

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

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

Ensemble means

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

Ensemble anomaly matrices and covariances

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

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

IES: Ensemble subspace version

Original cost functions

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

Solution is contained in the ensemble subspace, thus

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

and,

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

Reduces dimension of problem from state size to ensemble size.

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

Gradient and Hessian of cost function

Gradient

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

Hessian (approximate)

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

Gauss-Newton iterations

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

$$\Delta \mathbf{w}_j^i = \left\{ \mathbf{w}_j^i - (\mathbf{G}_j^i \mathbf{A})^T \left((\mathbf{G}_j^i \mathbf{A}) (\mathbf{G}_j^i \mathbf{A})^T + \mathbf{C}_{dd} \right)^{-1} \right. \\ \left. \times \left((\mathbf{G}_j^i \mathbf{A}) \mathbf{w}_j^i + \mathbf{d}_j - \mathbf{g}(\mathbf{x}_j^f + \mathbf{A} \mathbf{w}_j^i) \right) \right\}.$$

with

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

$G_j^i A$

Define the linear regression

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

$G_j^i A$

Define the linear regression

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

and write

$$G_j^i A \triangleq G_i A = C_{yx}^i (C_{xx}^i)^{-1} A \quad \text{Average sensitivity}$$

$G_j^i A$

Define the linear regression

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

and write

$$\begin{aligned} G_j^i A &\triangleq G_i A = C_{yx}^i (C_{xx}^i)^{-1} A \\ &\approx \bar{G}_i A = Y_i A_i^T (A_i A_i^T)^+ A \end{aligned}$$

Average sensitivity

Ensemble repr.

$G_j^i A$

Define the linear regression

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

and write

$$\begin{aligned} G_j^i A &\triangleq G_i A = C_{yx}^i (C_{xx}^i)^{-1} A \\ &\approx \bar{G}_i A = Y_i A_i^T (A_i A_i^T)^+ A \\ &= Y_i A_i^+ A \end{aligned}$$

Average sensitivity

Ensemble repr.

$$G_j^i A$$

Define the linear regression

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

and write

$$\begin{aligned} G_j^i A &\triangleq G_i A = C_{yx}^i (C_{xx}^i)^{-1} A \\ &\approx \bar{G}_i A = Y_i A_i^T (A_i A_i^T)^+ A \\ &= Y_i A_i^+ A \end{aligned}$$

Average sensitivity

Ensemble repr.

$$(A_i = A \Omega_i)$$

$G_j^i A$

Define the linear regression

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

and write

$$\begin{aligned} G_j^i A &\triangleq G_i A = C_{yx}^i (C_{xx}^i)^{-1} A \\ &\approx \bar{G}_i A = Y_i A_i^T (A_i A_i^T)^+ A \\ &= Y_i A_i^+ A \\ &= Y_i A_i^+ A_i \Omega_i^{-1} \end{aligned}$$

Average sensitivity

Ensemble repr.

$$(A_i = A \Omega_i)$$

$G_j^i A$

Define the linear regression

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

and write

$$\begin{aligned} G_j^i A &\triangleq G_i A = C_{yx}^i (C_{xx}^i)^{-1} A \\ &\approx \bar{G}_i A = Y_i A_i^T (A_i A_i^T)^+ A \\ &= Y_i A_i^+ A \\ &= Y_i A_i^+ A_i \Omega_i^{-1} \end{aligned}$$

Average sensitivity

Ensemble repr.

$$(A_i = A \Omega_i)$$

$$A_i^+ A_i = \Pi_{A^T} = \left(I - \frac{1}{N} \mathbf{1} \mathbf{1}^T \right)$$

$G_j^i A$

Define the linear regression

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

and write

$$G_j^i A \triangleq G_i A = C_{yx}^i (C_{xx}^i)^{-1} A$$

$$\approx \bar{G}_i A = Y_i A_i^T (A_i A_i^T)^+ A$$

$$= Y_i A_i^+ A$$

$$= Y_i A_i^+ A_i \Omega_i^{-1}$$

$$= S_i = \begin{cases} Y_i \Omega_i^{-1} & \text{linear case} \end{cases}$$

Average sensitivity

Ensemble repr.

$$(A_i = A \Omega_i)$$

$$A_i^+ A_i = \Pi_{A^T} = \left(I - \frac{1}{N} \mathbf{1} \mathbf{1}^T \right)$$

$G_j^i A$

Define the linear regression

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

and write

$$G_j^i A \triangleq G_i A = C_{yx}^i (C_{xx}^i)^{-1} A$$

$$\approx \bar{G}_i A = Y_i A_i^T (A_i A_i^T)^+ A$$

$$= Y_i A_i^+ A$$

$$= Y_i A_i^+ A_i \Omega_i^{-1}$$

$$= S_i = \begin{cases} Y_i \Omega_i^{-1} \\ Y_i \Omega_i^{-1} \end{cases}$$

linear case

$n \geq N - 1,$

Average sensitivity

Ensemble repr.

$$(A_i = A \Omega_i)$$

$$A_i^+ A_i = \Pi_{A^T} = \left(I - \frac{1}{N} \mathbf{1} \mathbf{1}^T \right)$$

$G_j^i A$

Define the linear regression

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

and write

$$G_j^i A \triangleq G_i A = C_{yx}^i (C_{xx}^i)^{-1} A$$

$$\approx \bar{G}_i A = Y_i A_i^T (A_i A_i^T)^+ A$$

$$= Y_i A_i^+ A$$

$$= Y_i A_i^+ A_i \Omega_i^{-1}$$

$$= S_i = \begin{cases} Y_i \Omega_i^{-1} & \text{linear case} \\ Y_i \Omega_i^{-1} & n \geq N - 1, \\ Y_i A_i^+ A_i \Omega_i^{-1} & n < N - 1, \end{cases}$$

Average sensitivity

Ensemble repr.

$$(A_i = A \Omega_i)$$

$$A_i^+ A_i = \Pi_{A^T} = \left(I - \frac{1}{N} \mathbf{1} \mathbf{1}^T \right)$$

Equation for W

Matrix form with $S_i = Y_i \Omega_i^{-1}$

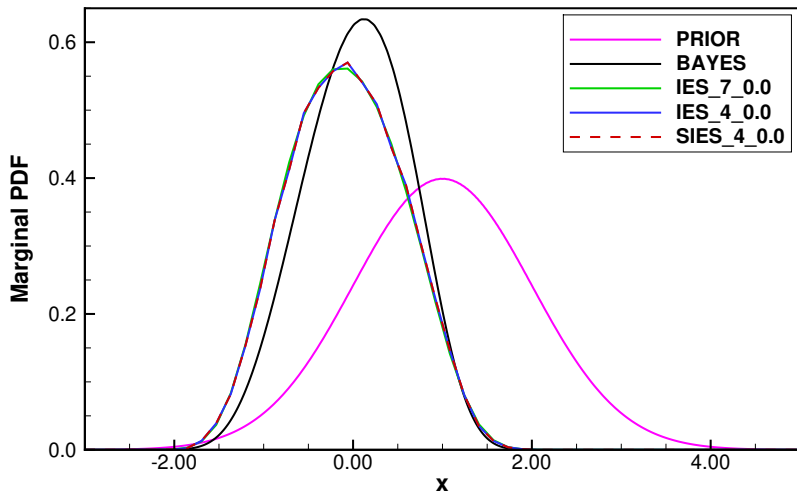
$$W_{i+1} = W_i -$$

$$\gamma \left(W_i - S_i^T (S_i S_i^T + C_{dd})^{-1} (S_i W_i - D + g(X_i)) \right)$$

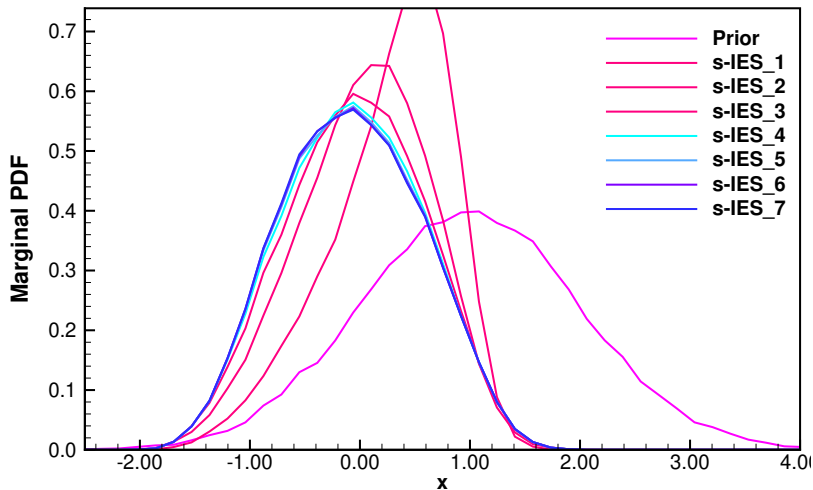
IES ensemble subspace algorithm

- 1: Inputs: \mathbf{X} , \mathbf{D} , (and \mathbf{C}_{dd})
 - 2: $\mathbf{W}_1 = \mathbf{0}$
 - 3: **for** $i = 1$, Convergence **do**
 - 4: $\mathbf{Y}_i = \mathbf{g}(\mathbf{X}_i)(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T)/\sqrt{N-1}$
 - 5: $\mathbf{\Omega}_i = \mathbf{I} + \mathbf{W}_i(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T)/\sqrt{N-1}$
 - 6: $\mathbf{\Omega}_i^T \mathbf{S}_i^T = \mathbf{Y}_i^T$ $\mathcal{O}(mN^2)$
 - 7: $\mathbf{H}_i = \mathbf{S}_i \mathbf{W}_i + \mathbf{D} - \mathbf{g}(\mathbf{X}_i)$ $\mathcal{O}(mN^2)$
 - 8: $\mathbf{W}_{i+1} = \mathbf{W}_i - \gamma(\mathbf{W}_i - \mathbf{S}_i^T(\mathbf{S}_i \mathbf{S}_i^T + \mathbf{C}_{dd})^{-1} \mathbf{H}_i)$ $\mathcal{O}(mN^2)$
 - 9: $\mathbf{X}_{i+1} = \mathbf{X}(\mathbf{I} + \mathbf{W}_{i+1}/\sqrt{N-1})$ $\mathcal{O}(nN^2)$
 - 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 W

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

$$W_{i+1} = W_i - \gamma \left(W_i - S_i^T (S_i S_i^T + C_{dd})^{-1} H_i \right)$$

From Woodbury, rewrite as

$$W_{i+1} = W_i - \gamma \left\{ W_i - (S_i^T C_{dd}^{-1} S_i + I_N)^{-1} S_i^T C_{dd}^{-1} H \right\}$$

For $C_{dd} = I_m$ we have ($\mathcal{O}(mN^2)$)

$$W_{i+1} = W_i - \gamma \left\{ W_i - (S_i^T S_i + I_N)^{-1} S_i^T H \right\}$$

Subspace inversion (*Evensen, 2004*)

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

$$\begin{aligned}
 & (\mathbf{S}\mathbf{S}^T + \mathbf{C}_{dd}) \\
 &= (\mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T + \mathbf{C}_{dd}) \\
 &\approx \mathbf{U}\mathbf{\Sigma}(\mathbf{I}_N + \mathbf{\Sigma}^+\mathbf{U}^T\mathbf{C}_{dd}\mathbf{U}(\mathbf{\Sigma}^+)^T)\mathbf{\Sigma}^T\mathbf{U}^T \\
 &\quad = \mathbf{S}\mathbf{S}^T + (\mathbf{S}\mathbf{S}^+)\mathbf{C}_{dd}(\mathbf{S}\mathbf{S}^+)^T \\
 &= \mathbf{U}\mathbf{\Sigma}(\mathbf{I}_N + \mathbf{Z}\mathbf{\Lambda}\mathbf{Z}^T)\mathbf{\Sigma}^T\mathbf{U}^T \\
 &= \mathbf{U}\mathbf{\Sigma}\mathbf{Z}(\mathbf{I}_N + \mathbf{\Lambda})\mathbf{Z}^T\mathbf{\Sigma}^T\mathbf{U}^T \\
 &(\mathbf{S}\mathbf{S}^T + \mathbf{C}_{dd})^{-1} \approx \mathbf{U}(\mathbf{\Sigma}^+)^T\mathbf{Z}(\mathbf{I}_N + \mathbf{\Lambda})^{-1}(\mathbf{U}(\mathbf{\Sigma}^+)^T\mathbf{Z})^T
 \end{aligned}$$

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

Subspace inversion with $C_{dd} \approx \mathbf{E}\mathbf{E}^T$.

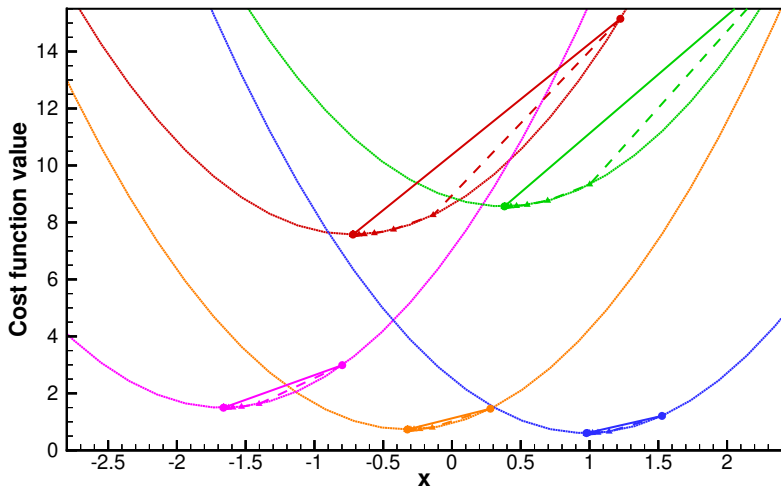
- ▶ Do not form C_{dd} but work directly with \mathbf{E} .

$$\begin{aligned}
 & (\mathbf{S}\mathbf{S}^T + \mathbf{E}\mathbf{E}^T) \\
 &= (\mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T + \mathbf{E}\mathbf{E}^T) \\
 &\approx \mathbf{U}\mathbf{\Sigma}(\mathbf{I}_N + \mathbf{\Sigma}^+\mathbf{U}^T\mathbf{E}\mathbf{E}^T\mathbf{U}(\mathbf{\Sigma}^+)^T)\mathbf{\Sigma}^T\mathbf{U}^T \\
 &\quad = \mathbf{S}\mathbf{S}^T + (\mathbf{S}\mathbf{S}^+)\mathbf{E}\mathbf{E}^T(\mathbf{S}\mathbf{S}^+)^T \\
 &= \mathbf{U}\mathbf{\Sigma}(\mathbf{I}_N + \mathbf{Z}\mathbf{\Lambda}\mathbf{Z}^T)\mathbf{\Sigma}^T\mathbf{U}^T \\
 &= \mathbf{U}\mathbf{\Sigma}\mathbf{Z}(\mathbf{I}_N + \mathbf{\Lambda})\mathbf{Z}^T\mathbf{\Sigma}^T\mathbf{U}^T
 \end{aligned}$$

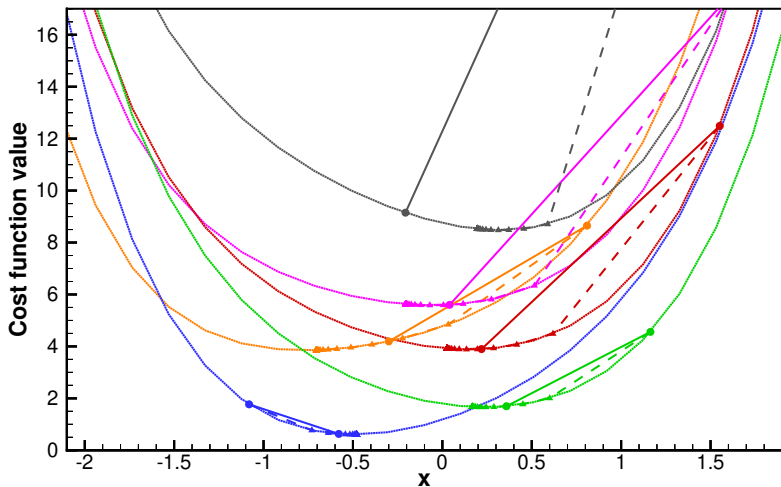
$$(\mathbf{S}\mathbf{S}^T + \mathbf{E}\mathbf{E}^T)^{-1} \approx \mathbf{U}(\mathbf{\Sigma}^+)^T\mathbf{Z}(\mathbf{I}_N + \mathbf{\Lambda})^{-1}(\mathbf{U}(\mathbf{\Sigma}^+)^T\mathbf{Z})^T$$

- ▶ Cost is $\mathcal{O}(mN^2)$.

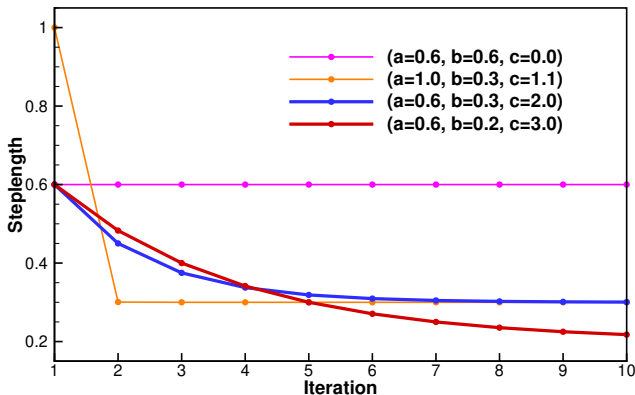
IES costfunctions: Linear case



IES costfunctions: Nonlinear case

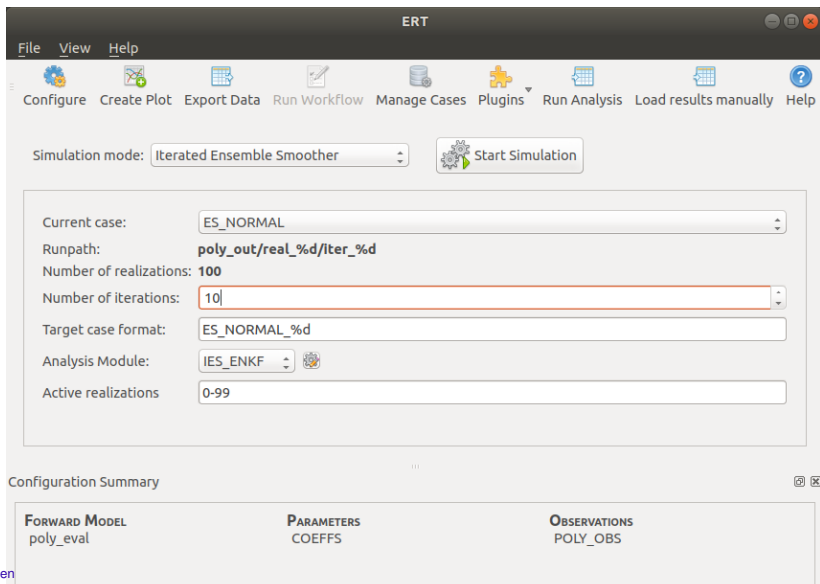


Steplength scheme



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

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



ERT

File View Help

Configure Create Plot Export Data Run Workflow Manage Cases Plugins Run Analysis Load results manually Help

Simulation mode: Iterated Ensemble Smoother Start Simulation

Current case: ES_NORMAL

Runpath: poly_out/real_%d/iter_%d

Number of realizations: 100

Number of iterations: 10

Target case format: ES_NORMAL_%d

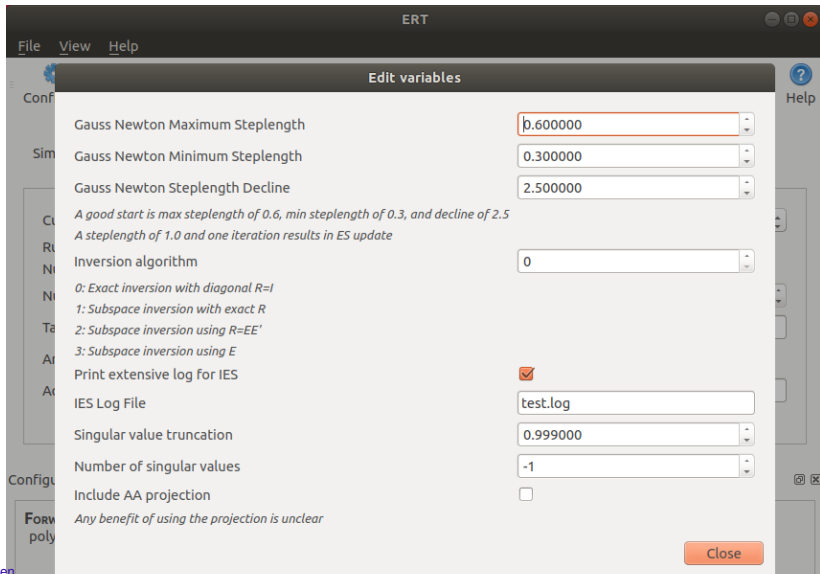
Analysis Module: IES_ENKF

Active realizations: 0-99

Configuration Summary

FORWARD MODEL	PARAMETERS	OBSERVATIONS
poly_eval	COEFFS	POLY_OBS

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



ERT

File View Help

Conf

Sim

CT

R

Ni

Ni

Ta

Ar

Ac

Config

FOR

poly

Help

Edit variables

Gauss Newton Maximum Steplength 0.600000

Gauss Newton Minimum Steplength 0.300000

Gauss Newton Steplength Decline 2.500000

*A good start is max steplength of 0.6, min steplength of 0.3, and decline of 2.5
A steplength of 1.0 and one iteration results in ES update*

Inversion algorithm 0

*0: Exact inversion with diagonal R=I
1: Subspace inversion with exact R
2: Subspace inversion using R=EE'
3: Subspace inversion using E*

Print extensive log for IES

IES Log File test.log

Singular value truncation 0.999000

Number of singular values -1

Include AA projection

Any benefit of using the projection is unclear

Close

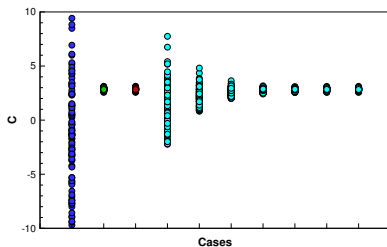
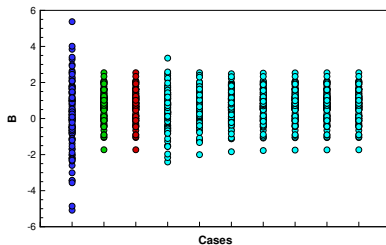
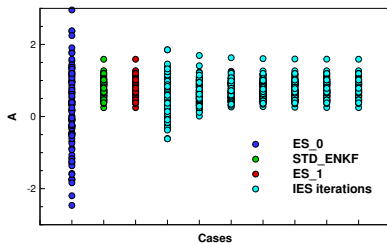
Poly case

Several simple tests are run using a “linear” model

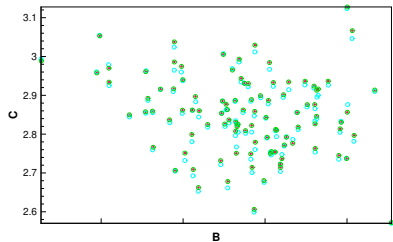
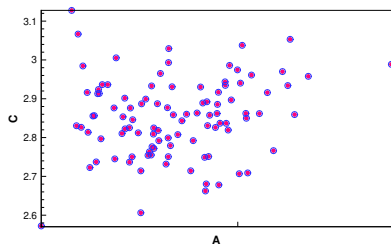
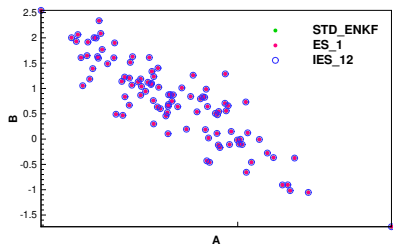
$$y(x) = ax^2 + bx + c \quad (1)$$

- ▶ Coefficients a , b , and c are random Gaussian variables.
- ▶ Measurements (d_1, \dots, 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.

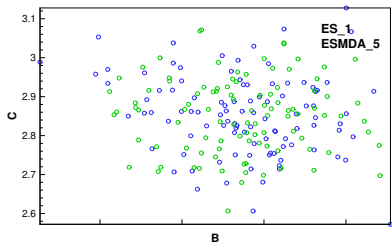
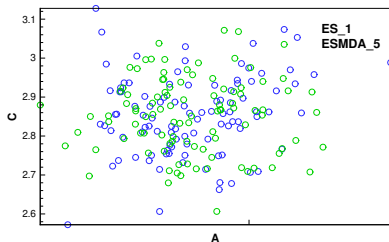
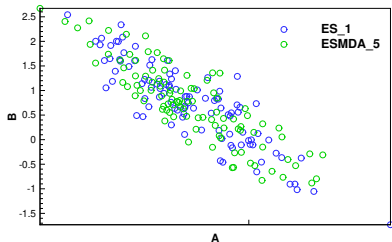
Subspace IES verification



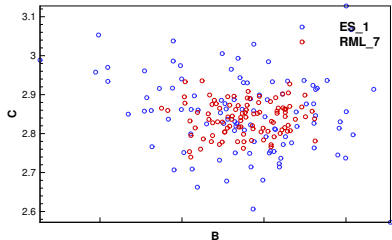
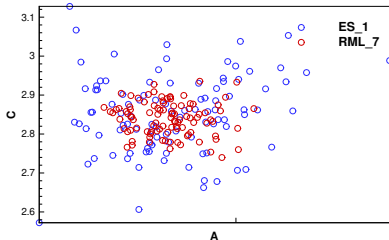
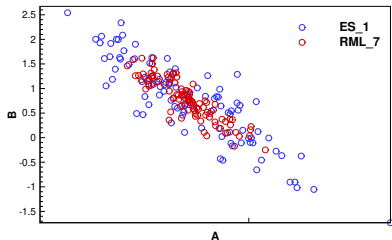
Subspace IES verification



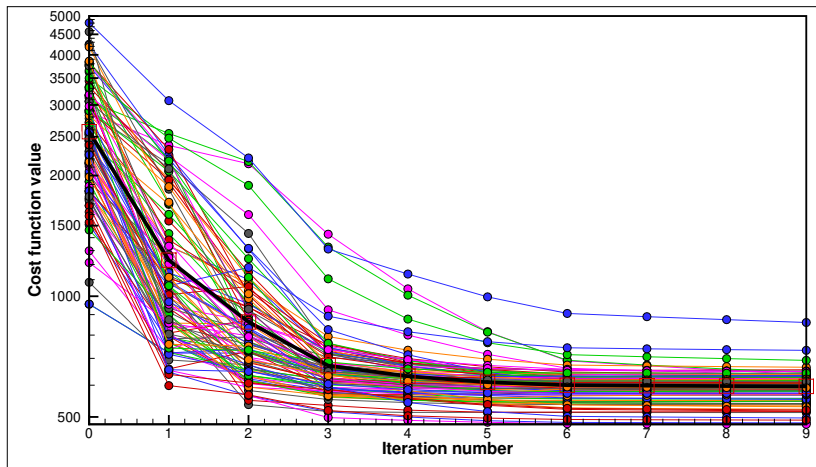
Subspace IES vs. ESM DA



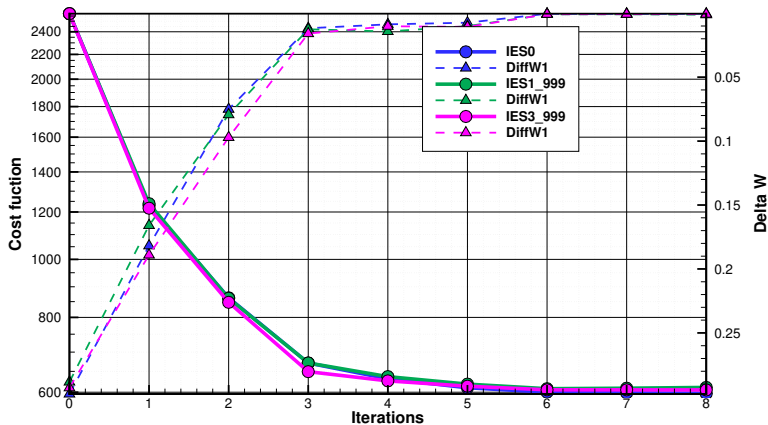
Subspace IES vs. EnRML implementation



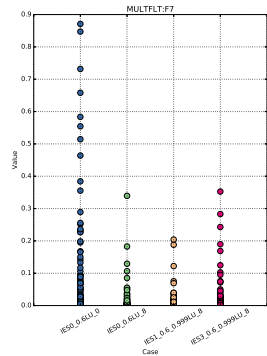
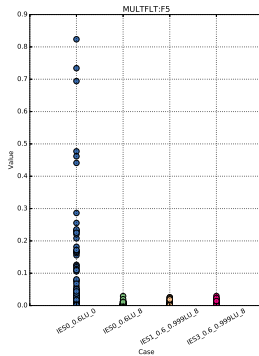
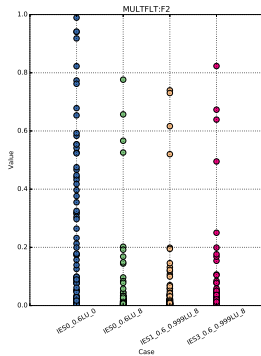
Reek case: Ensemble of cost functions



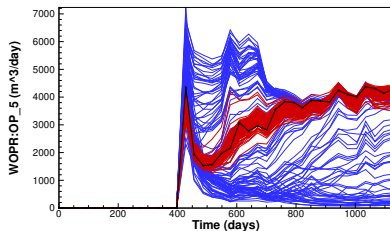
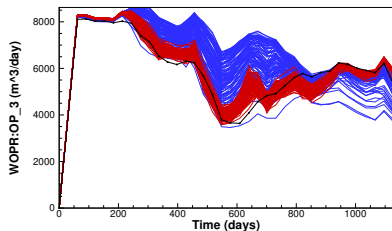
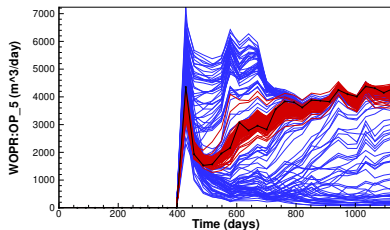
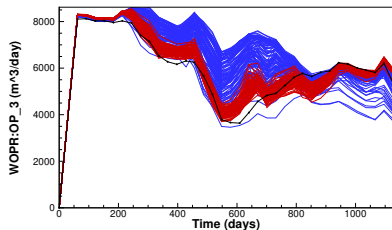
Reek case: Averaged cost function



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.