Multiscale Behavior of Groundwater Flow and Transport in Binary Media

Sean A. McKenna
Geoscience Research and Applications Group
Sandia National Laboratories, Albuquerque, New Mexico

Youssef Marzouk
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology, Cambridge, Massachusetts

Jaideep Ray
Predictive Simulation R&D Department
Sandia National Laboratories, Livermore, CA

Bart van Bloemen Waanders
Applied Math and Applications Department
Sandia National Laboratories, Albuquerque, New Mexico

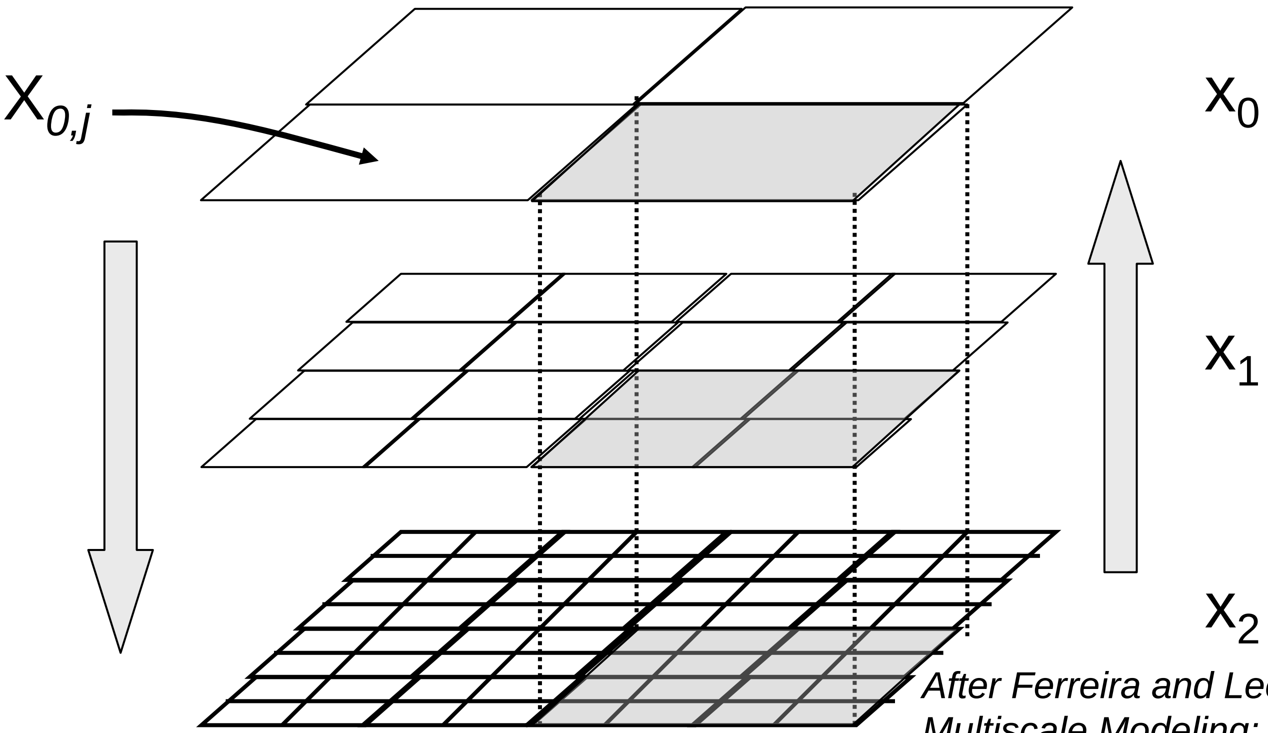
Sandia National Laboratories

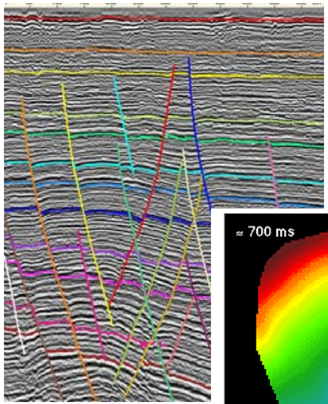
Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94-AL-85000

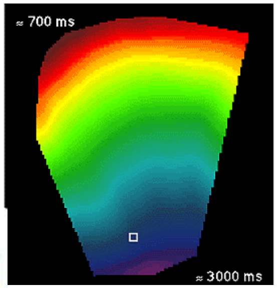
Multi-Scale Modeling Motivation

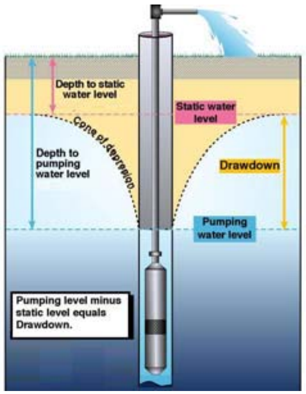
Data collected at one level informs values at other levels

Multiscale random fields with averaging "link" between them

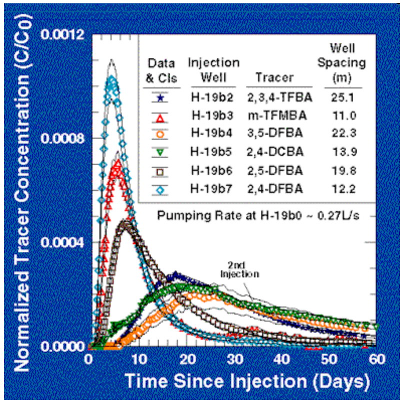








After Ferreira and Lee, 2007, Multiscale Modeling: A Bayesian Perspective



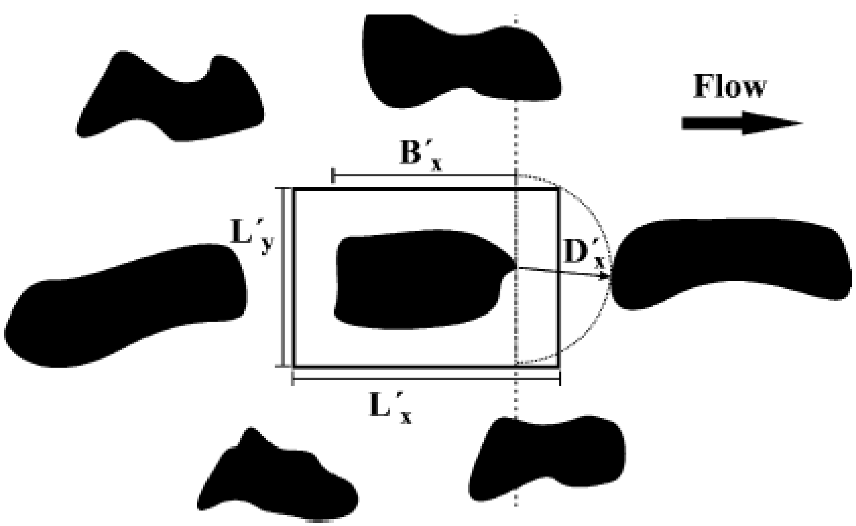
Two points of this presentation

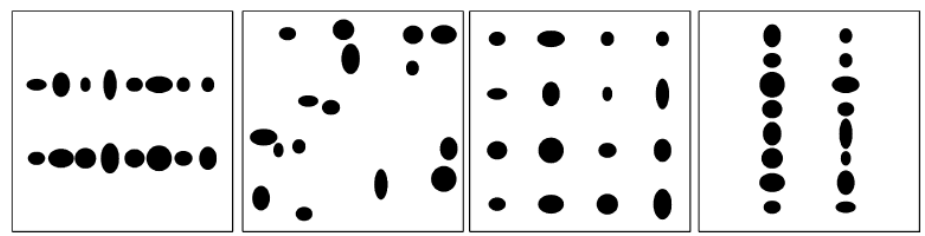
1) *More robust link function for binary media*
2) *Multiscale modeling using Bayesian inference*

Upscaling Binary Media

- Upscaling binary media is long-standing area of research interest in multiple fields
- Majority of approaches use information on modal permeabilities and proportions (+ inclusion shapes)
- *__Phase connectivity__ is not included in previous approaches*

Recent Development:
(Knudby et al., 2006)





$k_{eff} = f(k_1, k_2, proportion, \underline{distance})$
Average distance that a streamline travels through the matrix controls effective permeability

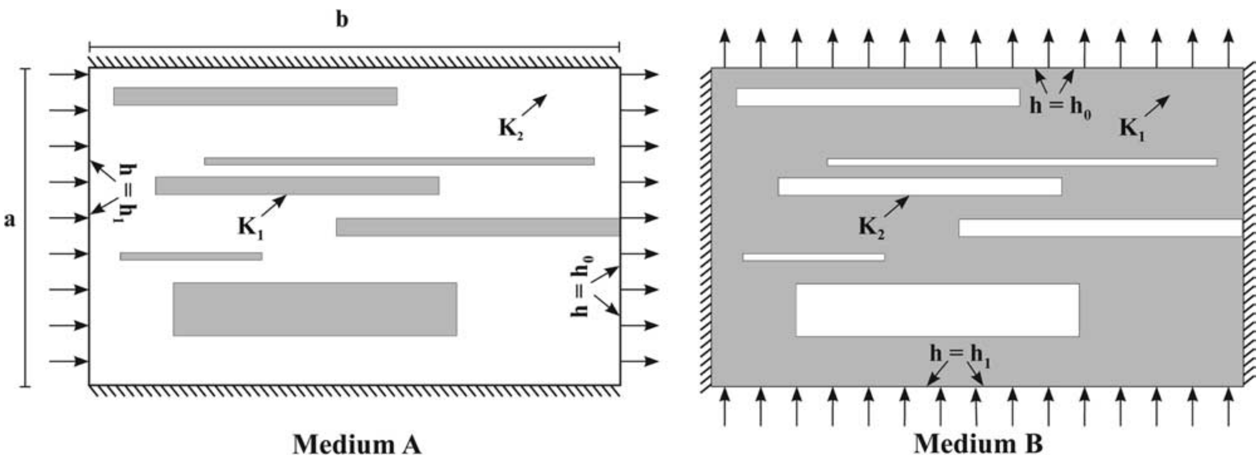
Figures from: Knudby et al., 2006, Advances in Water Resources

Upscaling Binary Media (Cont.)

- Phase Change Theorem
  - High perm inclusions in a low perm matrix *or vice versa*

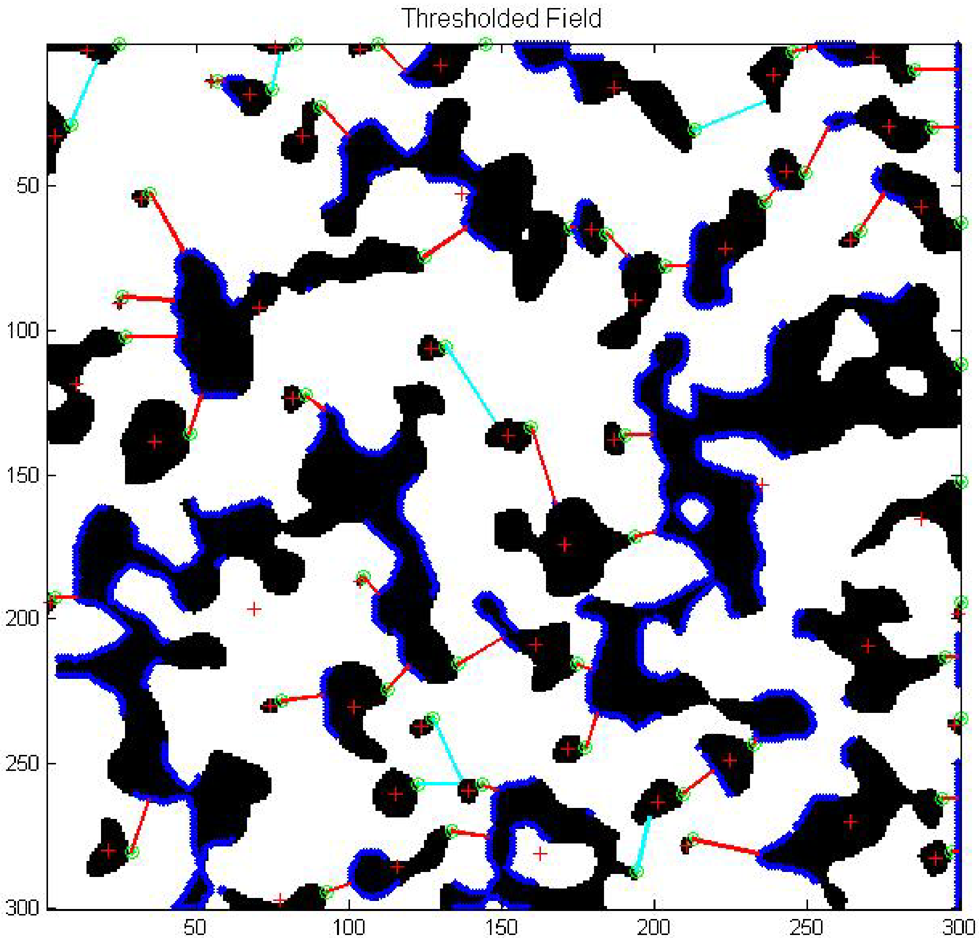
$$Q_A Q_B = k_1 k_2 (h_1 - h_0)^2$$

$$k_{eff,A} = \frac{Q_A b}{(h_1 - h_0) a} \qquad k_{eff,B} = \frac{Q_B a}{(h_1 - h_0) b} \quad \Longrightarrow \quad k_{eff,A} k_{eff,B} = k_1 k_2$$



From Knudby et al., 2006, AWR

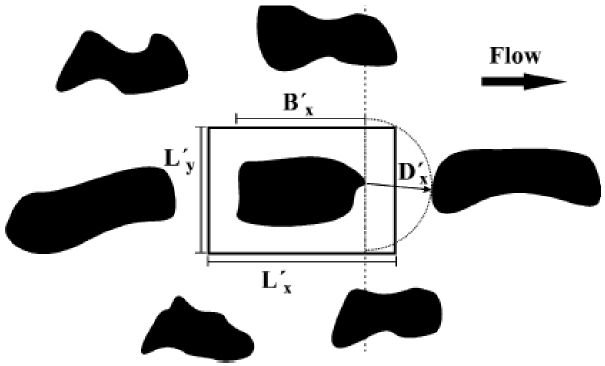
Object and Distance Calculations



Within Voronoi polygon

Detailed search

Need object area and average distance to closest "downstream" object



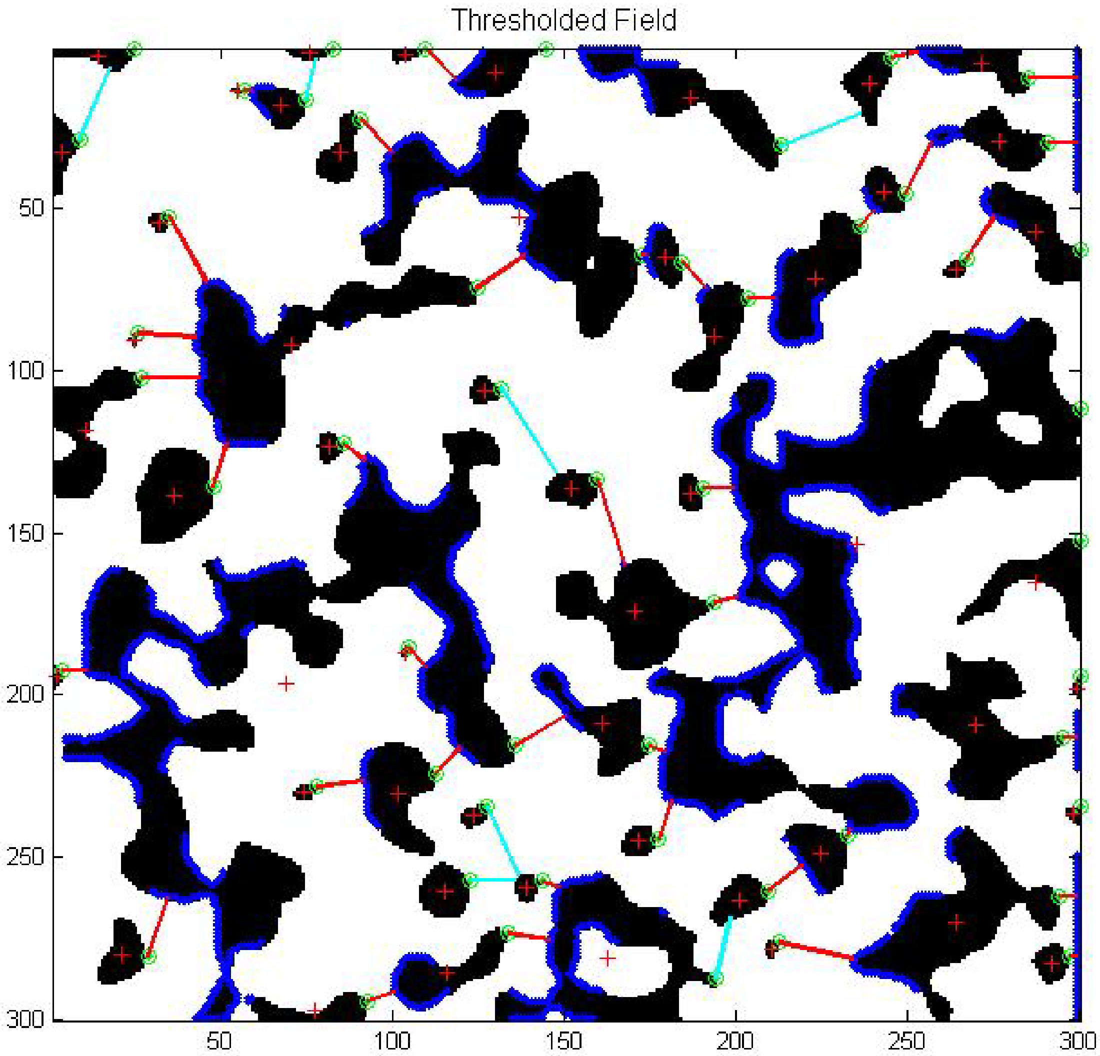
Identify downstream most point on each object ("from")

Search from there to closest perimeter point on nearby object ("to")

Two-stage approach for defining "nearby":

1) Voronoi polygons
2) Detailed search

Way Too Slow for Inversion!



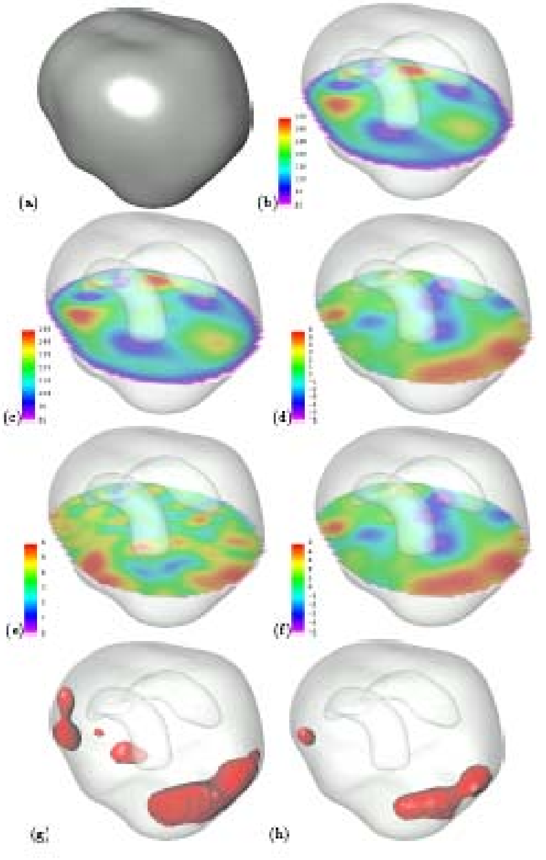
Subgrid Model

- Inexpensive means of estimating permeability within a region
  - No instantiation of binary field

Truncated Gaussian Fields!

Given two modal permeabilities, threshold value and information on the correlation of the field, estimate the permeability

Threshold crossing theory to get at size of excursions (inclusions)

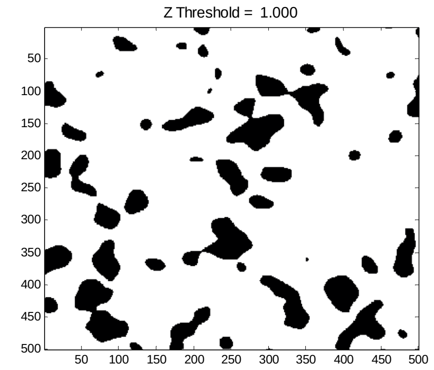


From: Adler et al., 2009, *Applications of Random Fields and Geometry, Foundations and Case Studies*

Truncated MultiGaussian Fields

N	The number of pixels above the truncation threshold, u.
m	The number of distinct regions (inclusions) above the threshold
n	The number of pixels in each region

Expectation relationship: *E[N] = E[m]·E[n]*.



For a threshold, *u*, and Gaussian distribution:

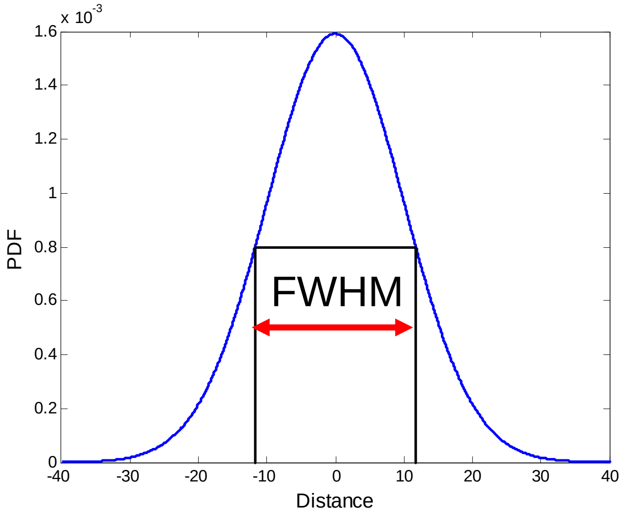
$$E[N] = S \cdot \Phi(-u) = S\int_{u}^{\infty}(2\pi)^{-1/2}e^{-z^2/2}dz$$

Euler Characteristic (EC) approximates expected *m* given threshold and structure of Gaussian field:

$$EC = E[m] \approx (2\pi)^{-(D+1)/2}W^{-D}u^{D-1}e^{-u^2/2}$$

$$W = FWHM/\sqrt{4\ln(2)}$$

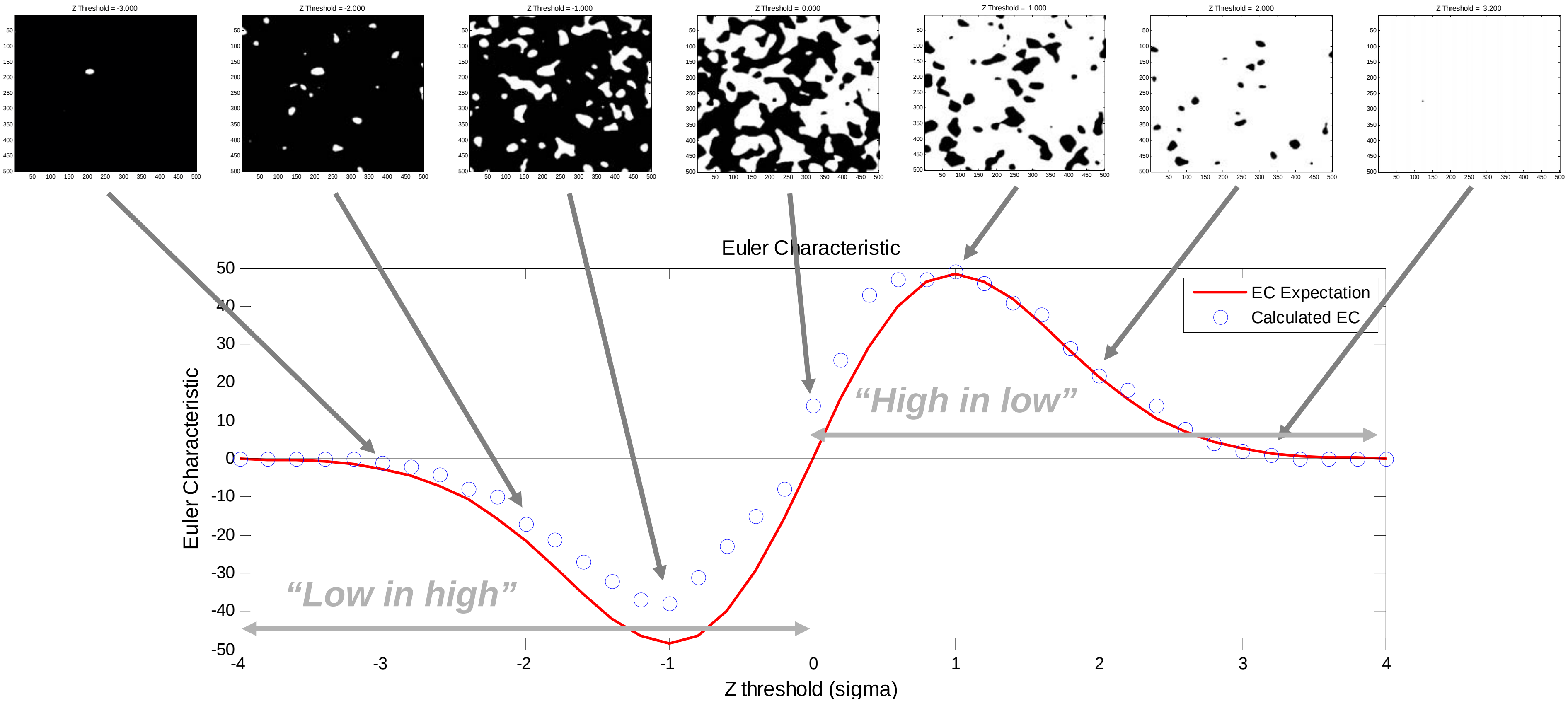
$$FWHM = \sigma\sqrt{8\ln(2)}$$



Euler Characteristic

Euler Characteristic = Euler Number = (#blobs - #holes), *at least in 2D*

Example field is 500x500 and filtered with an isotropic Gaussian kernel with a sigma = 10.0 distance units



Decreasing proportion of high permeability material

Average Distance in Background

Spatial Point Process theory gives distribution of nearest neighbor distances:

$$F(d) = 1 - \exp(-\pi\lambda d^2)$$

Quantity: πd^2 is exponentially distributed with mean = λ

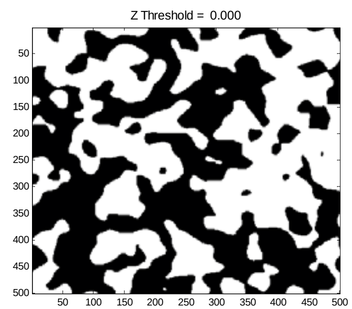
Correct for finite sized inclusions:

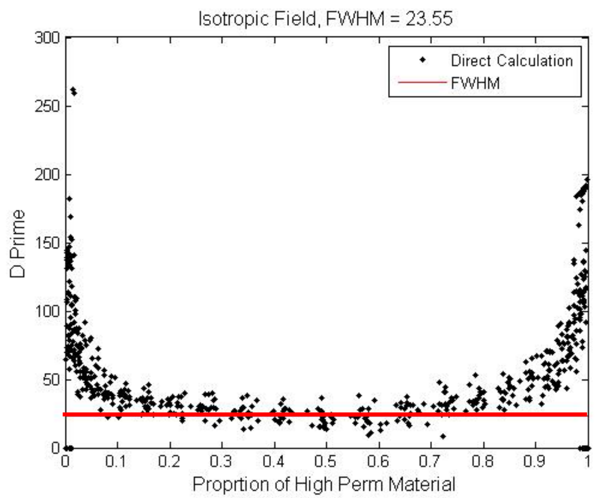
$$D' = \lambda - 2\sqrt{\frac{E(N)}{\pi \cdot EC}}$$

Subtract mean inclusion diameter (assuming circular shape)

Case when D' < inclusion diameter?

$$if \quad \left[\lambda - 2\sqrt{\frac{E(N)}{\pi \cdot EC}}\right] \leq 0.0, \quad D' = FWHM$$





Direct calculation of D' from 20 truncated fields with FWHM (23.55) as approximation of minimum D' value

SubGrid Model Testing

Initial test with two orders of magnitude difference in the permeabilities

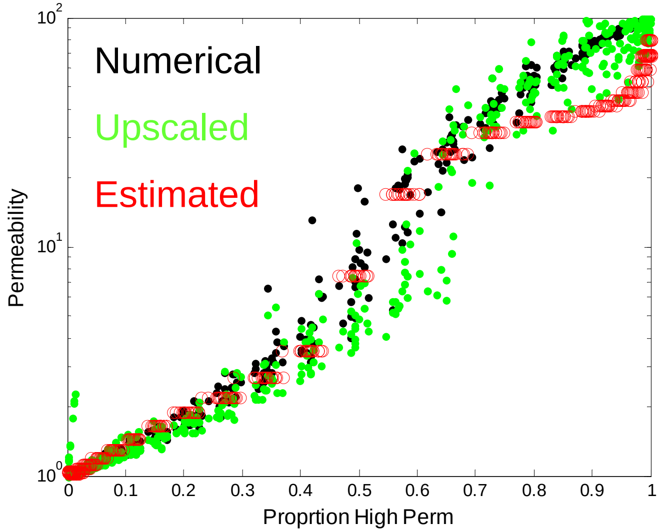
Under-estimation of permeability when high perm proportion > 0.70.

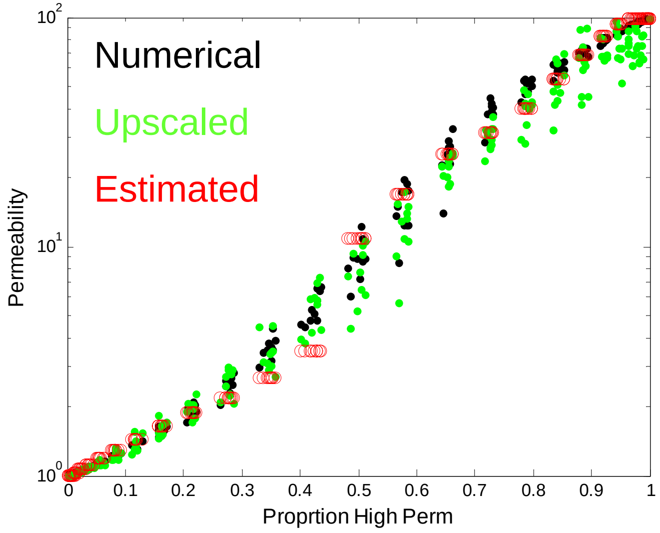
Spatial Point Process assumption not following constraints of complete spatial randomness (CSR) – Extremes are not random, they are dispersed due to maximum entropy nature of Gaussian fields

Added correction to decrease intensity of point process at high proportions:

$$\delta = \left[\frac{p - 0.7}{0.7} \right]^{\omega} \sqrt{FWHM} + 1$$

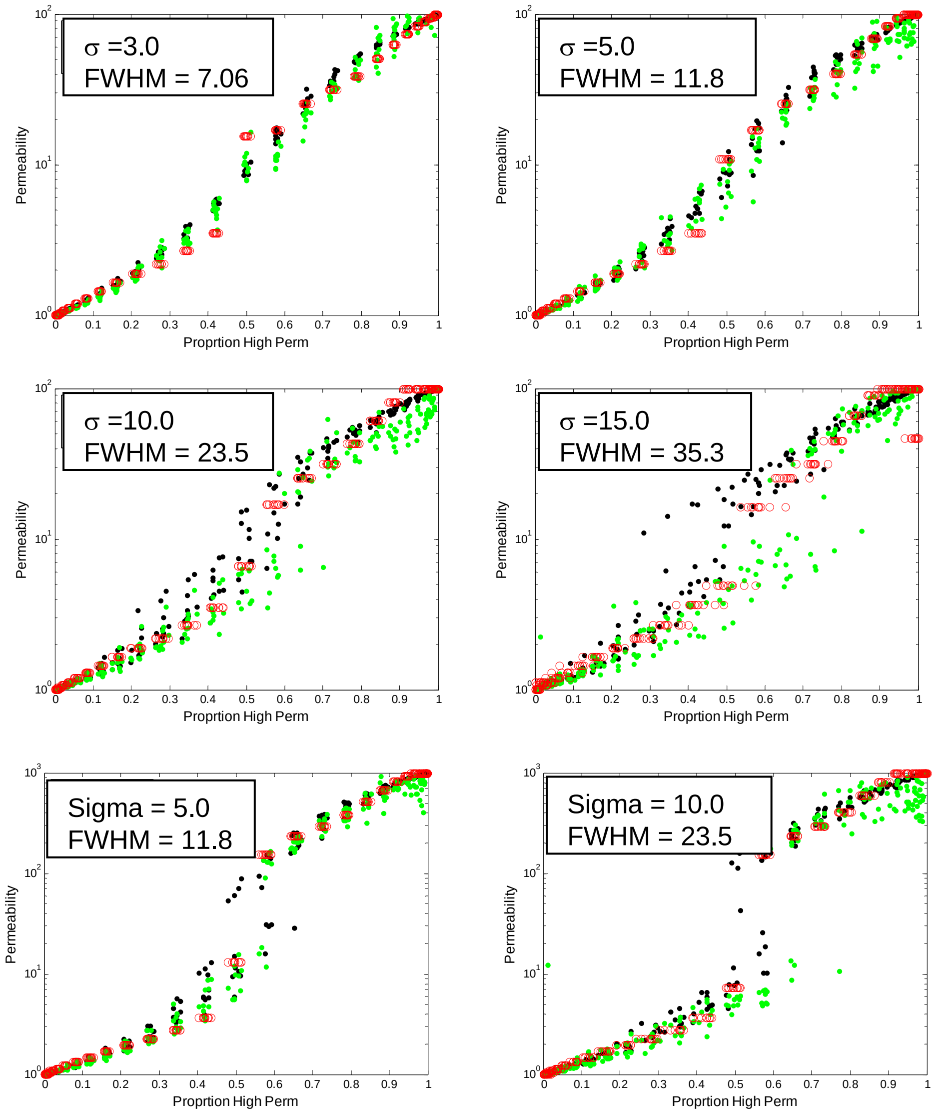
$$\lambda_c = \lambda / \delta$$





SubGrid Model Testing

Additional testing to check multiple kernel sizes and 2 and 3 orders of magnitude difference in permeabilities



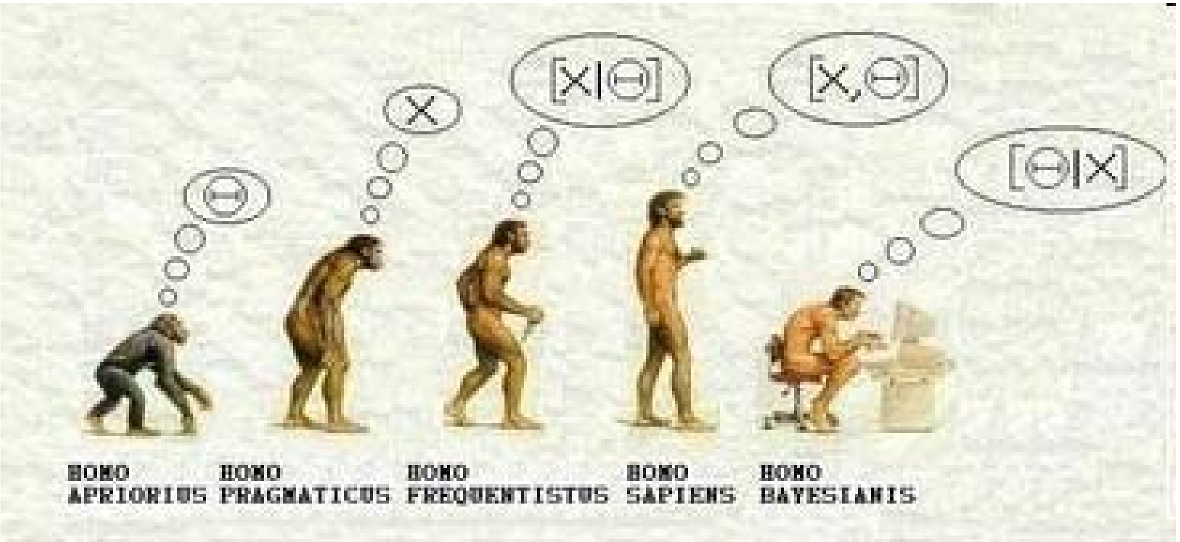
<u>Legend:</u>

Numerical result

Upscaled(Knudby et al., 2006)

Estimated (Subgrid model)

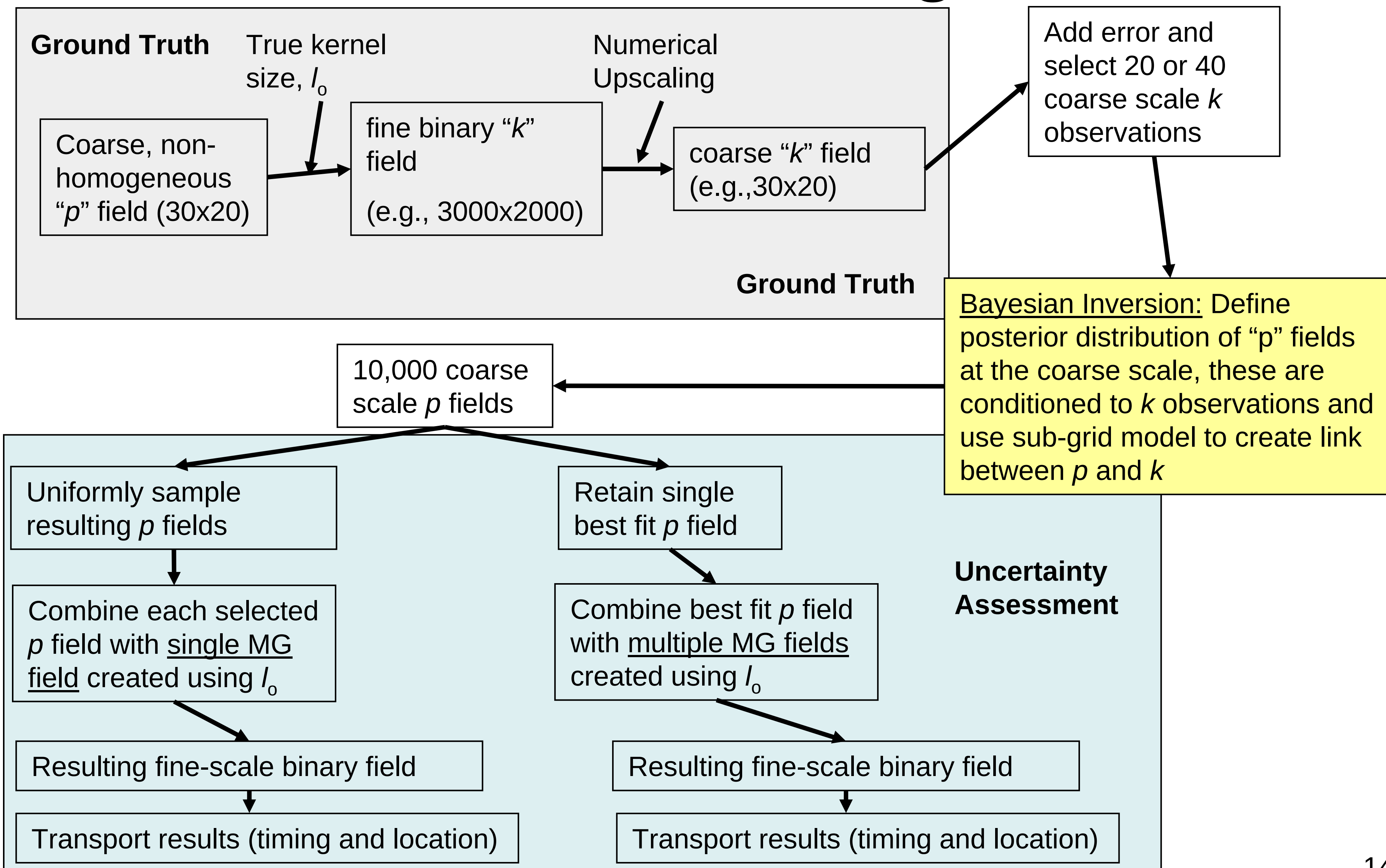
Bayesian Inversion



From: bayes.pl/bayesian.html

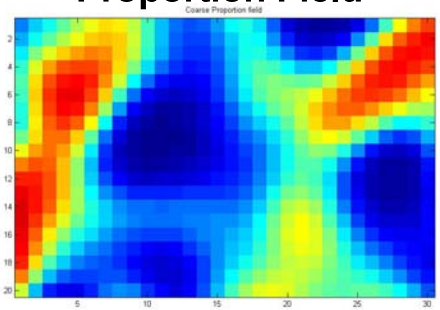
- Why Bayesian inversion?
  - Inverse solutions are non-unique – define uncertainty in solutions as a posterior distribution consistent observations
  - Latent variables – unobserved secondary variables connected to primary variable through a model relationship (example of proportion of high permeability material shown here)
  - More flexibility for incorporating probabilistic relationships between variables

Inversion and Testing Process

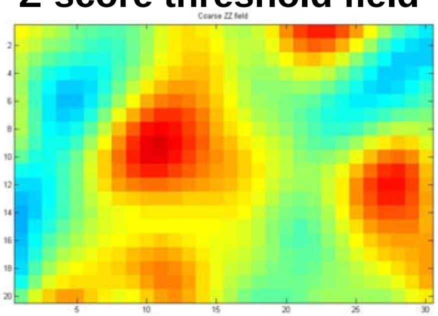


MultiScale Ground Truth

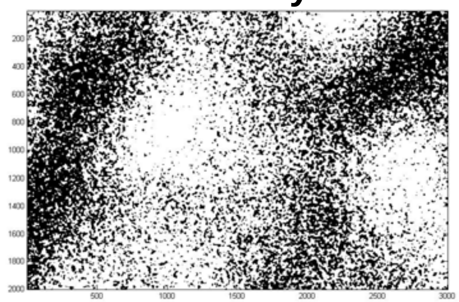
Proportion Field

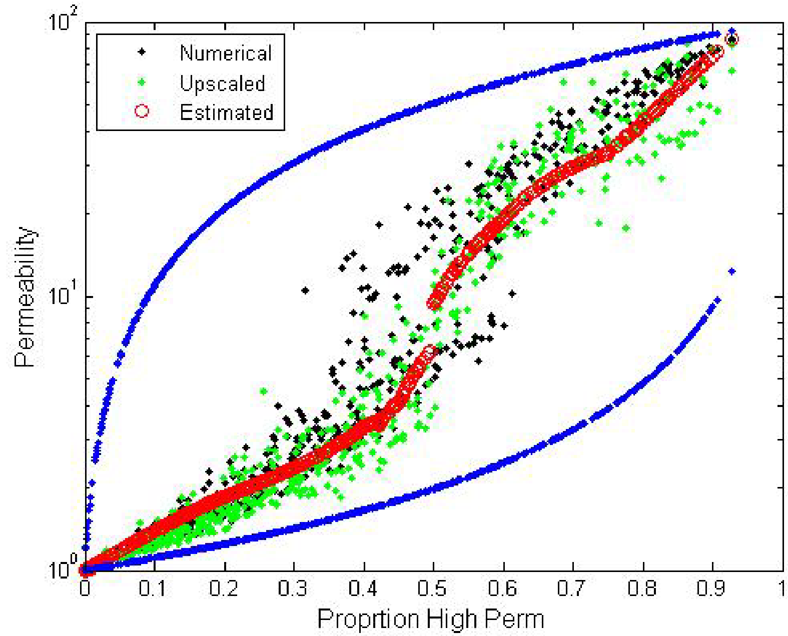


Z-score threshold field



Fine Scale Binary Medium





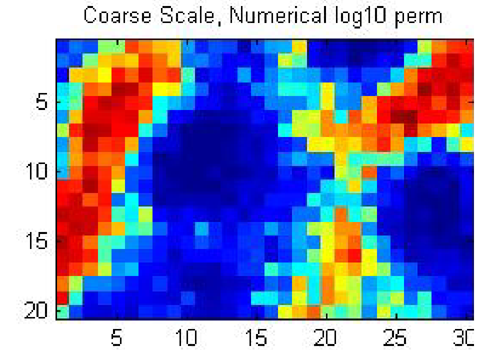
Kernel size (sigma = 5.0)

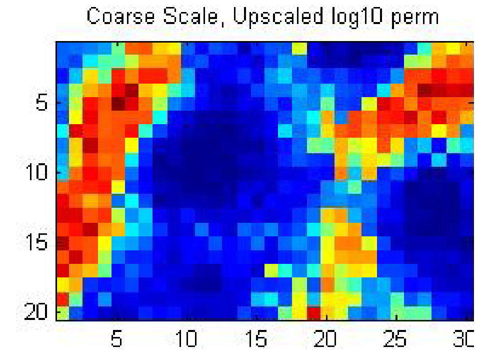
K1 = 100, k2 = 1

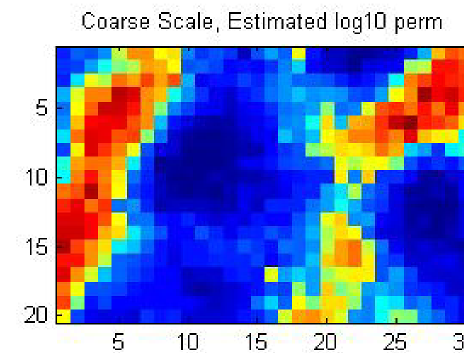
Coarse = 20x30

Fine = 2000x3000

Coarse block is 100x100 fine cells



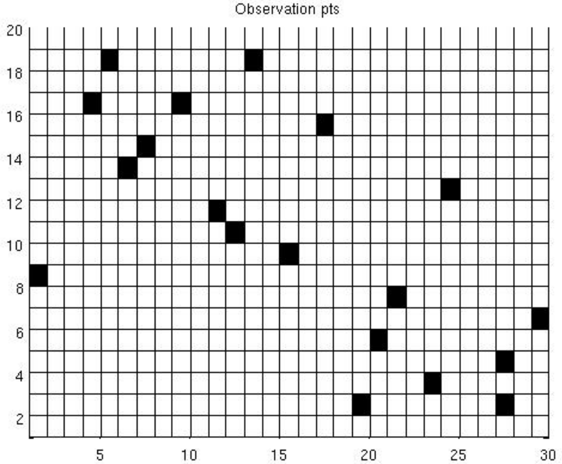




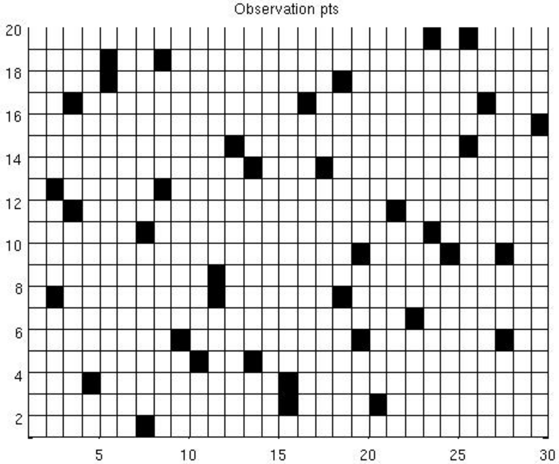
Observations

- Sample ground truth permeability (coarse, numerical) field at 20 and 40 locations
  - Sampling pattern from Latin Hypercube Sampling
- Given subgrid model to estimate *k* from *p*
  - Compare k_{est} to k_{num} at observation locations to define model error

20 obs. points



40 obs. points



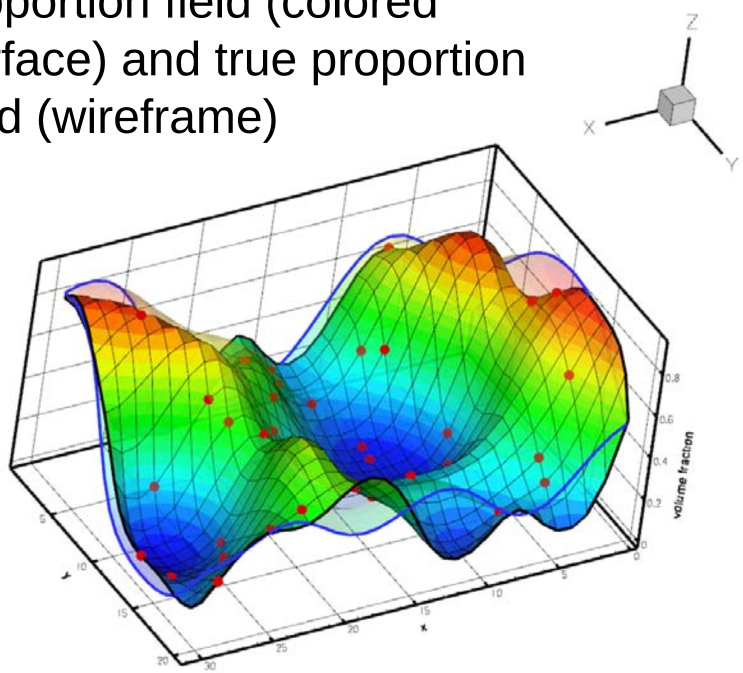
Karhunen-Loève Background

- Prior information is a Gaussian process *p* field with Gaussian covariance (correlation length is 10% of domain diagonal)
  - Convert realization of Gaussian process to proportion (bounded) using cdf and error function
- K-L expansion of Gaussian process with enough modes to capture 99% of variance
  - Decreased number of parameters are weights of eigenvectors of covariance matrix
  - Basis functions and rate of eigenvalue decay selected to maintain modeled covariance

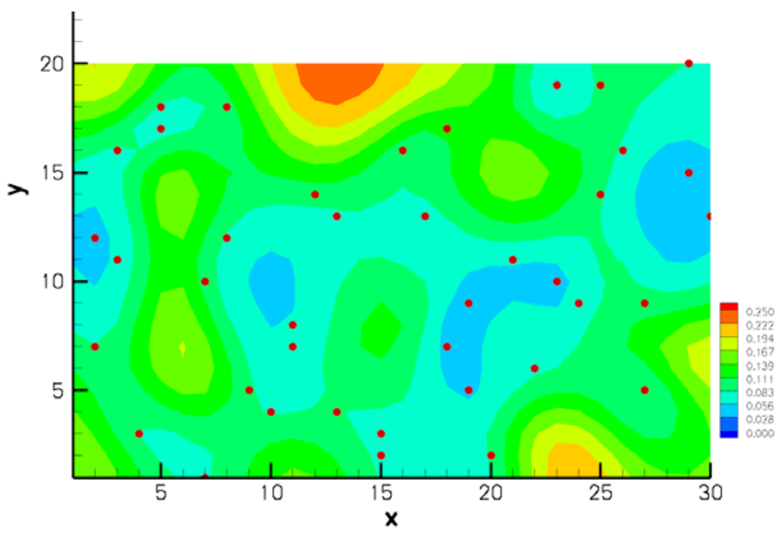
Results

Example results of inferring proportion field from 40 observations of permeability. Results were obtained using 30 KL modes and 300,000 iterations (run time approximately: 17 minutes)

Observed proportion values (points), posterior mean proportion field (colored surface) and true proportion field (wireframe)



Posterior standard deviation of inferred proportion values with observation locations

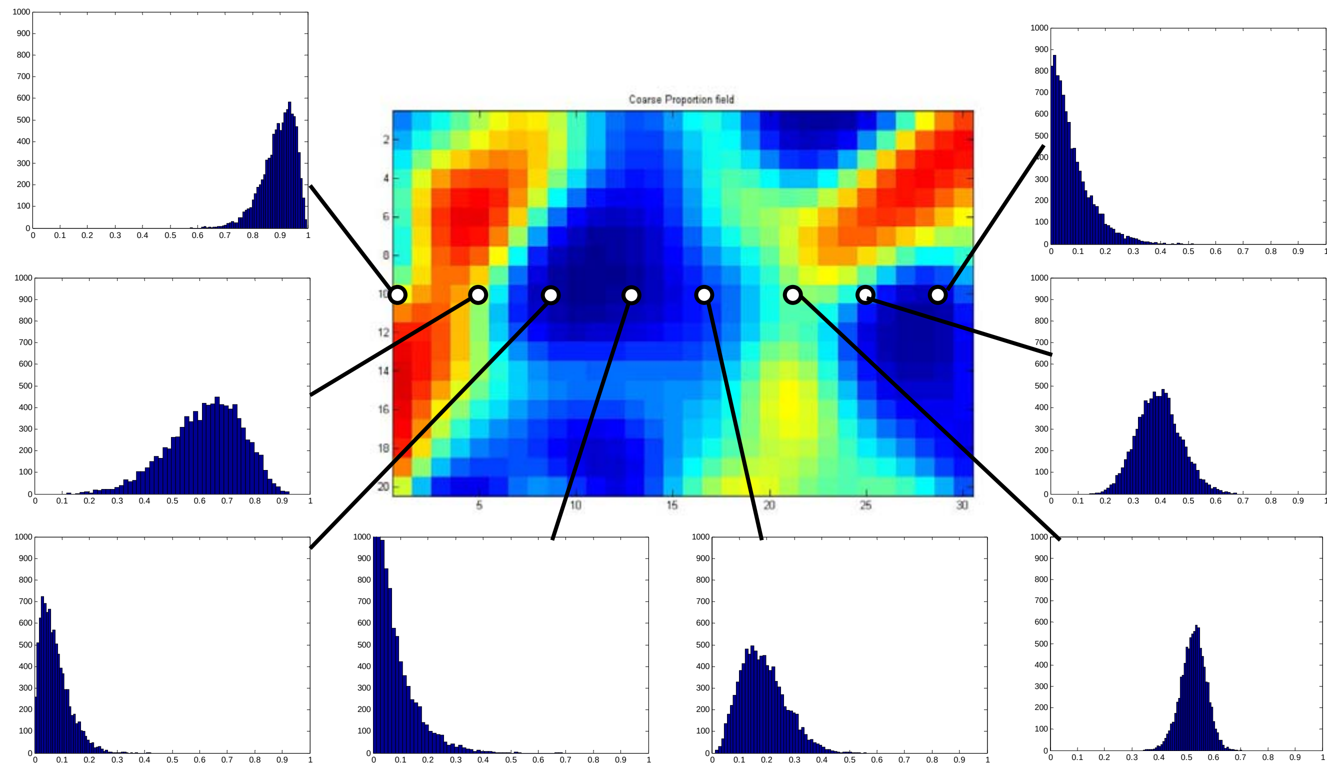


Mean and std. deviation calculated over 9000 realizations

Inference Performance

- MCMC runs met convergence diagnostics
- Posterior mean stabilizes with increasing KL modes (30-45)
- Results generally obtained with 300,000 iterations
  - Approximately 17 minutes on workstation
  - Results in 9000 realizations of proportion field

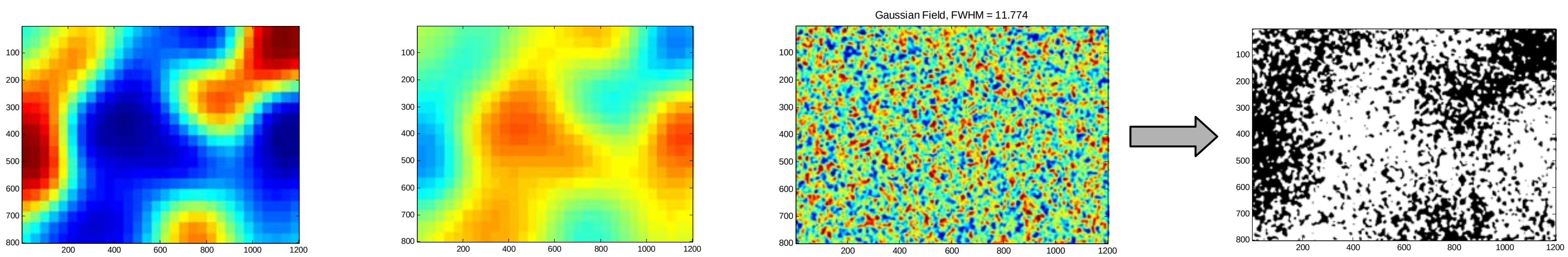
Comparison of posterior pdfs for eight points on proportion field



Flow and Transport Model

- Inferred proportion fields provide threshold for truncation of multiGaussian (MG) field
  - Results shown for 40 observations and 30 KL modes

P field | Z field | MG field | Binary field



*Z = (-1.0)*G^{-1}(p;0,1)*

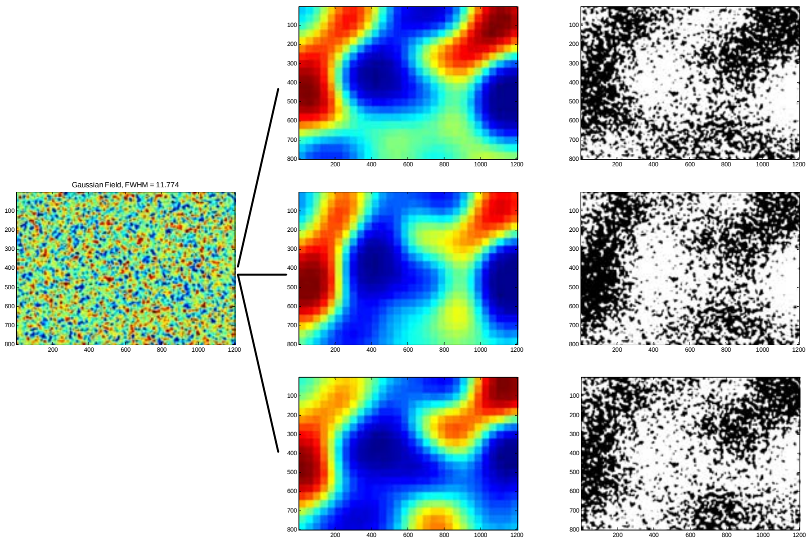
If (MG - Z > 0.0), Binary = 1, else 0

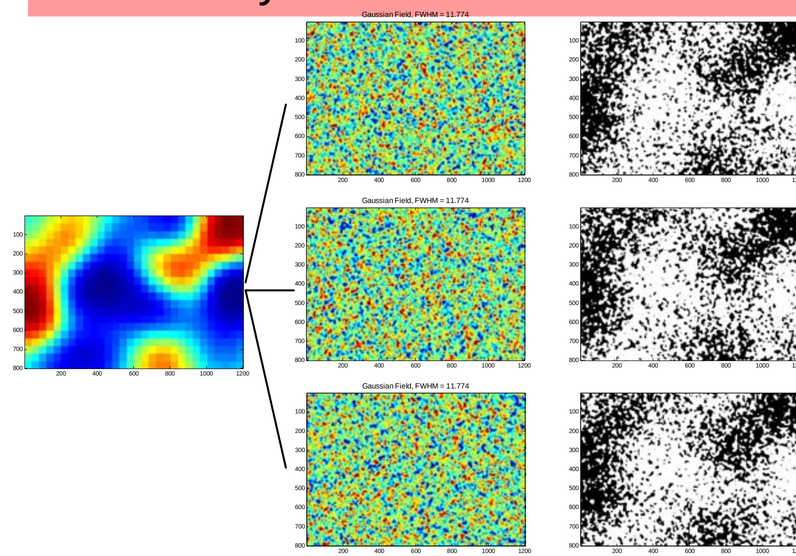
- Fine-scale (1200x800) binary fields are input to flow model
  - Two order of magnitude difference in perms
  - Steady state, single-phase permeameter BC’s
  - 2000 particles (streamlines) tracked across domain

Flow and Transport Results

One MG field and uncertainty in inferred proportion fields

One inferred proportion field and uncertainty in multi-Gaussian fields

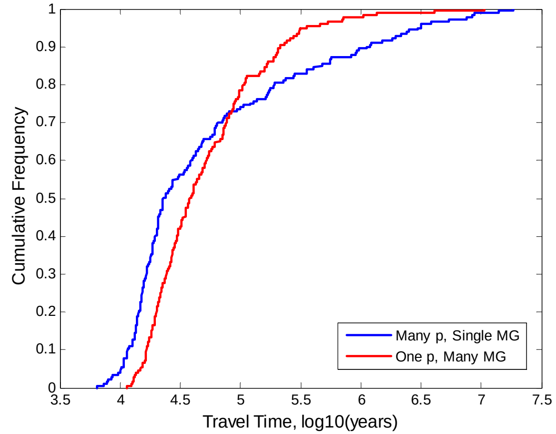




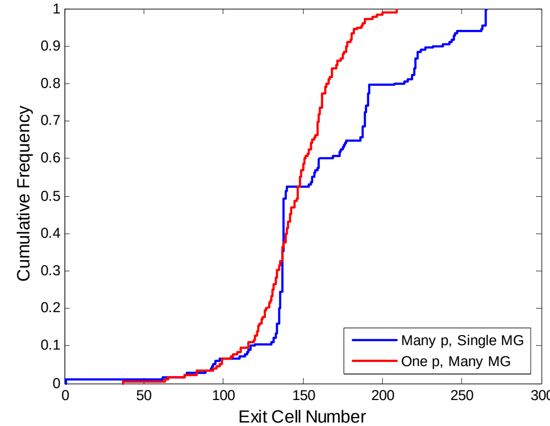
MultiGaussian and binary fields created at fine scale: here 1200x800

Single MG field with multiple proportion fields create greatest variation in travel time and exit location

Travel Time



Exit Coordinate



Summary

- Developed new subgrid model for binary media:
  - Exploits properties of truncated Gaussian fields
  - Based on recent upscaling algorithm (flow distances)
- Multiscale modeling with Bayesian inversion
  - Demonstrated practical approach to multiscale Bayesian inversion
  - Two scales with robust link function between them
  - Two-levels of uncertainty (fine scale from MG and coarse scale from realizations of proportionality)
- Future work
  - Extend subgrid model to include anisotropy in inclusions
  - Utilize dynamic data (pressure, transport) in Bayesian inference
  - Increase size difference between two scales
  - Incorporate uncertainty in upscaling algorithm into Bayesian inversion approach (variance of estimates from subgrid model)