An Iterative Ensemble-Smoother Solution of the HM Problem Formulated with Consistent Error Statistics

Geir Evensen

NORCE

Presentation is available from
https://github.com/geirev/Presentations

Background

- Conditioning reservoir models on rate data . . . (Evensen and Eikrem, 2018).
  - We must take correlated measurment errors into account.
  - We neglect rate errors when forcing the reservoir simulators.
- Accounting for model errors in iterative ensemble smoothers (Evensen, 2019).
  - Explained how to include model errors (like rate errors) in iterative smoothers.
- Ensemble subspace EnRML method (Evensen et al., 2019, Raanes et al., 2019).
  - The iterative ensemble smoother used in the current study.
- Formulating the history matching problem with consistent error statistics (Evensen, 2021).
  - The current presentation!

Standard formulation of the history-matching problem

Nonlinear model and measurements

$$\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x}) \qquad \boldsymbol{d} \leftarrow \boldsymbol{y} + \boldsymbol{e}$$

Bayesian formulation

$$f(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{d}) \propto f(\boldsymbol{d}|\,\boldsymbol{y}) f(\,\boldsymbol{y}|\boldsymbol{x}) f(\boldsymbol{x})$$

Model pdf

$$f(\,\boldsymbol{y}|\boldsymbol{x}) = \delta\big(\,\boldsymbol{y} - \boldsymbol{g}(\boldsymbol{x})\big)$$

Marginal pdf

$$f(\boldsymbol{x}|\boldsymbol{d}) \propto \int f(\boldsymbol{d}|\,\boldsymbol{y}) f(\,\boldsymbol{y}|\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{y} = f\big(\boldsymbol{d}|\boldsymbol{g}(\boldsymbol{x})\big) f(\boldsymbol{x})$$

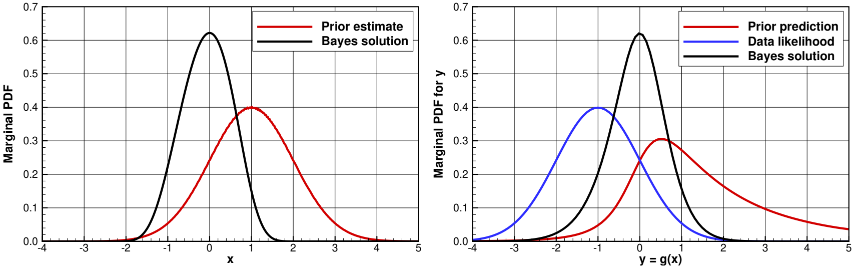
Standard Bayesian inverse problem

$$f(\boldsymbol{x}|\boldsymbol{d}) = f\big(\boldsymbol{d}|\boldsymbol{g}(\boldsymbol{x})\big) f(\boldsymbol{x})$$

Illustration of the parameter estimation problem

Bayes theorem gives posterior probability function for parameters $\boldsymbol{x}$

$$f(\boldsymbol{x}|\boldsymbol{d}) \propto f(\boldsymbol{d}\,|\,\boldsymbol{g}(\boldsymbol{x}))\,f(\boldsymbol{x})$$



- $\boldsymbol{x}$ represents model input parameters like, porosity, permeability, fault multipliers
- $\boldsymbol{y}$ could represent predicted production of oil, gas, and water
- Prior pdf represents uncertainty of $\boldsymbol{x}$.
- Prior prediction pdf represents uncertainty of $\boldsymbol{y} = g(\boldsymbol{x})$.
- Data likelihood represents uncertainty of measurement $\boldsymbol{d}$.

Approximate sampling of the posterior pdf

Gaussian priors and randomized-maximum-likelihood sampling of

$$f\left(\boldsymbol{x}|\boldsymbol{d}\right) \propto f\left(\boldsymbol{d}\,|\,\boldsymbol{g}(\boldsymbol{x})\right) f\left(\boldsymbol{x}\right)$$

by minimizing an ensemble of cost functions

$$\mathcal{J}(\boldsymbol{x}_j) = \underbrace{\left(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}}\right)^{\mathrm{T}} \boldsymbol{C}_{xx}^{-1} \left(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}}\right)}_{\text{Prior misfit}} + \underbrace{\left(\boldsymbol{g}(\boldsymbol{x}_j) - \boldsymbol{d}_j\right)^{\mathrm{T}} \boldsymbol{C}_{dd}^{-1} \left(\boldsymbol{g}(\boldsymbol{x}_j) - \boldsymbol{d}_j\right)}_{\text{data misfit}}.$$

Solutions methods:

1. Ensemble Smoother (ES) (van Leeuwen and Evensen, 1996) and (Evensen, 2009, Chap. 10).
2. Ensemble Randomized Likelihood (EnRML) (Chen and Oliver, 2013).
3. Ensemble Smoother with Multiple Data Assimilation (ESMDA) (Emerick and Reynolds, 2013).

Ensemble Smoother

Approximately solves $\nabla \mathcal{J}_j = 0$

$$\boldsymbol{C}_{xx}^{-1}(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}}) + \boldsymbol{G}_j^{\mathrm{T}} \boldsymbol{C}_{dd}^{-1} \big(\boldsymbol{g}(\boldsymbol{x}_j) - \boldsymbol{d}_j\big) = 0.$$

- Apply the linearization $\boldsymbol{g}(\boldsymbol{x}_j) = \boldsymbol{g}(\boldsymbol{x}_j^{\mathrm{f}}) + \boldsymbol{G}_j(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}})$.
- Replace model sensitivities by least-squares fit $\boldsymbol{C}_{yx} = \boldsymbol{G}\boldsymbol{C}_{xx}$.
- ES uses ensemble covariances $\overline{\boldsymbol{C}}_{xy}$, $\overline{\boldsymbol{C}}_{xx}$, and $\overline{\boldsymbol{C}}_{dd}$.

$$\begin{aligned}
\boldsymbol{x}_j^{\mathrm{f}} &\leftarrow \mathcal{N}(\boldsymbol{x}^{\mathrm{f}}, \boldsymbol{C}_{xx}^{\mathrm{f}}), \qquad \boldsymbol{d}_j \leftarrow \mathcal{N}(\boldsymbol{d}, \boldsymbol{C}_{dd}),\\
\boldsymbol{y}_j^{\mathrm{f}} &= \boldsymbol{g}\big(\boldsymbol{x}_j^{\mathrm{f}}\big),\\
\boldsymbol{x}_j^{\mathrm{a}} &= \boldsymbol{x}_j^{\mathrm{f}} + \overline{\boldsymbol{C}}_{xy}\Big(\overline{\boldsymbol{C}}_{yy} + \overline{\boldsymbol{C}}_{dd}\Big)^{-1}\Big(\boldsymbol{d}_j - \boldsymbol{y}_j\Big),\\
\boldsymbol{y}_j^{\mathrm{a}} &= \boldsymbol{g}\big(\boldsymbol{x}_j^{\mathrm{a}}\big).
\end{aligned}$$

EnRML

State-space formulation with Gauss-Newton iterations

$$\mathcal{J}(\boldsymbol{x}_j) = \left(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}}\right)^{\mathrm{T}} \boldsymbol{C}_{xx}^{-1} \left(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}}\right) + \left(\boldsymbol{g}(\boldsymbol{x}_j) - \boldsymbol{d}_j\right)^{\mathrm{T}} \boldsymbol{C}_{dd}^{-1} \left(\boldsymbol{g}(\boldsymbol{x}_j) - \boldsymbol{d}_j\right)$$

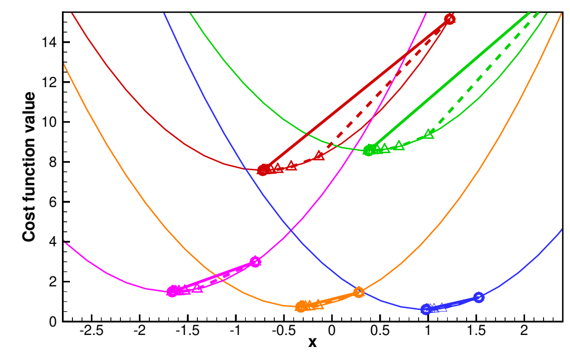
with gradient and Hessian

$$\begin{aligned} \nabla_x \mathcal{J} &= \boldsymbol{C}_{xx}^{-1}(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}}) + \boldsymbol{G}_j^{\mathrm{T}} \boldsymbol{C}_{dd}^{-1} \left(\boldsymbol{g}(\boldsymbol{x}_j) - \boldsymbol{d}_j\right), \\ \nabla_x \nabla_x \mathcal{J} &\approx \boldsymbol{C}_{xx}^{-1} + \boldsymbol{G}_j^{\mathrm{T}} \boldsymbol{C}_{dd}^{-1} \boldsymbol{G}_j \end{aligned}$$

Iterate

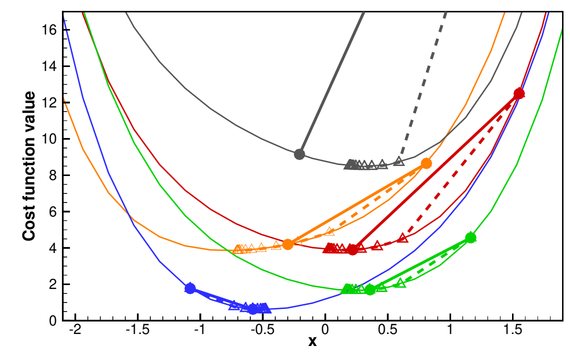
$$\begin{aligned} \boldsymbol{x}_j^{i+1} &= \boldsymbol{x}_j^i - \gamma \left(\nabla\nabla \mathcal{J}_i\right)^{-1} \nabla \mathcal{J}_j^i \\ \boldsymbol{y}_j^{i+1} &= \boldsymbol{g}\left(\boldsymbol{x}_j^{i+1}\right) \end{aligned}$$

ES and EnRML illustration: Linear model



- EnRML and ES both finds global minimum.
- Samples exactly posterior pdf.

ES and EnRML illustration: Non-linear model



- EnRML gets closer to minimum than ES
- Approximate sampling of posterior pdf.

ESMDA uses tapering of likelihood

Approximate sampling of $f\left(\boldsymbol{x}|\boldsymbol{d}\right)$ by gradually introducing the measurements (Neal, 1996)

$$
\begin{aligned}
f\left(\boldsymbol{x}|\boldsymbol{d}\right) &= f\left(\boldsymbol{d}|\,\boldsymbol{y}\right) f\left(\boldsymbol{x}\right) \\
&= f\left(\boldsymbol{d}|\,\boldsymbol{y}\right)^{\left(\sum_{i=1}^{N}\frac{1}{\alpha_i}\right)} f\left(\boldsymbol{x}\right) \quad \text{with} \quad \sum_{i=1}^{N}\frac{1}{\alpha_i} = 1 \\
&= f\left(\boldsymbol{d}|\,\boldsymbol{y}\right)^{\frac{1}{\alpha_N}} \cdots f\left(\boldsymbol{d}|\,\boldsymbol{y}\right)^{\frac{1}{\alpha_2}} f\left(\boldsymbol{d}|\,\boldsymbol{y}\right)^{\frac{1}{\alpha_1}} f\left(\boldsymbol{x}\right)
\end{aligned}
$$

We compute N ES steps with “inflated” observation errors.

- Small updates reduce impact of the linear approximation.
- ESMDA is identical to ES in the linear case.

Subspace EnRML: (Evensen et al., 2019, Raanes et al., 2019)

Original cost functions

$$\mathcal{J}(\boldsymbol{x}_j) = \left(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}}\right)^{\mathrm{T}} \boldsymbol{C}_{xx}^{-1} \left(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}}\right) + \left(\boldsymbol{g}(\boldsymbol{x}_j) - \boldsymbol{d}_j\right)^{\mathrm{T}} \boldsymbol{C}_{dd}^{-1} \left(\boldsymbol{g}(\boldsymbol{x}_j) - \boldsymbol{d}_j\right).$$

Solution is contained in the ensemble subspace, thus

$$\boldsymbol{x}_j^{\mathrm{a}} = \boldsymbol{x}_j^{\mathrm{f}} + \boldsymbol{A}\boldsymbol{w}_j,$$

and,

$$\mathcal{J}(\boldsymbol{w}_j) = \boldsymbol{w}_j^{\mathrm{T}} \boldsymbol{w}_j + \left(\boldsymbol{g}\left(\boldsymbol{x}_j^{\mathrm{f}} + \boldsymbol{A}\boldsymbol{w}_j\right) - \boldsymbol{d}_j\right)^{\mathrm{T}} \boldsymbol{C}_{dd}^{-1} \left(\boldsymbol{g}\left(\boldsymbol{x}_j^{\mathrm{f}} + \boldsymbol{A}\boldsymbol{w}_j\right) - \boldsymbol{d}_j\right)$$

Reduces dimension of problem from state size to ensemble size.

$$\boldsymbol{w}_j^{i+1} = \boldsymbol{w}_j^{i} - \gamma \left(\nabla\nabla\mathcal{J}_i\right)^{-1} \nabla\mathcal{J}_j^{i}$$

Rate-data dependency

Correlated errors in rate data?

- We often derive the observed production rates from a rate-allocation table.
- The rate-allocation tables are constructed from infrequent separator tests.
- This approach leads to strong time correlations of measurement errors.

Thus, we can not assume a diagonal $\boldsymbol{C}_{dd}$!

Neglecting correlations in measurement errors leads to:

- Too strong update of estimate.
- Too strong reduction of posterior variance.

Practical implementaion

Historical production rates have time-correlated errors acting as model errors.

1. Create an ensemble of historical OPR_j, GPR_j and WPR_j.
2. Force each realization by its corresponding RESV_j.
3. We augment OPR_j, GPR_j and WPR_j to the state ensemble for updating (Evensen, 2019).

We can include and estimate stochastic rates as part of the history-matching process.

Subspace EnRML algorithm: (Evensen, 2021)

Input: $\boldsymbol{X}_0 \in \Re^{n\times N}$ ▷ prior model ensemble

Input: $\boldsymbol{D} \in \Re^{m\times N}$ ▷ perturbed measurements

Input: $\boldsymbol{E}_0 \in \Re^{m_u\times N}$ ▷ initial rate perturbations

$\boldsymbol{W}_0 = 0$ ▷ $\boldsymbol{W} \in \Re^{N\times N}$

$\boldsymbol{\Pi} = \left(\boldsymbol{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^\mathrm{T}\right)/\sqrt{N-1}$ ▷ projection subtracting ensemble mean; $\boldsymbol{\Pi} \in \Re^{N\times N}$

$\boldsymbol{E} = \boldsymbol{D}\boldsymbol{\Pi}$ ▷ scaled measurement perturbations; $\boldsymbol{E} \in \Re^{m\times N}$

i=0

repeat

  $\boldsymbol{Y}_i = \boldsymbol{g}(\boldsymbol{X}_i, \boldsymbol{E}_i)\boldsymbol{\Pi}$ ▷ $\boldsymbol{Y} \in \Re^{m\times N}$

  $\boldsymbol{\Omega}_i = \boldsymbol{I} + \boldsymbol{W}_i\boldsymbol{\Pi}$ ▷ $\boldsymbol{\Omega} \in \Re^{N\times N}$

  $\boldsymbol{S}_i = \boldsymbol{Y}_i\boldsymbol{\Omega}_i^{-1}$ ▷ $\boldsymbol{S} \in \Re^{m\times N}$

  $\boldsymbol{H}_i = \boldsymbol{S}_i\boldsymbol{W}_i + \boldsymbol{D} - \boldsymbol{g}(\boldsymbol{X}_i, \boldsymbol{E}_i)$ ▷ $\boldsymbol{H} \in \Re^{m\times N}$

  $\boldsymbol{W}_{i+1} = \boldsymbol{W}_i - \gamma\left(\boldsymbol{W}_i - \boldsymbol{S}_i^\mathrm{T}\left(\boldsymbol{S}_i\boldsymbol{S}_i^\mathrm{T} + \boldsymbol{E}\boldsymbol{E}^\mathrm{T}\right)^{-1}\boldsymbol{H}_i\right)$

  $\boldsymbol{T}_i = \left(\boldsymbol{I} + \boldsymbol{W}_{i+1}\Big/\sqrt{N-1}\right)$ ▷ $\boldsymbol{T} \in \Re^{N\times N}$

  $\boldsymbol{X}_{i+1} = \boldsymbol{X}\boldsymbol{T}_i$

  $\boldsymbol{E}_{i+1} = \boldsymbol{E}_0\boldsymbol{T}_i$

  i=i+1

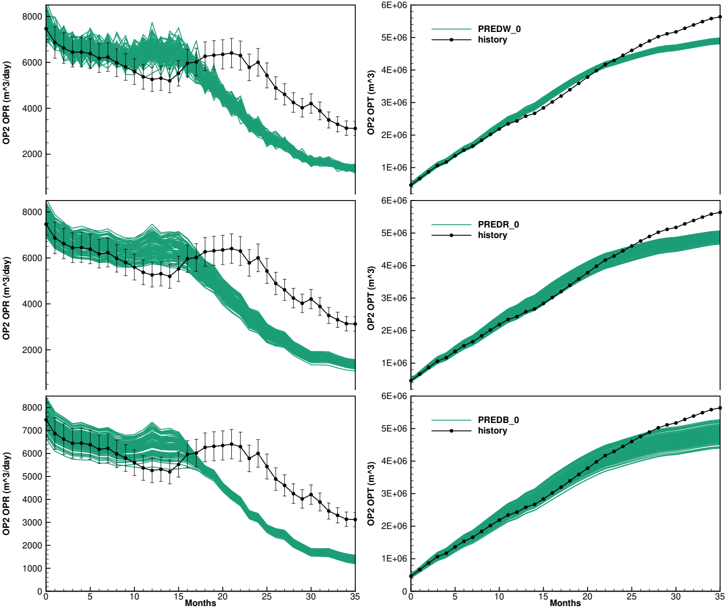
until convergence

Reservoir examples

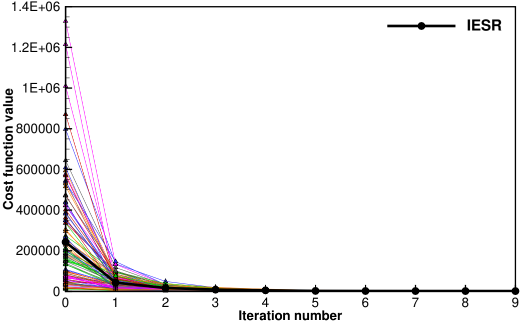
Case	$\boldsymbol{C}_{dd}$	Noise Model	Update E
PREDW		White	
PREDR		Red	
PREDB		Bias	
IES0	$\boldsymbol{I}$	White	no
IESnd	$\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}$	Red	no
IESR	$\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}$	Red	yes
IESB	$\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}$	Bias	yes

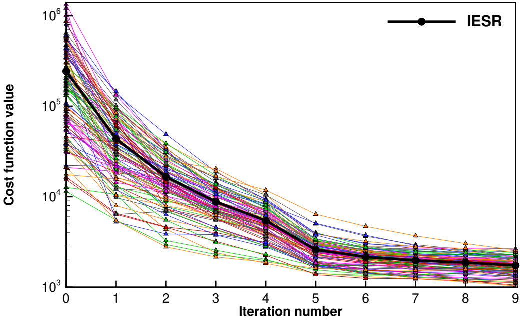
1. IES0 is the standard case neglecting error correlations and errors in controls.
2. IESnd includes error correlations in inversion.
3. IESR adds time-correlated perturbations to rates and updates them.
4. IESB adds perturbation biases to rates and updates them.

Predictions

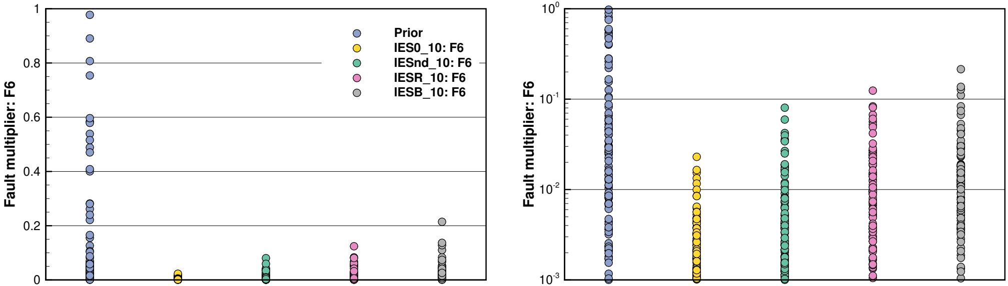


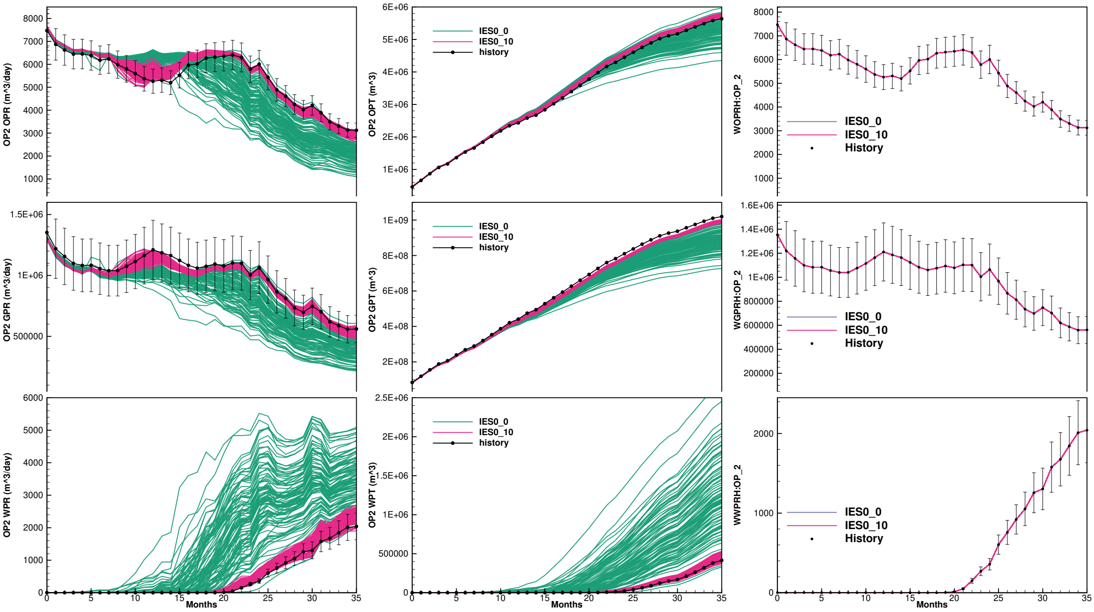
IESR: Ensemble of cost functions

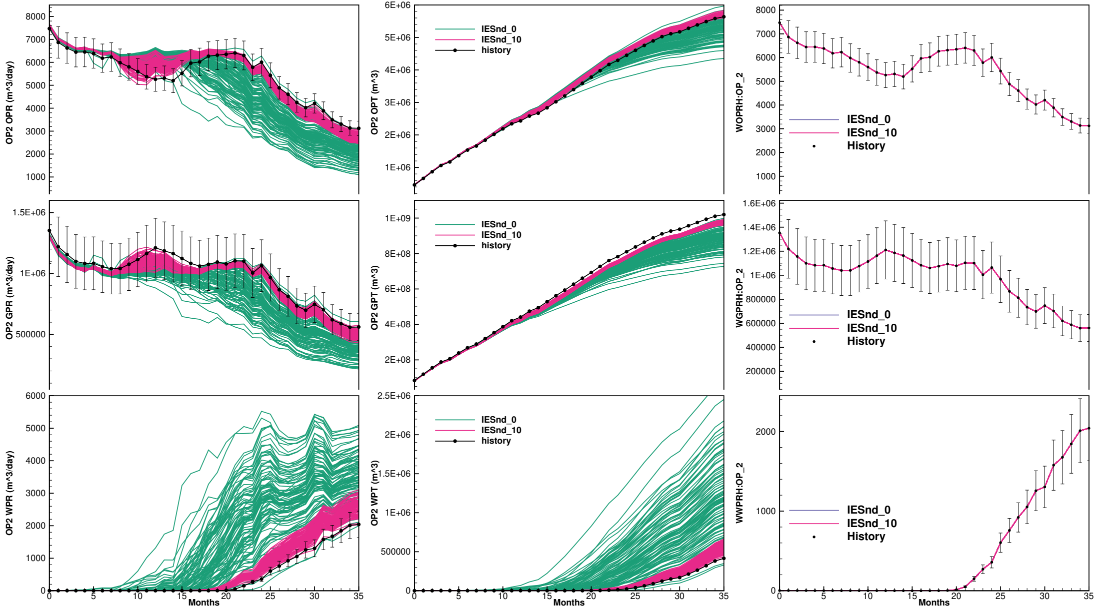


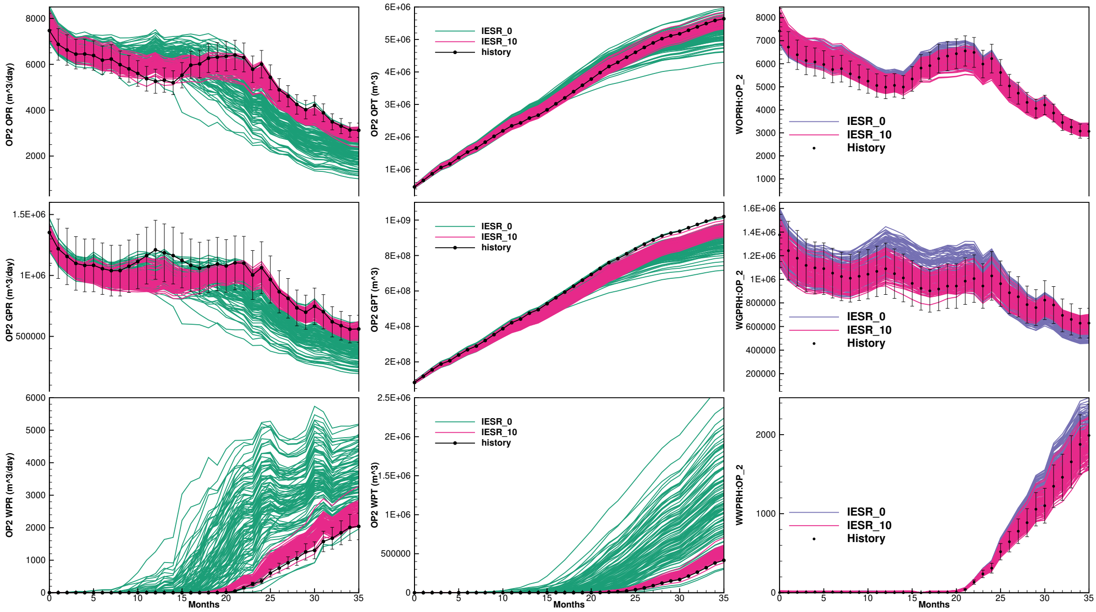


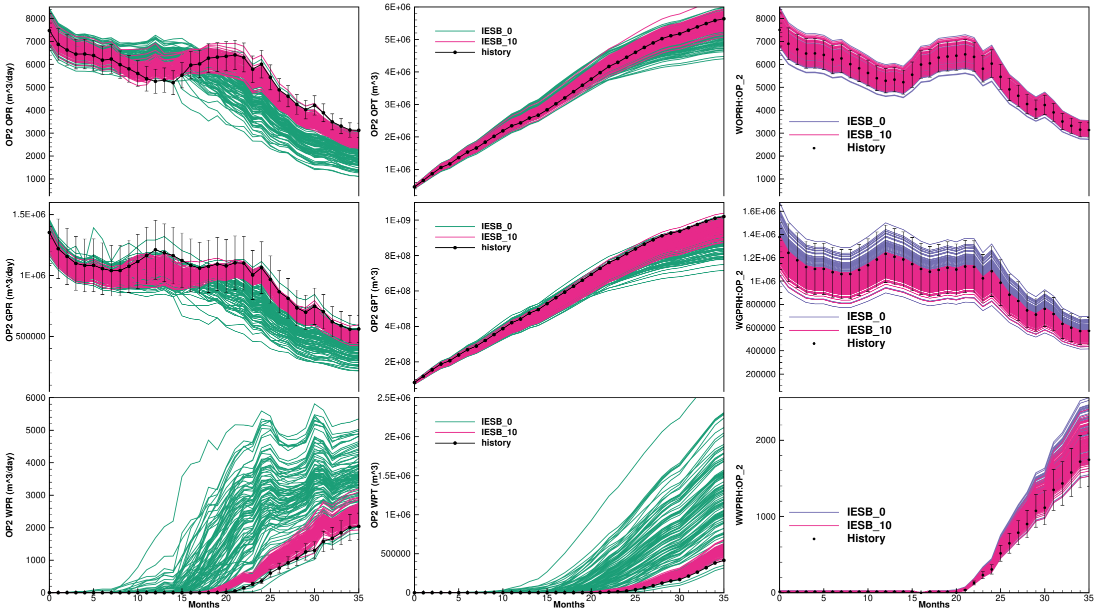
Fault multiplier F6




OP2 OPR (m^3/day)
OP2 OPT (m^3)
WOPRH:OP_2
OP2 GPR (m^3/day)
OP2 GPT (m^3)
WGPRH:OP_2
OP2 WPR (m^3/day)
OP2 WPT (m^3)
WWPRH:OP_2
IES0_0
IES0_10
history
History
Months


OP2 OPR (m^3/day)
OP2 OPT (m^3)
WOPRH:OP_2
OP2 GPR (m^3/day)
OP2 GPT (m^3)
WGPRH:OP_2
OP2 WPR (m^3/day)
OP2 WPT (m^3)
WWPRH:OP_2
IESnd_0
IESnd_10
history
History
Months


OP2 OPR (m^3/day)
OP2 OPT (m^3)
WOPRH:OP_2
OP2 GPR (m^3/day)
OP2 GPT (m^3)
WGPRH:OP_2
OP2 WPR (m^3/day)
OP2 WPT (m^3)
WWPRH:OP_2
IESR_0
IESR_10
history
History
Months


OP2 OPR (m^3/day)
OP2 OPT (m^3)
WOPRH:OP_2
OP2 GPR (m^3/day)
OP2 GPT (m^3)
WGPRH:OP_2
OP2 WPR (m^3/day)
OP2 WPT (m^3)
WWPRH:OP_2
IESB_0
IESB_10
history
History
Months

Summary

Rederives consistent HM formulation from Bayes'.

- Condition on rate data while allowing for time-correlated errors.
- Force the model using historical rates with stochastic errors.
- Update stochastic rates as part of the state vector.

Leads to:

- Realistic posterior error statistics.
- Avoids underestimated posterior errors.

Chen, Y. and D. S. Oliver. Levenberg-Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Computat Geosci*, 17:689–703, 2013. doi:10.1007/s10596-013-9351-5.

Emerick, A. A. and A. C. Reynolds. Ensemble smoother with multiple data assimilation. *Computers and Geosciences*, 55:3–15, 2013. doi:10.1016/j.cageo.2012.03.011.

Evensen, G. *Data Assimilation: The Ensemble Kalman Filter*. Springer, 2nd edition, 2009. doi:10.1007/978-3-642-03711-5.

Evensen, G. Accounting for model errors in iterative ensemble smoothers. *Computat Geosci*, 23(4):761–775, 2019. doi:10.1007/s10596-019-9819-z.

Evensen, G. Formulating the history matching problem with consistent error statistics. *Computat Geosci*, page 26, 2021. doi:10.1007/s10596-021-10032-7.

Evensen, G. and K. S. Eikrem. Strategies for conditioning reservoir models on rate data using ensemble smoothers. *Computat Geosci*, 22(5): 1251–1270, 2018. doi:10.1007/s10596-018-9750-8.

Evensen, G., P. Raanes, A. Stordal, and J. Hove. Efficient implementation of an iterative ensemble smoother for data assimilation and reservoir history matching. *Frontiers in Applied Mathematics and Statistics*, 5:47, 2019. doi:10.3389/fams.2019.00047.

Neal, R. M. Sampling from multimodal distributions using tempered transitions. *Statistics and Computing*, 6(4):353–366, 1996. doi:10.1007/BF00143556.

Neto, G. M. S., R. V. Soares, G. Evensen, A. Davolioa, and D. J. Schiozer. Subspace ensemble randomized maximum likelihood with local analysis for time-lapse-seismic-data assimilation. *SPE Journal*, SPE-205029-PA:21, 2020. doi:10.2118/205029-PA.

Raanes, P. N., A. S. Stordal, and G. Evensen. Revising the stochastic iterative ensemble smoother. *Nonlin. Processes Geophys*, 26:325–338, 2019. doi:10.5194/npg-2019-10.

van Leeuwen, P. J. and G. Evensen. Data assimilation and inverse methods in terms of a probabilistic formulation. *Mon. Weather Rev.*, 124: 2898–2913, 1996. doi:10.1175/1520-0493(1996)124<2898:DAAIMI>2.0.CO;2.