Implementation of an iterative ensemble smoother for big-data assimilation and reservoir history matching

Geir Evensen

NORCE–Norwegian Research Center
Nansen Environmental and Remote Sensing Center

University of Potsdam, August 2019

Parameter-estimation problem

Nonlinear model and measurements

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

Bayesian formulation

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

- Standard History-Matching problem for oil-reservoir models.

“Indirect” DA update

Nonlinear model

$$\boldsymbol{x}_{i+1} = \boldsymbol{m}(\boldsymbol{x}_i)$$

Nonlinear measurement operator and measurements

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}_{i+1}) = \boldsymbol{h}\big(\boldsymbol{m}(\boldsymbol{x}_i)\big) = \boldsymbol{g}(\boldsymbol{x}_i), \qquad \boldsymbol{d} \leftarrow \boldsymbol{y} + \boldsymbol{e}$$

Bayesian formulation

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

- Smoother update step in sequential data assimilation

Marginal pdf

Nonlinear model and measurements

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

Model pdf

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

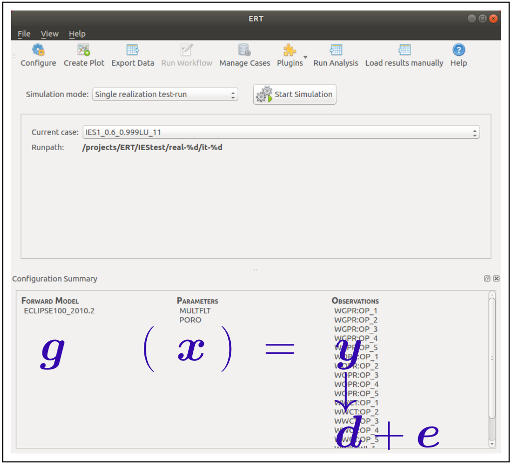
Bayesian formulation

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

Marginal pdf

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

ERT



$$g(x) = y \;\longrightarrow\; d + e$$

Gaussian priors

Maximizing $f(\boldsymbol{x}|\boldsymbol{d})$ is equivalent to minimizing

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

Exact direct solution in the case when $\boldsymbol{g}(\boldsymbol{x})$ is linear (KF update)

$$\boldsymbol{x}^{\mathrm{a}} = \boldsymbol{x}^{\mathrm{f}} + \boldsymbol{K}(\boldsymbol{d} - \boldsymbol{g}(\boldsymbol{x})), \qquad \boldsymbol{C}_{xx}^{\mathrm{a}} = (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\boldsymbol{C}_{xx}^{\mathrm{f}}.$$

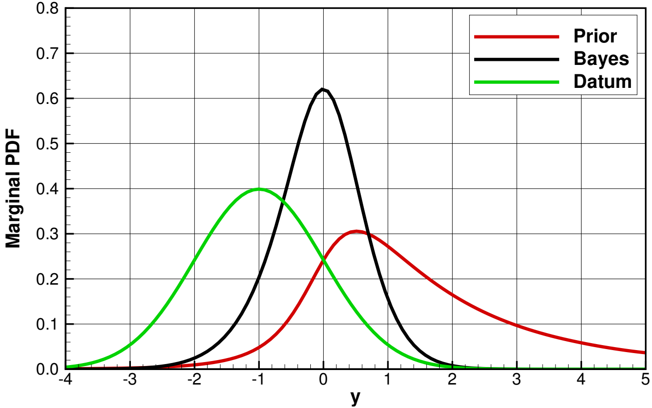
Ensemble representation (EnKF)

$$\boldsymbol{x}_j^{\mathrm{a}} = \boldsymbol{x}_j^{\mathrm{f}} + \overline{\boldsymbol{K}}(\boldsymbol{d}_j - \boldsymbol{g}(\boldsymbol{x}_j)), \qquad \overline{\boldsymbol{x}}^{\mathrm{a}} \rightarrow \boldsymbol{x}^{\mathrm{a}}, \quad \overline{\boldsymbol{C}}_{xx}^{\mathrm{a}} \rightarrow \boldsymbol{C}_{xx}^{\mathrm{a}}.$$

Equivalent to minimizing (linear case)

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

Bayesian update



EnKF update

Prior ensemble

$$\boldsymbol{y}_j^{\mathrm{f}} = \boldsymbol{g}(\boldsymbol{x}_j^{\mathrm{f}})$$

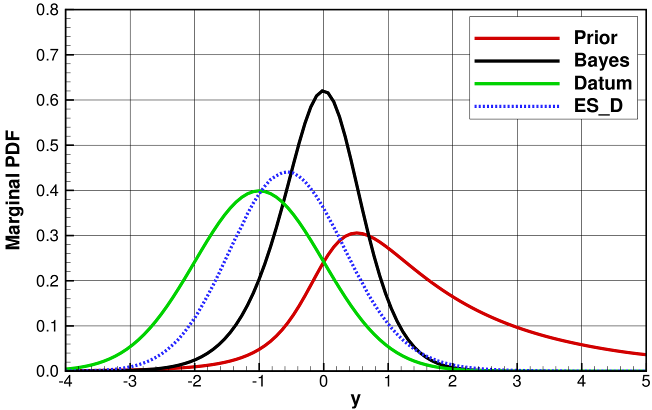
Ensemble of measurements of $\boldsymbol{y}$

$$\boldsymbol{d}_j = \boldsymbol{d} + \boldsymbol{e}_j$$

Direct update of $\boldsymbol{y}_j$

$$\boldsymbol{y}_j^{\mathrm{a}} = \boldsymbol{y}_j^{\mathrm{f}} + \overline{\boldsymbol{C}}_{yy}^{\mathrm{f}}\big(\overline{\boldsymbol{C}}_{yy}^{\mathrm{f}} + \boldsymbol{C}_{dd}\big)^{-1}\Big(\boldsymbol{d}_j - \boldsymbol{y}_j^{\mathrm{f}}\Big)$$

EnKF (direct) update



Ensemble smoother (ES) indirect update

Prior

$$\boldsymbol{y}_j^{\mathrm{f}} = \boldsymbol{g}(\boldsymbol{x}_j^{\mathrm{f}})$$

Ensemble of measurements of $\boldsymbol{y}$

$$\boldsymbol{d}_j = \boldsymbol{d} + \boldsymbol{e}_j$$

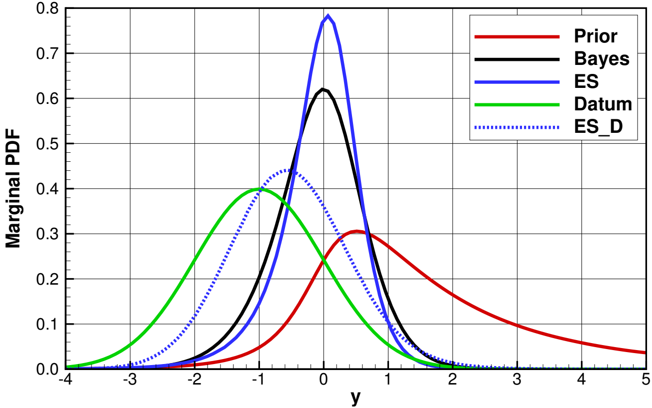
Smoother update of $\boldsymbol{x}$

$$\boldsymbol{x}_j^{\mathrm{a}} = \boldsymbol{x}_j^{\mathrm{f}} + \overline{\boldsymbol{C}}_{xy}^{\mathrm{f}} \left(\overline{\boldsymbol{C}}_{yy}^{\mathrm{f}} + \boldsymbol{C}_{dd}\right)^{-1} \left(\boldsymbol{d}_j - \boldsymbol{y}_j^{\mathrm{f}}\right)$$

Indirect update of $\boldsymbol{y}$

$$\boldsymbol{y}_j^{\mathrm{a}} = \boldsymbol{g}(\boldsymbol{x}_j^{\mathrm{a}})$$

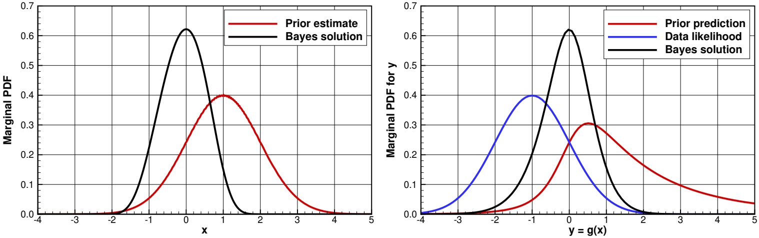
ES (indirect) update



History matching problem

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

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



- $\boldsymbol{x}$ represents static parameters like porosity, permeability, ...
- $\boldsymbol{y}$ could represent predicted OPR from Eclipse.
- Prior pdf represents uncertainty of $\boldsymbol{x}$.
- Prior prediction pdf represents uncertainty of $\boldsymbol{y} = g(\boldsymbol{x})$.
- Data likelihood represents uncertainty of measurement $\boldsymbol{d}$.

ES and IES

Gaussian assumption!

Approximate sampling of posterior pdf

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

by minimizing an ensemble of cost functions

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

Prior misfit — data misfit

Solved by

1. Ensemble Smoother
2. Iterative Ensemble Smoother

Ensemble Smoother

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

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

- Define $\boldsymbol{G}_j^{\mathrm{T}} = \nabla_x \boldsymbol{g}(\boldsymbol{x}_j)$.
- ES applies a linearization $\boldsymbol{g}(\boldsymbol{x}_j) = \boldsymbol{x}_j^{\mathrm{f}} + \boldsymbol{G}_j(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}})$.
- Model sensitivities replaced 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 $\overline{\boldsymbol{C}}_{dd}$

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

Iterative Ensemble Smoother

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

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

$$\nabla_x \nabla_x \mathcal{J} \approx \boldsymbol{C}_{xx}^{-1} + \nabla_x \boldsymbol{g}(\boldsymbol{x}_j) \boldsymbol{C}_{dd}^{-1} (\nabla_x \boldsymbol{g}(\boldsymbol{x}_j))^{\mathrm{T}}$$

Iterate

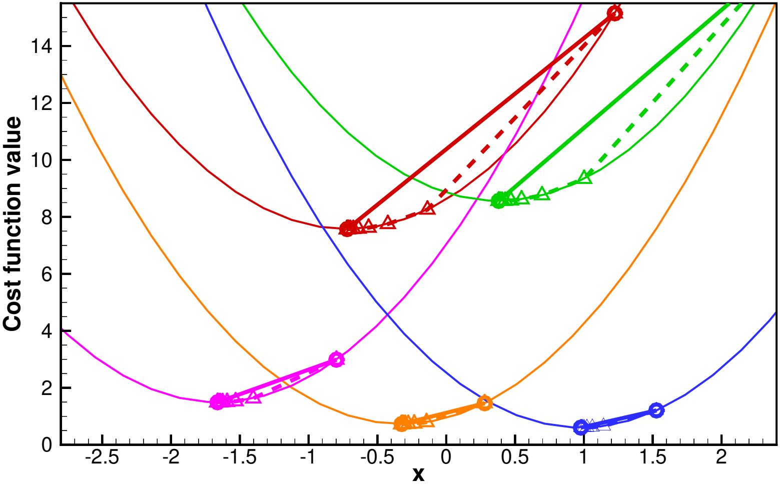
$$\boldsymbol{x}_j^{i+1} = \boldsymbol{x}_j^i - \gamma \left(\nabla\nabla \mathcal{J}_i\right)^{-1} \nabla \mathcal{J}_j^i$$
$$\boldsymbol{y}_j^{i+1} = \boldsymbol{g}\left(\boldsymbol{x}_j^{i+1}\right)$$

IES gradient

$$
\begin{aligned}
\left(\nabla\nabla\mathcal{J}_i\right)^{-1}\nabla\mathcal{J}_{j,i} = {} & \overline{\boldsymbol{C}}^{i}_{xx}\overline{\boldsymbol{C}}^{-1}_{xx}(\boldsymbol{x}_{j,i}-\boldsymbol{x}^{\mathrm{f}}_j) \\
& - \overline{\boldsymbol{C}}^{i}_{xy}\left(\overline{\boldsymbol{C}}^{i}_{yx}(\overline{\boldsymbol{C}}^{i}_{xx})^{-1}\overline{\boldsymbol{C}}^{i}_{xy} + \overline{\boldsymbol{C}}_{dd}\right)^{-1} \\
& \times \left(\overline{\boldsymbol{C}}^{i}_{yx}\overline{\boldsymbol{C}}^{-1}_{xx}(\boldsymbol{x}_{j,i}-\boldsymbol{x}^{\mathrm{f}}_j) - \big(\boldsymbol{g}(\boldsymbol{x}_{j,i}) - \boldsymbol{d}_j\big)\right).
\end{aligned}
$$

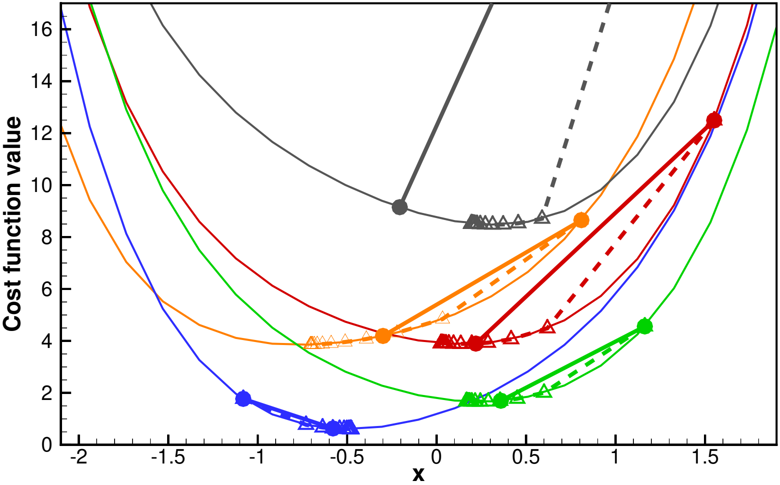
- Model sensitivities replaced with averaged least squares fit.
- Covariances replaced with ensemble representations.
- Leftmost factor of gradient is ensemble anomaly matrix.
- Solution is searched for in the ensemble subspace

ES and IES illustration: Linear model



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

ES and IES illustration: Non-linear model



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

ESMDA: Rewriting likelihood

Approximate sampling of

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

by gradually introducing the measurements (**?**)

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

with

$$\sum_{i=1}^{N} \frac{1}{\alpha_i} = 1$$

Some definitions

Prior ensemble and perturbed measurements

$$\boldsymbol{X} = \left(\boldsymbol{x}_1^{\mathrm{f}},\ \boldsymbol{x}_2^{\mathrm{f}},\ \dots\ ,\ \boldsymbol{x}_N^{\mathrm{f}}\right) \quad \text{and} \quad \boldsymbol{D} = \left(\boldsymbol{d}_1,\ \boldsymbol{d}_2,\ \dots\ ,\ \boldsymbol{d}_N\right)$$

Ensemble means

$$\overline{\boldsymbol{x}} = \frac{1}{N}\sum_{j=1}^{N} \boldsymbol{x}_j \quad \text{and} \quad \overline{\boldsymbol{d}} = \frac{1}{N}\sum_{j=1}^{N} \boldsymbol{d}_j$$

Ensemble anomaly matrices and covariances

$$\boldsymbol{A} = \boldsymbol{X}\left(\boldsymbol{I}_N - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\right)\Big/\sqrt{N-1} \quad \rightarrow \quad \overline{\boldsymbol{C}}_{xx} = \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$$

$$\boldsymbol{E} = \boldsymbol{D}\left(\boldsymbol{I}_N - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\right)\Big/\sqrt{N-1} \quad \rightarrow \quad \overline{\boldsymbol{C}}_{dd} = \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}$$

IES: Ensemble subspace version

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,

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

Reduces dimension of problem from state size to ensemble size.

$$\boldsymbol{w}_j^{i+1} = \boldsymbol{w}_j^{i} - \gamma \nabla \mathcal{J}_j^{i}$$

Gradient and Hessian of cost function

Gradient

$$\nabla \mathcal{J}(\boldsymbol{w}_j) = 2\boldsymbol{w}_j + 2\big(\boldsymbol{G}_j\boldsymbol{A}\big)^{\mathrm{T}}\boldsymbol{C}_{dd}^{-1}\big(\boldsymbol{g}\big(\boldsymbol{x}_j^{\mathrm{f}} + \boldsymbol{A}\boldsymbol{w}_j\big) - \boldsymbol{d}_j\big),$$

Hessian (approximate)

$$\nabla\nabla \mathcal{J}(\boldsymbol{w}_j) \approx 2\boldsymbol{I} + 2\big(\boldsymbol{G}_j\boldsymbol{A}\big)^{\mathrm{T}}\boldsymbol{C}_{dd}^{-1}\big(\boldsymbol{G}_j\boldsymbol{A}\big)$$

Gauss-Newton iterations

$$\boldsymbol{w}_j^{i+1} = \boldsymbol{w}_j^i - \gamma \Delta \boldsymbol{w}_j^i$$

$$\Delta \boldsymbol{w}_j^i = \Bigg\{ \boldsymbol{w}_j^i - \left(\boldsymbol{G}_j^i \boldsymbol{A}\right)^{\mathrm{T}} \Big(\left(\boldsymbol{G}_j^i \boldsymbol{A}\right)\left(\boldsymbol{G}_j^i \boldsymbol{A}\right)^{\mathrm{T}} + \boldsymbol{C}_{dd} \Big)^{-1}$$

$$\times \Big(\left(\boldsymbol{G}_j^i \boldsymbol{A}\right) \boldsymbol{w}_j^i + \boldsymbol{d}_j - \boldsymbol{g}\left(\boldsymbol{x}_j^{\mathrm{f}} + \boldsymbol{A}\boldsymbol{w}_j^i\right) \Big) \Bigg\}.$$

with

$$\boldsymbol{G}_j^i = \left(\nabla \boldsymbol{g}|_{\boldsymbol{x}_j^{\mathrm{f}} + \boldsymbol{A}\boldsymbol{w}_j^i} \right)^{\mathrm{T}}.$$

$\boldsymbol{G}^i_j \boldsymbol{A}$

Replace $\boldsymbol{G}^i_j$ with average sensitivity $\boldsymbol{G}^i$

$G_j^i A$

Replace G_j^i with average sensitivity G^i

Define the linear regression

$$\overline{G}_i = Y_i A_i^+$$

$$Y_i = g(X_i)\left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^\mathrm{T}\right)\Big/\sqrt{N-1}$$

$\boldsymbol{G}_j^i \boldsymbol{A}$

Replace $\boldsymbol{G}_j^i$ with average sensitivity $\boldsymbol{G}^i$

Define the linear regression

$$\overline{\boldsymbol{G}}_i = \boldsymbol{Y}_i \boldsymbol{A}_i^+$$

Write

$$\boldsymbol{G}_j^i \boldsymbol{A} \approx \overline{\boldsymbol{G}}^i \boldsymbol{A} = \boldsymbol{Y}_i \boldsymbol{A}_i^+ \boldsymbol{A}$$

$$\boldsymbol{Y}_i = \boldsymbol{g}\big(\boldsymbol{X}_i\big)\big(\boldsymbol{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^\mathrm{T}\big)\Big/\sqrt{N-1}$$

$G_j^i A$

Replace G_j^i with average sensitivity G^i

Define the linear regression

$$\overline{\boldsymbol{G}}_i = \boldsymbol{Y}_i \boldsymbol{A}_i^+$$

Write

$$\begin{aligned} \boldsymbol{G}_j^i \boldsymbol{A} \approx \overline{\boldsymbol{G}}^i \boldsymbol{A} &= \boldsymbol{Y}_i \boldsymbol{A}_i^+ \boldsymbol{A} \\ &= \boldsymbol{Y}_i \boldsymbol{A}_i^+ \boldsymbol{A}_i \boldsymbol{\Omega}_i^{-1} \qquad \boldsymbol{A} = \boldsymbol{A}_i \boldsymbol{\Omega}_i^{-1} \end{aligned}$$

$$\boldsymbol{Y}_i = \boldsymbol{g}(\boldsymbol{X}_i)\left(\boldsymbol{I} - \frac{1}{N}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\right)\Big/\sqrt{N-1}$$

$$\boldsymbol{\Omega}_i = \boldsymbol{I} + \boldsymbol{W}_i\left(\boldsymbol{I} - \frac{1}{N}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\right)\Big/\sqrt{N-1}$$

$\boldsymbol{G}_j^i \boldsymbol{A}$

Replace $\boldsymbol{G}_j^i$ with average sensitivity $\boldsymbol{G}^i$

Define the linear regression

$$\overline{\boldsymbol{G}}_i = \boldsymbol{Y}_i \boldsymbol{A}_i^+$$

Write

$$\begin{aligned}
\boldsymbol{G}_j^i \boldsymbol{A} \approx \overline{\boldsymbol{G}}^i \boldsymbol{A} &= \boldsymbol{Y}_i \boldsymbol{A}_i^+ \boldsymbol{A} \\
&= \boldsymbol{Y}_i \boldsymbol{A}_i^+ \boldsymbol{A}_i \boldsymbol{\Omega}_i^{-1} \qquad \boldsymbol{A} = \boldsymbol{A}_i \boldsymbol{\Omega}_i^{-1} \\
&= \boldsymbol{Y}_i \boldsymbol{\Omega}_i^{-1} = \boldsymbol{S}_i \qquad \text{when } n \geq N-1
\end{aligned}$$

$$\boldsymbol{Y}_i = \boldsymbol{g}\big(\boldsymbol{X}_i\big)\big(\boldsymbol{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\big)\Big/\sqrt{N-1}$$

$$\boldsymbol{\Omega}_i = \boldsymbol{I} + \boldsymbol{W}_i\big(\boldsymbol{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\big)\Big/\sqrt{N-1}$$

Orthogonal projection $\boldsymbol{Y}_i \boldsymbol{A}_i^+ \boldsymbol{A}_i$

Nonlinear case with $n \geq N-1$

$$\boldsymbol{Y}_i \boldsymbol{A}_i^+ \boldsymbol{A}_i = \boldsymbol{Y}_i \big(\boldsymbol{I} - \frac{1}{N}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\big) = \boldsymbol{Y}_i$$

Nonlinear case with $n < N-1$

$$\boldsymbol{Y}_i \boldsymbol{A}_i^+ \boldsymbol{A}_i = \boldsymbol{Y}_i \boldsymbol{V} \begin{pmatrix} \boldsymbol{I}_{n\times n} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{V}^{\mathrm{T}}.$$

Linear case

$$\begin{aligned} \boldsymbol{Y}_i \boldsymbol{A}_i^+ \boldsymbol{A}_i &= \boldsymbol{H}\boldsymbol{M}\boldsymbol{X}_i \big(\boldsymbol{I} - \frac{1}{N}\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}\big) \Big/ \sqrt{N-1}\, \boldsymbol{A}_i^+ \boldsymbol{A}_i \\ &= \boldsymbol{H}\boldsymbol{M}\boldsymbol{A}_i \boldsymbol{A}_i^+ \boldsymbol{A}_i \\ &= \boldsymbol{H}\boldsymbol{M}\boldsymbol{A}_i = \boldsymbol{Y}_i \end{aligned}$$

Equation for $\boldsymbol{W}$

Matrix form with $\boldsymbol{S}_i = \boldsymbol{Y}_i \boldsymbol{\Omega}_i^{-1}$

$$\boldsymbol{W}_{i+1} = \boldsymbol{W}_i - \gamma\Big(\boldsymbol{W}_i - \boldsymbol{S}_i^{\mathrm{T}}\big(\boldsymbol{S}_i\boldsymbol{S}_i^{\mathrm{T}} + \boldsymbol{C}_{dd}\big)^{-1}\big(\boldsymbol{S}_i\boldsymbol{W}_i - \boldsymbol{D} + \boldsymbol{g}(\boldsymbol{X}_i)\big)\Big)$$

IES ensemble subspace algorithm

1: Inputs: $\boldsymbol{X}$, $\boldsymbol{D}$, (and $\boldsymbol{C}_{dd}$)

2: $\boldsymbol{W}_1 = 0$

3: **for** $i = 1, \text{Convergence}$ **do**

4: $\quad \boldsymbol{Y}_i = \boldsymbol{g}(\boldsymbol{X}_i)\big(\boldsymbol{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^\mathrm{T}\big)/\sqrt{N-1}$

5: $\quad \boldsymbol{\Omega}_i = \boldsymbol{I} + \boldsymbol{W}_i\big(\boldsymbol{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^\mathrm{T}\big)/\sqrt{N-1}$

6: $\quad \boldsymbol{\Omega}_i^\mathrm{T}\boldsymbol{S}_i^\mathrm{T} = \boldsymbol{Y}_i^\mathrm{T}$ $\qquad \mathcal{O}(mN^2)$

7: $\quad \boldsymbol{H}_i = \boldsymbol{S}_i\boldsymbol{W}_i + \boldsymbol{D} - \boldsymbol{g}(\boldsymbol{X}_i)$ $\qquad \mathcal{O}(mN^2)$

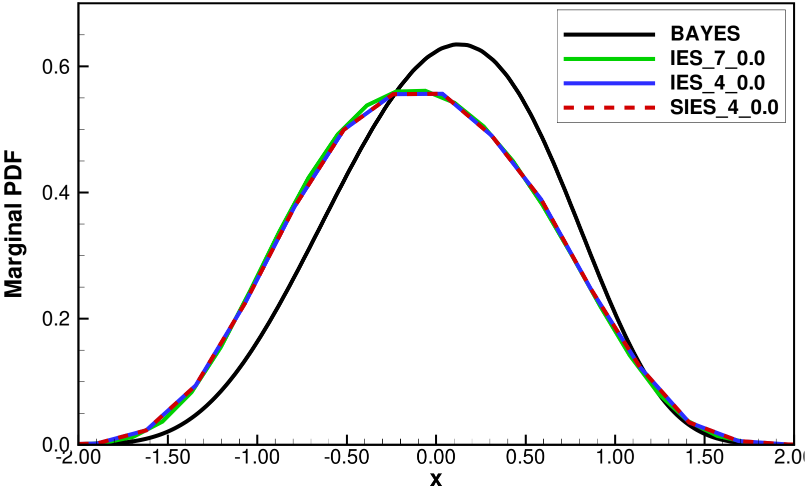
8: $\quad \boldsymbol{W}_{i+1} = \boldsymbol{W}_i - \gamma\Big(\boldsymbol{W}_i - \boldsymbol{S}_i^\mathrm{T}\big(\boldsymbol{S}_i\boldsymbol{S}_i^\mathrm{T} + \boldsymbol{C}_{dd}\big)^{-1}\boldsymbol{H}_i\Big)$ $\qquad \mathcal{O}(mN^2)$

9: $\quad \boldsymbol{X}_{i+1} = \boldsymbol{X}\big(\boldsymbol{I} + \boldsymbol{W}_{i+1}/\sqrt{N-1}\big)$ $\qquad \mathcal{O}(nN^2)$

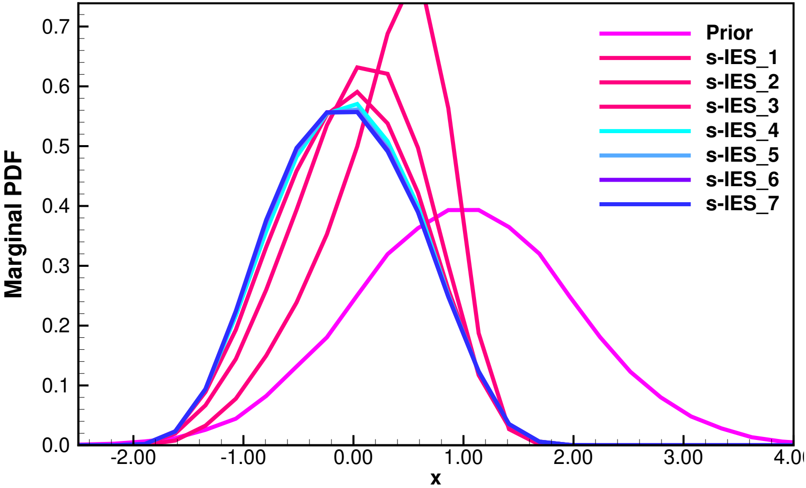
10: **end for**

- Order $\mathcal{O}(mN^2)$ and $\mathcal{O}(nN^2)$
- No pseudo inversions of large matrices.

Example nonlinear model



Iterations nonlinear model



Equation for $\boldsymbol{W}$

Standard form $(\mathcal{O}(m^3))$

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

From Woodbury, rewrite as

$$\boldsymbol{W}_{i+1} = \boldsymbol{W}_i - \gamma\Big\{\boldsymbol{W}_i - \left(\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{C}_{dd}^{-1}\boldsymbol{S}_i + \boldsymbol{I}_N\right)^{-1}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{C}_{dd}^{-1}\boldsymbol{H}\Big\}$$

For $\boldsymbol{C}_{dd} = \boldsymbol{I}_m$ we have $(\mathcal{O}(mN^2))$

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

Subspace inversion (?)

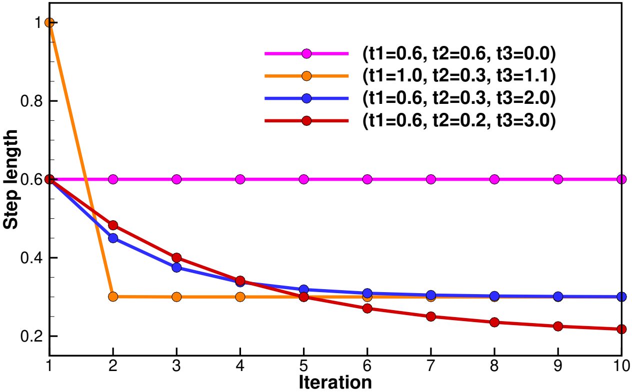
- Why invert m-dimensional matrix when solving for N coefficients?
- Do not form $\boldsymbol{C}_{dd} \approx \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}$ but work directly with $\boldsymbol{E}$.

$$
\begin{aligned}
&\big(\boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}\big) \\
&\qquad \approx \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + (\boldsymbol{S}\boldsymbol{S}^{+})\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}(\boldsymbol{S}\boldsymbol{S}^{+})^{\mathrm{T}} \\
&\qquad = \boldsymbol{U}\boldsymbol{\Sigma}\big(\boldsymbol{I}_N + \boldsymbol{\Sigma}^{+}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}\boldsymbol{U}(\boldsymbol{\Sigma}^{+})^{\mathrm{T}}\big)\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}} \\
&\qquad = \boldsymbol{U}\boldsymbol{\Sigma}\big(\boldsymbol{I}_N + \boldsymbol{Z}\boldsymbol{\Lambda}\boldsymbol{Z}^{\mathrm{T}}\big)\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}} \\
&\qquad = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{Z}\big(\boldsymbol{I}_N + \boldsymbol{\Lambda}\big)\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}.
\end{aligned}
$$

$$
\big(\boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}\big)^{-1} \approx \boldsymbol{U}(\boldsymbol{\Sigma}^{+})^{\mathrm{T}}\boldsymbol{Z}\big(\boldsymbol{I}_N + \boldsymbol{\Lambda}\big)^{-1}\big(\boldsymbol{U}(\boldsymbol{\Sigma}^{+})^{\mathrm{T}}\boldsymbol{Z}\big)^{\mathrm{T}}
$$

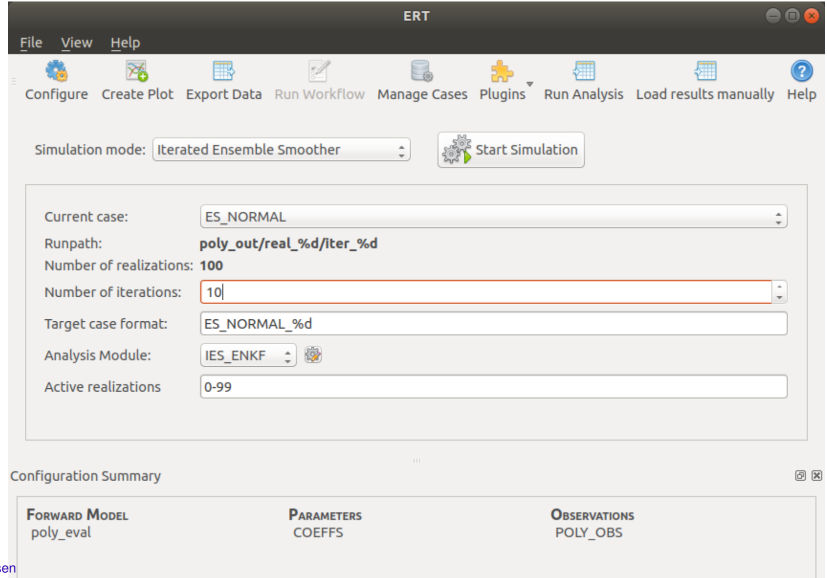
- Cost is $\mathcal{O}(mN^2)$.

Steplength scheme

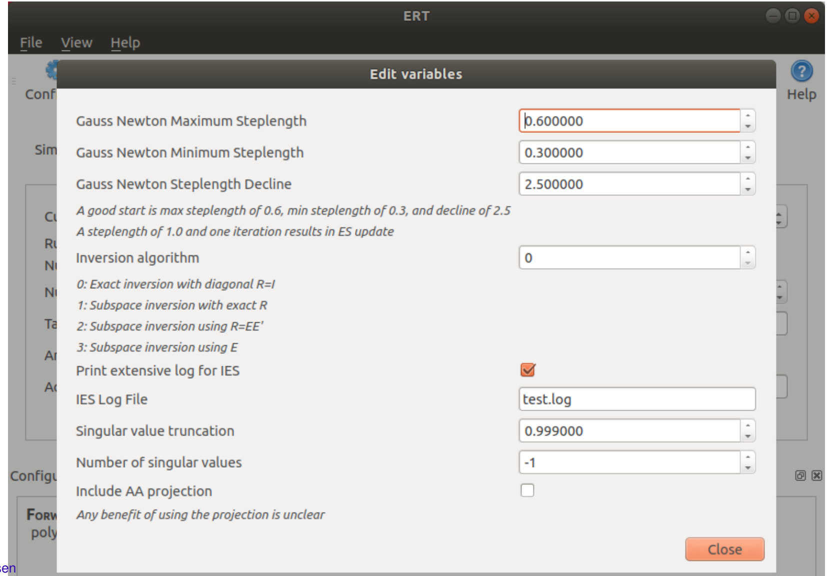


$$\gamma_i = b + (a - b)2^{\left(-(i-1)/(c-1)\right)}$$

ERT: https://github.com/equinor/ert



ERT: https://github.com/equinor/ert



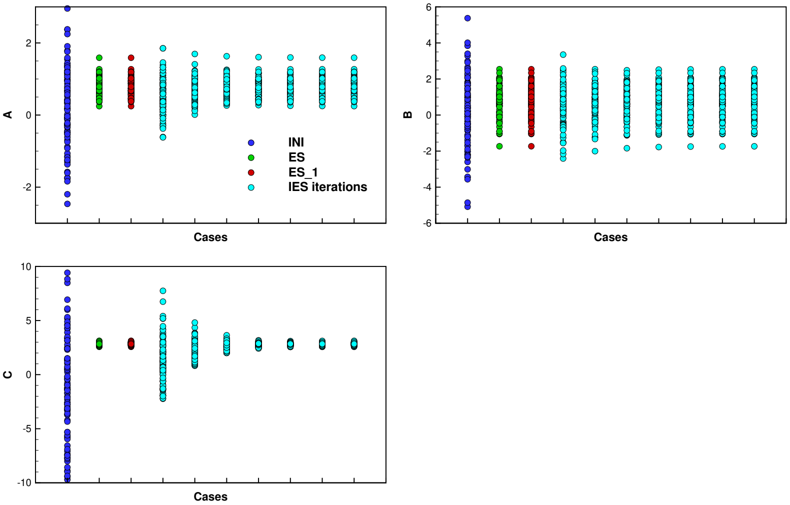
Poly case

Several simple tests are run using a “linear” model

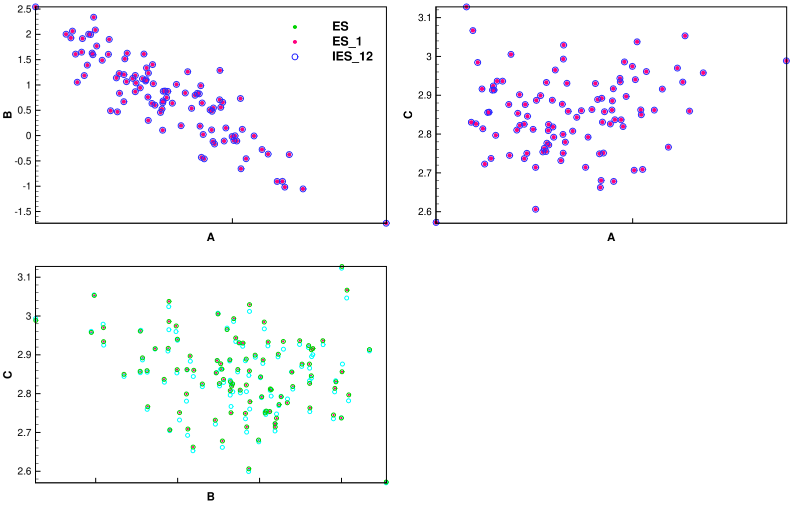
$$y(x) = ax^2 + bx + c \tag{1}$$

- Coefficients a, b, and c are random Gaussian variables.
- Measurements $(d_1, \ldots, d_5)$ at $x = (0, 2, 4, 6, 8)$.
- Polynomial curve fitting to the 5 data points.
- Gauss-linear problem solved exactly by the ES.

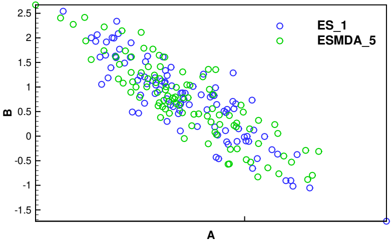
Subspace IES verification

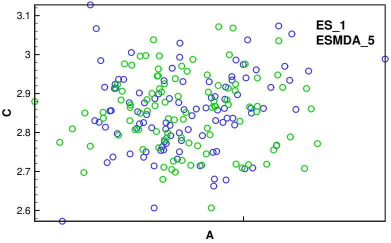


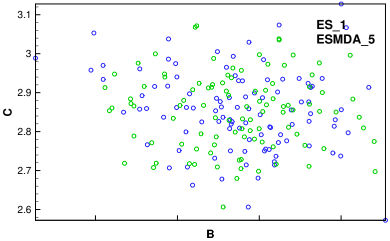
Subspace IES verification



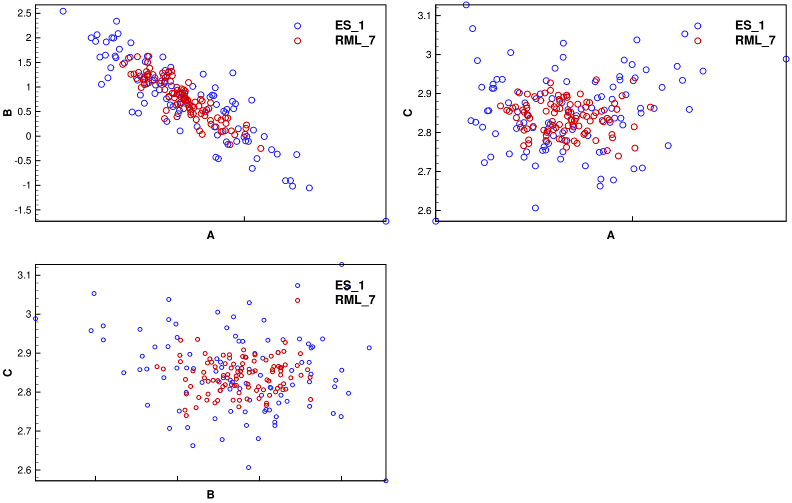
Subspace IES vs. ESMDA



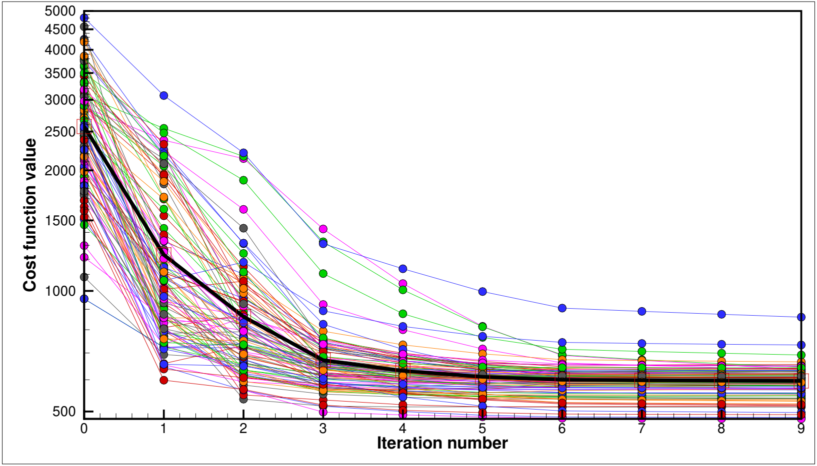




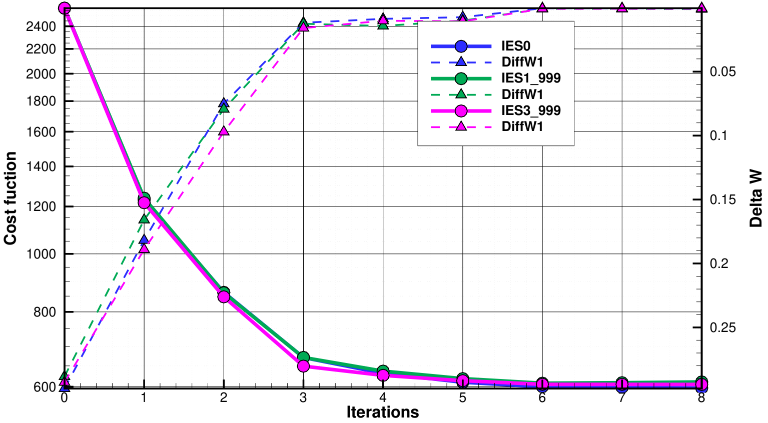
Subspace IES vs. EnRML implementation



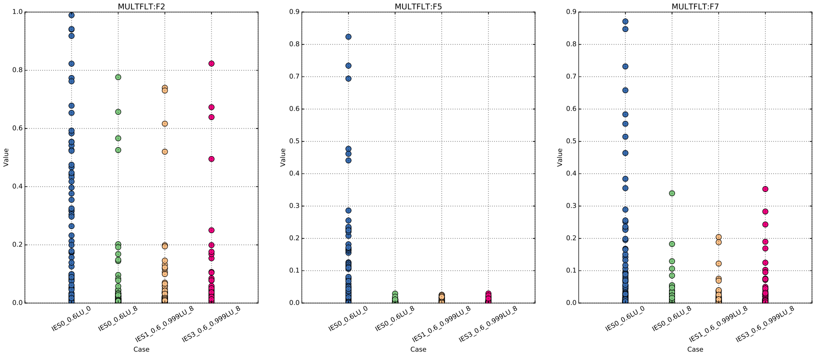
Reek case: Ensemble of cost functions



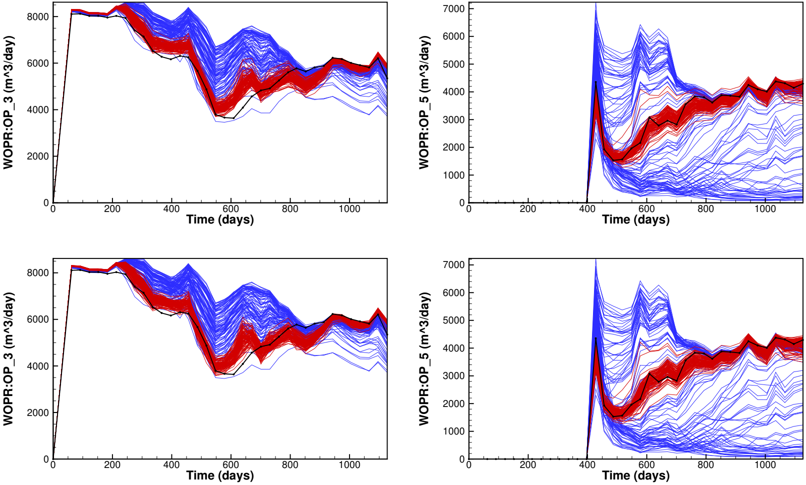
Reek case: Averaged cost function



Reek case: Fault multiplier



Reek case: Oil production



Summary

- Robust implementation of a robust IES formulation in ERT.
- IES algorithm formulated for big data and big models.
- Convergence properties meet requirements for operational use.
- Pointed out the value of test-based code development.
- Missing: module for simulating correlated measurement errors.
- Missing: guidelines for consistent conditioning on production data.

References

http://digires.no
http://digires.no/research-/publications