SAND 2015-10966C

Multifidelity Surrogates of Groundwater Flow

M. Asher[1], J.D. Jakeman[2] and A.J. Jakeman[1]
[1]Australian National University, ACT, AUS
[2]Sandia National Laboratories, NM, USA
jdjakem@sandia.gov

Bi-fidelity algorithm

1. Evaluate the low-fidelity (LF) model u^L on a candidate set of model parameter vectors $\Gamma = \{\mathbf{z}_m\}_{m\in(1,2,\cdots,M)} \subset I_Z$, ie. compute $u^L(\Gamma) = \{u^L(\mathbf{z}_m)\}_{m\in(1,2,\cdots,M)}$.
2. Choose a subset $\gamma = \{\mathbf{z}^\gamma_n\}_{n\in(1,2,\cdots,N)} \subset \Gamma$ using a greedy algorithm to find a linearly independent set of LF solutions $u^L(z^\gamma_n, \boldsymbol{x}, t)$.
   - We can approximate u^L as a combination of these solutions, i.e.

$$u^L(z,\boldsymbol{x},t) \approx u^L_N(z,\boldsymbol{x},t) = \sum_{n=1}^{N} c(u^L(z,\boldsymbol{x},t))u^L(z^\gamma_n,\boldsymbol{x},t)$$

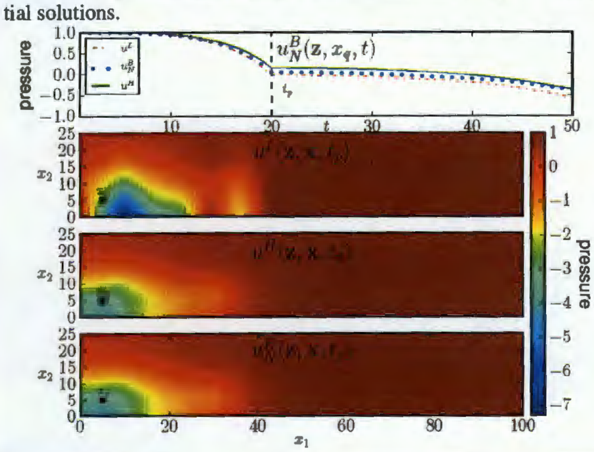
3. Evaluate the high-fidelity (HF) model u^H for the selected model parameter vectors in γ, ie. compute $u^H(\gamma) = \{u^H(\mathbf{z}^\gamma_n)\}_{n\in(1,2,\cdots,N)}$.
4. For any new input $\mathbf{z} \in I_z$, the bi-fidelity (BF) surrogate is found by projecting $u^L(\mathbf{z})$ onto $u^L(\gamma)$, then using these coordinates to estimate $u^H(\mathbf{z})$ in terms of $u^H(\gamma)$.

$$u^B_N(z,\boldsymbol{x},t) = \sum_{n=1}^{N} c(u^L(z,\boldsymbol{x},t))u^H(z^\gamma_n,\boldsymbol{x},t)$$

   - To evaluate c, the LF model must be evaluated at the new vector z

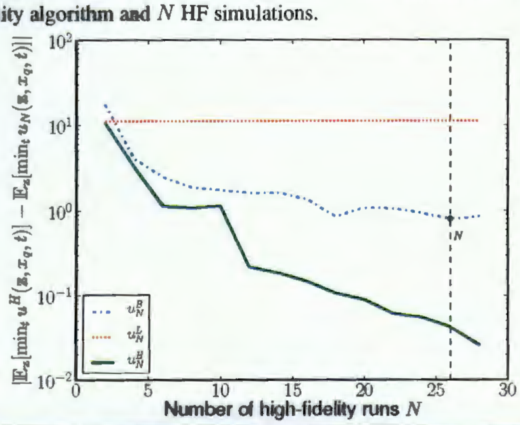
Approximating spatial and temporal solutions

- Bi-fidelity algorithm can be used to approximate both temporal and spatial solutions.



Convergence

- We care about the maximum drawdown at a location x_q.
- u^L is very inaccurate, however the error in u^B_N decreases as N increases.
- u^H_N represents the best possible approximation obtained using the bi-fidelity algorithm and N HF simulations.



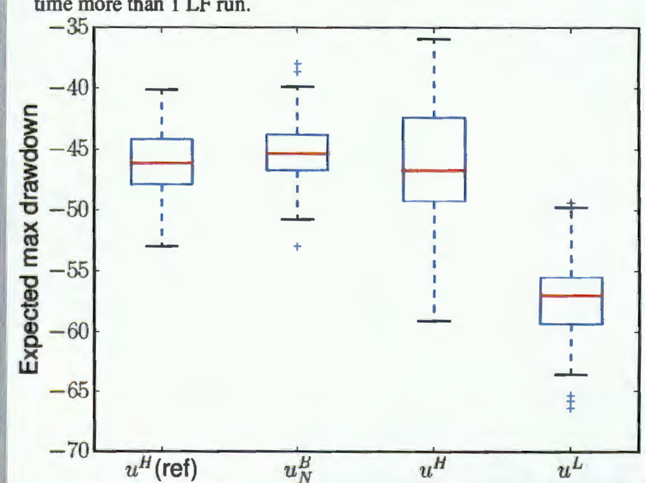
Model

Case study of a groundwater flow model designed to quantify drawdown (decrease in pressure) from coal seam gas development.

- Implemented in the 2.5D finite difference groundwater flow model code, MODFLOW.
- 1200m thick, 34 layers including coal seams and confining units.
- 30 year development period with pumping to keep production wells in lower layers dry and further 60 years without pumping.
- $\mathbf{z} = (z_1, z_2, \cdots z_{10})$ contains storage aquifer and fault properties and linear parameterization of conductivity.
- Quantity of interest is the drawdown in pressure in the top layer at a particular location (a timeseries).
- 10×40km area horizontal cell size is 400×400m for HF, 2400×2400m for LF.
- The runtime of HF is 50 times longer than the run time of LF.

Statistics

- Reference u^H(ref): use 1000 HF runs to compute mean max drawdown.
- Low-fidelity u^L: use 2300 LF runs compute mean.
- Bi-fidelity u^B_N: use 1000 LF runs and $N = 26$ HF runs to compute mean.
- High-fidelity u^H: use 226 HF runs to compute mean.
- Cost to compute u^B_N, u^H and u^L are the same because 1 HF runs cost 50 time more than 1 LF run.



Conclusions

- Bi-fidelity surrogate can be used for demanding applications such as sensitivity analysis, uncertainty propagation, integrated modeling, optimization and decision support.
- Bi-fidelity surrogate can approximate spatially and temporally varying output, from which particular quantities of interest can be computed.
- Despite inaccuracy of LF model, BF model can achieve similar accuracy to HF using only $O(10)$ HF simulations and many LF simulations.
- Accuracy of BF dependent on discrepancy between LF and HF. For a given N, as the discrepancy decreases the accuracy of BF increases.

References

[1] A. Narayan, C. Gittelson, and D. Xiu. A stochastic collocation algorithm with multifidelity models. *SIAM Journal on Scientific Computing*, 36(2):A495–A521, 2014.