Importance Weighting in Hybrid Iterative Ensemble Smoothers for Data Assimilation[a]

Dean S. Oliver – NORCE Norwegian Research Centre
Yuming Ba – Guangdong Polytechnic Normal University

25 June 2024

[a]Hybrid refers to a combination of two methods of estimation gradients.

Parameter estimation and UQ in porous media flow

Two-phase incompressible flow in a porous medium.

$$\phi \frac{\partial}{\partial t} S_\alpha - \nabla \cdot [\lambda_\alpha(S_w) K \nabla p] = q_\alpha \qquad \text{for } \alpha = o,\ w$$

and

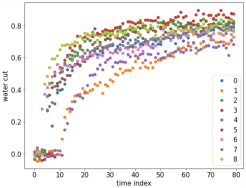
$$S_o + S_w = 1.$$

where permeability $K = e^m$ and porosity ϕ are uncertain.

Observe historical water production rates at a few locations.

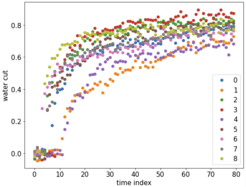
- Typically high-dimensional, relatively few data
- minimization, but no derivatives available
- ensemble-Kalman methods sometimes fail

History matching

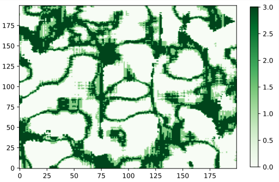


production data

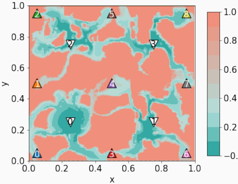
History matching



production data

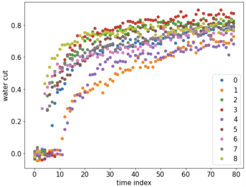


data-generating model

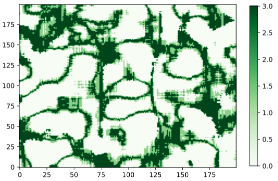


S_w at $t = 10$

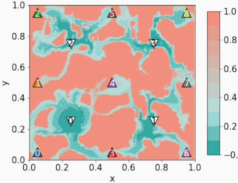
History matching



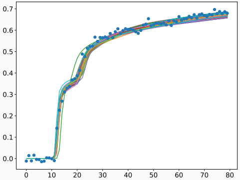
production data



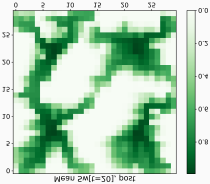
data-generating model



S_w at $t = 10$



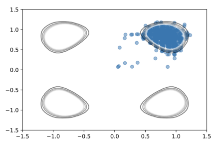
history-matched data



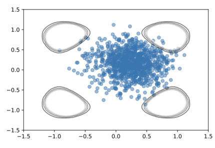
calibrated model

Good data match, but calibrated model not similar to the “truth”.

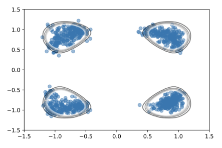
For problems with several similar minima, standard ensemble-based sampling methods will fail[1]



MCMC



ensemble smoother



random minimizer

Random minimizer potentially samples all modes. Example is overly simplistic.

[1] Dunbar et al. (2022) "Ensemble Inference Methods for Models With Noisy and Expensive Likelihoods"

Data assimilation – Minimization for sampling

1. Sample gaussian random variables $x^* \sim N(x^{\mathrm{pr}}, C_x)$
2. Sample the observation error $\epsilon^* \sim N(0, C_d)$
3. Compute $\operatorname{argmin}_x \|d^o - (g(x) + \epsilon^*)\|^2_{C_d^{-1}} + \|x - x^*\|^2_{C_x^{-1}}$

Gauss-Newton minimization

$$\delta x_\ell = x^* - x_\ell - C_x G_\ell^T \left[C_d + G_\ell C_x G_\ell^T \right]^{-1} \times \left[(g(x_\ell) + \epsilon^* - d^o) - G_\ell (x_\ell - x^*) \right].$$

where $G^{\mathrm{T}} = \nabla_x g^{\mathrm{T}}$. (For subsurface flow problems, $\nabla_x g^{\mathrm{T}}$ is often unavailable.)

Hybrid iterative ensemble smoother

Consider composite mappings $\boldsymbol{d} = \boldsymbol{g}(\boldsymbol{m}(\boldsymbol{x}))$

For the hybrid IES, the derivatives of $\boldsymbol{m}$ with respective to $\boldsymbol{x}$ and of data $\boldsymbol{g}$ with respective to $\boldsymbol{m}$ are required.

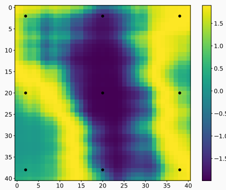
$$\boldsymbol{G} = \nabla_x(g^T) = \boldsymbol{G}_m.(\nabla_x(\boldsymbol{m}^T))^T = \boldsymbol{G}_m\boldsymbol{M}_x$$

A hybrid approximation of the Gauss-Newton update:

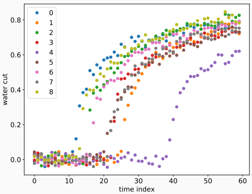
$$\begin{aligned}\delta\boldsymbol{x}_{\ell+1} = -(\boldsymbol{x}_\ell - \boldsymbol{x}^*) - \boldsymbol{C}_x\boldsymbol{M}_x^T\boldsymbol{G}_m^T\Big(\boldsymbol{C}_d + \boldsymbol{G}_m\boldsymbol{M}_x\boldsymbol{C}_x\boldsymbol{M}_x^T\boldsymbol{G}_m^T\Big)^{-1} \\ \times\left(g(\boldsymbol{m}_\ell) + \boldsymbol{\epsilon}^* - \boldsymbol{d}^o - \boldsymbol{G}_m\boldsymbol{M}_x(\boldsymbol{x}_\ell - \boldsymbol{x}^*)\right),\end{aligned}$$

where $\boldsymbol{G}_m = (\Delta\boldsymbol{d}_\ell)(\Delta\boldsymbol{m}_\ell)^{-1}$.

Example: flow in porous medium with multiple rock types[2]



log-permeability, $m(x)$



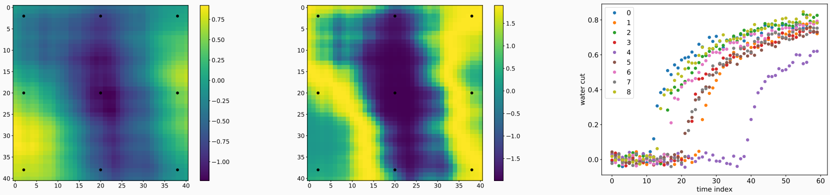
water rate, $g(m(x))$

$$m = 2\tanh(4x+2) + \tanh(2-4x) - 1$$

Generates channel-like features of high permeability.

[2] Ba and Oliver (2024)

Example: flow in porous medium with multiple rock types[2]



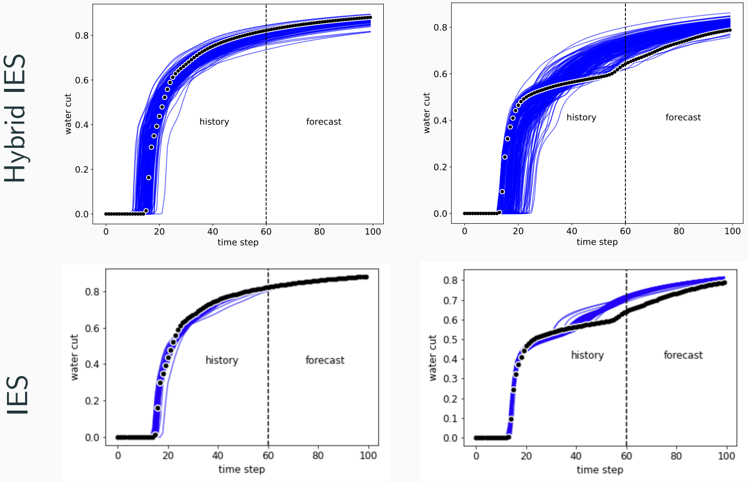
Gaussian variable, x log-permeability, $m(x)$ water rate, $g(m(x))$

$$m = 2\tanh(4x + 2) + \tanh(2 - 4x) - 1$$

Generates channel-like features of high permeability.

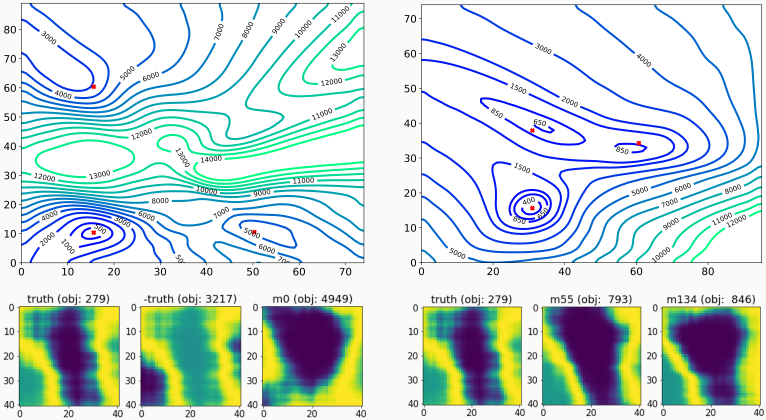
[2] Ba and Oliver (2024)

Difficulties with this problem



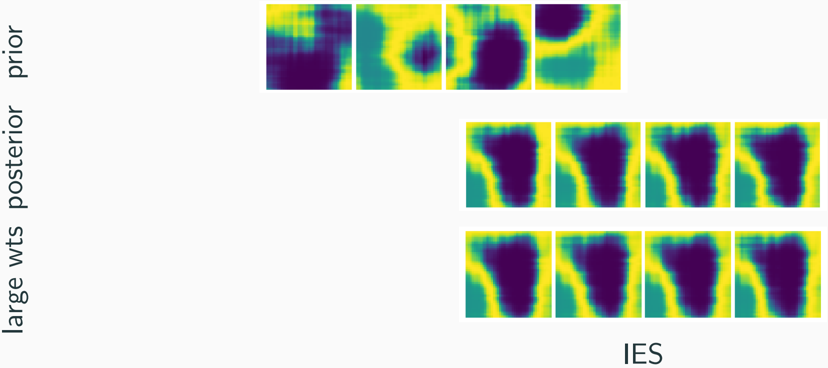
Observations shown by circles. Better match to data with IES.

Fitness landscape



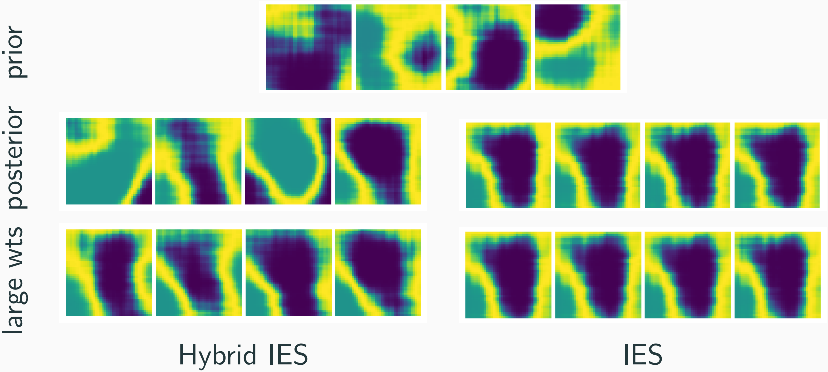
Data misfit function in 2D subspace that includes truth (Oliver, 2022). Many history-matched models end up in “unimportant” local minima.

Posterior sampling



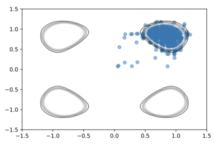
Model realizations for non-monotonic transform of log-permeability using hybrid IES and IES, respectivly. Lack of diversity in IES samples.

Posterior sampling

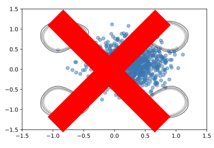


Model realizations for non-monotonic transform of log-permeability using hybrid IES and IES, respectivly. Lack of diversity in IES samples.

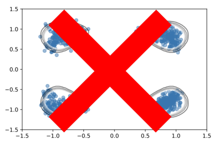
Posterior sampling



MCMC



IES



random minimizer

- Modes are probably not equally deep.
- IES likely to migrate to one mode.
- Random minimizer still likely to sample all – but not all of equal importance.

Importance weights (Ba et al., 2022)

The importance weight for a sample obtained using a "minimization" approach is

$$\omega = \frac{\text{Target probability}}{\text{Proposal probability}} \propto \frac{\pi_X(\boldsymbol{x})\pi_\Delta(\boldsymbol{\delta}|\boldsymbol{x})}{p_{X\Delta}(\boldsymbol{x},\boldsymbol{\delta})},$$

Proposal density for samples from the prior pdf:

$$q_{X'\Delta'}(\boldsymbol{x}',\boldsymbol{\delta}') = q_{X'}(\boldsymbol{x}')\, q_{\Delta'}(\boldsymbol{\delta}')$$

An approximate posterior sample is then generated by computing the critical points of the cost functional

$$Q(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}')^T \boldsymbol{C}_x^{-1}(\boldsymbol{x}-\boldsymbol{x}') + \frac{1}{2}(\boldsymbol{g}(\boldsymbol{m})-\delta')^T \boldsymbol{C}_d^{-1}(\boldsymbol{g}(\boldsymbol{m})-\delta').$$

Importance weights

Solving $\nabla_x Q(\boldsymbol{x}) = 0$ leads to a map from $(\boldsymbol{x}, \boldsymbol{\delta})$ to $(\boldsymbol{x}', \boldsymbol{\delta}')$,

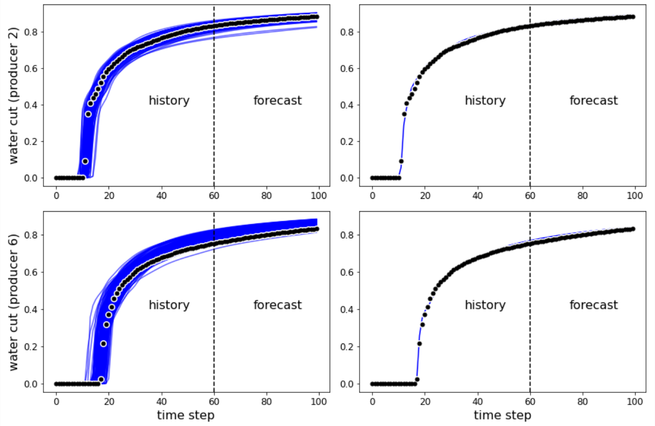
$$\begin{cases} \boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{C}_x \boldsymbol{G}^T \boldsymbol{C}_d^{-1} \big(\boldsymbol{g}(\boldsymbol{m}) - \boldsymbol{\delta} \big) \\ \boldsymbol{\delta}' = \boldsymbol{\delta}. \end{cases}$$

so the proposal pdf is given by[3]

$$p_{X\Delta}(\boldsymbol{x}, \boldsymbol{\delta}) = n(\boldsymbol{x}')^{-1} q_{X'} \Big(\boldsymbol{x} + \boldsymbol{C}_x \boldsymbol{G}^T \boldsymbol{C}_d^{-1} \big(g(\boldsymbol{m}) - \boldsymbol{\delta} \big) \Big) q_{\Delta'}(\boldsymbol{\delta}) J(\boldsymbol{x}, \boldsymbol{\delta}),$$

[3] If $\nabla_x Q(\boldsymbol{x}) = 0$ has multiple solutions, we should either obtain all critical points, or randomly sample.

Effect of weighting (highly non-linear)



Unweighted samples Weighted samples

The posterior predictions for wells 2 and 6 using hybrid IES for the non-monotonic transform. Black points show observations. Effective sampling efficiency is about 1.7%.

Summary

- Possibly many local minima in posterior pdf for parameter estimation in porous media.
- If iterate long enough, IES generally converges to a single minima (but often a “good one”).
- Hybrid IES has much greater diversity, but includes minima with low probability mass.
- Importance weighting of particles from hybrid-IES improves forecasts.

References

Yuming Ba and Dean S. Oliver. Importance weighting in hybrid iterative ensemble smoothers for data assimilation. *Mathematical Geosciences*, 2024. doi: 10.1007/s11004-023-10127-0.

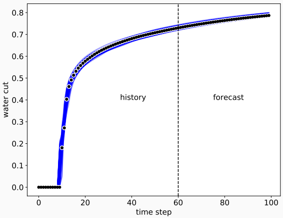
Yuming Ba, Jana de Wiljes, Dean S. Oliver, and Sebastian Reich. Randomized maximum likelihood based posterior sampling. *Computat. Geosci.*, 26(1):217–239, 2022.

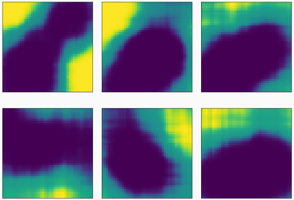
Oliver R. A. Dunbar, Andrew B. Duncan, Andrew M. Stuart, and Marie-Therese Wolfram. Ensemble inference methods for models with noisy and expensive likelihoods. *SIAM Journal on Applied Dynamical Systems*, 21(2):1539–1572, 2022.

Dean S. Oliver. Hybrid iterative ensemble smoother for history matching of hierarchical models. *Mathematical Geosciences*, 54 (8):1289–1313, 2022.

Ensemble collapse?

posterior





Very small spread of data predictions does not imply “collapse” of the ensemble of model realizations.