Scalable spatio-temporal modeling using a fast multipole method for 3D tracer concentration breakthrough data with magnetic resonance imaging

Jonghyun Harry Lee[1], Chao Chen[2], Toru Takahashi[3], Eric Darve[4], Hongkyu Yoon[5]

[1]University of Hawaii at Manoa, [2]University of Texas at Austin, [3]Nagoya University, Japan, [4]Stanford University, [5]Sandia National Laboratories

Motivations

With recent advances in sensing technology, large volumes of hydrogeophysical and geochemical data can be obtained to achieve continuous tracking of the movement of a fluid or a plume in the subsurface. However, characterization with such a large amount of information requires prohibitive computational and storage costs associated with matrix construction, matrix-matrix multiplication, and linear system solution. To tackle such challenges, we present a spatio-temporal modeling without explicit construction of the covariance matrix, and take advantage of the parallel black-box fast multipole method (FMM) and the parallel inverse Fast Multipole Method (IFMM) for matrix-vector multiplication and linear system solution, respectively. Overall, our approach requires O(N) computation and storage. For an illustrative example, we use 6 million transient tracer concentration measurements in a laboratory-scale 3-D sand-box obtained using magnetic resonance imaging to monitor real-time tracer plume migration. The sand-box was filled with discrete patterns of 5 different sizes of sands at 1 cubic centimeter, creating relatively heterogeneous patterns of permeability distribution and associated complex transport of the tracer. We demonstrate that the spatio-temporal modeling can be performed with a big data set such as this concentration data for real-time plume tracking.

Methods

1. Spatiotemporal modeling [1]

In this work, we focus on the simple (spatio-temporal) Kriging type method for spatiotemporal interpolation of a tracer plume:

$$\hat{\mathbf{s}} = \mathbf{Q}_{sy}\mathbf{Q}_{yy}^{-1}\mathbf{y}$$

$$\mathbf{Q}_{ss|y} = \mathbf{Q}_{ss} - \mathbf{Q}_{sy}\mathbf{Q}_{yy}^{-1}\mathbf{Q}_{ys}$$

where:

$\hat{\mathbf{s}}(x,t)$ = the best spatiotemporal estimate

$\mathbf{y}(x,t)$ = observation

$\mathbf{Q}_{ss}$ = auto-covariance matrix of the estimate s with kernel K(x,t) = Cov(s(x),s(t))

$\mathbf{Q}_{sy}$ = cross-covariance matrix between estimate **s** and observation **y**

$\mathbf{H}$ = sensing matrix whose columns are a subset of identity matrix at obs. locations

Thus, $\mathbf{Q}_{sy} = \mathbf{Q}_{ss}\mathbf{H}^T, \mathbf{Q}_{ys} = \mathbf{H}\mathbf{Q}_{ss}$

$\mathbf{Q}_{yy}$ = auto-covariance matrix of the estimate s, $\mathbf{Q}_{yy} = \mathbf{H}\mathbf{Q}_{ss}\mathbf{H}^T$

$\mathbf{Q}_{ss|y}$ = posterior covariance matrix

Computational costs for the estimate and posterior covariance are O(N²) and O(N³), respectively.

2. PBBFMM3D: Parallel Black-box Fast Multipole Method [2]

PBBFMM3D accelerates non-oscillatory kernel matrix-vector multiplications

$$\mathbf{Q}_{yy}\mathbf{y} = \sum_{j=1}^{N} K(x_i, x_j)\mathbf{y}_j, \qquad i = 1, \dots, N$$

by hierarchical separation of the problem domain to separate the interactions into near-field and far-field interactions:

1. the near-field interactions are computed exactly
2. the far-field interactions are approximated using low-rank techniques

Therefore,

- PBBFMM3D reduces the costs from O(N²) to O(N) complexity in time/storage.
- PBBFMM3D only requires the ability to evaluate the kernel without the explicit formula.
- PBBFMM3D further accelerate the computation using OpenMP on shared-memory machines.

Methods (cont.)

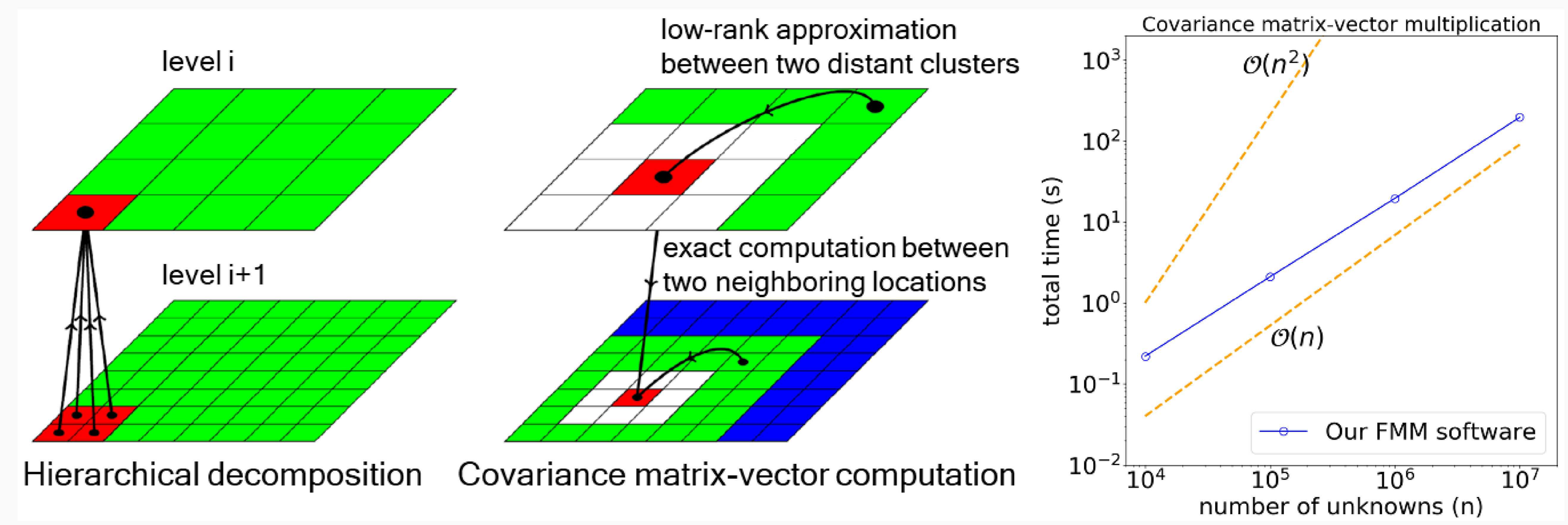


Figure: Schematic illustration of PBBFMM operations (left) and computational scaling for mat-vec operations with different kernels and interpolation functions (right)

- Spatiotemporal problem can be solved efficiently using the Conjugate Gradient (CG) method that only requires mat-vec operations.
- However, for high-dimensional problems, condition number of Q becomes large and the number of iterations will increase.

3. PBBIFMM: Parallel Black-box Inverse Fast Multipole Method [3]

The inverse FMM (IFMM) uses the same hierarchical decomposition structure as in the FMM and construct an approximate direct solver for dense linear systems. IFMM keeps a compact representation of the far-field throughout the factorization and is able to retain the same asymptotic computation and memory complexities as the FMM. Based on greedy coloring algorithm. IFMM is further parallelized using OpenMP on shared-memory machine. In this work, PBBIFMM constructs a preconditioner to reduce the iterations in the CG method.

Data

The entire flowcell has dimensions of 21.5 x 9 x 8.5 cm, and is packed with 1 cm cubes of five different sand type [4]:

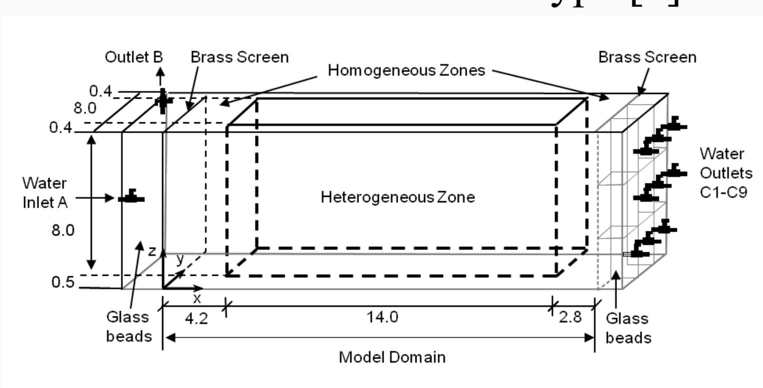


Figure: Illustration of 3-D flowcell

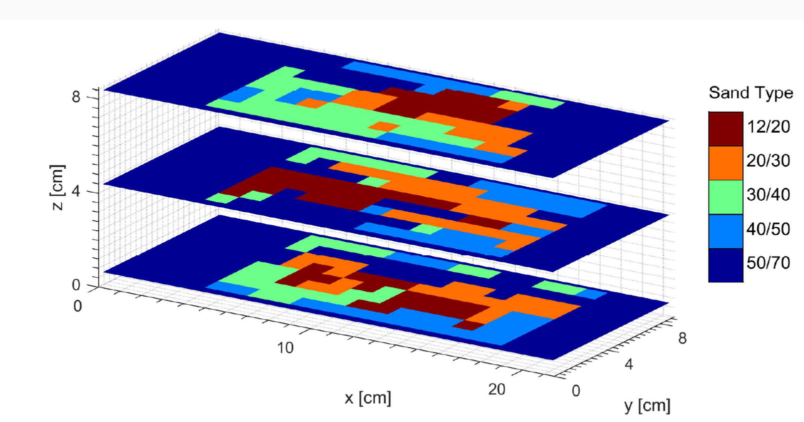


Figure: Sand packing in 3 layers (out of 8 layers)

- Sand distribution was created using SISIM in SGLIB to construct a heterogeneous K filed for the central portion (14 x 8 x 8 cm)
- Constant water flow rate with a uniform tracer concentration
- ~ 6 million transient tracer concentrations were imaged using MRI at a resolution of 0.25³ cm² at a regular interval time over the central region.

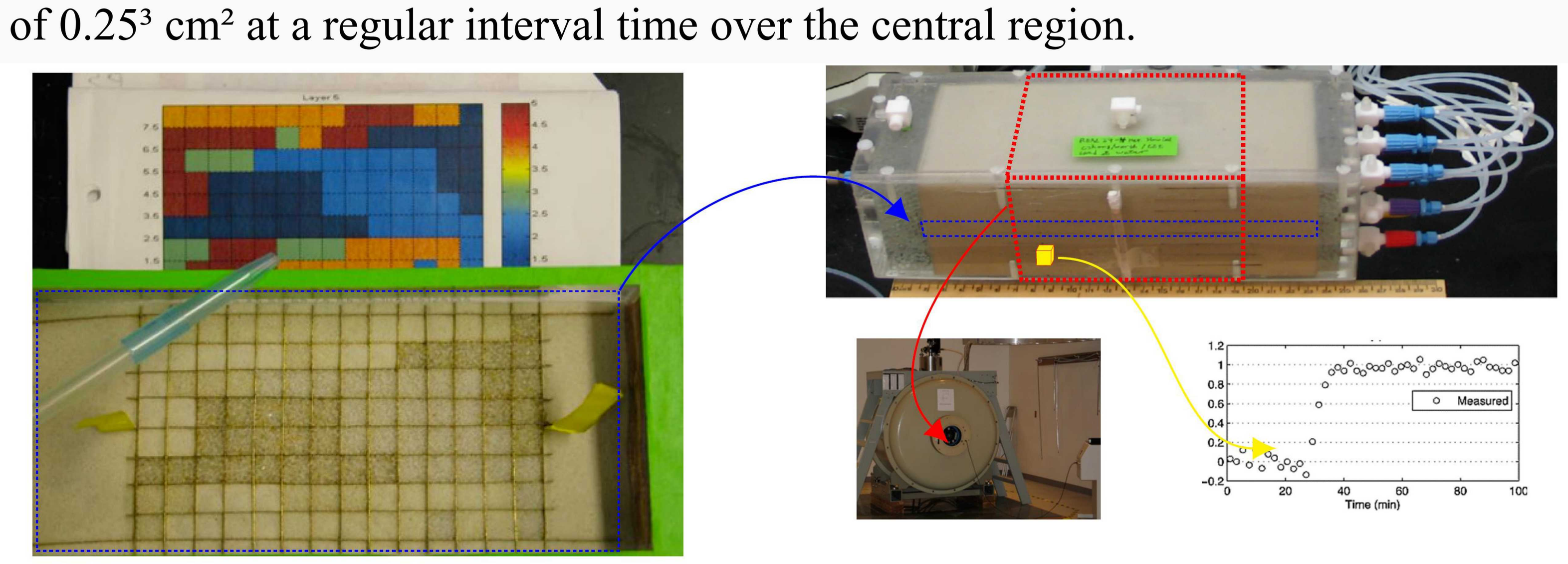
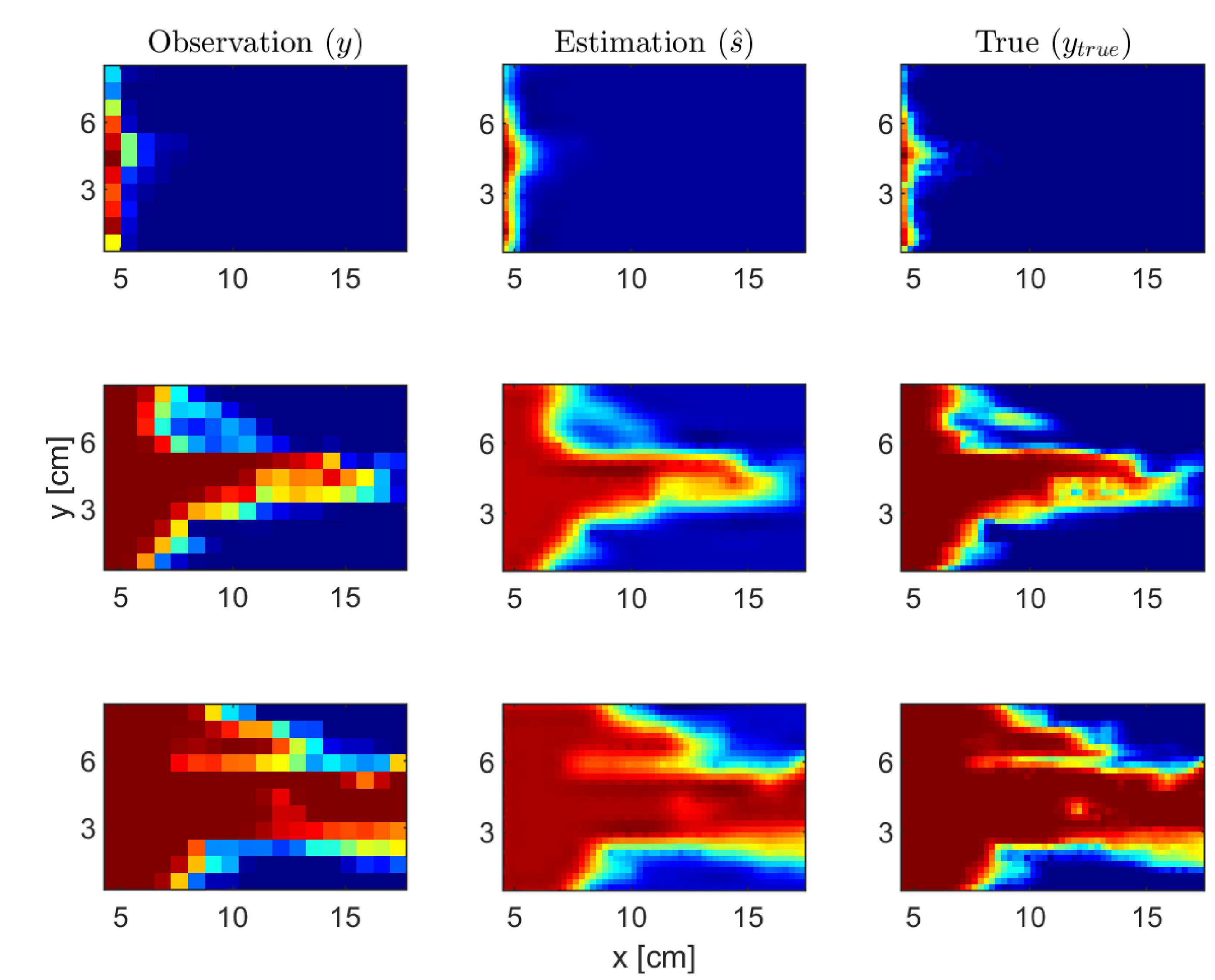


Figure: Packing with a brass divider with 1cm³ openings (left), flowcell (upper right), MRI magnet (lower left) and normalized signal intensity at a voxel (lower right).

Preliminary Results

- We test with an exponential kernel with spatial and temporal stationarity assumption with anisotropy (perhaps valid for this application).
- 82,764 data (y) out of measured 7 million observations (y_{true}) are used for spatio-temporal modeling; currently only layer-wise interpolation (x,y,t) is presented here.



- Currently, the entire estimation of contaminant plume took less than 2 minutes on an Intel 48 core workstation.
- The solution with a CG solver converges within ~20 iterations only requiring PBBFMM3D implementation.
- PBBIFFM will be beneficial if we use more spatiotemporal data points.

Concluding Remarks and Future Direction

- We use FMM and IFMM to accelerate the large-scale spatio-temporal problems.
- Numerical tests show that our proposed method produces reasonable plume reconstruction.
- Our framework will be extended for full 4D application.
- Ordinary and Universal Kriging methods will be applied for better modeling of spatiotemporal modeling of subsurface plumes.
- Optimal hyper parameter estimation through empirical Bayesian approach (i.e., MAP estimate of marginalized distribution of hyperparameters) will be applied.

References

[1] Cressie and Wikle. Statistics for Spatio-Temporal Data. New York: Wiley; 2011.

[2] Wang, Chen, Lee, and Darve, PBBFMM3D: a Parallel Black-Box Fast Multipole Method for Non-oscillatory Kernels, arXiv:1903.02153

[3] Takahashi, Chen, and Darve, Parallelization of the IFMM with an application to boundary element method, Computer Physics Communications, 247: 106975, 2020

[4] Yoon, Zhang, Werth, Valocchi, Webb, Numerical simulation of water flow in three-dimensional heterogeneous porous media observed in a magnetic resonance imaging experiment, Water Resources Research, 44(6), 2008

This study was supported by Hawai'i Experimental Program to Stimulate Competitive Research (EPSCoR) and the Faculty Research Participation Program at the U.S. Engineer Research and Development Center, Coastal and Hydraulics Laboratory administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. department of Energy and ERDC. This work was also supported by the Laboratory Directed Research and Development program at Sandia National Laboratories This poster describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U S Department of Energy or the United States Government.