approximation and posterior approximation**



true observation c = 0.842
observation error $e \sim N(0, 0.01)$

true parameter $\theta = 0.2$

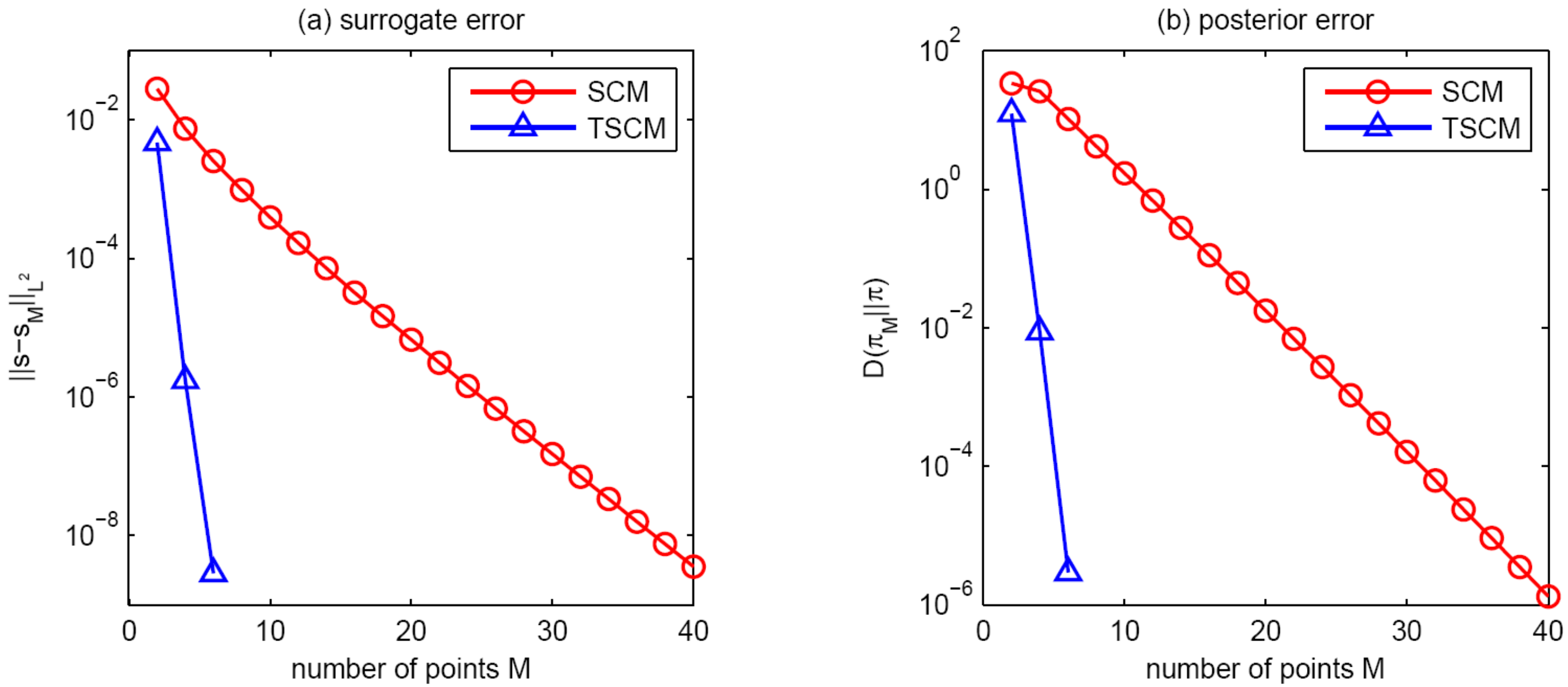
1D example

- **Convergence rate**

Marzouk & Xiu (2009)

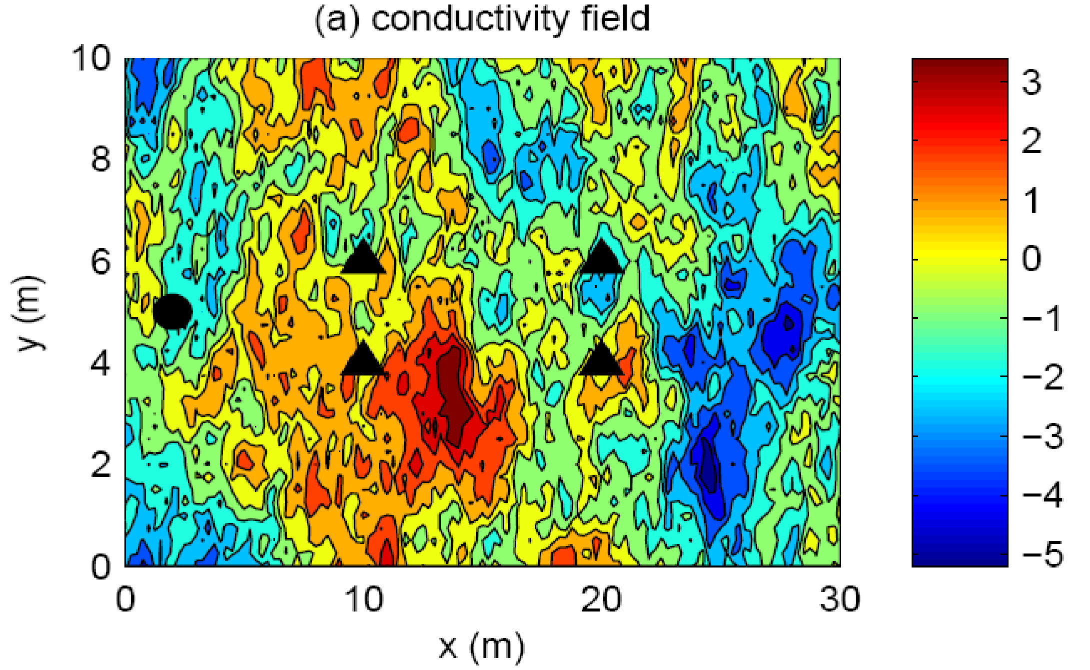
surrogate model: $s_M(\theta) = \sum_{i=1}^{M} f(\xi_i) L_i(\theta)$ error: $\left\| s - s_M \right\|_{L^2} = \int \left(s(\theta) - s_M(\theta) \right)^2 p(\theta) d\theta \to 0,\ M \to \infty$

posterior: $\pi(\theta) \equiv p(\theta \mid s)$ Kullback-Leibler divergence: $D\left(\pi_M \,\|\, \pi\right) \equiv \int \pi_M(\theta) \log \frac{\pi_M(\theta)}{\pi(\theta)} d\theta \to 0,\ M \to \infty$



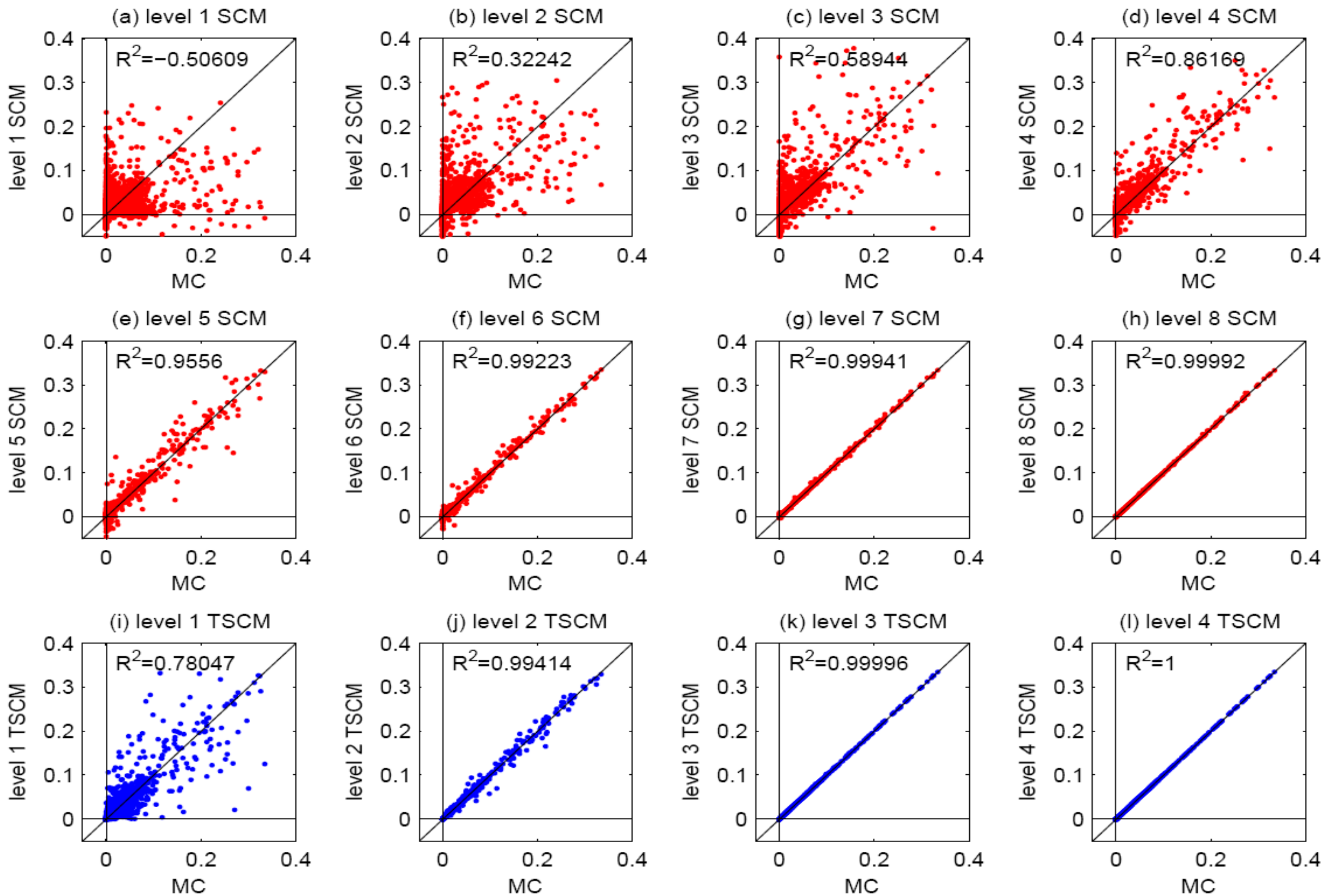
2D example

- Assume conductivity field is known
- Top and bottom are no-flow boundaries , right head h_2=0
- One instantaneous release location (circle), four observation wells (triangles)
- Input parameters: release time $t_0 \in [0,20]$, mass $m_0 \in [1,2]$, left head $h_1 \in [3,10]$
- Output responses: concentration c from t = 0 to 80, observation error $e \sim N(0, 0.001)$



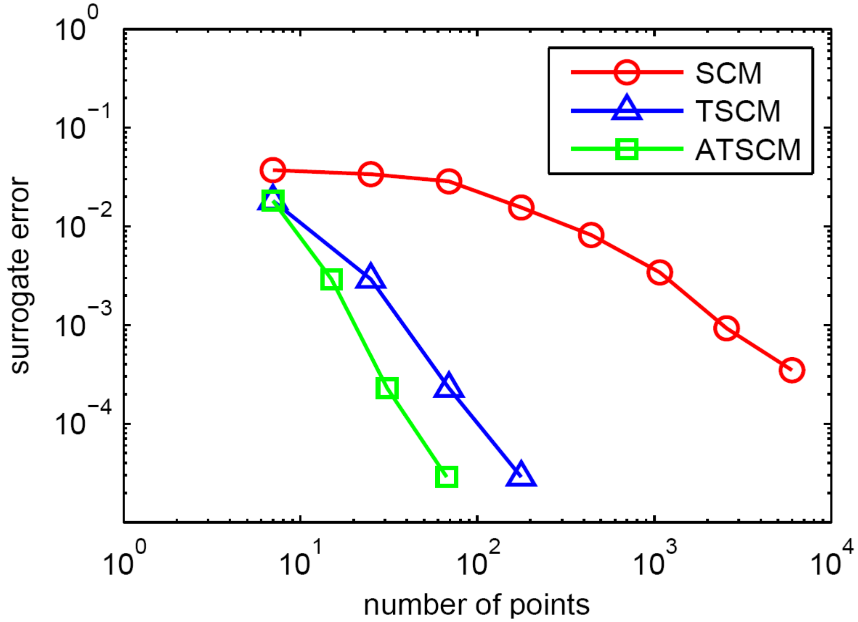
2D example

- **Compare true concentration and approximated concentration**



2D example

- **Surrogate approximation error**

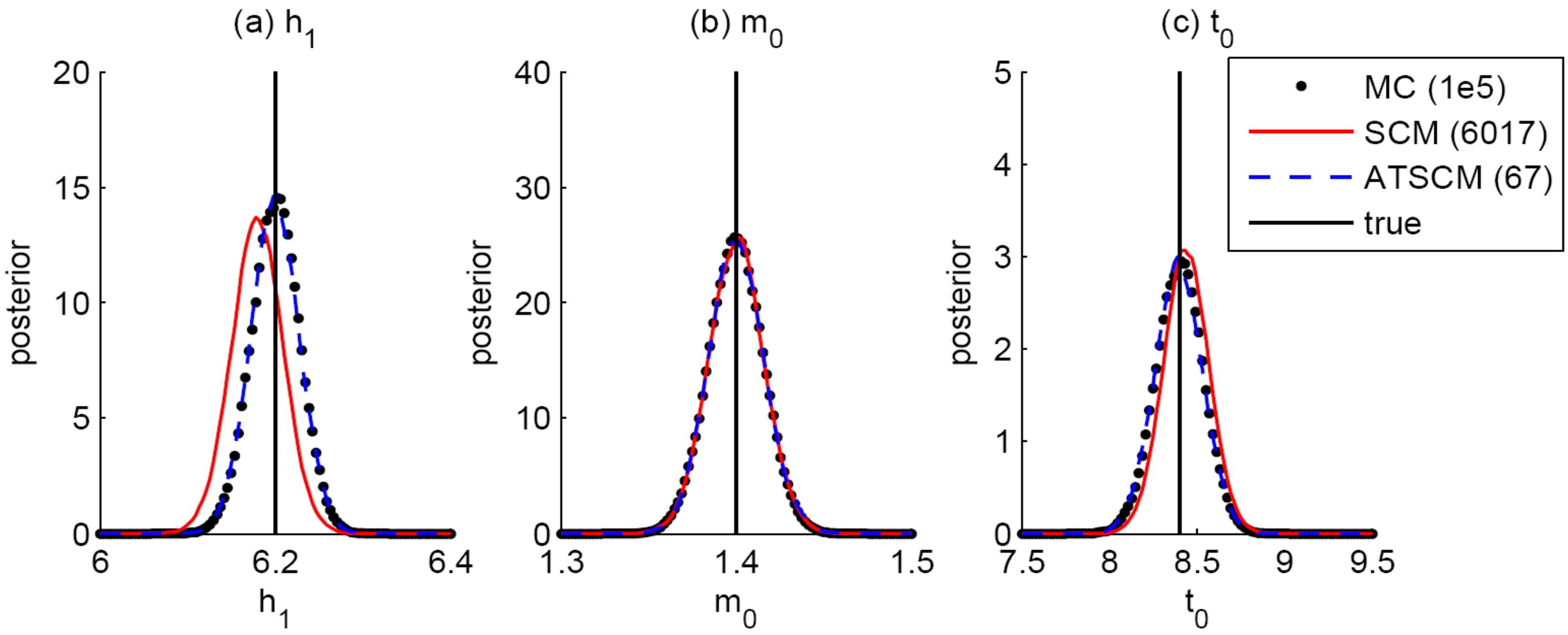


- **Adaptive transformed SCM (ATSCM)**
  - Dimension-adaptive: automatically select important dimensions
  - Further reduce the number of collocation points

Klimke (2006)
Liao et al. (JCP, 2016)

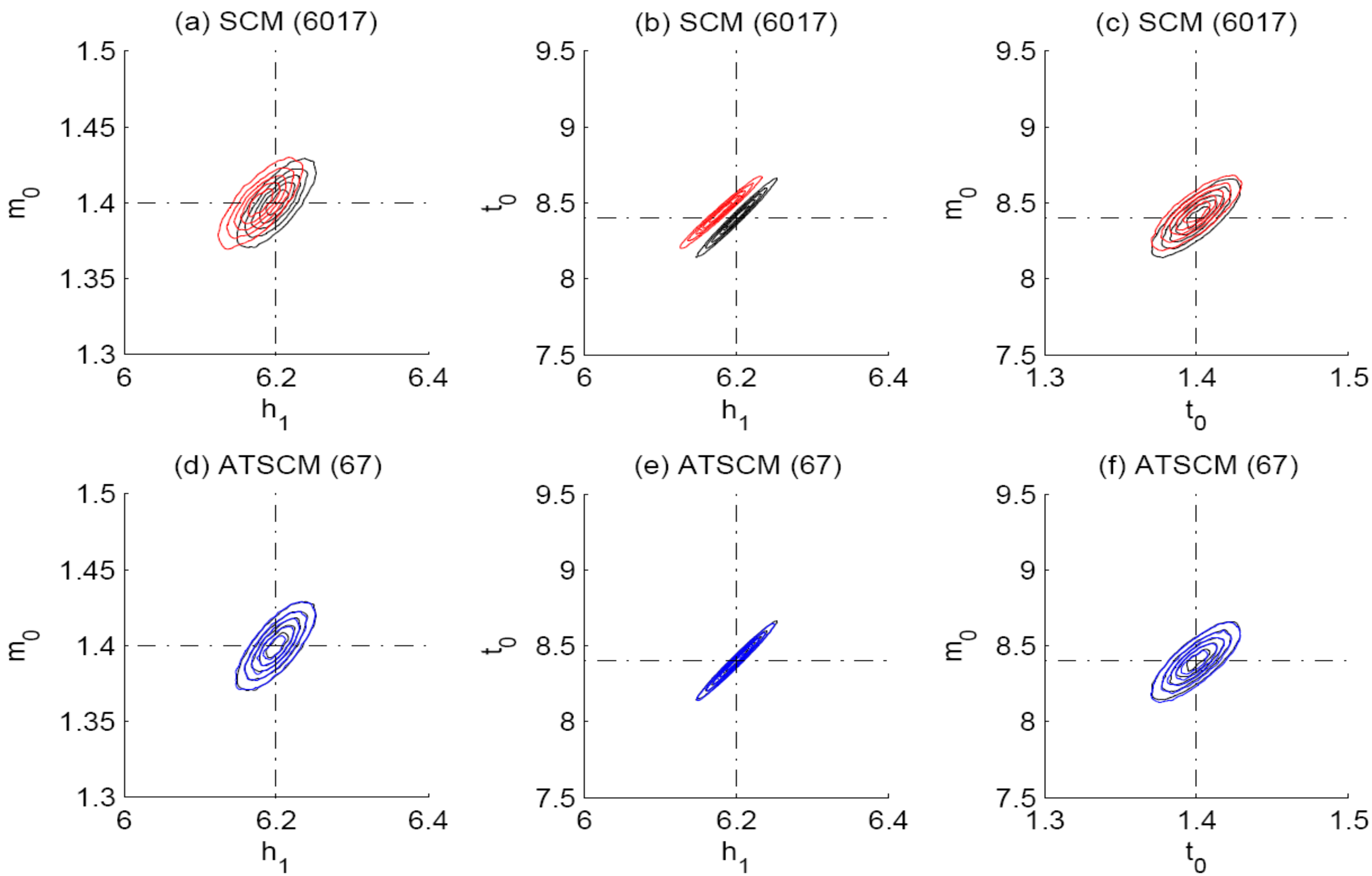
2D example

- **Marginal PDF**
  - MCMC with 10^5 model runs as a reference
  - ATSCM with 67 model runs is more accurate than SCM with 6017 model runs



2D example

- **Marginal PDF**
  - Black: MCMC, red: SCM, blue: ATSCM



Inverse modeling

- **Maximizing the posterior PDF**

$$p(\boldsymbol{m} \mid \boldsymbol{d}^{obs}) = \frac{p(\boldsymbol{m})\,p(\boldsymbol{d}^{obs} \mid \boldsymbol{m})}{p(\boldsymbol{d}^{obs})}$$

- **For Gaussian prior and error, minimizing an objective function**

$$J(\boldsymbol{m}) = \frac{1}{2}(g(\boldsymbol{m}) - \boldsymbol{d}^{obs})^T C_D^{-1}(g(\boldsymbol{m}) - \boldsymbol{d}^{obs}) + \frac{1}{2}(\boldsymbol{m} - \boldsymbol{m}^{pr})^T C_M^{-1}(\boldsymbol{m} - \boldsymbol{m}^{pr}).$$

- **Ensemble based methods**

EnKF
- Sequentially assimilate the data
- One step method
- Moderate simulation effort (restart required)

ES
- Simultaneously assimilate all the data
- One step method
- Small simulation effort (no restart)

Iterative ES
- Simultaneously assimilate all the data
- Multi-step method
- Large simulation effort (no restart, iteration)
- Suitable for highly non-linear problems

Iterative ensemble smoother

- **Gauss-Newton method**

$$\boldsymbol{m}_{l+1} = \boldsymbol{m}_l - \left[(1+\lambda_l)C_M^{-1} + G_l^T C_D^{-1} G_l\right]^{-1} \times \left[C_M^{-1}\left(\boldsymbol{m}_l - \boldsymbol{m}^{pr}\right) + G_l^T C_D^{-1}\left(g(\boldsymbol{m}_l) - \boldsymbol{d}^{obs}\right)\right]$$

Size: Nm x Nm

- **Equivalent form**

$$\boldsymbol{m}_{l+1} = \boldsymbol{m}_l - \frac{1}{1+\lambda_l}\left[C_M - C_M G_l^T \left((1+\lambda_l)C_D + G_l C_M G_l^T\right)^{-1} G_l C_M\right] C_M^{-1}\left(\boldsymbol{m}_l - \boldsymbol{m}^{pr}\right) - C_M G_l^T \left((1+\lambda_l)C_D + G_l C_M G_l^T\right)^{-1}\left(g(\boldsymbol{m}_l) - d^{obs}\right).$$

Size: Nd x Nd

Can be further approximated

- **Ensemble based implementation**

$$\boldsymbol{m}_{l+1,j} = \boldsymbol{m}_{l,j} - \frac{1}{1+\lambda_l}\left[C_M - C_M G_{l,j}^T \left((1+\lambda_l)C_D + G_{l,j} C_M G_{l,j}^T\right)^{-1} G_{l,j} C_M\right] C_M^{-1}\left(\boldsymbol{m}_{l,j} - \boldsymbol{m}_j^{pr}\right) - C_M G_{l,j}^T \left((1+\lambda_l)C_D + G_{l,j} C_M G_{l,j}^T\right)^{-1}\left(g(\boldsymbol{m}_{l,j}) - \boldsymbol{d}_j^{obs}\right), \quad j = 1, \ldots, N_e,$$

Iterative ensemble smoother

- **Further modification**

Change with iteration — Mean sensitivity — Can be approximated by $C_{D_l D_l}$

$$\boldsymbol{m}_{l+1,j} = \boldsymbol{m}_{l,j} - \frac{1}{1+\lambda_l}\left[C_{M_l} - C_{M_l}\bar{G}_l^T\left((1+\lambda_l)C_D + \bar{G}_l C_{M_l}\bar{G}_l^T\right)^{-1}\bar{G}_l C_{M_l}\right]C_M^{-1}\left(\boldsymbol{m}_{l,j} - \boldsymbol{m}_j^{pr}\right) - C_{M_l}\bar{G}_l^T\left((1+\lambda_l)C_D + \bar{G}_l C_{M_l}\bar{G}_l^T\right)^{-1}\left(g(\boldsymbol{m}_{l,j}) - \boldsymbol{d}_j^{obs}\right),\quad j = 1,\ldots,N_e,$$

Can be approximated by $C_{M_l D_l}$

- **One has** (Chen & Oliver, 2013)

$$\boldsymbol{m}_{l+1,j} = \boldsymbol{m}_{l,j} - \frac{1}{1+\lambda_l}\left[C_{M_l} - C_{M_l D_l}\left((1+\lambda_l)C_D + C_{D_l D_l}\right)^{-1}C_{D_l M_l}\right]C_M^{-1}\left(\boldsymbol{m}_{l,j} - \boldsymbol{m}_j^{pr}\right) - C_{M_l D_l}\left((1+\lambda_l)C_D + C_{D_l D_l}\right)^{-1}\left(g(\boldsymbol{m}_{l,j}) - \boldsymbol{d}_j^{obs}\right),\quad j = 1,\ldots,N_e.$$

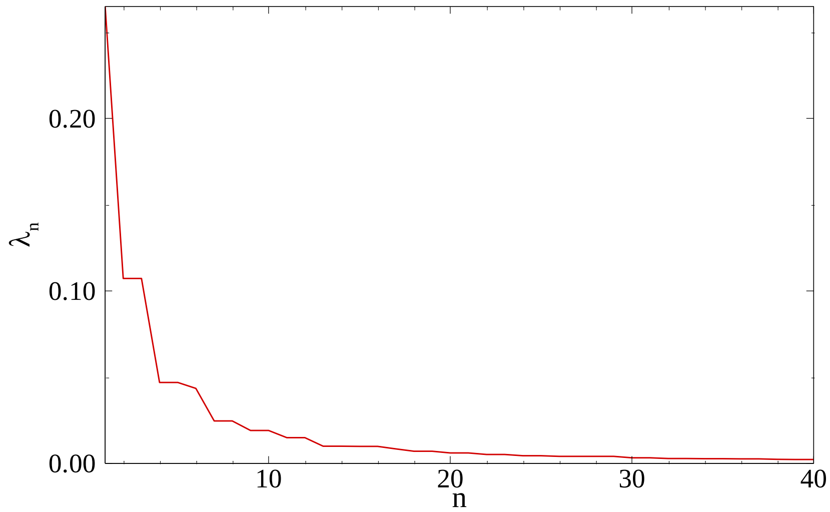
Surrogate model based iterative ES

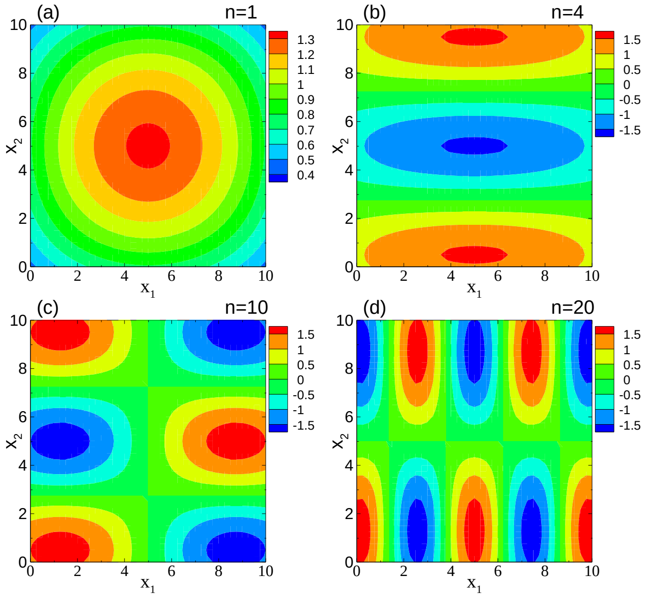
- **Independent model parameters**

$$\boldsymbol{m} : \boldsymbol{\xi} = \{\xi^1, \xi^2, \ldots, \xi^N\}^T$$

- **Random fields can be represented via** Karhunen-Loeve Expansion (KLE) **:**

$$Y'(\mathbf{x}, \omega) = \sum_{n=1}^{N} \xi_n(\omega) \sqrt{\lambda_n} f_n(\mathbf{x})$$





Surrogate model based iterative ES

- **Independent model parameters**

$$\boldsymbol{m}:\boldsymbol{\xi}=\{\xi^1,\xi^2,\ldots,\xi^N\}^T$$

- **Independent parameter based iterative ES**

$$\boldsymbol{\xi}_{l+1,j}=\boldsymbol{\xi}_{l,j}-\frac{1}{1+\lambda_l}\left[C_{\xi_l}-\underline{C_{\xi_l D_l}}\left((1+\lambda_l)C_D+\underline{C_{D_l D_l}}\right)^{-1}\underline{C_{D_l\xi_l}}\right]\left(\boldsymbol{\xi}_{l,j}-\boldsymbol{\xi}_j^{pr}\right)$$
$$-\underline{C_{\xi_l D_l}}\left((1+\lambda_l)C_D+\underline{C_{D_l D_l}}\right)^{-1}\left(g(\boldsymbol{\xi}_{l,j})-\boldsymbol{d}_j^{obs}\right),\quad j=1,\ldots,N_e.$$

Can be obtained from surrogate

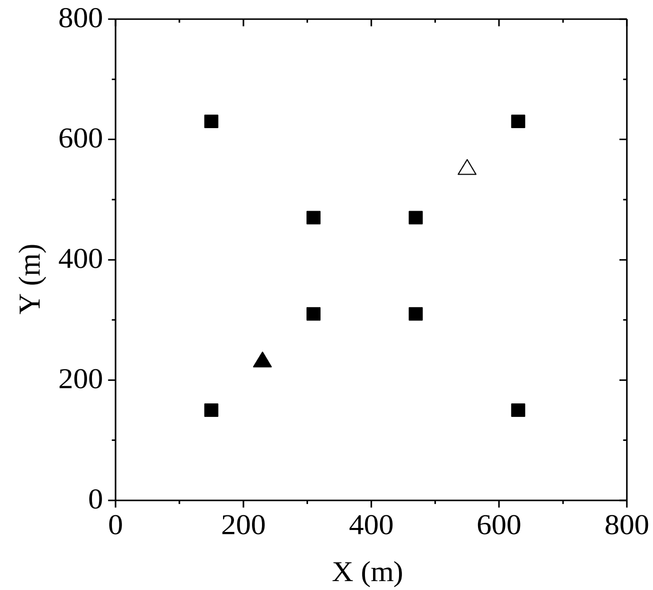
- **Surrogate model based iterative ES (Chang, Liao & Zhang, AWR, 2016)**

$$\boldsymbol{\xi}_{l+1,j}=\boldsymbol{\xi}_{l,j}-\frac{1}{1+\lambda_l}\left[C_{\xi_l}-C_{\xi_l D_l}^{surr}\left((1+\lambda_l)C_D+C_{D_l D_l}^{surr}\right)^{-1}C_{D_l\xi_l}^{surr}\right]\left(\boldsymbol{\xi}_{l,j}-\boldsymbol{\xi}_j^{pr}\right)$$
$$-C_{\xi_l D_l}^{surr}\left((1+\lambda_l)C_D+C_{D_l D_l}^{surr}\right)^{-1}\left(\boldsymbol{d}^{surr}(\boldsymbol{\xi}_{l,j})-\boldsymbol{d}_j^{obs}\right),\quad j=1,\ldots,N_e.$$

Case study: single-phase flow

$$S_s \frac{\partial h(\mathbf{x},t)}{\partial t} - \nabla \cdot \left(K(\mathbf{x}) \nabla h(\mathbf{x},t) \right) = q(\mathbf{x},t)$$

- **Model size:** 800 m x 800 m x 1 m
- **Grid:** 40 x 40 x 1
- **Boundary:** constant head (left and right);
  no flow (two lateral boundaries).
- **Well:** pumping well at block (12, 12);
  injecting well at block (28, 28).
- **Observation:** hydraulic heads
- **Observation duration:**
  start from day 0.2, at every 0.6 day, up to 5 days



Uncertain random field

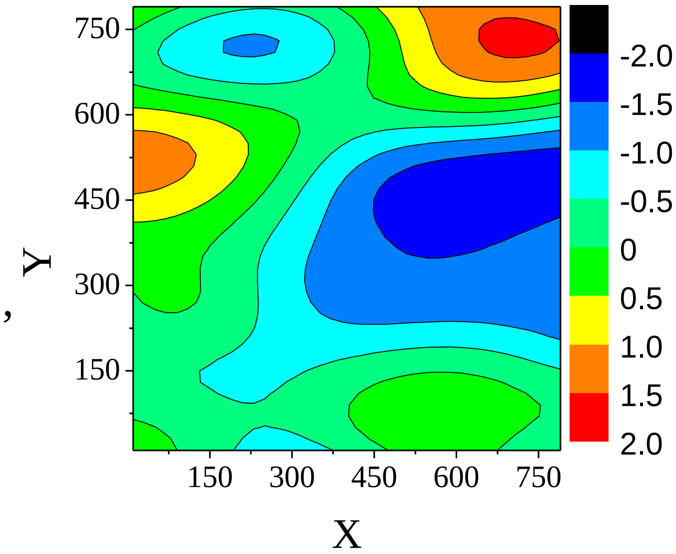
- **log-transformed conductivity is Gaussian random field**

$$< \ln K(\boldsymbol{x}) >= 0 + \ln(\text{m}/\text{day}),$$

$$\sigma^2_{\ln K}(\boldsymbol{x}) = 0.75,$$

$$C_{\ln K}(\boldsymbol{x}_{i_1,j_1}, \boldsymbol{x}_{i_2,j_2}) = \sigma^2_{\ln K} \exp\left\{ -\left[\left(\frac{|x_{i_1} - x_{i_2}|}{\eta_x} \right)^2 + \left(\frac{|y_{j_1} - y_{j_2}|}{\eta_y} \right)^2 \right] \right\},$$

$$\eta_x / L_x = 0.4, \eta_y / L_y = 0.2,$$



- **Retained terms in KLE: 11**
- **Define errors:**

Surrogate error: $e_{surr} = \frac{1}{N_1 N_d} \sum_{j=1}^{N_1} \sum_{i=1}^{N_d} \left| \frac{d_j^{i,sim} - d_j^{i,surr}}{d_j^{i,sim}} \right|,$

data match error: $e_d = \frac{1}{N_e N_d} \sum_{j=1}^{N_e} \sum_{i=1}^{N_d} \left| \frac{d_j^{i,obs} - d_j^{i,update}}{d_j^{i,obs}} \right|,$

Parameter estimation error: $e_m = \sqrt{\frac{1}{N_e N_m} \sum_{j=1}^{N_e} \sum_{i=1}^{N_m} \left(m^{i,ref} - m_j^{i,update} \right)^2}.$

Results

- **Updated mean fields:**

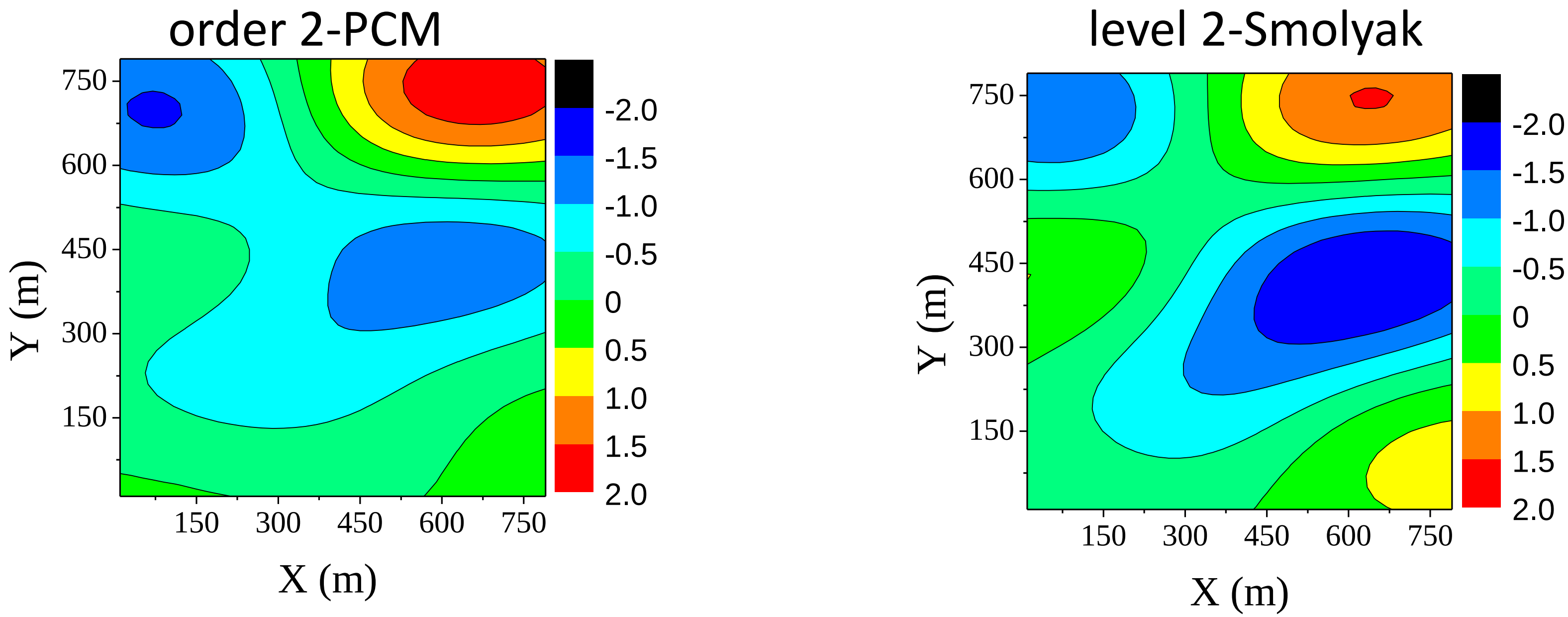


Table 2. Error analysis of case 2

Surrogate model	Simulation number, R	Surrogate model error, e_{surr}	Data match error, e_d	Parameter estimation error, e_m
Order 2-PCM	78	3.47E-3	5.92E-3	0.606
Level 2-Smolyak	287	1.45E-3	2.16E-3	0.445

Traditional iterative ES

- **Updated mean fields:**

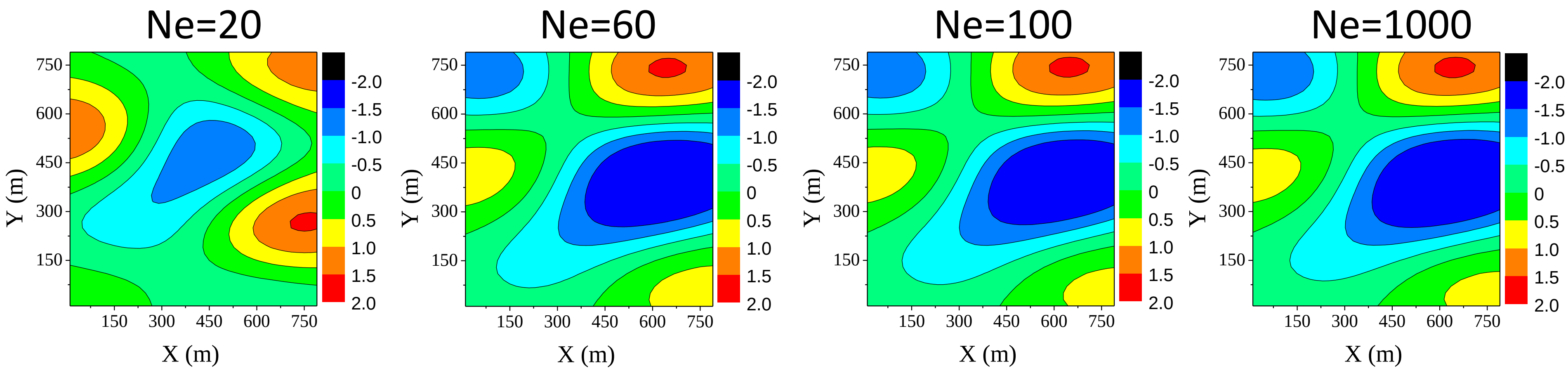
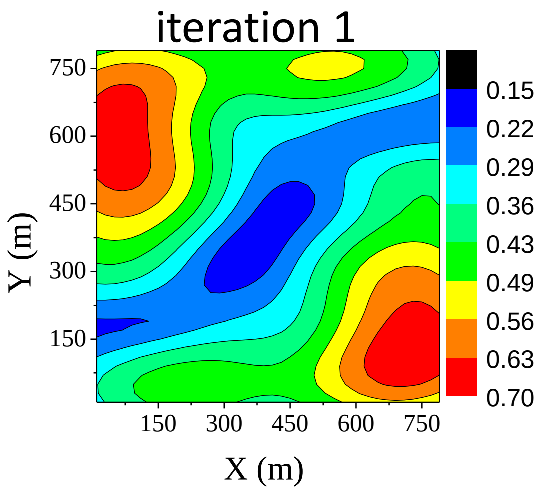


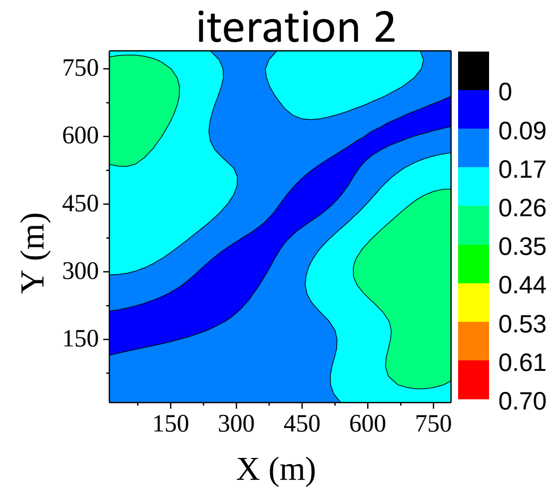
Table 3. Error analysis of traditional iterative ES

Realization number, *Ne*	Iteration number, N_I	Simulation number, *R*	Parameter estimation error, e_m
20	4	80	0.901
60	4	240	0.452
100	4	400	0.433
1000	4	4000	0.435

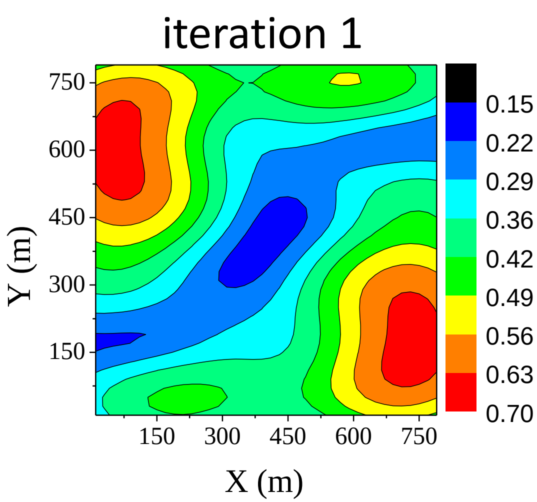
Standard deviation comparison

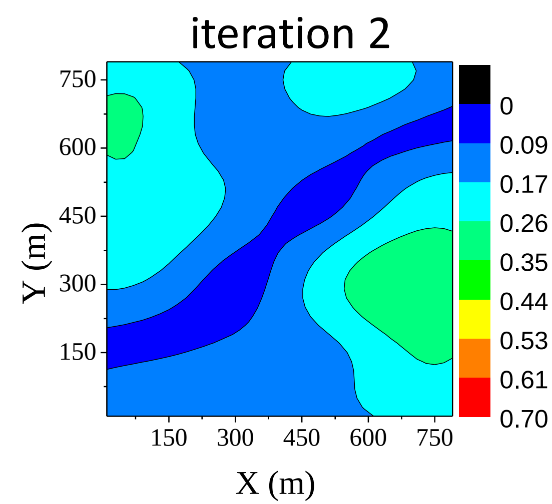
- **Level 2-Smolyak based iterative ES:**





- **Traditional iterative ES with Ne being 1000:**



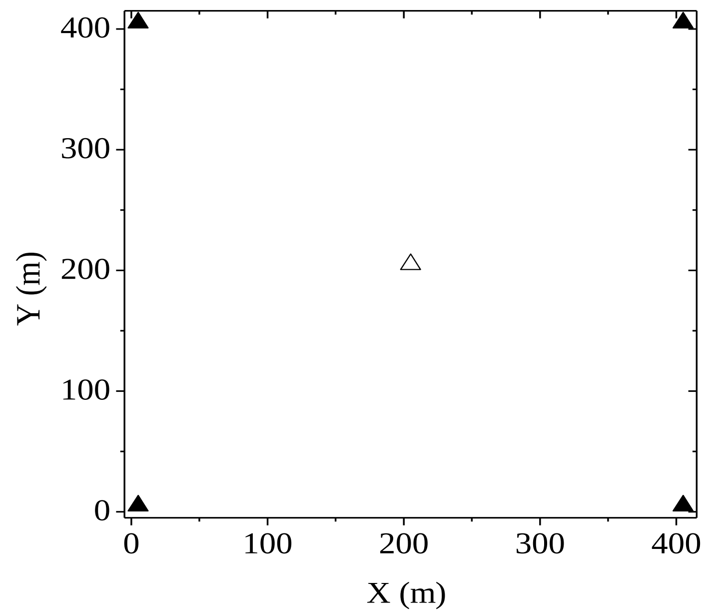


Case study: Multi-phase flow

Water-oil two-phase system:

$$\nabla \cdot \left[\frac{k(\mathbf{x})k_{ri}}{\mu_i} \left(\nabla p_i - \rho_i g \nabla z \right) \right] + q_i = \phi(\mathbf{x}) \frac{\partial S_i}{\partial t}, \ i = w, o$$

- **Model size:** 410 m x 410 m x 1 m
- **Grid:** 41 x 41 x 1
- **Boundary:** no flow
- **Well:** injector at the center;
  producers at the corners.
- **Observation:** WCT and OPR
- **Observation duration:**
  every 30 day, up to 510 days



Uncertain random field

- **log-transformed permeability is Gaussian random field**

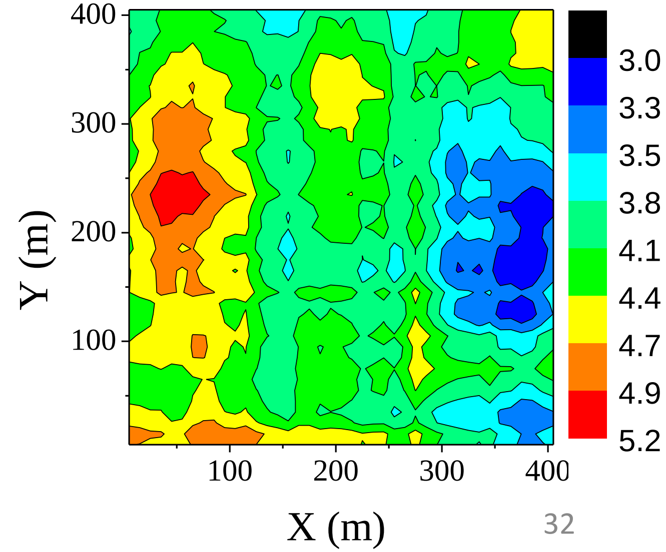
$$< \ln k(\boldsymbol{x}) >= 4 + \ln(\mathrm{mD}),$$

$$\sigma^2_{\ln k}(\boldsymbol{x}) = 0.16,$$

$$C_{\ln k}(\boldsymbol{x}_{i_1,j_1}, \boldsymbol{x}_{i_2,j_2}) = \sigma^2_{\ln k} \exp\left\{-\left[\frac{\left|x_{i_1} - x_{i_2}\right|}{\eta_x} + \frac{\left|y_{j_1} - y_{j_2}\right|}{\eta_y}\right]\right\},$$

$$\eta_x / L_x = 0.4, \eta_y / L_y = 0.4.$$

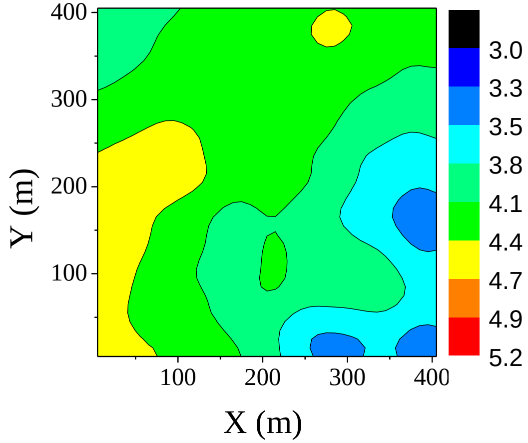
- **Retained terms in KLE:** 55
- **Surrogate model:** tTPCM (Liao & Zhang, 2016)
- **Simulation number:** 111



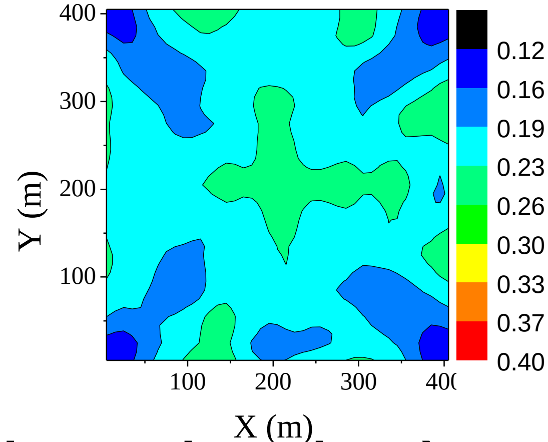
Results

- **tTPCM based iterative ES**, $(e_m = 0.27)$

Mean

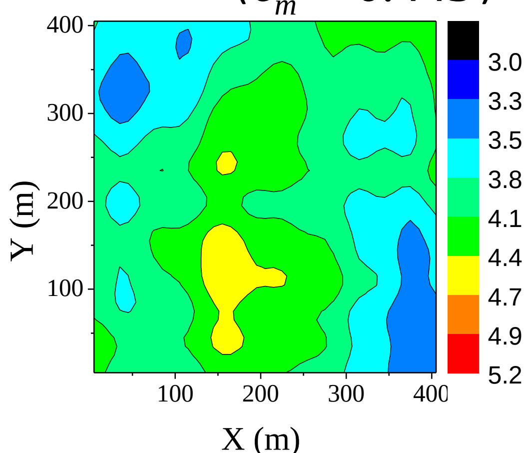


standard deviation

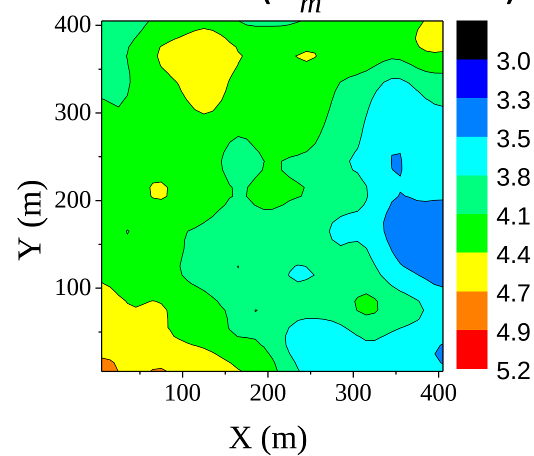


- **Traditional iterative ES (each run takes 10 iterations):**

Ne=20 $(e_m = 0.445)$

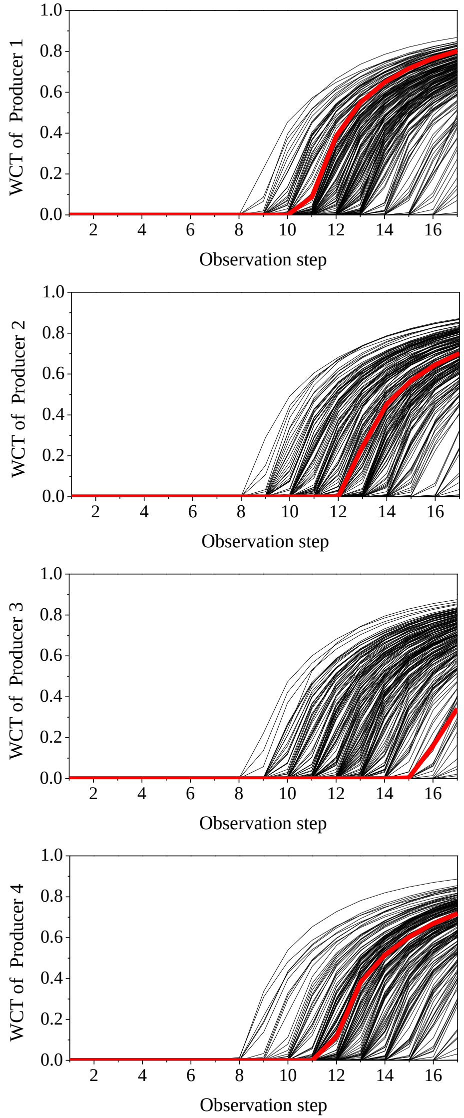


Ne=100 $(e_m = 0.268)$

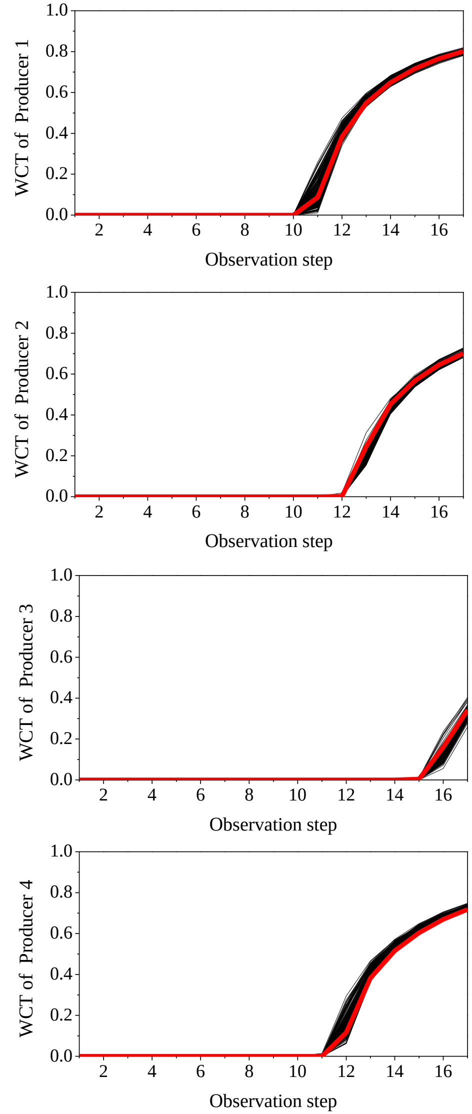


Data match of water cut from tTPCM

Initial ensemble



updated ensemble



Conclusions

- **Inverse problems may be solved efficiently with the aid of surrogate models**
  - Use surrogate model to approximate the forward solution
  - Apply transformation to address the low-regularity
  - Select the important dimensions adaptively to reduce the points
- **Performance**
  - Fast convergence of the surrogate solution to the exact forward solution
  - Fast convergence of the surrogate posterior to the true posterior

Thanks !

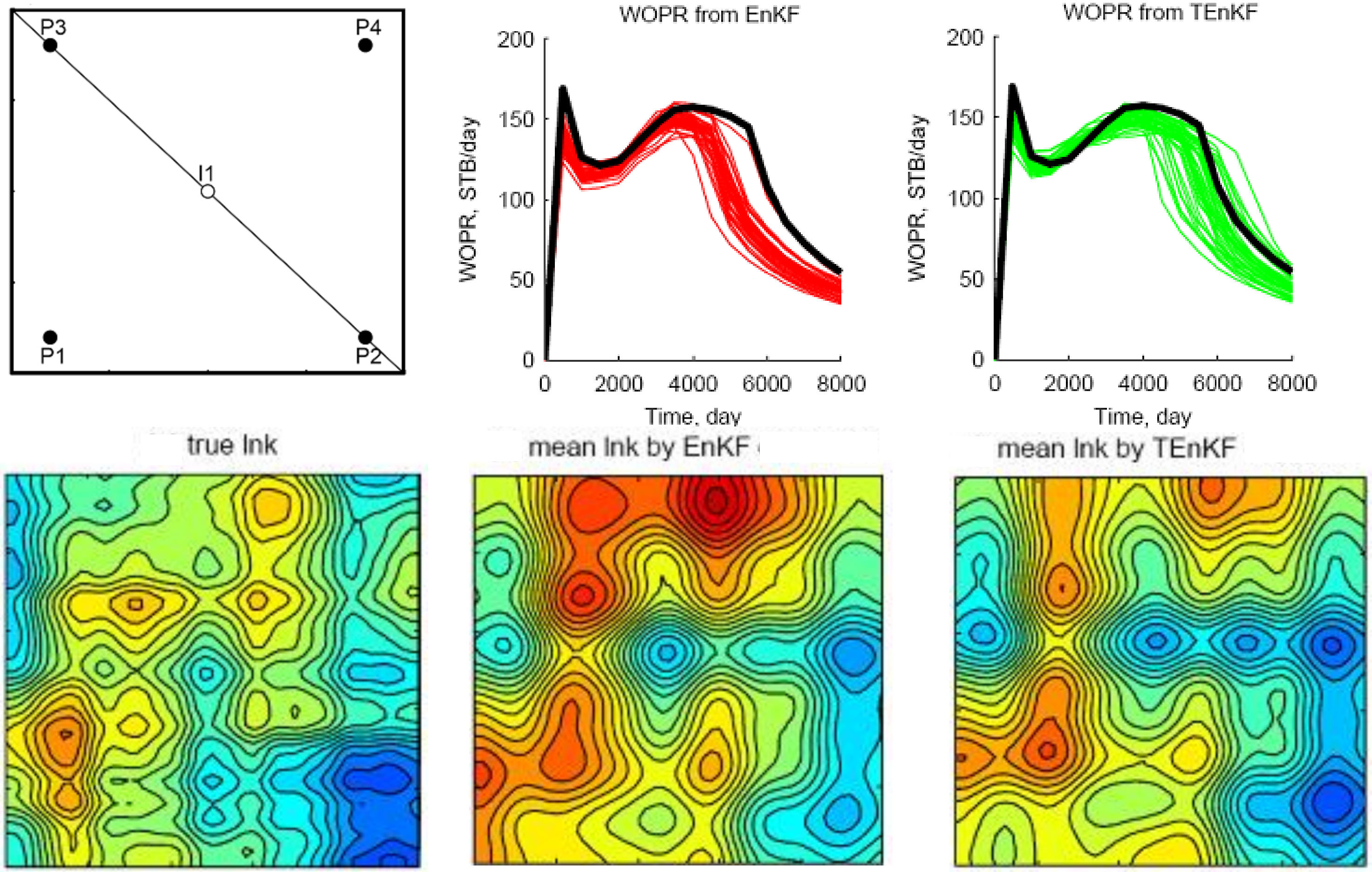
Q&A

References

- Chang, H., & Zhang, D. (2009). A comparative study of stochastic collocation methods for flow in spatially correlated random fields. *Communications in Computational Physics, 6*(3), 509-535.
- Chang, H., Liao, Q., Zhang, D., 2016. Surrogate model based iterative ensemble smoother for subsurface flow data assimilation. Advances in Water Resources. Doi:10.1016/j.advwatres.2016.12.001
- Liao, Q., & Zhang, D. (2016). Probabilistic collocation method for strongly nonlinear problems: 3. transform by time. *Water Resources Research,52*(3), 7911–7928.
- Liao, Q., Zhang, D., & Tchelepi, H. (2016). A two-stage adaptive stochastic collocation method on nested sparse grids for multiphase flow in randomly heterogeneous porous media. *Journal of Computational Physics*. In press.
- Lin, G., and A. M. Tartakovsky (2009), An efficient, high-order probabilistic collocation method on sparse grids for three-dimensional flow and solute transport in randomly heterogeneous porous media, *Adv. Water Resour., 32*(5), 712–722.
- Marzouk, Y., and D. Xiu (2009), A stochastic collocation approach to Bayesian inference in inverse problems, *Commun. Comput. Phys.*, 6(4), 826–847.
- W.A. Klimke (2006), Uncertainty modeling using fuzzy arithmetic and sparse grids, Citeseer.
- Xiu, D., and J. S. Hesthaven (2005), High-order collocation methods for differential equations with random inputs, *SIAM J. Sci. Comput., 27*(3), 1118-1139.
- Zhang, D., L. Shi, H. Chang, and J. Yang (2010), A comparative study of numerical approaches to risk assessment, *Stoch. Environ. Res. Risk. Assess. 24*, 971–984.

Backup: transformed EnKF (TEnKF)

- 2D synthetic case: 20 random variables after Karhunen-Loeve expansion



Liao, Q., & Zhang, D. (2015). Data assimilation for strongly nonlinear problems by transformed ensemble kalman filter. *SPE Journal, 20*(1), 202-221.