Multiscale Bayesian inversion of binary permeability fields from static and dynamic observations

J. Ray, S. A. McKenna, Y. M. Marzouk and
B. van Bloemen Waanders

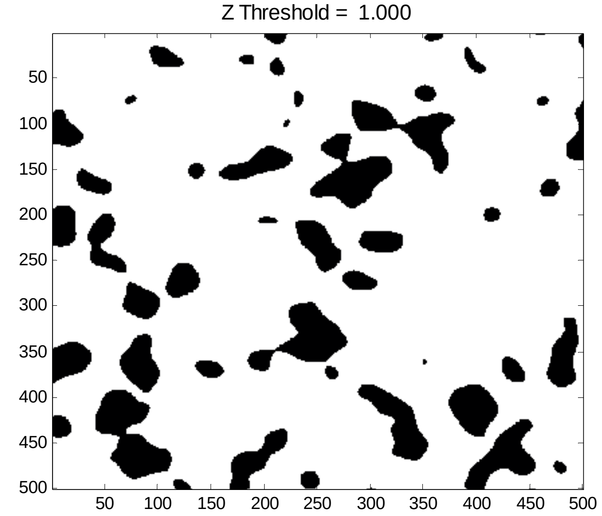
jairay [at] somnet [dot] sandia [dot] gov

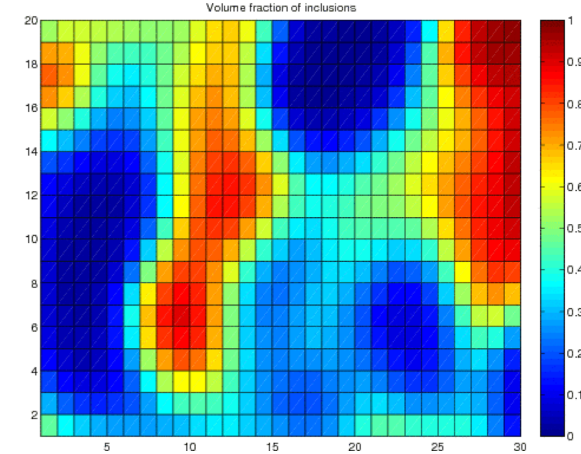
Sandia National Laboratories, Livermore CA,
Sandia National Laboratories, Albuquerque, NM
Massachusetts Institute of Technology, Cambridge, MA

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,
for the United States Department of Energy's National Nuclear Security Administration
under contract DE-AC04-94AL85000.

Problem statement

- **Characterize a binary medium (BM) from sparse observations**
- **Binary medium**
  - **Low permeability matrix with high permeability inclusions**
  - **Volume fraction, *F*, of inclusion varies in space, *F(x)***
  - **2-3 orders of magnitude difference between matrix and inclusion permeability**
- **Domain ~4-5 orders of magnitude bigger than inclusions**
  - **Characterization = statistical summary of fine-scale (inclusion scale), including variation in space**
    - **Non-unique reconstruction !**
  - **Fine-scale reconstruction comes with quantified uncertainty**





Sandia National Laboratories

The inverse problem

- **Sources of data**
  - Static data, i.e. permeability measurements from core samples
  - Dynamic data, i.e. tracer breakthrough times
- **Objects of inference, $\Theta = \{F(x), \delta, \alpha\}$**
  - *F(x),* spatially variable distribution of inclusion volume fraction
  - Geometrical parameters used to describe inclusions (δ, α)
- **Bayesian inverse problem**

$$-2\log\pi(\Theta) \propto \frac{\{T_B - M_T(\Theta)\}^2}{\sigma^2{}_T} + \frac{\{K_{obs} - M_K(\Theta)\}^2}{\sigma_K{}^2} + \frac{\{\Theta - \Theta_p\}^2}{\sigma^2{}_\Theta}$$

  - M_T, model to relate Θ to breakthrough times T_B
  - M_K, model to relate Θ to observed permeability at certain sampling points
  - Θ_p, prior beliefs regarding the values of Θ
  - $\sigma_{\{K, T\}}$, std. dev. of various measurement errors
- **$\pi(\Theta)$ evaluated by Markov Chain Monte Carlo sampling**
  - Particular algorithm called DRAM
  - Uses samples collected so far to adjust the proposal distribution

Sandia National Laboratories

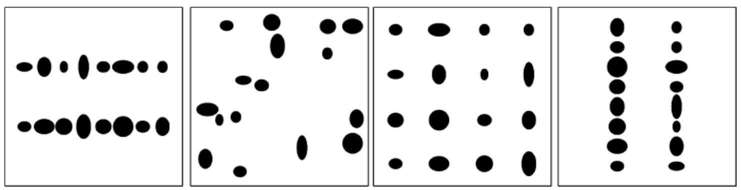
Sub-grid Model

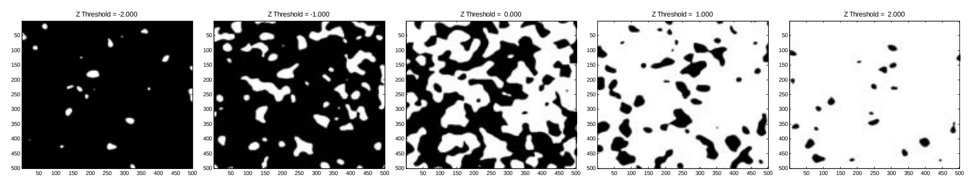
<u>GOAL</u>: $k_{eff} = f(k_1,k_2,f,FWHM)$

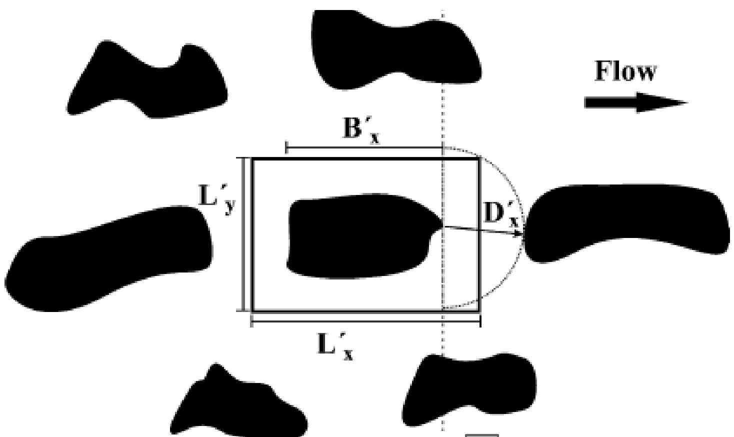
Distance-Based Upscaling + Truncated Gaussian Field Theory

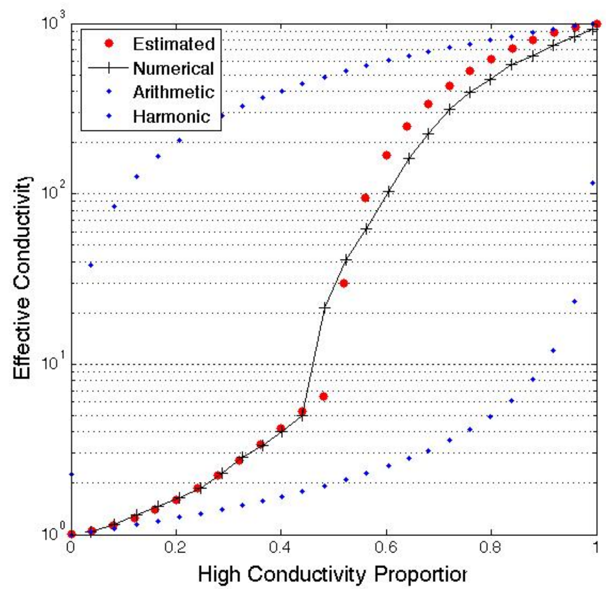
Knudby, et al., 2006

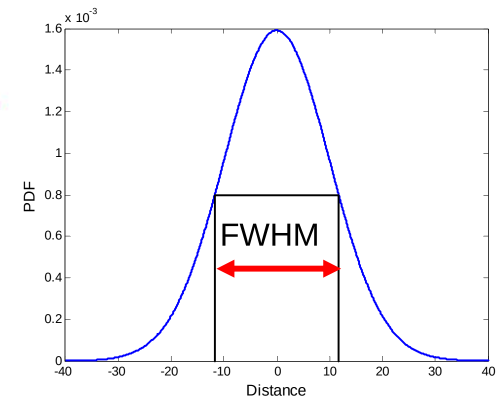
e.g., Adler et al., 2009





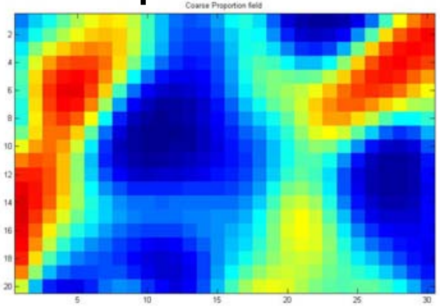




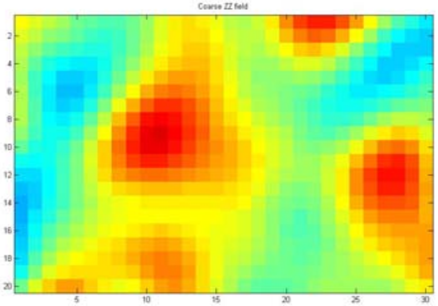


Multiscale Ground Truth

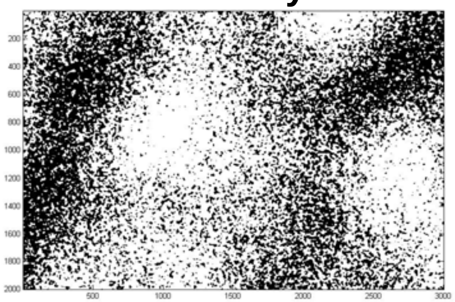
Proportion Field

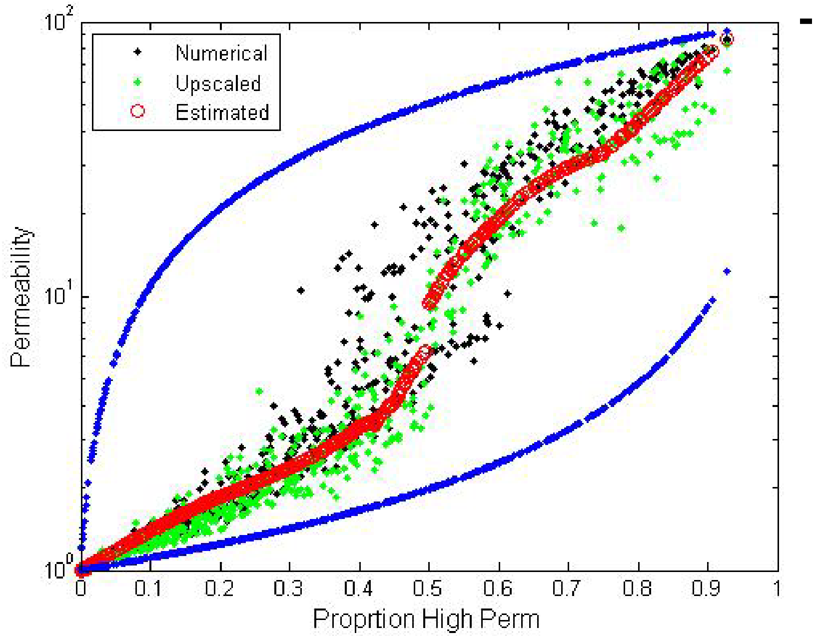


Z-score threshold field



Fine Scale Binary Medium





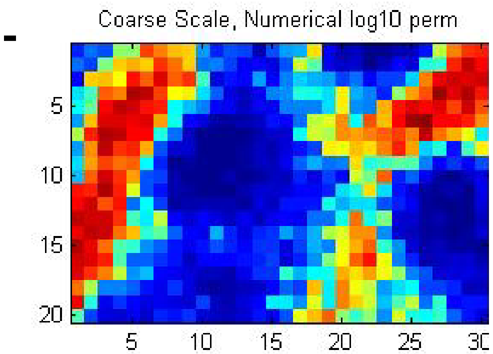
Kernel size (sigma = 5.0)

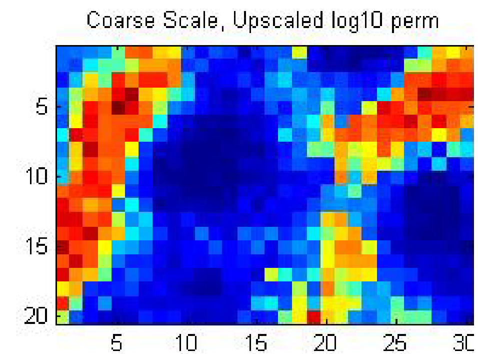
K1 = 100, K2 = 1

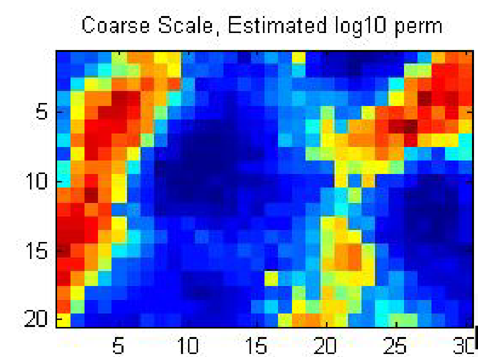
Coarse = 20x30

Fine = 2000x3000

Coarse block is 100x100 fine cells

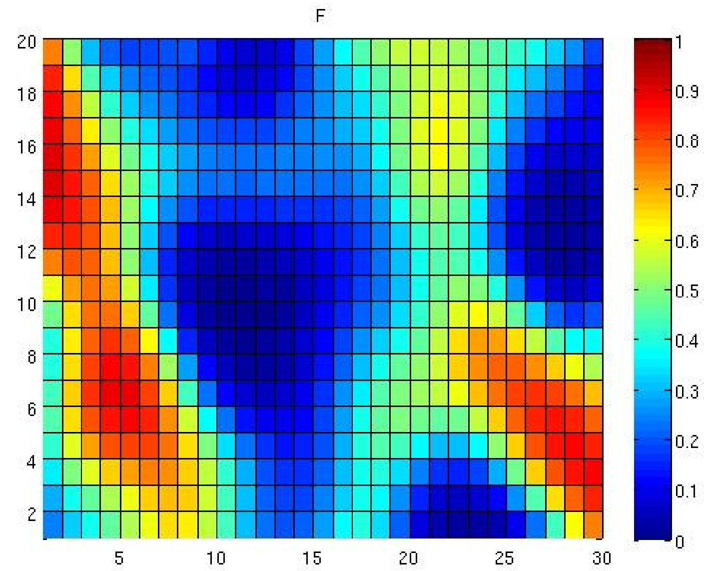


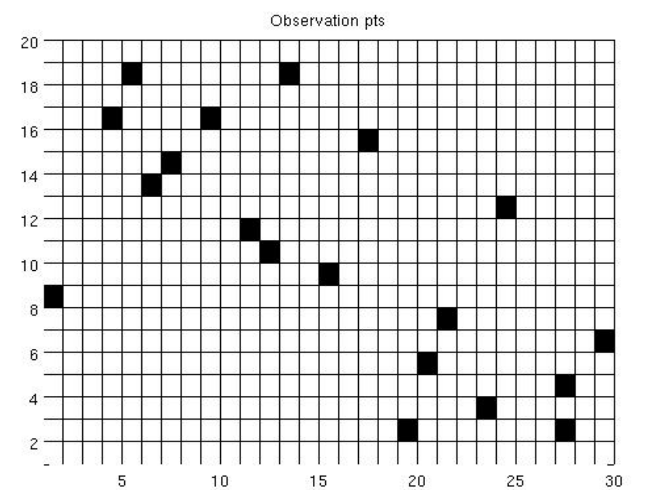




Test A – an isotropic case

- **Rectangular domain discretized by 30 x 20 grid**
  - **Sufficient to resolve a smoothly varying *F(x)***
  - ***F(x)* characterized by a known semi-variogram**
    - **Correlation length ~10% of the domain hypotenuse**
- ***F(x)* expressed as a summation of Karhunen-Loeve (KL) modes**
  - **Delinks accuracy of *F(x)* from the resolution of the mesh**
- **Synthetic data**
  - **Develop a true *F*, $F_{true}(x)$, using a distribution of Gaussian processes**
    - **Isotropic, so Gaussians have circular cross-section**
  - **Permeability contrast of 100; δ = 11.78**
  - **True permeability field developed using truncated Gaussian processes on a 3000 x 2000 mesh**
  - **Breakthroughs generated using MODFLOW**
- **Infer**
  - **Weights, c_i, of KL modes; δ, the size of GP used to generate the fine scale**
  - **Correlation structure of $F_{infer}(x)$ assumed known**
  - **Reconstruct $F_{infer}(x)$; compare against $F_{true}(x)$**





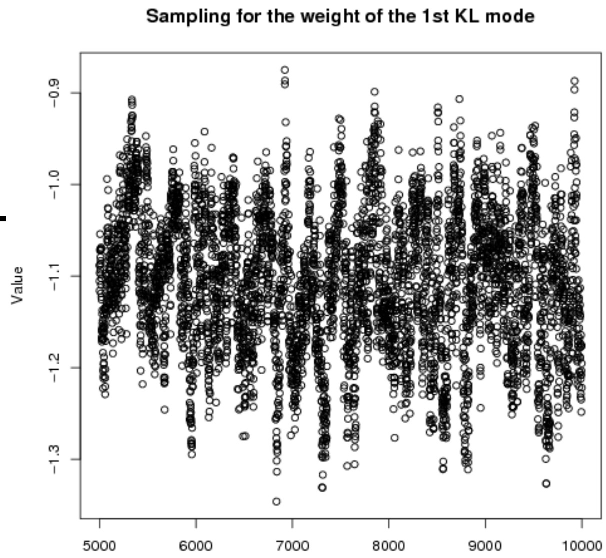
Simplified view of the MCMC procedure

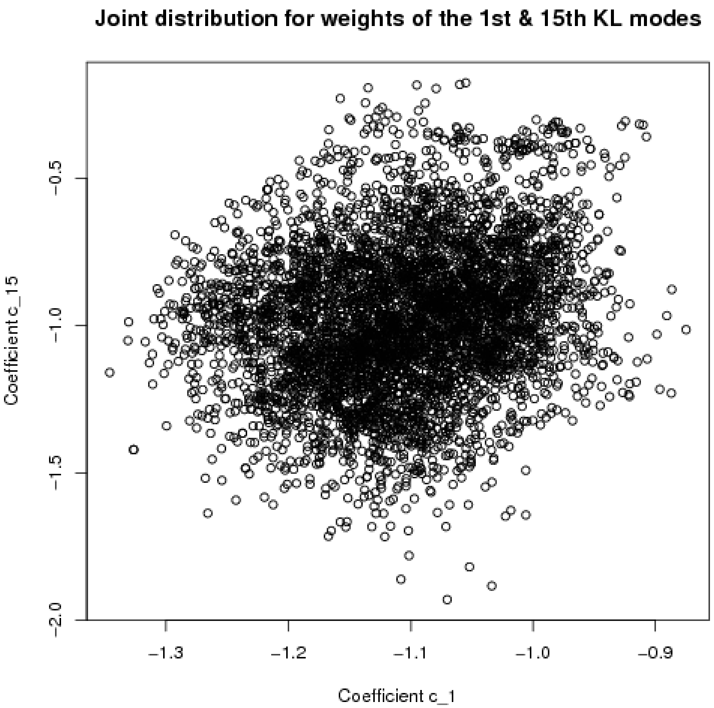
- Propose $\Theta = \{c_i, \delta\}$, c_i in lieu of $F(x)$
- Create $F(x)$ = summation of KL modes with weights c_i
  - Called "proposed" volume fraction field
- Create a permeability field K from "proposed" $F(x)$
  - Use the sub-grid model i.e. $K = M_K(c_i, \delta)$
  - Called the "proposed" permeability field
- Push water and tracer and find breakthrough times
  - "proposed" breakthrough times, T_B
- Compare "proposed" K and T_B to observed K, T_B at 20 obs. pts
- Accept "proposed" $\{K, T_B\}$ with a probability proportional to how close they are to observations
- Save $\Theta = \{c_i, \delta\}$ to file
  - These are the samples from the posterior distribution $\pi(\Theta)$

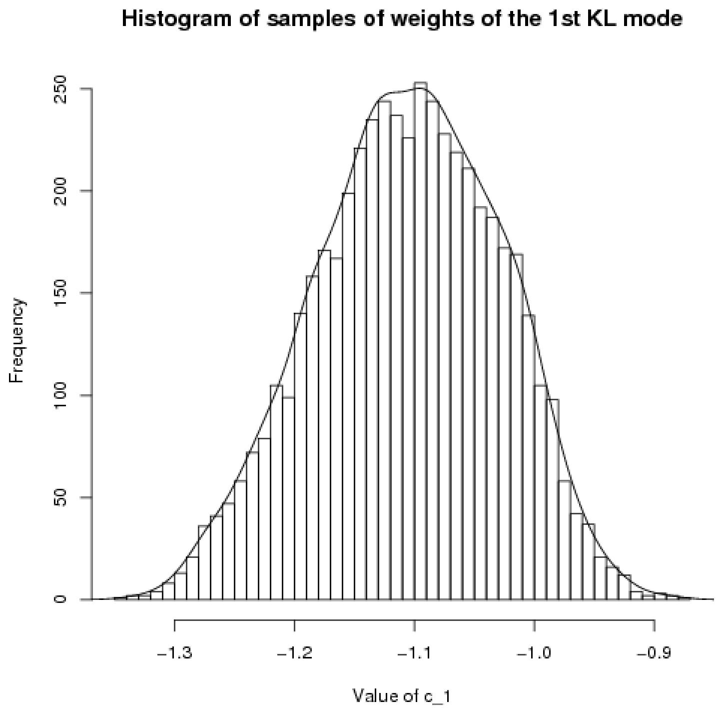
Sandia National Laboratories

Results from Test A

- **Construct posterior PDFs for c_i, δ**
- **~2M DRAM iterations**
  - **Expensive (72 hrs)**

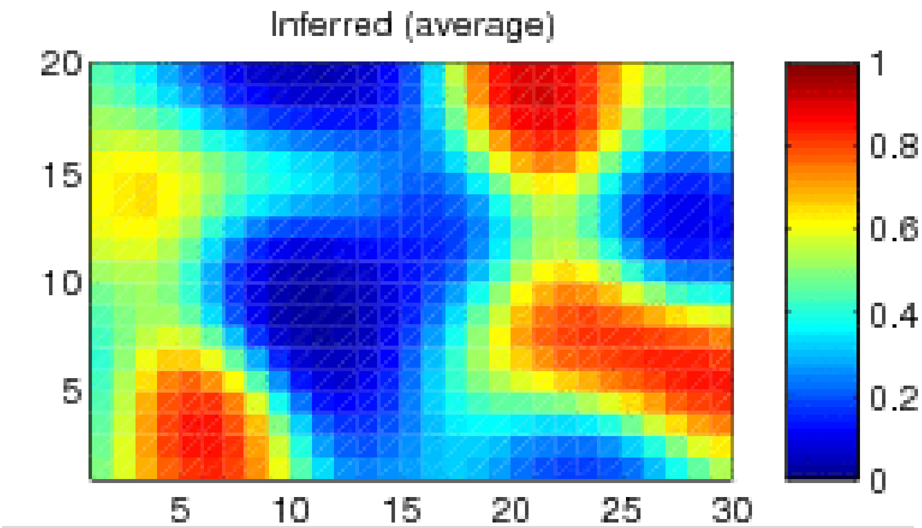


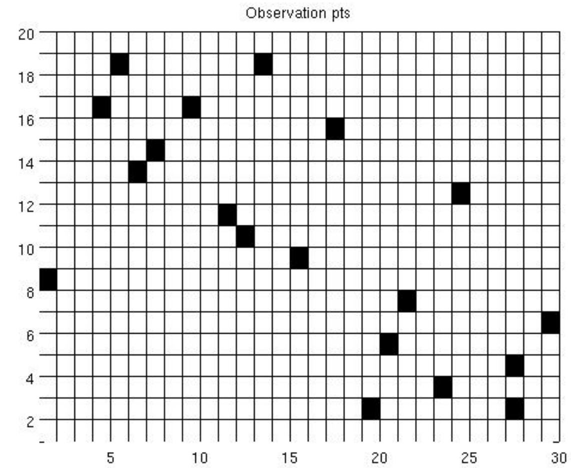




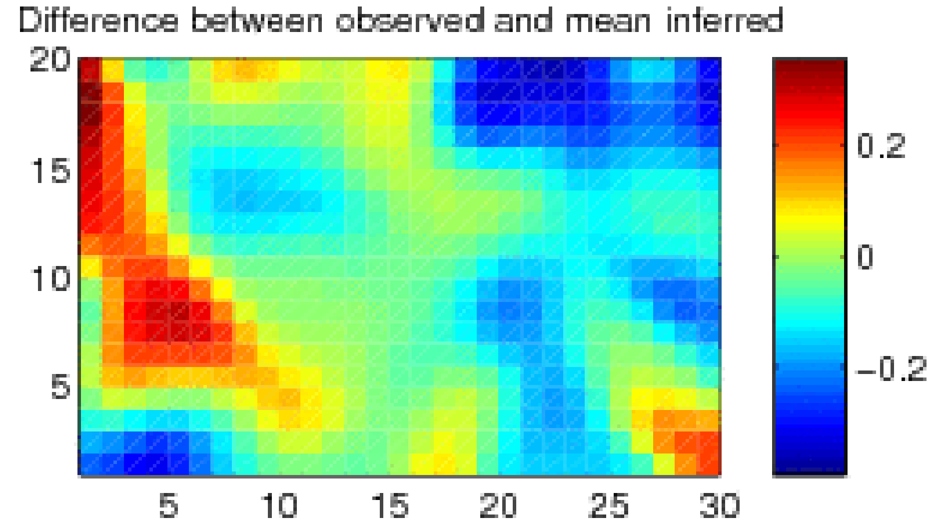
Sandia National Laboratories

Reconstructions from Test A

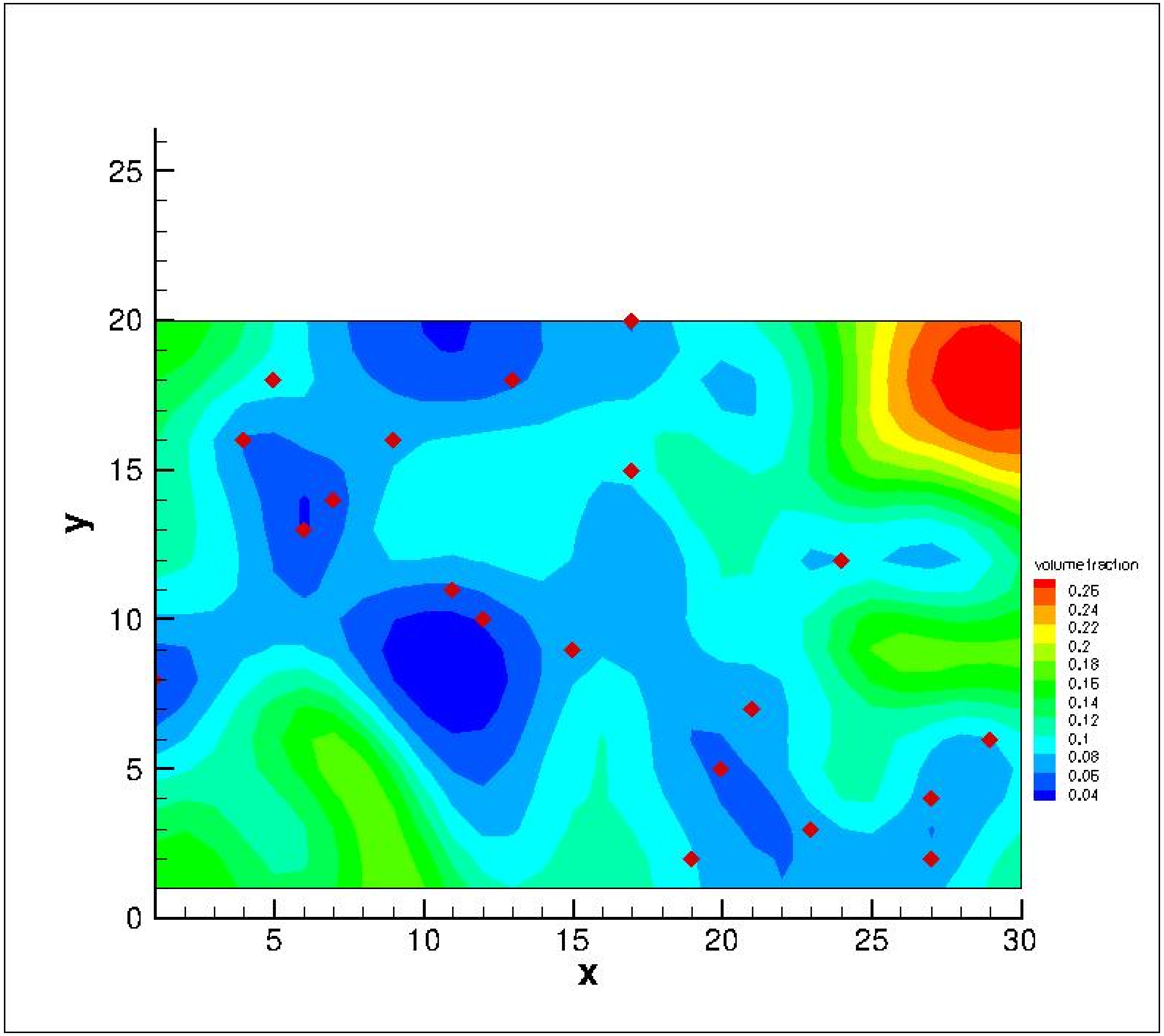




- **RMS error in *F* = 0.145**
  - **0.19 with static data only**

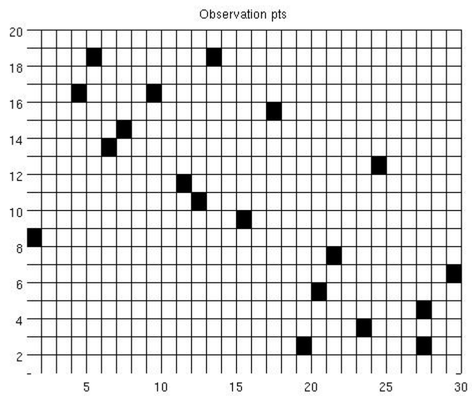


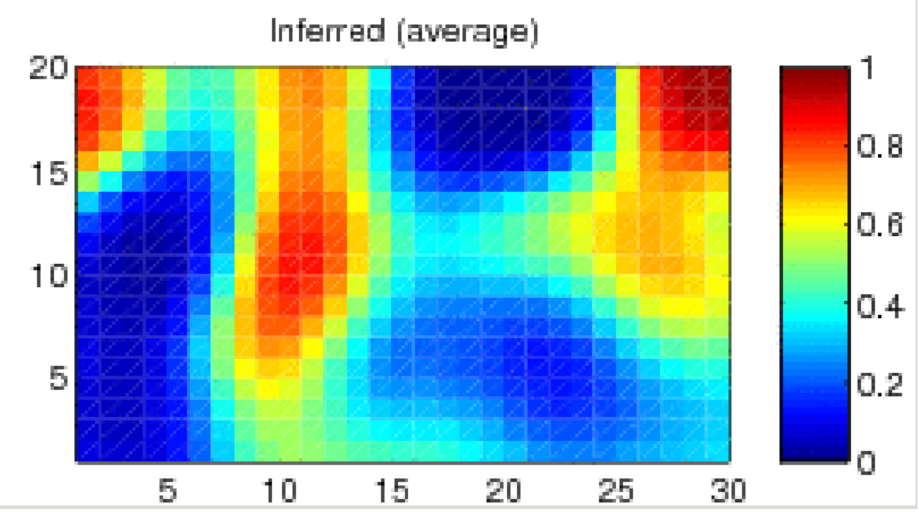
Uncertainty and observations

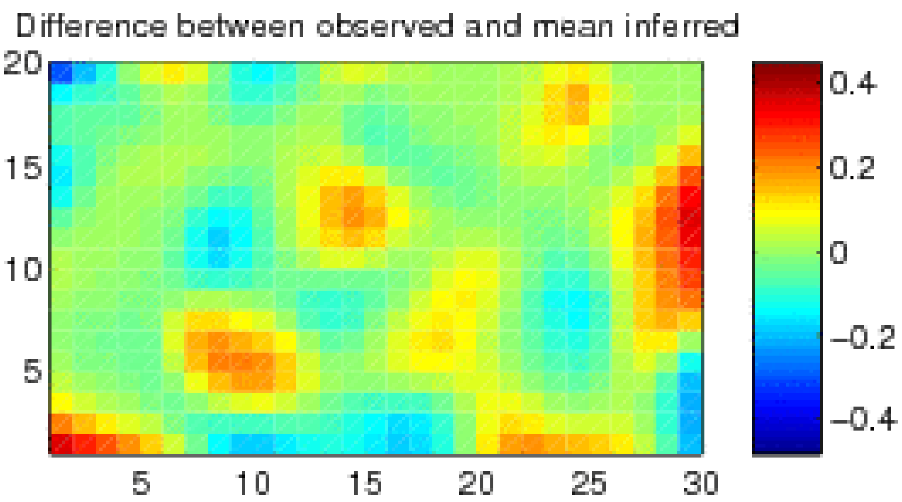


Test B – anisotropic case

- **Similar to Test A, but**
  - **Anisotropic permeability field**
  - **Gaussian kernels have elliptical cross-section**
    - ***AR* = 1.5, *δ* = 10**
- **Observables**
  - **K_{xx} and K_{yy} at 20 obs. pts; breakthrough times**
- **RMS error: 0.108**
  - **0.157 with static data only**







Conclusions

- It may be possible to reconstruct latent BM fields from sparse data
- Crucial elements
  - A sub-grid model, necessarily parameterized, that enables transfer of information from coarse scale (breakthrough times) to fine-scale
  - Smooth *F(x),* which enables its modeling with KL modes
    - And drastically drops the dimensionality of the problem
- Next steps
  - Increase domain size
  - 2-level inference – coarse-scale + (intermediate + fine-scale)
    - Coarse scale seeks out “good” parts of the Θ space
    - (intermediate + fine-scale) similar to now
  - What is the link function for coarse and intermediate scales?
  - What MCMC scheme? Implicitly coupled chains?
  - Parallelism needed.