Hybrid data assimilation

Dean S. Oliver
NORCE Norwegian Research Centre

25 August 2022

Parameter estimation

Noisy observations, d^o related to model parameters, m:

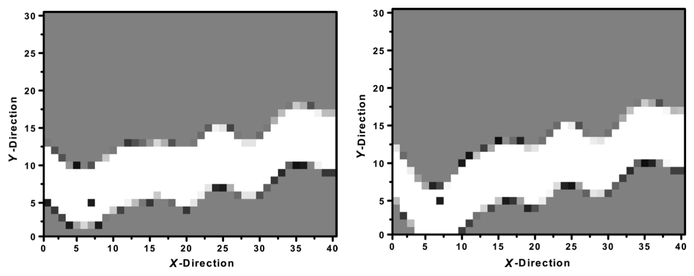
$$d^o = g(m) + \epsilon$$

and prior

$$m \sim N[m^{pr}, C_m]$$

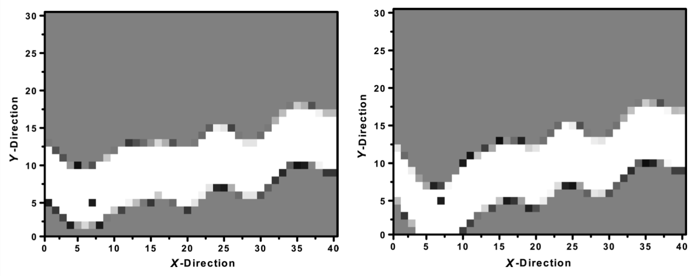
- m might be log-permeability and porosity on a reservoir simulation grid.
- $g(m)$ might be water production rate at well locations.
- Objective is to make inference about m (so that we can forecast).

The posterior landscape[1]



[1] **?** "The impact of upscaling errors on conditioning a stochastic channel to pressure data"

The posterior landscape[1]



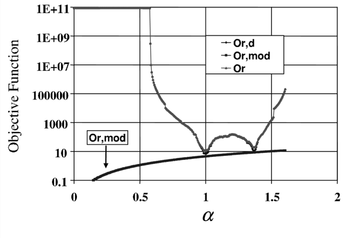
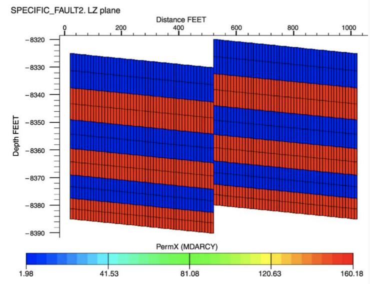


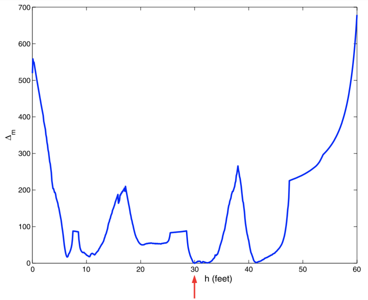
Fig. 4—Objective function computed along a particular line in the model space.

- Posterior pdf is bi-modal
- Seems to be common with facies models.

[1] **?** “The impact of upscaling errors on conditioning a stochastic channel to pressure data”

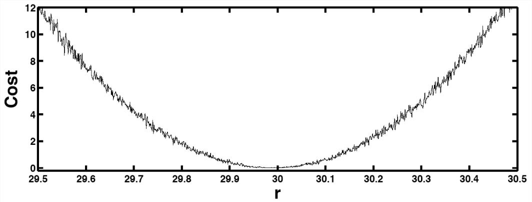
The posterior landscape[2]





[2] **?** "An analysis of history matching errors "

Sometimes, posterior is “rough” but roughness at finer scale

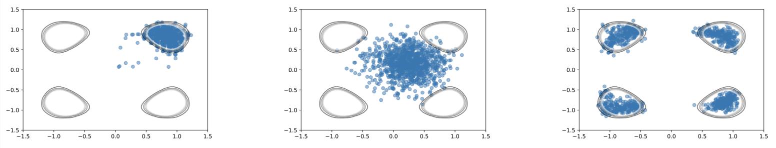


Parameter estimation in Lorentz (1963) model.[3]

Iterative ensemble smoothers work well on this type of problem.
Exact derivatives would be useless.

[3] **?** “Efficient parameter estimation for a highly chaotic system”

For problems with several similar minima, standard ensemble methods will fail[4]



MCMC | ensemble smoother | RML-like

[4] **?** "Ensemble Inference Methods for Models With Noisy and Expensive Likelihoods"

Data assimilation – Randomized maximum likelihood

1. Sample a gaussian random variables x^*
2. Sample the observation error ϵ
3. Compute $\operatorname{argmin}_x \|d^o - g(m(x)) - \epsilon^*\|^2_{C_d^{-1}} + \|x - x^*\|^2_{C_x^{-1}}$

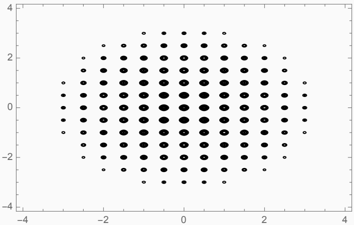
Gauss-Newton minimization[5]

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

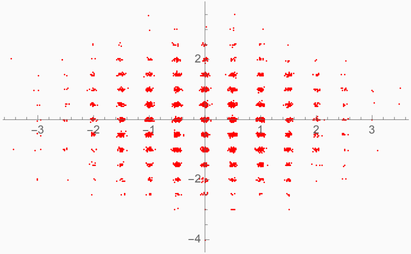
where $G^{\mathrm{T}} = \nabla_x g^{\mathrm{T}}$.

[5]We almost always use Levenberg-Marquardt but omitted for clarity.

Example: RML sampling with many modes[6]



True posterior pdf for x.



RML samples.

Two model variables and two nonlinear observations.

$$g[x_1, x_2] = \begin{bmatrix} \sin[2\pi x_1] \\ \sin[2\pi x_2] \end{bmatrix}$$

$\sigma_D = 0.2$, $x_{\mathrm{pr}} = (0.0, 0.0)$ and $\sigma_X = 1.$, $d_{\mathrm{obs}} = (0., 0.)$

[6] **?** "Metropolized randomized maximum likelihood for improved sampling from multimodal distributions"

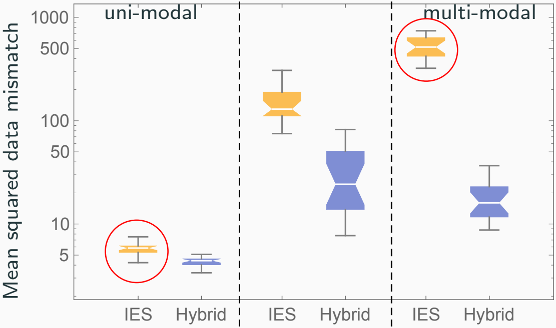
Data assimilation – Iterative ensemble smoother (IES)

To avoid computation of G,[7] we may use an IES approach:

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

where $\Delta \boldsymbol{x}_\ell = \frac{(\boldsymbol{X}_\ell - \bar{\boldsymbol{X}}_\ell)}{\sqrt{(N-1)}}$ and similar for $\Delta \boldsymbol{d}_\ell$.

[7] Computation of G requires the solution of the adjoint system for the reservoir flow and transport equations.



History matching of truncated plurigaussian models.[8] Iterative ensemble smoother works well for monotonic threshold maps, but poorly for symmetric (channel-like).

[8] ?

RML:

$$\boldsymbol{C}_x \boldsymbol{G}_\ell^T \left[\boldsymbol{C}_d + \boldsymbol{G}_\ell \boldsymbol{C}_x \boldsymbol{G}_\ell^T \right]^{-1}$$

IES:

$$\Delta \boldsymbol{x}_\ell (\Delta \boldsymbol{d}_\ell)^T \left(\boldsymbol{C}_d + \Delta \boldsymbol{d}_\ell (\Delta \boldsymbol{d}_\ell)^T \right)^{-1}$$

same effective $\boldsymbol{G}$ for all ensemble members

Hybrid iterative ensemble smoother

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

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

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

Then RML update can be written as

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

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

Two types of problems:

1. Hierarchical parameterization (hyper parameters of the model such as orientation of anisotropy are uncertain):
   - $\boldsymbol{M}_x$ is potentially full but $\boldsymbol{C}_x$ is diagonal
     example $\boldsymbol{m} = \boldsymbol{m}_{pr} + \boldsymbol{C}_m^{1/2}\boldsymbol{x}$ in which case $\boldsymbol{M}_x = \boldsymbol{C}_m^{1/2}$ (potentially dense) and $\boldsymbol{C}_x = \boldsymbol{I}$.
2. Permeability is a nonlinear function of Gaussian random variable (e.g. truncated plurigaussian)
   - $\boldsymbol{C}_x$ is potentially full but $\boldsymbol{M}_x$ is diagonal
     example $\boldsymbol{x} \sim N[0, \boldsymbol{C}_x]$ but $\boldsymbol{M}_x = \text{diag}\, \partial m_i / \partial x_i, \quad i = i, \ldots N_x$.

Compute terms like $\boldsymbol{C}_x \boldsymbol{M}_x^T \boldsymbol{G}_m^T$

Application to large models

Compute terms like $\boldsymbol{C}_x \boldsymbol{M}_x^T \boldsymbol{G}_m^T$ (multiply either $\boldsymbol{C}_x \boldsymbol{v}$ or $\boldsymbol{M}_x^T \boldsymbol{v}$)

Assume $\boldsymbol{C}_x$ and $\boldsymbol{M}_x^T$ are (block) Toeplitz, in which case there are very fast methods for multiplication.[9] Recall

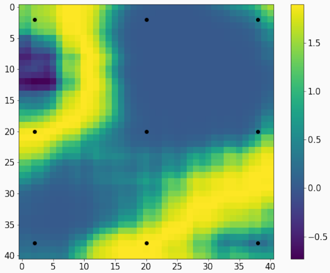
$$A = \begin{bmatrix} a_0 & a_{-1} & \cdots & \cdots & a_{-(n-1)} \\ a_1 & a_0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & a_{-1} & a_{-2} \\ \vdots & \ddots & a_1 & a_0 & a_{-1} \\ a_{n-1} & \cdots & a_2 & a_1 & a_0 \end{bmatrix}$$

is a Toeplitz matrix.

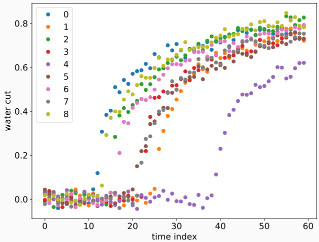
[9] **?** "Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix"

Examples

log-permeability transformation[10]



“true log k



observed water cut

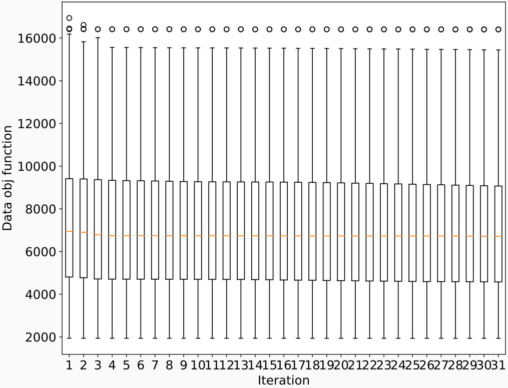
$$m = 2\tanh(4x+2) + \tanh(2-4x) - 1$$

has “connectivity” of high permeability regions.

$$dm/dx = 4 - 8\tanh^2(4x+2) + 4\tanh^2(2-4x)$$

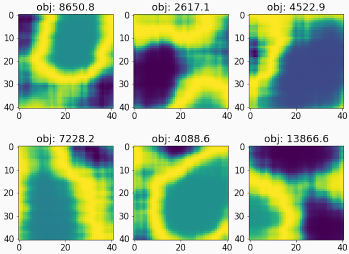
[10] Ongoing work with Yuming Ba, building on **?**

Iterative ensemble smoother with Levenberg-Marquardt minimization

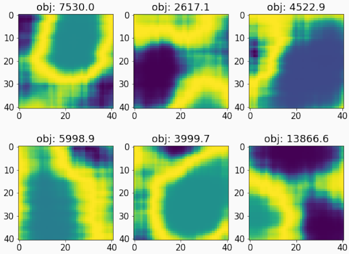


Problem has multiple main minima. Iterative ensemble smoother (with localization) is essentially useless.

Six realizations from prior and posterior



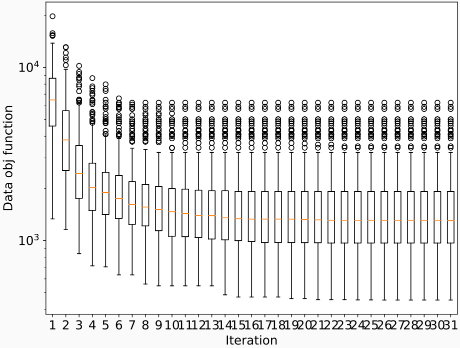
prior realizations



posterior realizations

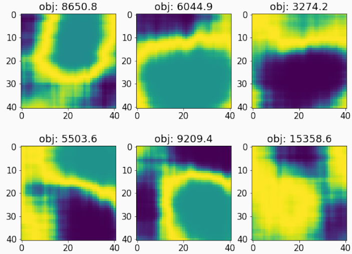
Reminder: This is using somewhat state-of-the-art IES.

Hybrid Iterative ensemble smoother with Levenberg-Marquardt minimization

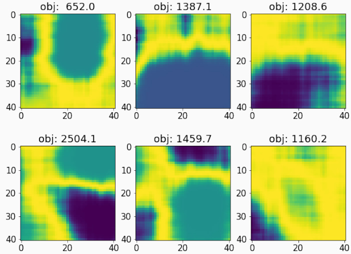


Greedy minimization. Many local minima.

Hybrid IES: Six realizations from prior and posterior



prior realizations



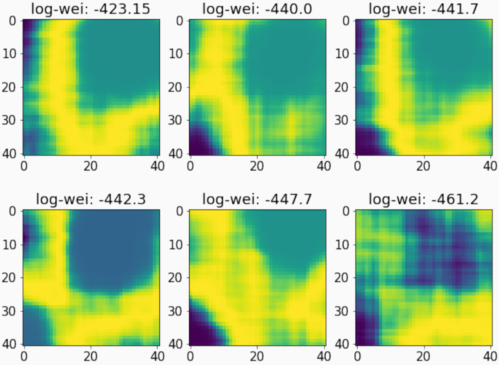
posterior realizations

Each ensemble member has a distinct Kalman gain matrix.

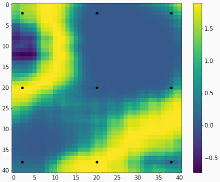
Less likely to collapse.

Can use independent LM.

Six realizations from posterior with largest weights



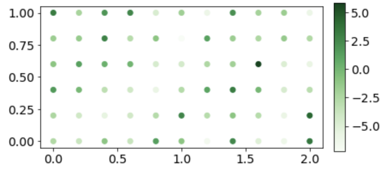
posterior realizations



"true model"

Hierarchical model with unknown orientation of anisotropy[11]

Spatially distributed observations of the permeability field.



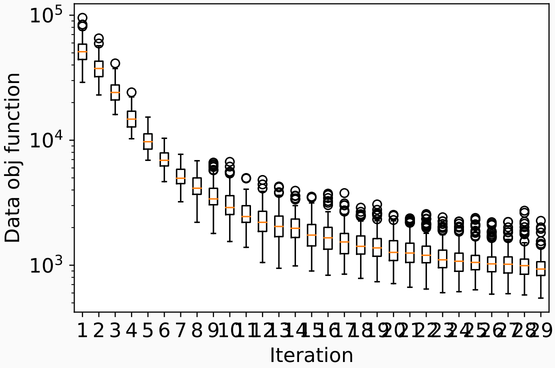
The sensitivity of model parameters (the GRF) to the latent independent standard normal variates are computed analytically.

$$G_i = G_m \left[L_i \quad (\tfrac{\partial}{\partial \phi} L_i) z_i \right]$$

where $L_i = C_{m,i}^{1/2}$ (different covariance matrix for each ensemble member).

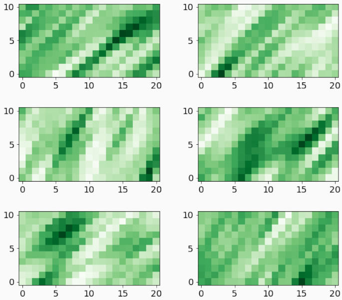
[11] **?** “Hybrid iterative ensemble smoother for history matching of hierarchical models ”

Reduction in data mismatch

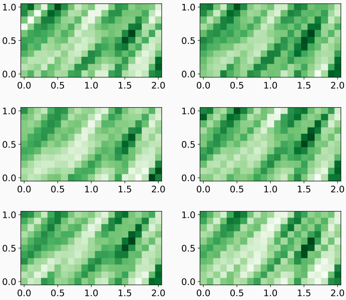


Realizations

prior realizations

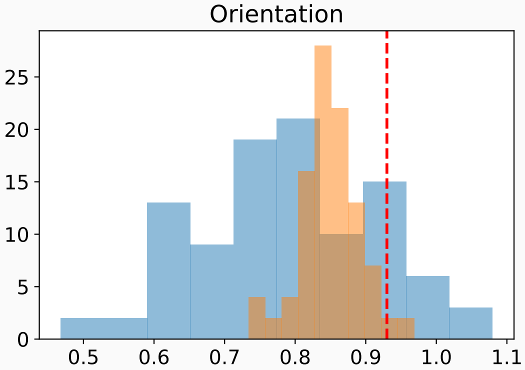


posterior realizations



Reduction in uncertainty of orientation

Comparison of prior and posterior pdf for orientation



Summary

- Allows approximate sampling for some problems where IES fails (similar to RML)
- Approximate sampling in problems with multiple modes.
- Allows individual minimization control
- Analysis step is slower than standard IES
    - requires some large matrix multiplications (use circulant embedding when appropriate)
    - requires the prior covariance or correlation function (not just a prior ensemble)

References

Annan, J. D. and Hargreaves, J. C. (2004). Efficient parameter estimation for a highly chaotic system. *Tellus A*, 56(5):520–526.

Ba, Y., de Wiljes, J., Oliver, D. S., and Reich, S. (2022). Randomized maximum likelihood based posterior sampling. *Computat. Geosci.*, 26(1):217–239.

Dietrich, C. R. and Newsam, G. N. (1997). Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM Journal on Scientific Computing*, 18(4):1088–1107.

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

Oliver, D. S. (2017). Metropolized randomized maximum likelihood for improved sampling from multimodal distributions.

SIAM/ASA Journal on Uncertainty Quantification, 5(1):259–277.

Oliver, D. S. (2022). Hybrid iterative ensemble smoother for history matching of hierarchical models. *Mathematical Geosciences*. online 11 August 2022.

Oliver, D. S. and Chen, Y. (2018). Data assimilation in truncated plurigaussian models: impact of the truncation map. *Mathematical Geosciences*, 50(8):867–893.

Tavassoli, Z., Carter, J. N., and King, P. R. (2005). An analysis of history matching errors. *Comput. Geosci.*, 9(2):99–123.

Zhang, F., Reynolds, A. C., and Oliver, D. S. (2003). An initial guess for the Levenberg-Marquardt algorithm for conditioning a stochastic channel to pressure data. *Math. Geol.*, 35(1):67–88.