Implementation of IES in ERT and validation

Geir Evensen, Patrick Raanes, and Andreas Stordal

NORCE–Norwegian Research Center
Nansen Environmental and Remote Sensing Center

EnKF Workshop

Voss, June 3–6, 2019

https://www.nonlin-processes-geophys-discuss.net/npg-2019-10/

Some definitions

Prior ensemble and perturbed measurements

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

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

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

$$\Delta \boldsymbol{w}_j^i = \Bigg\{ \boldsymbol{w}_j^i - \big(\boldsymbol{G}_j^i \boldsymbol{A}\big)^{\mathrm{T}} \Big(\big(\boldsymbol{G}_j^i \boldsymbol{A}\big)\big(\boldsymbol{G}_j^i \boldsymbol{A}\big)^{\mathrm{T}} + \boldsymbol{C}_{dd} \Big)^{-1} \times \Big(\big(\boldsymbol{G}_j^i \boldsymbol{A}\big) \boldsymbol{w}_j^i + \boldsymbol{d}_j - \boldsymbol{g}\big(\boldsymbol{x}_j^{\mathrm{f}} + \boldsymbol{A}\boldsymbol{w}_j^i\big) \Big) \Bigg\}.$$

with

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

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

Define the linear regression

$$\boldsymbol{C}^i_{yx} = \boldsymbol{G}_i \boldsymbol{C}^i_{xx} \qquad \text{or} \qquad \boldsymbol{G}_i = \boldsymbol{C}^i_{yx}\left(\boldsymbol{C}^i_{xx}\right)^{-1}$$

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

Define the linear regression

$$\boldsymbol{C}_{yx}^i = \boldsymbol{G}_i\boldsymbol{C}_{xx}^i \qquad \text{or} \qquad \boldsymbol{G}_i = \boldsymbol{C}_{yx}^i\big(\boldsymbol{C}_{xx}^i\big)^{-1}$$

and write

$$\boldsymbol{G}_j^i\boldsymbol{A} \triangleq \boldsymbol{G}_i\boldsymbol{A} = \boldsymbol{C}_{yx}^i(\boldsymbol{C}_{xx}^i)^{-1}\boldsymbol{A} \qquad \text{Average sensitivity}$$

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

Define the linear regression

$$\boldsymbol{C}^i_{yx} = \boldsymbol{G}_i\boldsymbol{C}^i_{xx} \qquad \text{or} \qquad \boldsymbol{G}_i = \boldsymbol{C}^i_{yx}\left(\boldsymbol{C}^i_{xx}\right)^{-1}$$

and write

$$\begin{aligned} \boldsymbol{G}^i_j\boldsymbol{A} &\triangleq \boldsymbol{G}_i\boldsymbol{A} = \boldsymbol{C}^i_{yx}(\boldsymbol{C}^i_{xx})^{-1}\boldsymbol{A} && \text{Average sensitivity} \\ &\approx \overline{\boldsymbol{G}}_i\boldsymbol{A} = \boldsymbol{Y}_i\boldsymbol{A}_i^{\mathrm{T}}\left(\boldsymbol{A}_i\boldsymbol{A}_i^{\mathrm{T}}\right)^{+}\boldsymbol{A} && \text{Ensemble repr.} \end{aligned}$$

$G^i_j A$

Define the linear regression

$$C^i_{yx} = G_i C^i_{xx} \qquad \text{or} \qquad G_i = C^i_{yx} \left(C^i_{xx}\right)^{-1}$$

and write

$$\begin{aligned} G^i_j A &\triangleq G_i A = C^i_{yx} (C^i_{xx})^{-1} A && \text{Average sensitivity} \\ &\approx \overline{G}_i A = Y_i A_i^{\mathrm{T}} \left(A_i A_i^{\mathrm{T}}\right)^{+} A && \text{Ensemble repr.} \\ &= Y_i A_i^{+} A \end{aligned}$$

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

Define the linear regression

$$\boldsymbol{C}_{yx}^i = \boldsymbol{G}_i \boldsymbol{C}_{xx}^i \qquad \text{or} \qquad \boldsymbol{G}_i = \boldsymbol{C}_{yx}^i \left(\boldsymbol{C}_{xx}^i\right)^{-1}$$

and write

$$\begin{aligned}
\boldsymbol{G}_j^i\boldsymbol{A} &\triangleq \boldsymbol{G}_i\boldsymbol{A} = \boldsymbol{C}_{yx}^i(\boldsymbol{C}_{xx}^i)^{-1}\boldsymbol{A} && \text{Average sensitivity} \\
&\approx \overline{\boldsymbol{G}}_i\boldsymbol{A} = \boldsymbol{Y}_i\boldsymbol{A}_i^{\mathrm{T}}\left(\boldsymbol{A}_i\boldsymbol{A}_i^{\mathrm{T}}\right)^{+}\boldsymbol{A} && \text{Ensemble repr.} \\
&= \boldsymbol{Y}_i\boldsymbol{A}_i^{+}\boldsymbol{A} && (\boldsymbol{A}_i = \boldsymbol{A}\boldsymbol{\Omega}_i)
\end{aligned}$$

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

Define the linear regression

$$\boldsymbol{C}_{yx}^i = \boldsymbol{G}_i \boldsymbol{C}_{xx}^i \qquad \text{or} \qquad \boldsymbol{G}_i = \boldsymbol{C}_{yx}^i \left(\boldsymbol{C}_{xx}^i\right)^{-1}$$

and write

$$\begin{aligned}
\boldsymbol{G}_j^i \boldsymbol{A} &\triangleq \boldsymbol{G}_i \boldsymbol{A} = \boldsymbol{C}_{yx}^i (\boldsymbol{C}_{xx}^i)^{-1} \boldsymbol{A} && \text{Average sensitivity} \\
&\approx \overline{\boldsymbol{G}}_i \boldsymbol{A} = \boldsymbol{Y}_i \boldsymbol{A}_i^{\mathrm{T}} \left(\boldsymbol{A}_i \boldsymbol{A}_i^{\mathrm{T}}\right)^{+} \boldsymbol{A} && \text{Ensemble repr.} \\
&= \boldsymbol{Y}_i \boldsymbol{A}_i^{+} \boldsymbol{A} && (\boldsymbol{A}_i = \boldsymbol{A}\boldsymbol{\Omega}_i) \\
&= \boldsymbol{Y}_i \boldsymbol{A}_i^{+} \boldsymbol{A}_i \boldsymbol{\Omega}_i^{-1}
\end{aligned}$$

$G^i_j A$

Define the linear regression

$$C^i_{yx} = G_i C^i_{xx} \qquad \text{or} \qquad G_i = C^i_{yx}\left(C^i_{xx}\right)^{-1}$$

and write

$$\begin{aligned}
G^i_j A &\triangleq G_i A = C^i_{yx}(C^i_{xx})^{-1} A && \text{Average sensitivity} \\
&\approx \overline{G}_i A = Y_i A_i^{\mathrm{T}} \left(A_i A_i^{\mathrm{T}}\right)^{+} A && \text{Ensemble repr.} \\
&= Y_i A_i^{+} A && (A_i = A\Omega_i) \\
&= Y_i A_i^{+} A_i \Omega_i^{-1} && A_i^{+} A_i = \Pi_{A^{\mathrm{T}}} = \left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\right)
\end{aligned}$$

$G^i_j A$

Define the linear regression

$$C^i_{yx} = G_i C^i_{xx} \qquad \text{or} \qquad G_i = C^i_{yx}\left(C^i_{xx}\right)^{-1}$$

and write

$$\begin{aligned}
G^i_j A &\triangleq G_i A = C^i_{yx}(C^i_{xx})^{-1} A && \text{Average sensitivity} \\
&\approx \overline{G}_i A = Y_i A_i^{\mathrm{T}}\left(A_i A_i^{\mathrm{T}}\right)^{+} A && \text{Ensemble repr.} \\
&= Y_i A_i^{+} A && (A_i = A\Omega_i) \\
&= Y_i A_i^{+} A_i \Omega_i^{-1} && A_i^{+} A_i = \Pi_{A^{\mathrm{T}}} = \left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\right) \\
&= S_i = \begin{cases} Y_i \Omega_i^{-1} & \text{linear case} \\ & \end{cases}
\end{aligned}$$

$G^i_j A$

Define the linear regression

$$C^i_{yx} = G_i C^i_{xx} \qquad \text{or} \qquad G_i = C^i_{yx}\big(C^i_{xx}\big)^{-1}$$

and write

$$\begin{aligned}
G^i_j A &\triangleq G_i A = C^i_{yx}(C^i_{xx})^{-1} A && \text{Average sensitivity} \\
&\approx \overline{G}_i A = Y_i A_i^{\mathrm{T}}\big(A_i A_i^{\mathrm{T}}\big)^{+} A && \text{Ensemble repr.} \\
&= Y_i A_i^{+} A && (A_i = A\Omega_i) \\
&= Y_i A_i^{+} A_i \Omega_i^{-1} && A_i^{+} A_i = \Pi_{A^{\mathrm{T}}} = \Big(I - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\Big) \\
&= S_i = \begin{cases} Y_i \Omega_i^{-1} & \text{linear case} \\ Y_i \Omega_i^{-1} & n \geq N-1, \\ \end{cases}
\end{aligned}$$

$G_j^i A$

Define the linear regression

$$C_{yx}^i = G_i C_{xx}^i \qquad \text{or} \qquad G_i = C_{yx}^i \left(C_{xx}^i\right)^{-1}$$

and write

$$\begin{aligned}
G_j^i A &\triangleq G_i A = C_{yx}^i (C_{xx}^i)^{-1} A && \text{Average sensitivity} \\
&\approx \overline{G}_i A = Y_i A_i^{\mathrm{T}} \left(A_i A_i^{\mathrm{T}}\right)^{+} A && \text{Ensemble repr.} \\
&= Y_i A_i^{+} A && (A_i = A\Omega_i) \\
&= Y_i A_i^{+} A_i \Omega_i^{-1} && A_i^{+} A_i = \Pi_{A^{\mathrm{T}}} = \left(I - \frac{1}{N} \mathbf{1}\mathbf{1}^{\mathrm{T}}\right) \\
&= S_i = \begin{cases} Y_i \Omega_i^{-1} & \text{linear case} \\ Y_i \Omega_i^{-1} & n \geq N-1, \\ Y_i A_i^{+} A_i \Omega_i^{-1} & n < N-1, \end{cases}
\end{aligned}$$

Equation for $\boldsymbol{W}$

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

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

IES ensemble subspace algorithm

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

$\boldsymbol{W}_1 = 0$

for $i = 1, \mathrm{Convergence}$ **do**

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

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

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

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

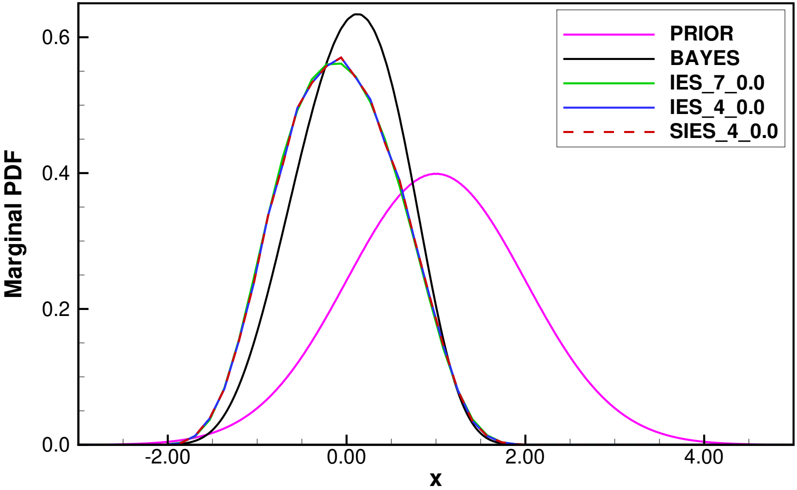
$\quad \boldsymbol{W}_{i+1} = \boldsymbol{W}_i - \gamma\Big(\boldsymbol{W}_i - \boldsymbol{S}_i^{\mathrm{T}}\big(\boldsymbol{S}_i\boldsymbol{S}_i^{\mathrm{T}} + \boldsymbol{C}_{dd}\big)^{-1}\boldsymbol{H}_i\Big)$ $\qquad \mathcal{O}(mN^2)$

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

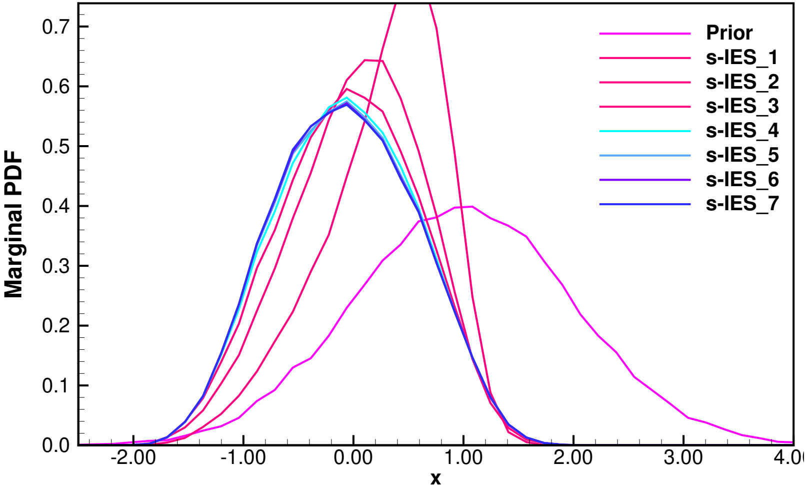
end for

- 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 ($\mathcal{O}(m^3)$)

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

From Woodbury, rewrite as

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

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

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

Subspace inversion (*Evensen*, 2004)

- Why invert m-dimensional matrix when solving for N coefficients?

$$
\begin{aligned}
&\left(\boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + \boldsymbol{C}_{dd}\right) \\
&= \left(\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}} + \boldsymbol{C}_{dd}\right) \\
&\approx \boldsymbol{U}\boldsymbol{\Sigma}\left(\boldsymbol{I}_N + \boldsymbol{\Sigma}^{+}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{C}_{dd}\boldsymbol{U}(\boldsymbol{\Sigma}^{+})^{\mathrm{T}}\right)\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}} \\
&\qquad = \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + (\boldsymbol{S}\boldsymbol{S}^{+})\boldsymbol{C}_{dd}(\boldsymbol{S}\boldsymbol{S}^{+})^{\mathrm{T}} \\
&= \boldsymbol{U}\boldsymbol{\Sigma}\left(\boldsymbol{I}_N + \boldsymbol{Z}\boldsymbol{\Lambda}\boldsymbol{Z}^{\mathrm{T}}\right)\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}} \\
&= \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{Z}\left(\boldsymbol{I}_N + \boldsymbol{\Lambda}\right)\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}
\end{aligned}
$$

$$
\left(\boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + \boldsymbol{C}_{dd}\right)^{-1} \approx \boldsymbol{U}(\boldsymbol{\Sigma}^{+})^{\mathrm{T}}\boldsymbol{Z}\left(\boldsymbol{I}_N + \boldsymbol{\Lambda}\right)^{-1}\left(\boldsymbol{U}(\boldsymbol{\Sigma}^{+})^{\mathrm{T}}\boldsymbol{Z}\right)^{\mathrm{T}}
$$

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

Subspace inversion with $\boldsymbol{C}_{dd} \approx \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}$.

NORCE

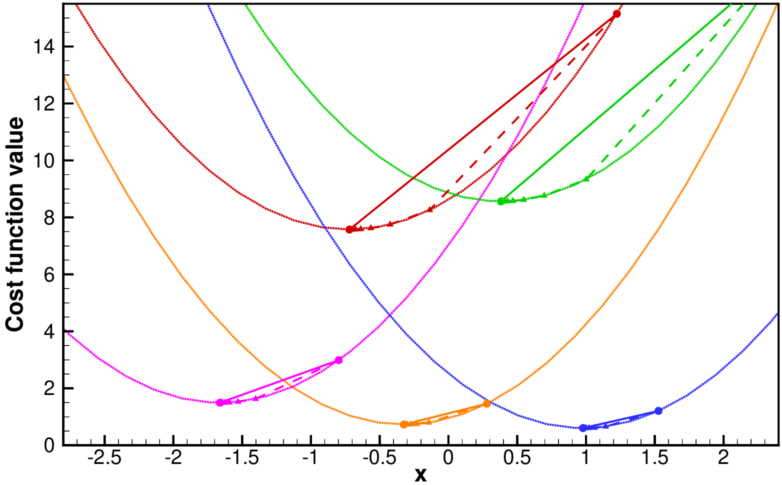
- Do not form $\boldsymbol{C}_{dd}$ but work directly with $\boldsymbol{E}$.

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

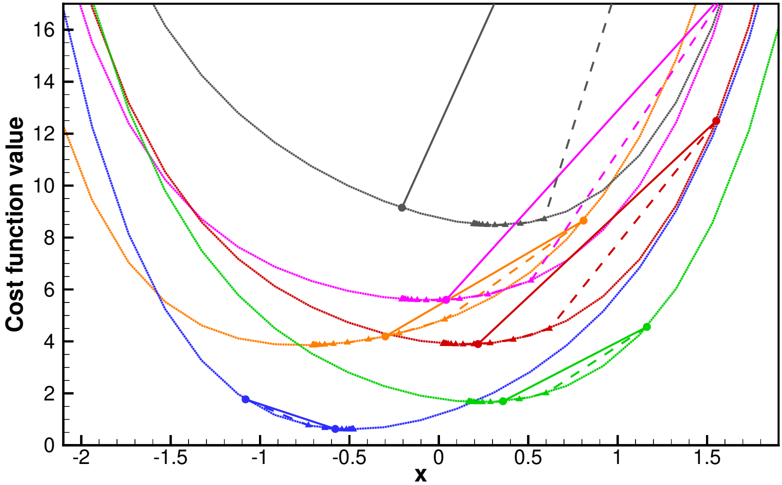
$$
\big(\boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}\big)^{-1} \approx \boldsymbol{U}(\boldsymbol{\Sigma}^{+})^{\mathrm{T}}\boldsymbol{Z}\big(\boldsymbol{I}_N + \boldsymbol{\Lambda}\big)^{-1}\big(\boldsymbol{U}(\boldsymbol{\Sigma}^{+})^{\mathrm{T}}\boldsymbol{Z}\big)^{\mathrm{T}}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

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

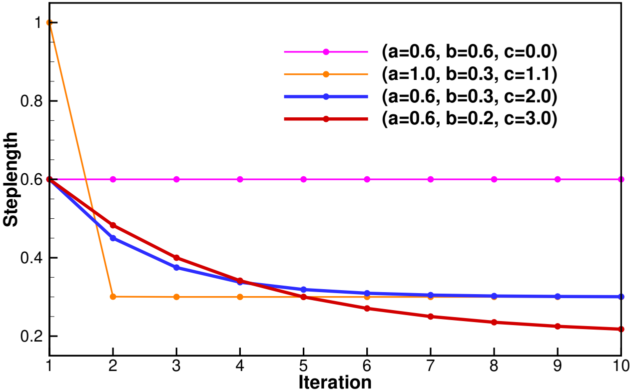
IES costfunctions: Linear case



IES costfunctions: Nonlinear case

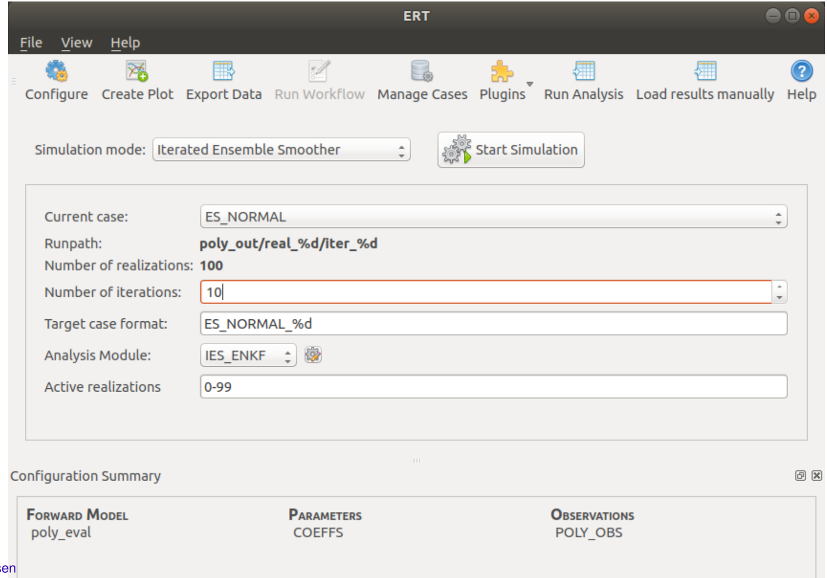


Steplength scheme

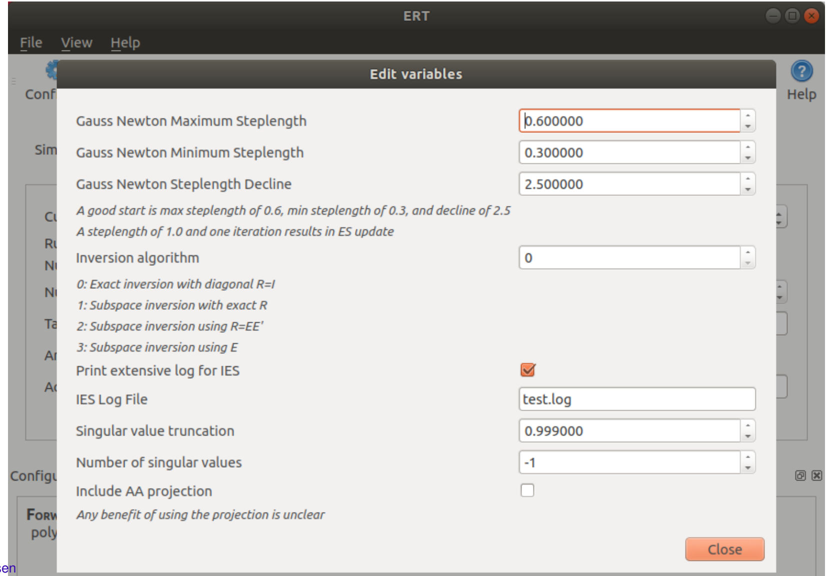


$$\gamma_i = b + (a-b)2^{\left(-(i-1)/(c-1)\right)}$$

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



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



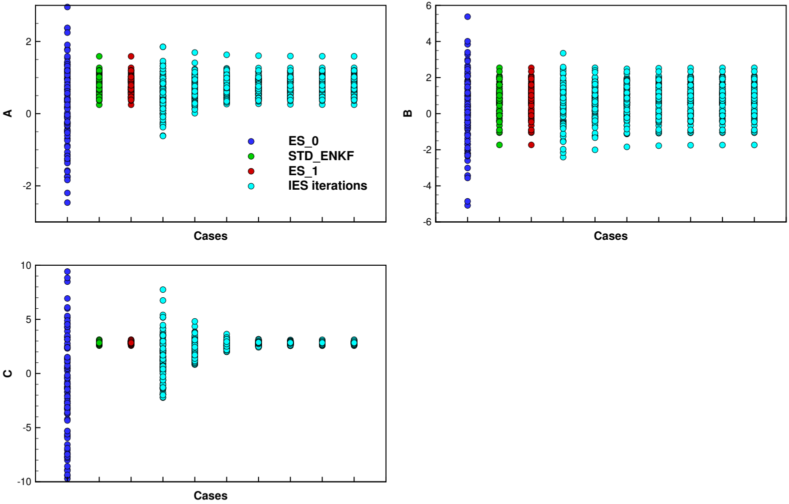
Poly case

Several simple tests are run using a "linear" model

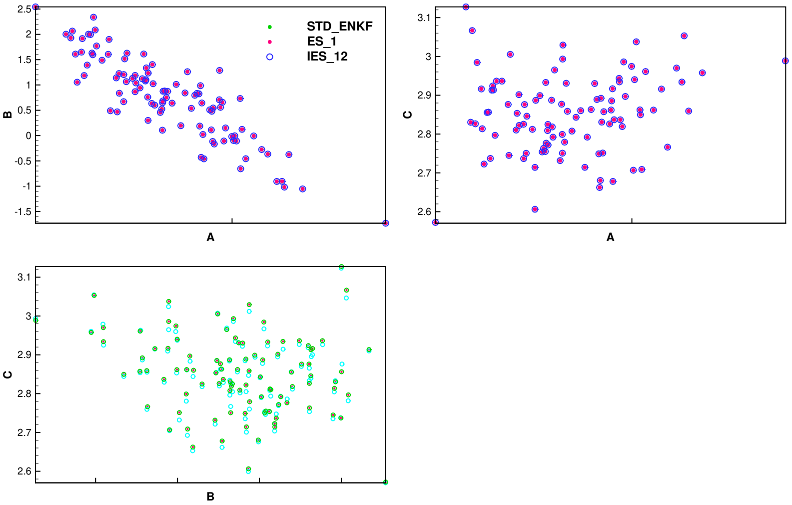
$$y(x) = ax^2 + bx + c \tag{1}$$

- 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.

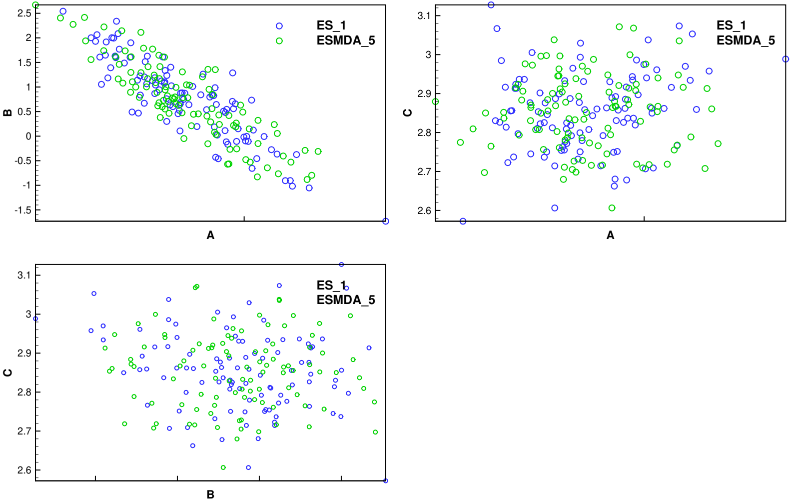
Subspace IES verification



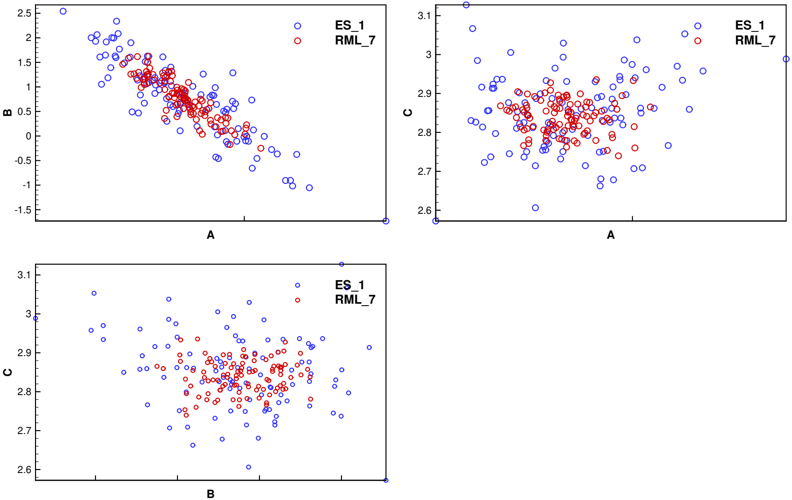
Subspace IES verification



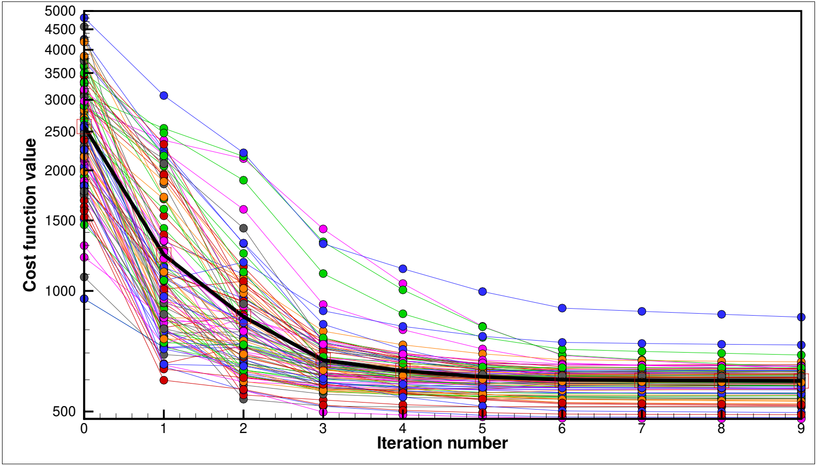
Subspace IES vs. ESMDA



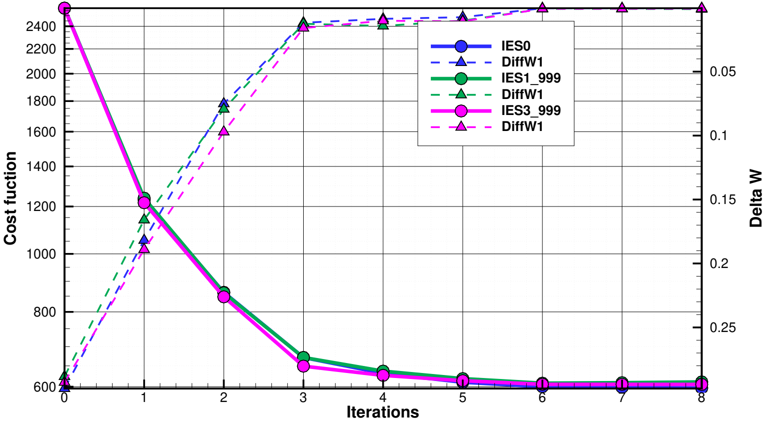
Subspace IES vs. EnRML implementation



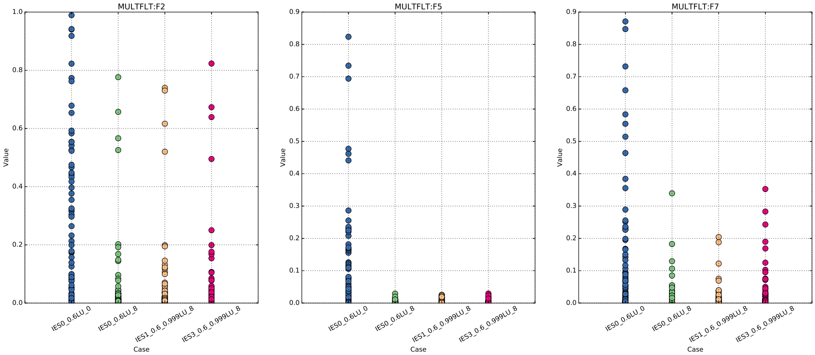
Reek case: Ensemble of cost functions



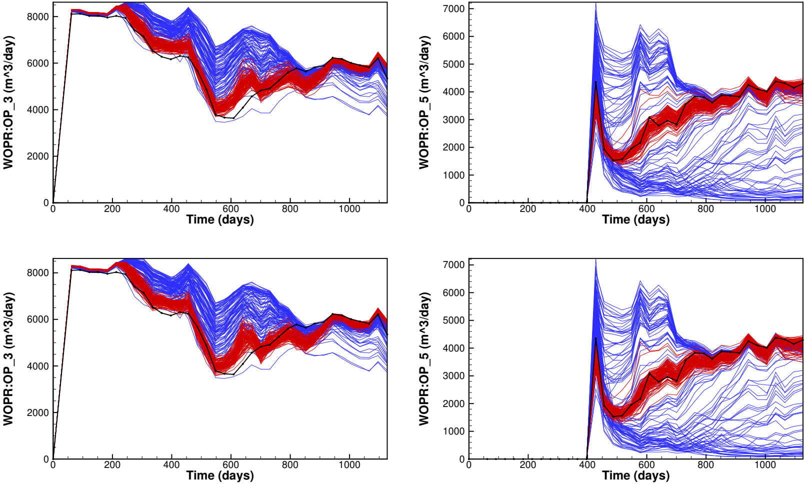
Reek case: Averaged cost function



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.
- ERT is a flexible tool for reservoir HM.

References

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

Chen, Y., and D. S. Oliver, Ensemble randomized maximum likelihood method as an iterative ensemble smoother, *Math. Geosci.*, *44*, 1–26, 2012.

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.

Evensen, G., Sampling strategies and square root analysis schemes for the EnKF, *Ocean Dynamics*, *54*, 539–560, 2004.

Evensen, G., Analysis of iterative ensemble smoothers for solving inverse problems, *Computat Geosci*, *22*, pp. 885–908, 2018, https://doi.org/10.1007/s10596-018-9731-y.

Evensen, G., Accounting for model errors in iterative ensemble smoothers, *Computat Geosci*, 2019, https://doi.org/10.1007/s10596-019-9819-z.

Raanes, P. N., A. S. Stordal, and G. Evensen, Revising the stochastic iterative ensemble smoother, *Nonlinear Processes in Geophysics Discussions*, *2019*, 1–22, 2019.