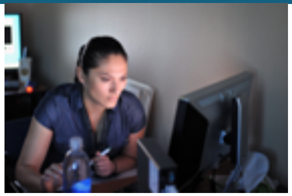
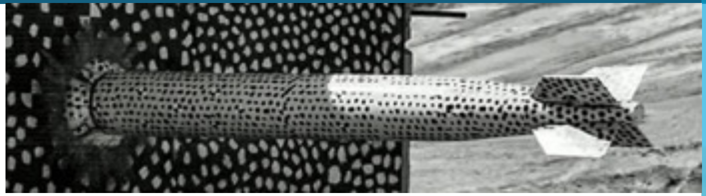
SAND2021-2125 C

Application of Multifidelity Uncertainty Quantification Methods to a Subsurface Transport Model





SIAM CSE21

T. Portone, M.S. Eldred, G. Geraci, and L.P. Swiler

This work has been funded by the Geologic Disposal Safety Assessment (GDSA) program, under the Spent Fuel and Waste Science and Technology Campaign of the Office of Nuclear Energy in the U.S. Dept. of Energy.

UQ for geologic disposal safety assessment (GDSA)

Performance assessment for nuclear waste repository site.

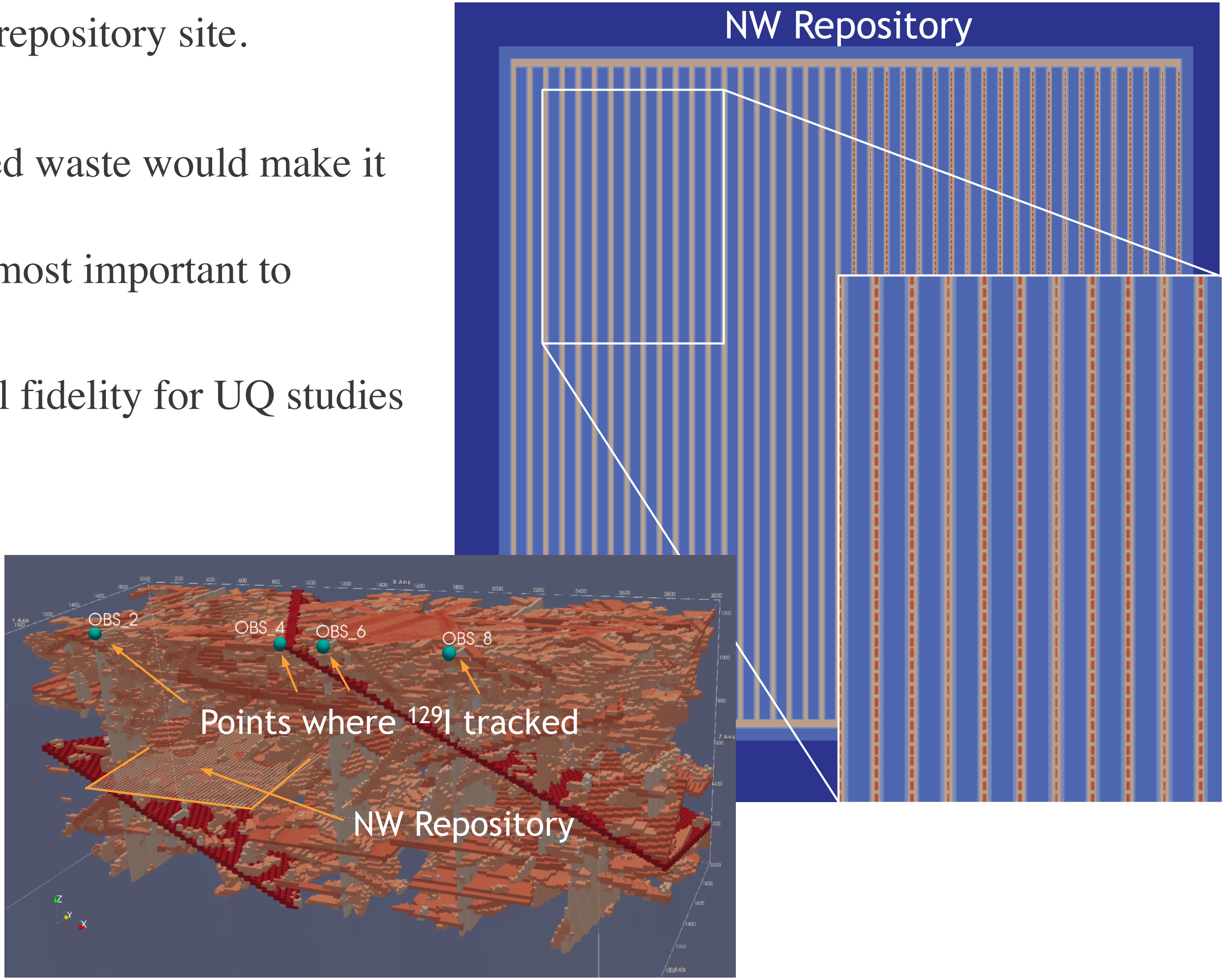
Want to understand things like:

1) probability dangerous levels of leaked waste would make it to a water supply;
2) subsurface/waste storage properties most important to repository performance.

Deploying an unprecedented level of model fidelity for UQ studies in this application area.

Simulations very costly (~1.5 hours on 512 cores per simulation): $\mathcal{O}(1000)$ model evaluations for production UQ studies.

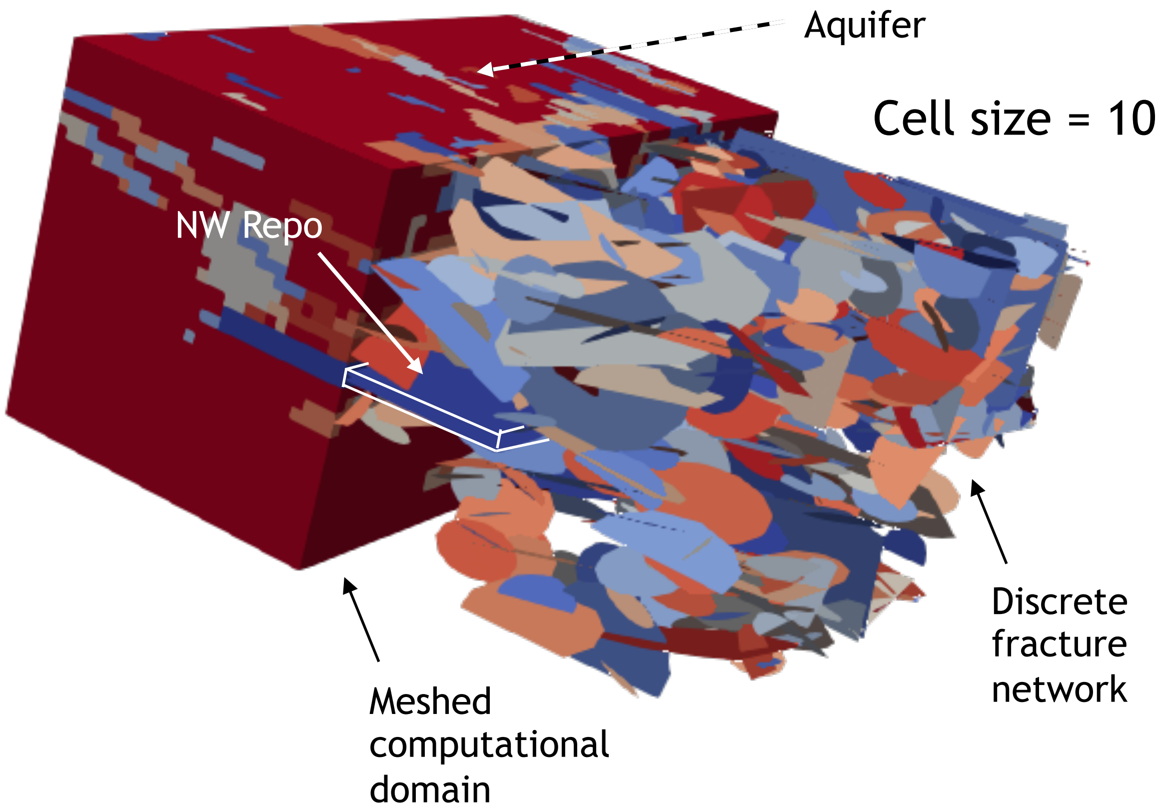
Goal: explore feasibility of using multifidelity polynomial chaos expansion (PCEs) for global sensitivity analysis (GSA) in this problem.

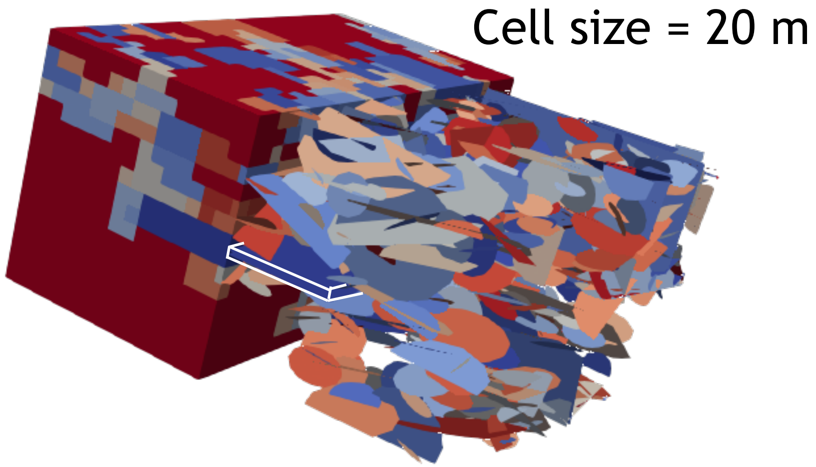


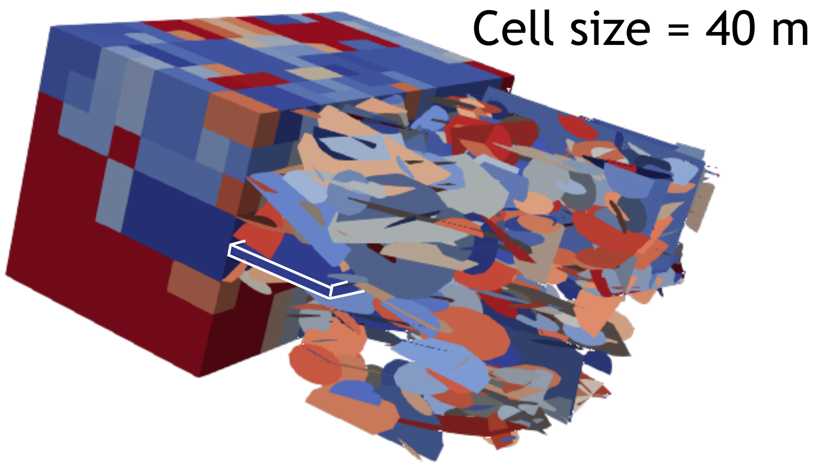
Fractured subsurface uncertainty treatment and model hierarchy

- Can't observe all fractures in subsurface; represent uncertainty using random instantiations of discrete fracture networks (DFNs).

- DFNs mapped to equivalent continuous porous media (ECPMs).
  - Define grid, sweep through cells and map fracture permeability and porosity to equivalent continuum values.



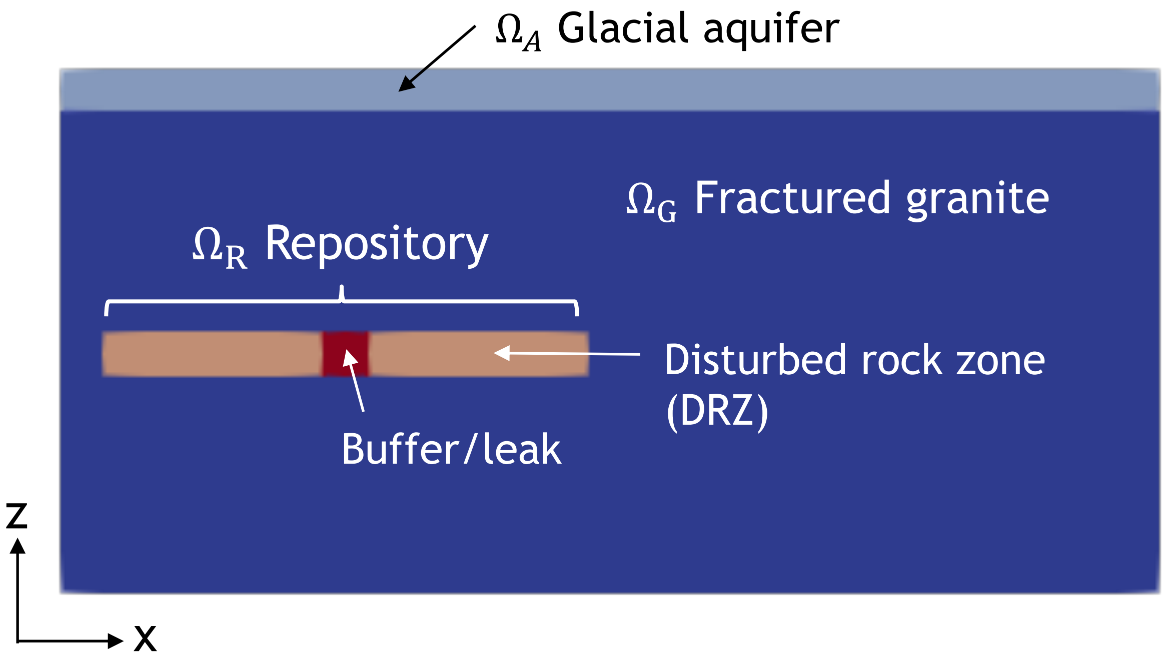




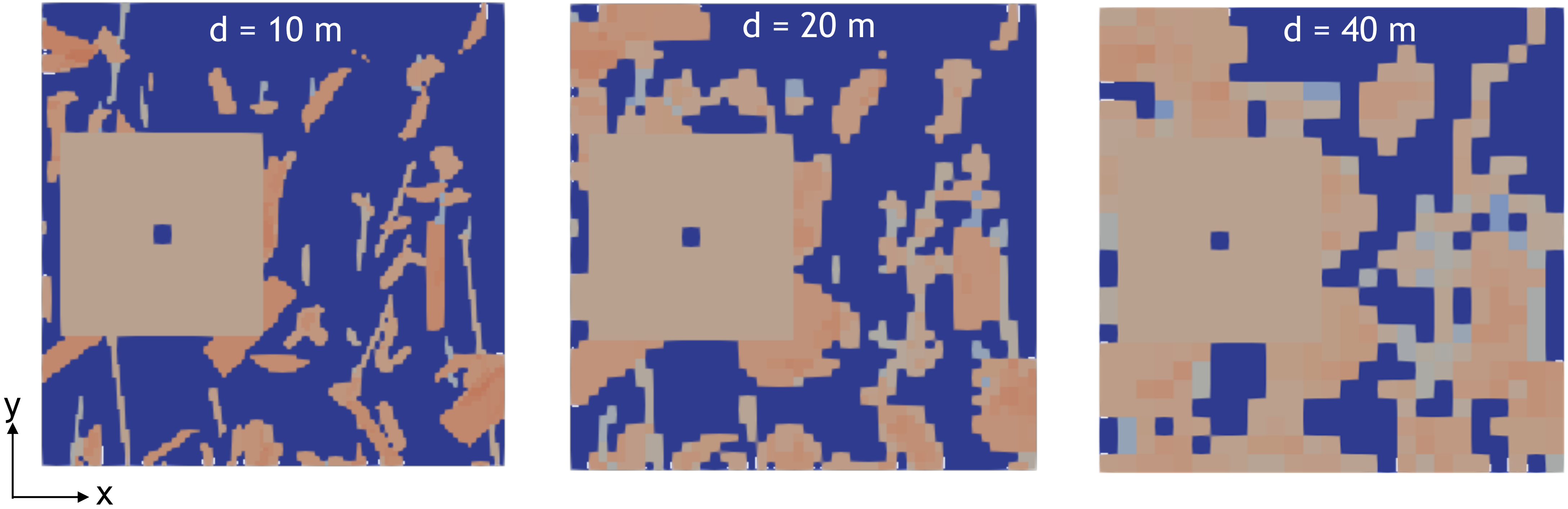
Domains generated using https://dfnworks.lanl.gov.

Simplified crystalline problem

- 1000 x 1000 x 500 m
- 1 waste package in buffer region
- Same underlying DFN statistics as production problem.
- Same QoIs as production problem.

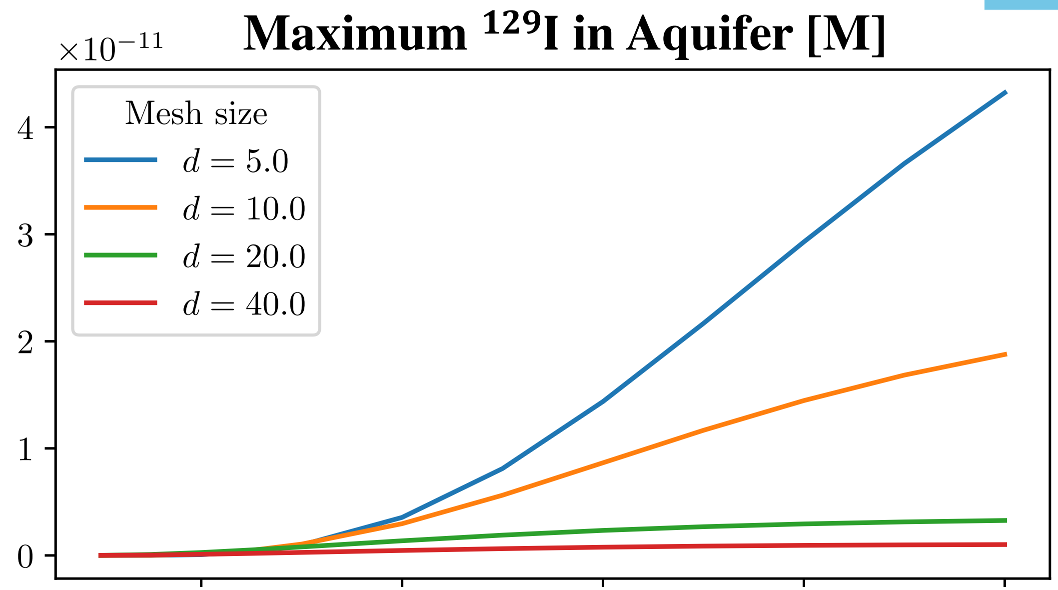


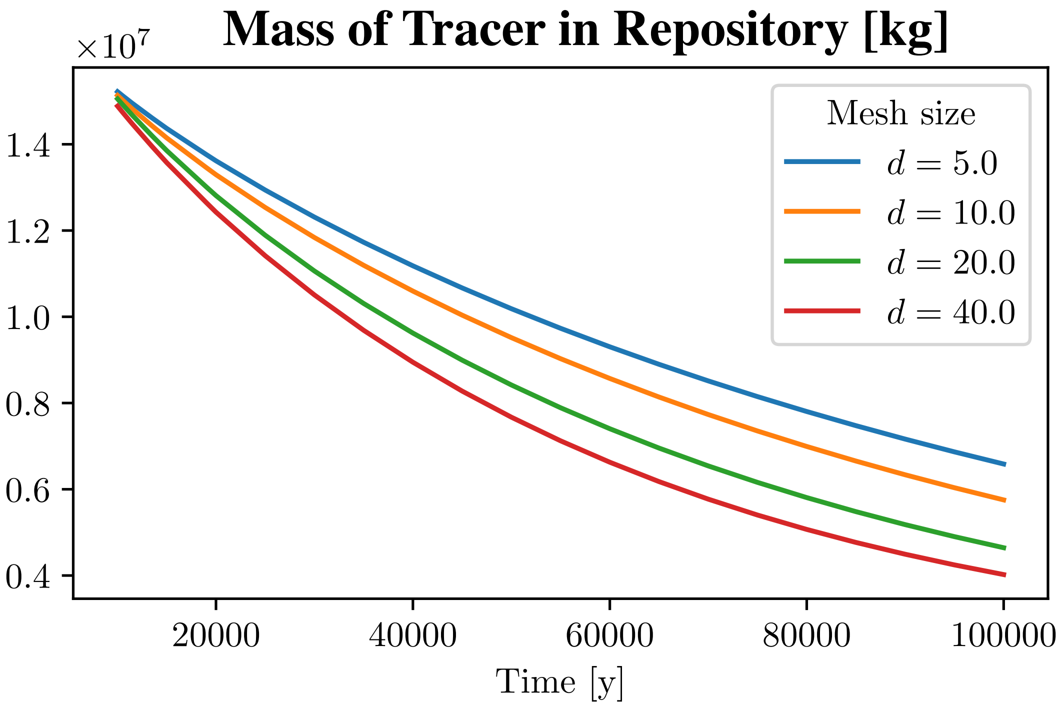
K_{xx} view from above:



Challenge: mesh dependence of quantities of interest (QoIs)

- Behavior of several QoIs converges as mesh is refined, but some do not.
- Max ^{129}I concentration in aquifer affected significantly by numerical diffusion/dispersion.
  - No correlation in location of max across meshes.
  - Some derived quantities depend on location of max.
- Mass in repository affected by increased flow from false connections in coarse mesh.
- Using the finest mesh we can afford.





Multifidelity Polynomial Chaos Expansion (MF PCE) for GSA

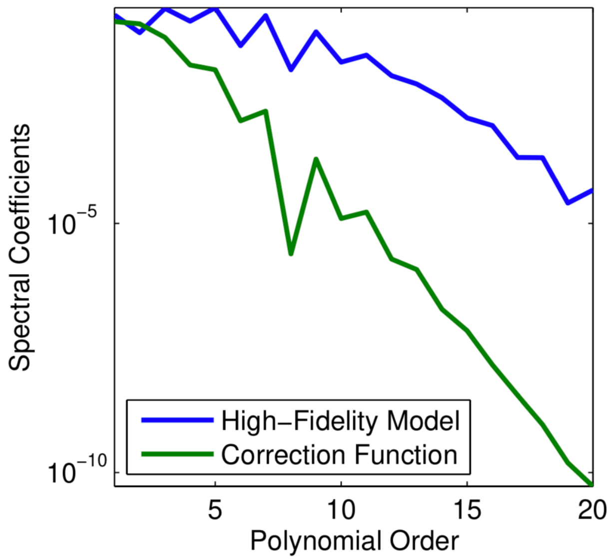
- Discrepancy-based, hierarchical multifidelity PCE (Ng, Eldred 2012).
- MF PCE coefficients can be recombined and used to compute Sobol indices in postprocessing (Sudret 2008).
- (R – model response; P – PCE expansion):

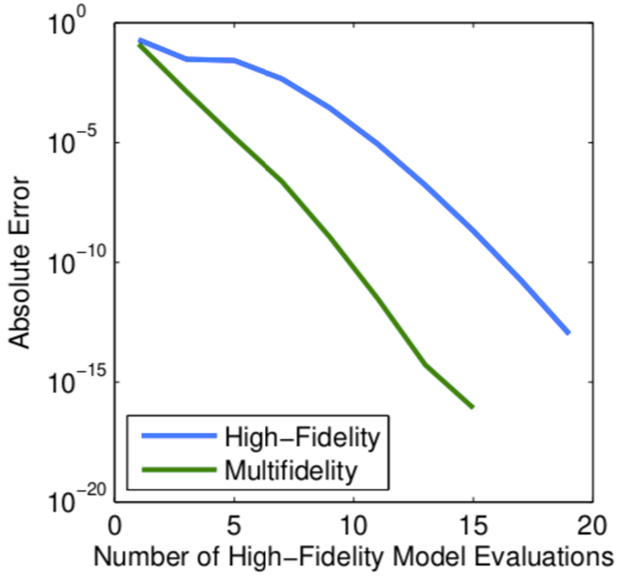
$$\begin{aligned} R_{10} &= R_{10} + (R_{20} - R_{20}) + (R_{40} - R_{40}) \\ &= R_{40} + (R_{20} - R_{40}) + (R_{10} - R_{20}) \\ &\approx P_{40} + (P_{20-40} + P_{10-20}) \end{aligned}$$

- If spectral content of $P_{10-20} \ll P_{10}$, need fewer 10 m evaluations to derive accurate expansion of R_{10}.
- In this case using a recursive approach to construct discrepancy PCE:

$$P_{20-40} \approx R_{20} - P_{40} \text{ rather than } \mathrm{P}_{20-40} \approx R_{20} - R_{40}.$$

- Random (MC) sampling + coefficients estimated with regression + cross validation for PCE order.
- Sampling and PCE construction performed with Dakota.





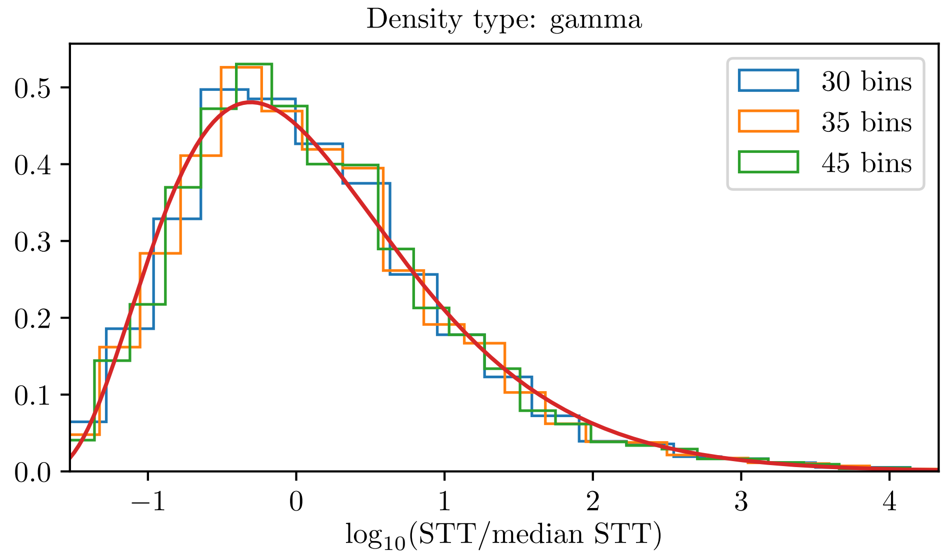
Characterizing subsurface uncertainty

- How do subsurface properties affect quantities of interest through GSA?
  - Can't impose subsurface properties directly-–have to compute them after generating a DFN.

- Compute proxy variables based on graph/network representation of each DFN (fractures = nodes, intersections = edges)
- Looked for proxy variables that were correlated with QoIs.
  - Number intersections with repository (NIwR)
  - Average degree (average number of intersections a fracture is part of)
  - Shortest travel time (STT) between repo and aquifer (rough approx. based on fracture area and pressure differentials)

- **These are derived quantities and therefore can't be prescribed by adaptive algorithms**; use regression.
  - Fitted analytical distributions for proxy variables from ~3000 sample DFNs.
  - Augment build point input space with proxy variables for each DFN

- **This is the first time we have been able to account for subsurface variability in GSA**
  - However, these are only proxies and cannot completely capture the induced variance from subsurface.



Preliminary multifidelity PCE study with subsurface uncertainty

- Each sample model evaluation was generated with a different DFN (different subsurface instantiation)
  - Input parameter space augmented with proxy variables
  - Constructed regression PCE over augmented input space

- Constructed PCE with 828 samples from 10 m model for comparison

- Hand-selected number of samples at d = 10, 20, and 40 for preliminary studies.

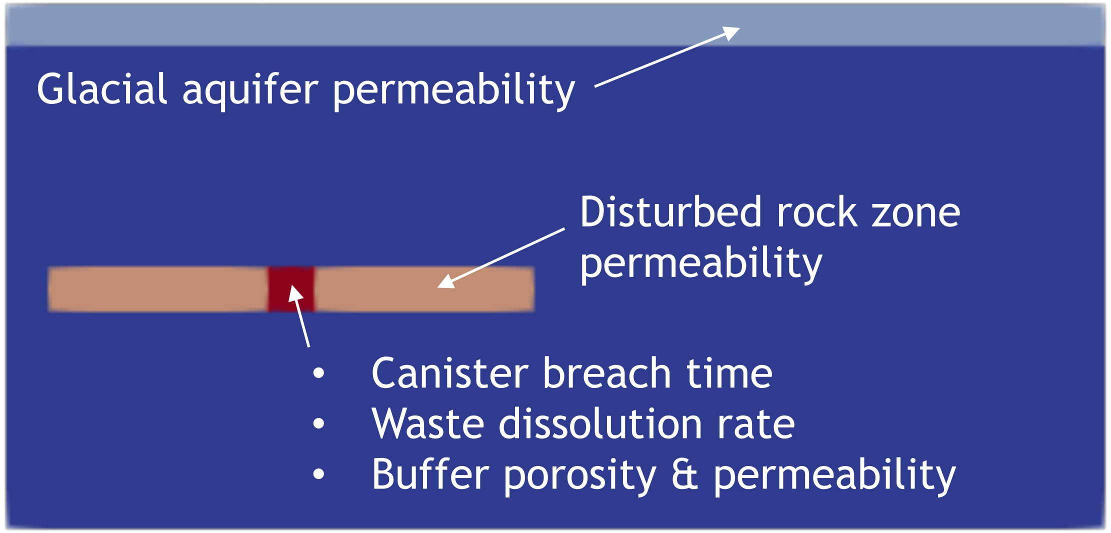
- Interested in potential benefits of augmenting small number of high-fidelity (HF) samples with samples from coarser meshes.

- Generated single-fidelity PCE with 18 HF samples.

- Generated multifidelity PCE augmenting HF samples with 828 samples from d = 20, d = 40.

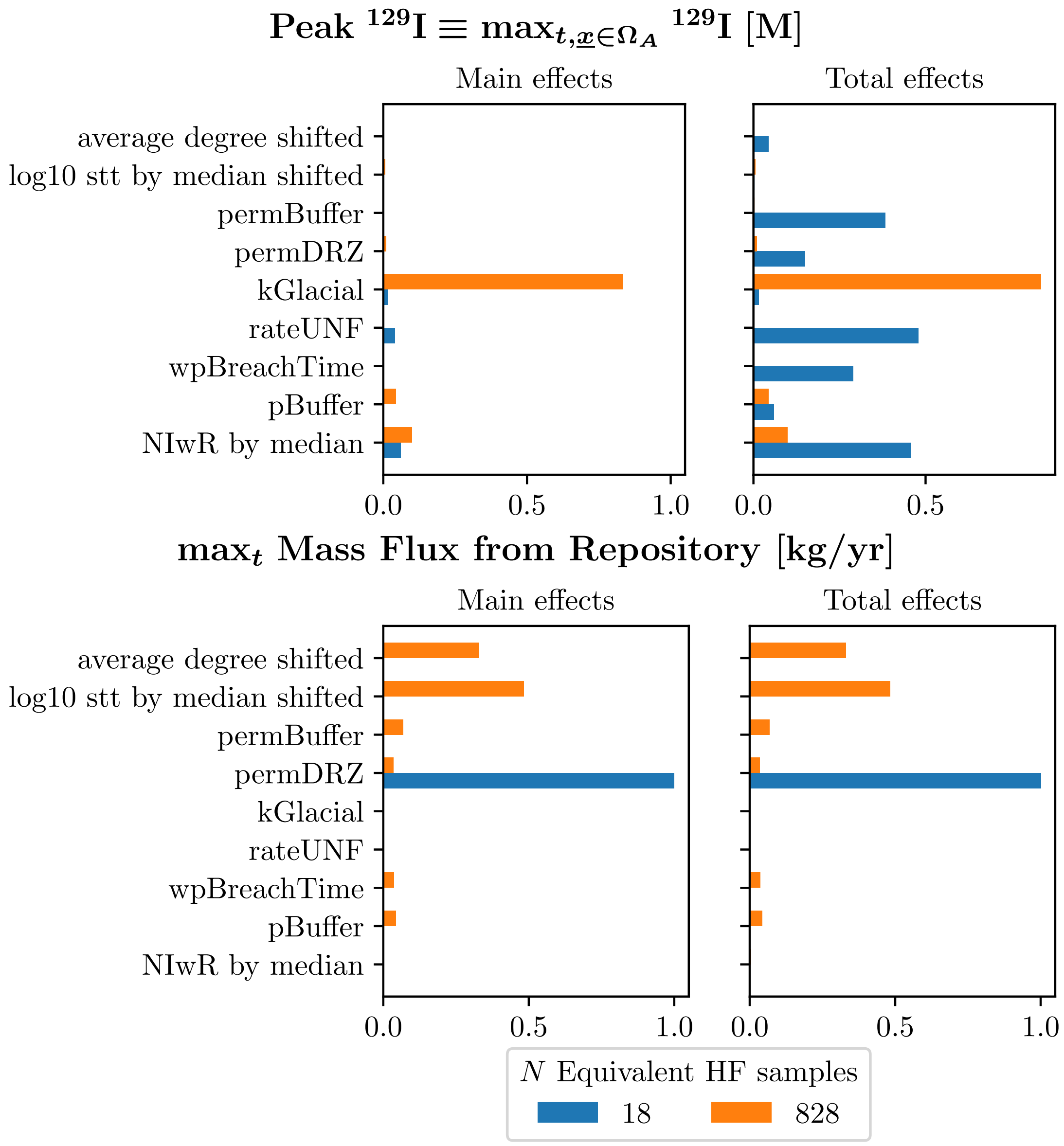
	d = 10	d = 20	d = 40
N samples	18	828	828
Relative costs	1	0.02	0.006

Uncertain parameters:

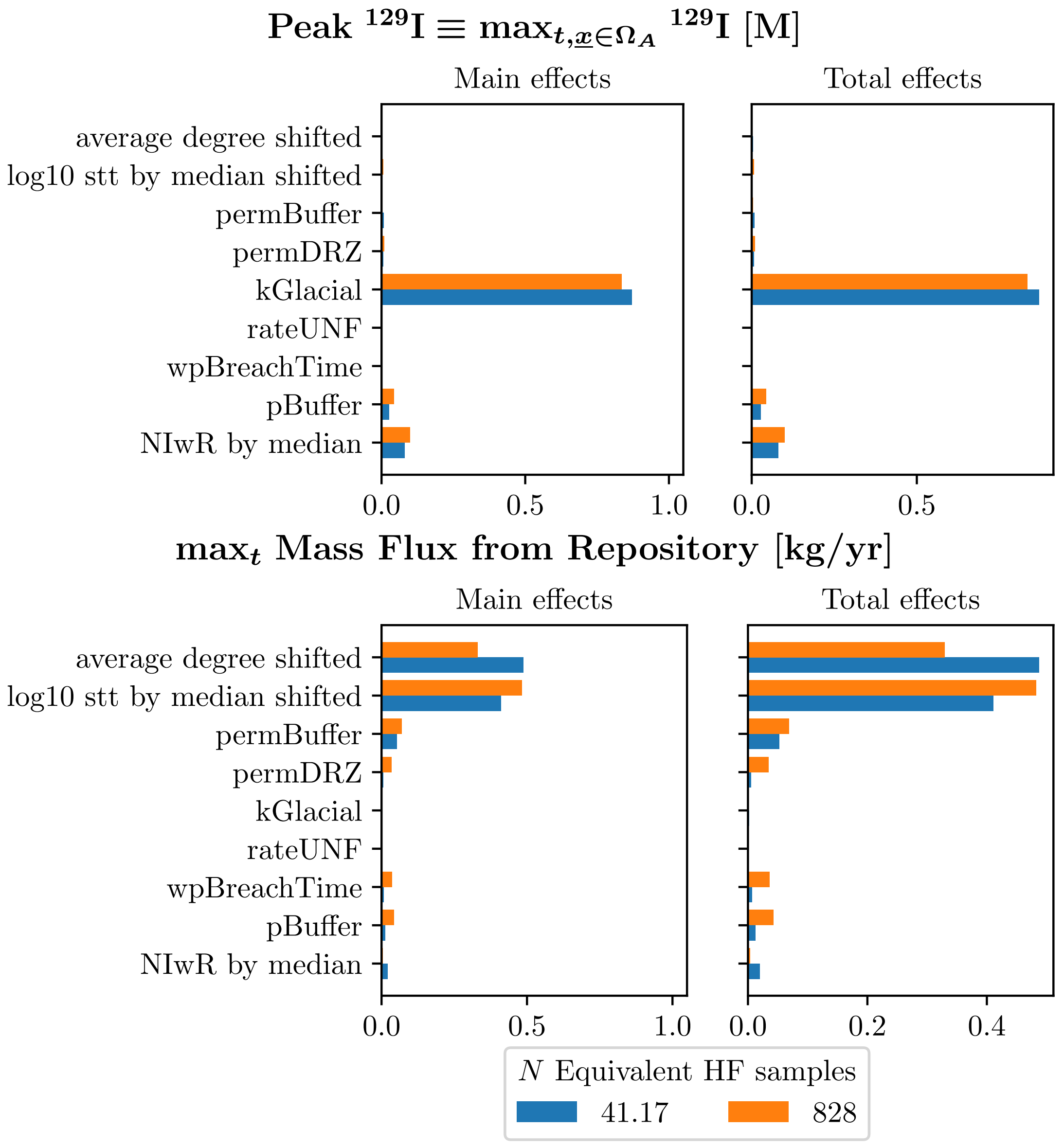


For all but the QOIs depending on location of peak ^{129}I, estimation of Sobol indices was drastically improved with the multifidelity PCE.

Single-fidelity PCE, HF model

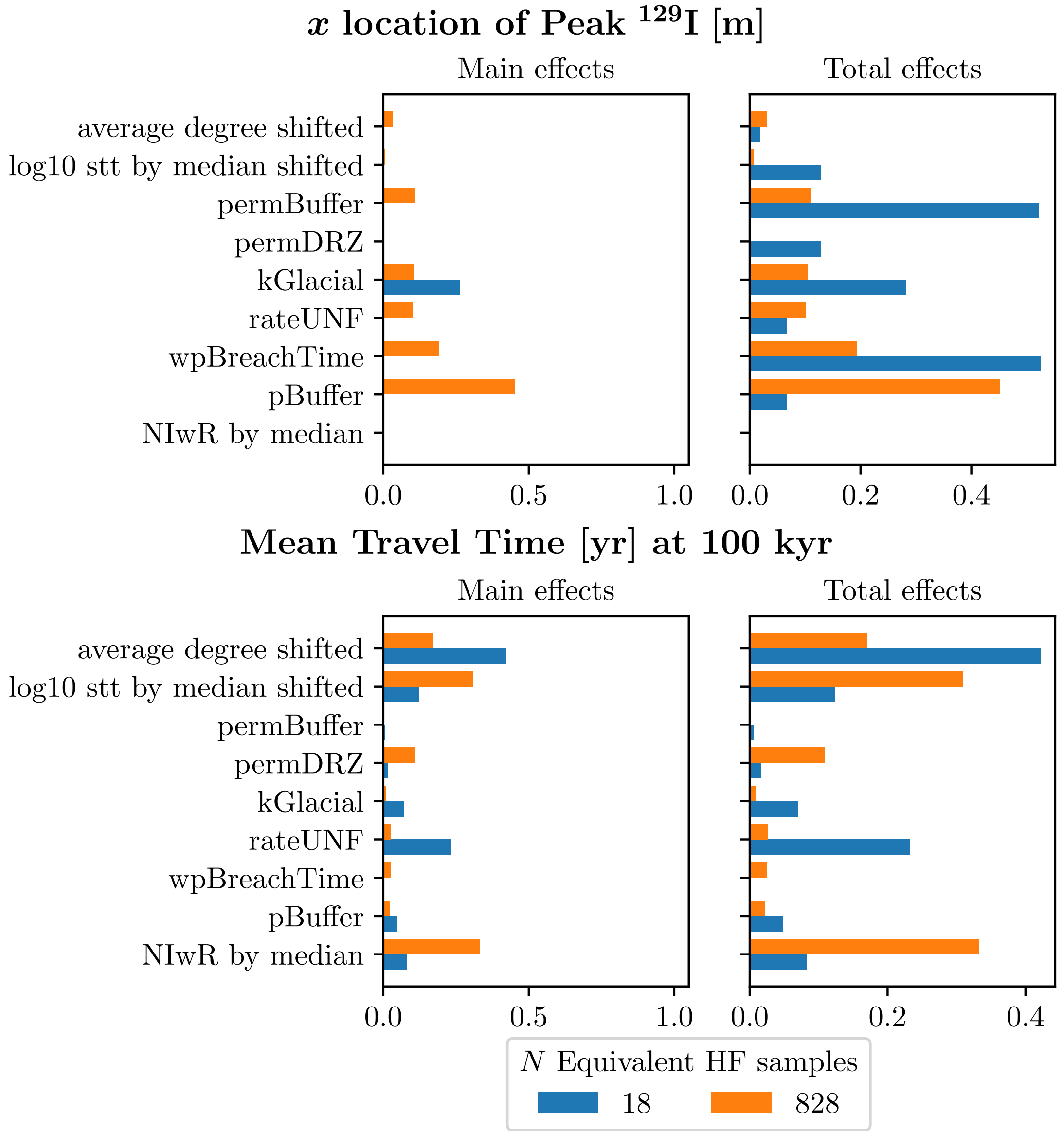


Multifidelity PCE

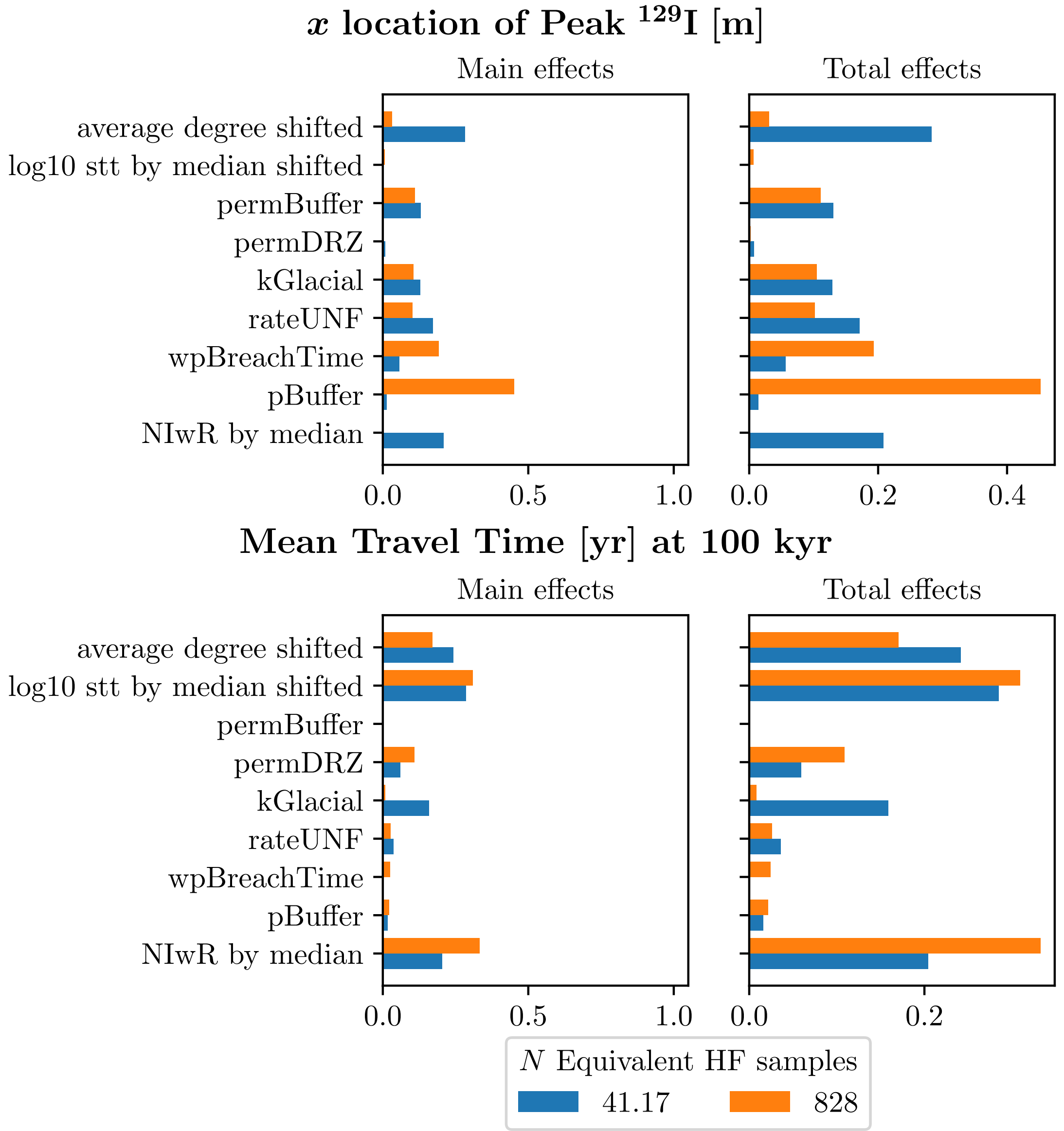


For QOIs depending on location of peak ^{129}I, multifidelity PCE Sobol indices were not significantly improved.

Single-fidelity PCE, HF model



Multifidelity PCE



Takeaways/future directions

- Overall, these results show there is promise in pursuing MF PCE for GSA in this problem.

- However, performance for QoIs depending significantly max ^{129}I location was poor.
  - Most important QoIs for performance assessment not fully settled yet; some of these QoIs may not factor into eventual performance assessment.

- Need to determine reasonable model hierarchy and reasonable QoIs.

- Challenge: how to take advantage of adaptive algorithms/optimize sample profiles in presence of proxy variables?
  - Optimal sample allocations for multifidelity approximation of Sobol indices with PCE – see Michael Merritt's talk later in the session!