Accounting for model errors in iterative ensemble smoothers

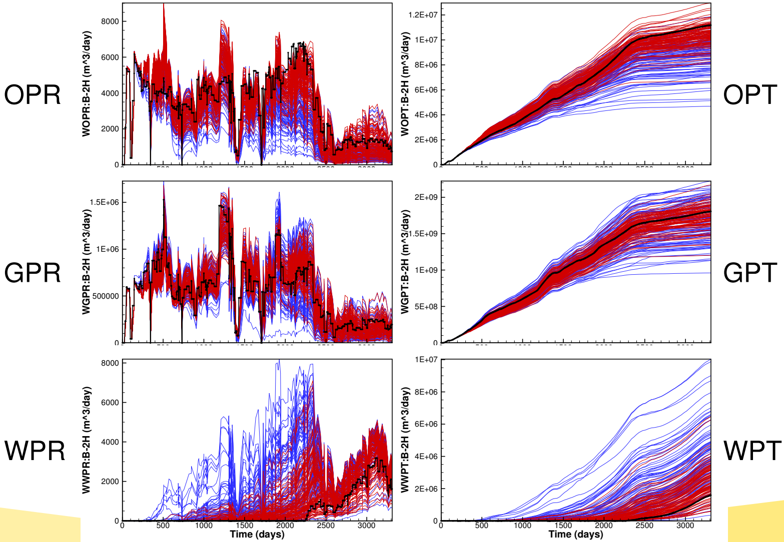
Geir Evensen

International Research Institute of Stavanger

Nansen Environmental and Remote Sensing Center

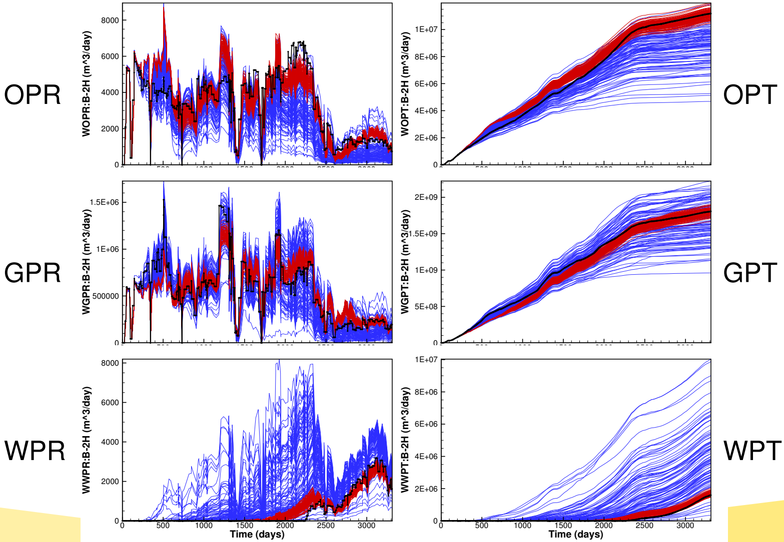
The 13th EnKF Data Assimilation Workshop, Bergen, May 28–30, 2018

Example: ES



OPR
WOPR:B-2H (m^3/day)
OPT
WOPT:B-2H (m^3/day)
GPR
WGPR:B-2H (m^3/day)
GPT
WGPT:B-2H (m^3/day)
WPR
WWPR:B-2H (m^3/day)
WPT
WWPT:B-2H (m^3/day)
Time (days)
Time (days)

Example: An iterative ES



Inverse problem

Find $\boldsymbol{x} \in \Re^n$ given a model

$$\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$$

given measurements $\boldsymbol{d} \in \Re^m$ of $\boldsymbol{y} \in \Re^m$

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

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

Data-assimilation problem

Find $\boldsymbol{x}^{\mathrm{a}}_{i+1} \in \Re^n$ given a model prediction

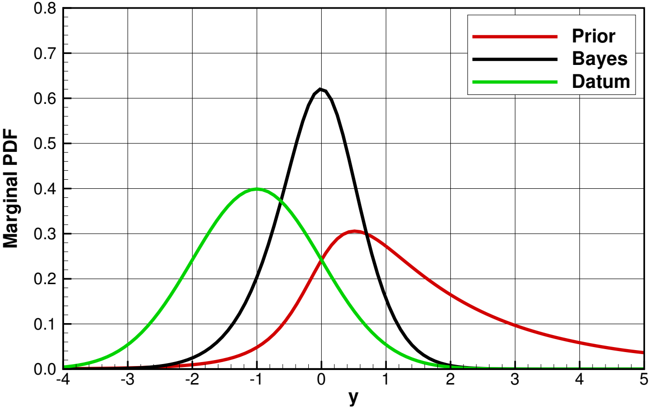
$$\boldsymbol{x}_{i+1} = \boldsymbol{g}(\boldsymbol{x}_i)$$

and measurements $\boldsymbol{d} \in \Re^m$ of $\boldsymbol{x}_{i+1} \in \Re^n$

$$\boldsymbol{d} = \boldsymbol{h}(\boldsymbol{x}_{i+1}) + \boldsymbol{e}$$

- Standard update step in sequential data assimilation

Bayesian update



Ensemble filter update

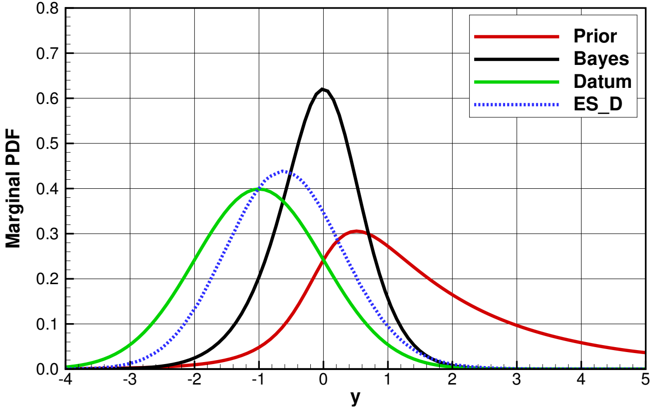
Prior

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

Direct update of $\boldsymbol{y}$

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

EnKF (direct) update



Ensemble smoother update

Prior

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

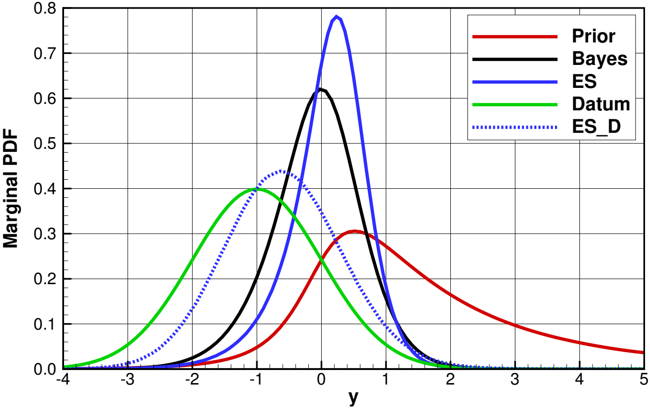
Smoother update of $\boldsymbol{x}$

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

Indirect update of $\boldsymbol{y}$

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

ES (indirect) update



Indirect update

Inverse problems: Measure $\boldsymbol{y}$ $\rightarrow$ Update $\boldsymbol{x}$ $\rightarrow$ predict $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$

Data assimilation: Measure $\boldsymbol{x}_{i+1}$ $\rightarrow$ Update $\boldsymbol{x}_i$ $\rightarrow$ predict $\boldsymbol{x}_{i+1} = \boldsymbol{g}(\boldsymbol{x}_i)$

- The indirect update allows for the use of iterative methods.

Nonlinearity

Sensitivity of smoothers to nonlinearity

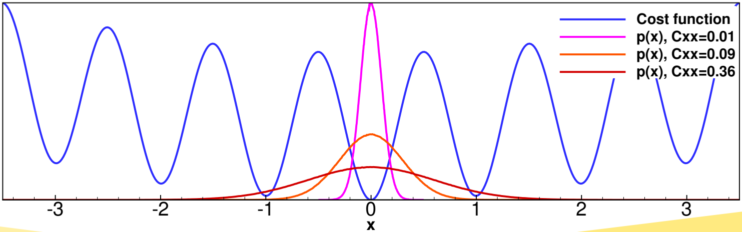
Model

$$y = 1 + \sin(\pi x)$$

Case 1: $x_j^{\mathrm{f}} \leftarrow \mathcal{N}(0.0, C_{xx} = 0.01)$ $\quad d_j \leftarrow \mathcal{N}(1.0, C_{dd} = 0.01)$

Case 2: $x_j^{\mathrm{f}} \leftarrow \mathcal{N}(0.0, C_{xx} = 0.09)$ $\quad d_j \leftarrow \mathcal{N}(1.0, C_{dd} = 0.01)$

Case 3: $x_j^{\mathrm{f}} \leftarrow \mathcal{N}(0.0, C_{xx} = 0.36)$ $\quad d_j \leftarrow \mathcal{N}(1.0, C_{dd} = 0.01)$



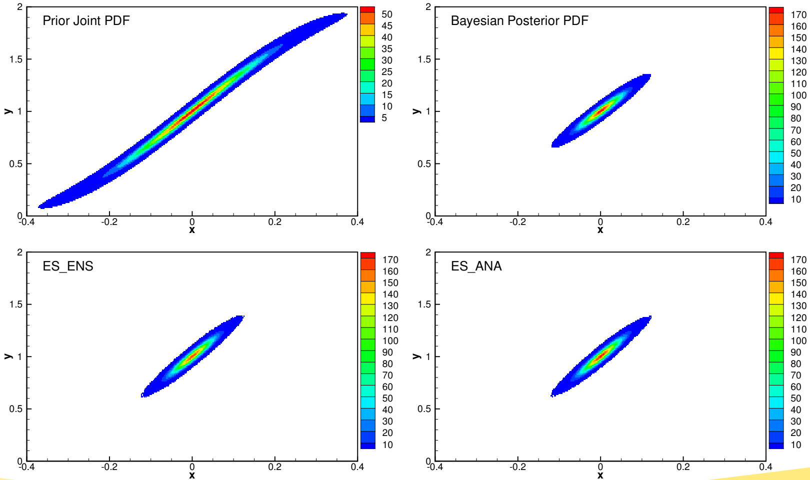
Ensemble vs analytic gradient

$$x_j^{\mathrm{a}} = x_j^{\mathrm{f}} + g'(x_j^{\mathrm{f}}) C_{xx} \left(g'(x_j^{\mathrm{f}}) C_{xx} g'(x_j^{\mathrm{f}}) + C_{dd} \right)^{-1} \left(d_j - g(x_j^{\mathrm{f}}) \right)$$

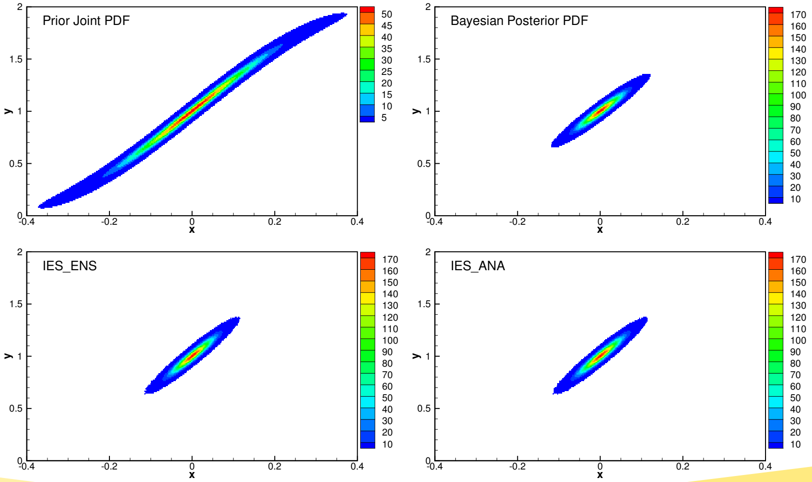
$$x_j^{\mathrm{a}} = x_j^{\mathrm{f}} + \overline{C}_{xy} (\overline{C}_{yy} + \overline{C}_{dd})^{-1} (d_j - g(x_j^{\mathrm{f}}))$$

$$y_j^{\mathrm{a}} = g(x_j^{\mathrm{a}})$$

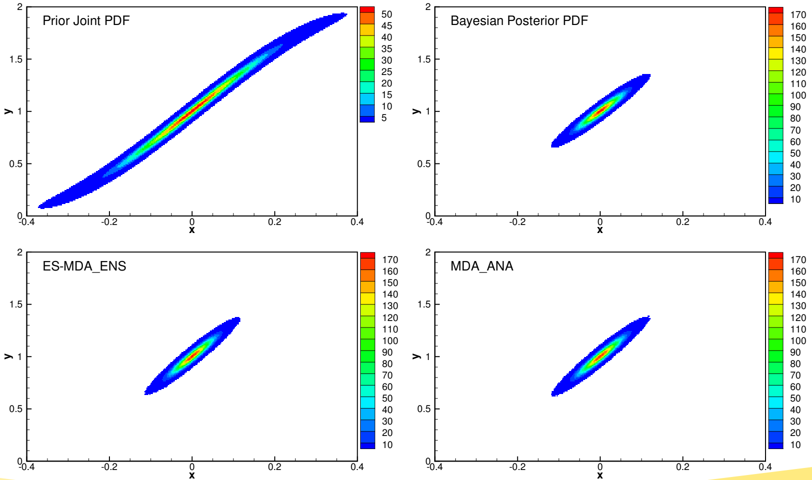
Case 1: ES


Prior Joint PDF
Bayesian Posterior PDF
ES_ENS
ES_ANA
x
y
-0.4 -0.2 0 0.2 0.4
0 0.5 1 1.5 2
50 45 40 35 30 25 20 15 10 5
170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10

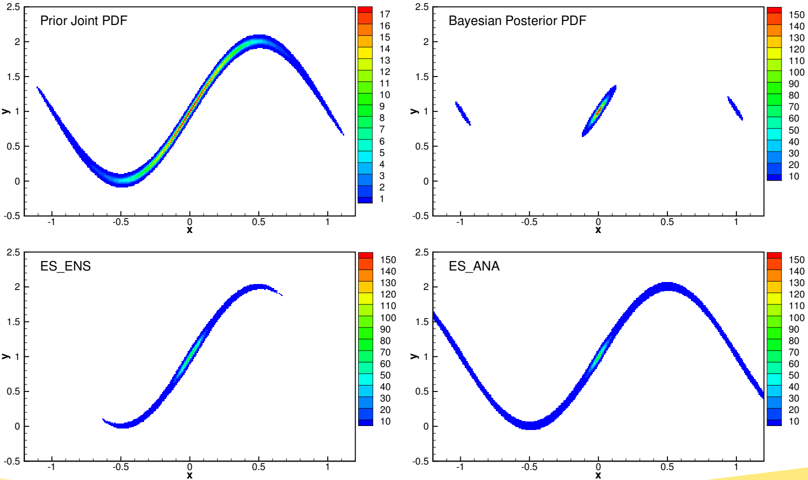
Case 1: IES


Prior Joint PDF
Bayesian Posterior PDF
IES_ENS
IES_ANA
x
y
-0.4
-0.2
0
0.2
0.4
0
0.5
1
1.5
2
50
45
40
35
30
25
20
15
10
5
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

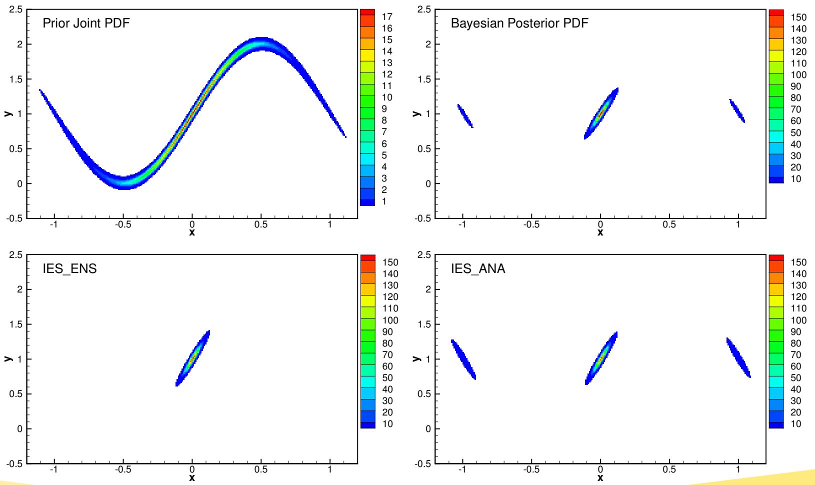
Case 1: ESMDA


Prior Joint PDF
Bayesian Posterior PDF
ES-MDA_ENS
MDA_ANA
x
y
-0.4
-0.2
0
0.2
0.4
0.5
1
1.5
2
50
45
40
35
30
25
20
15
10
5
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

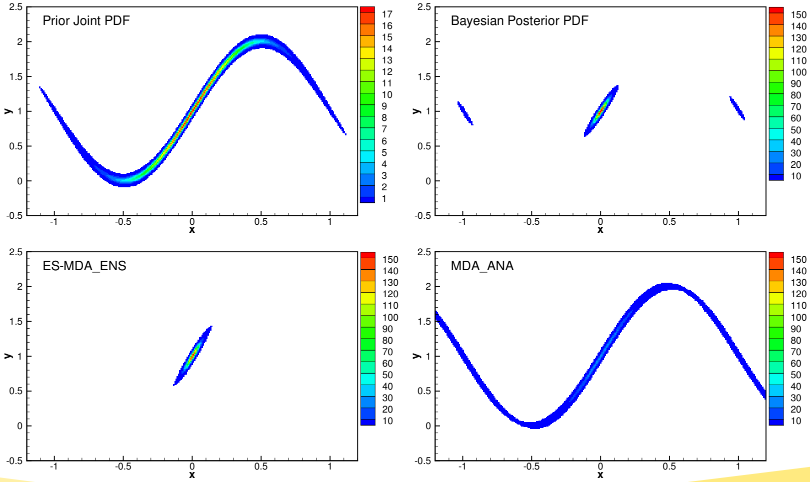
Case 2: ES


Prior Joint PDF
Bayesian Posterior PDF
ES_ENS
ES_ANA
x
y
2.5
2
1.5
1
0.5
0
-0.5
-1
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
150
140
130
120
110
100
90
80
70
60
50
40
30
20

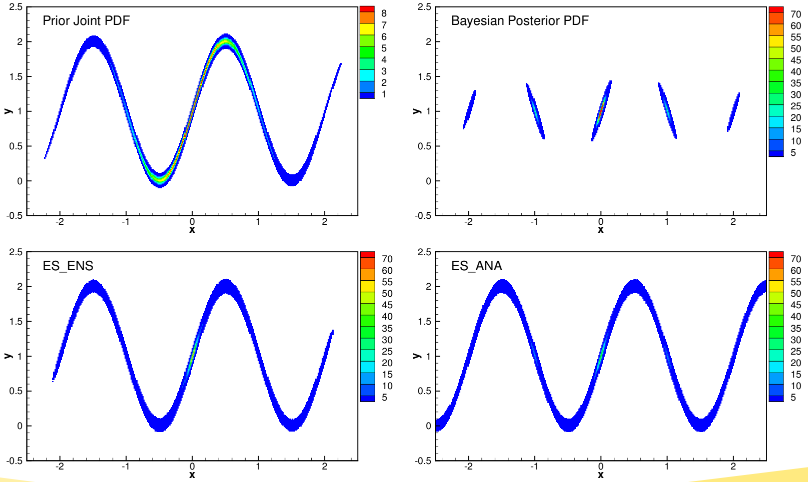
Case 2: IES



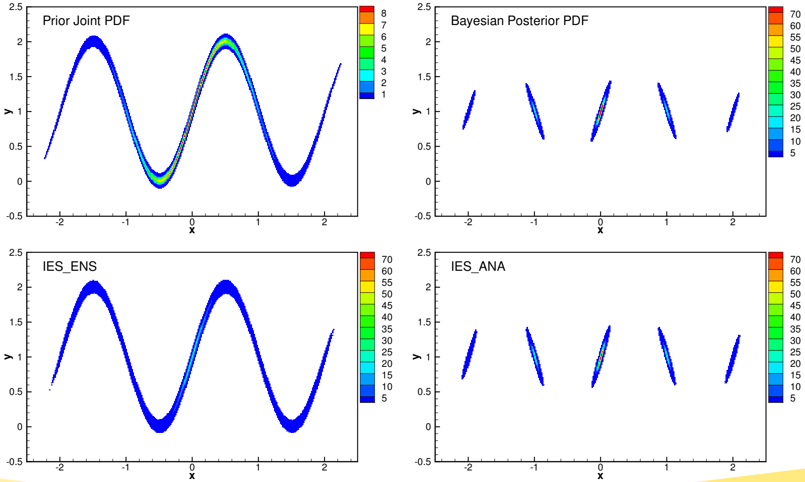
Case 2: ESMDA



Case 3: ES

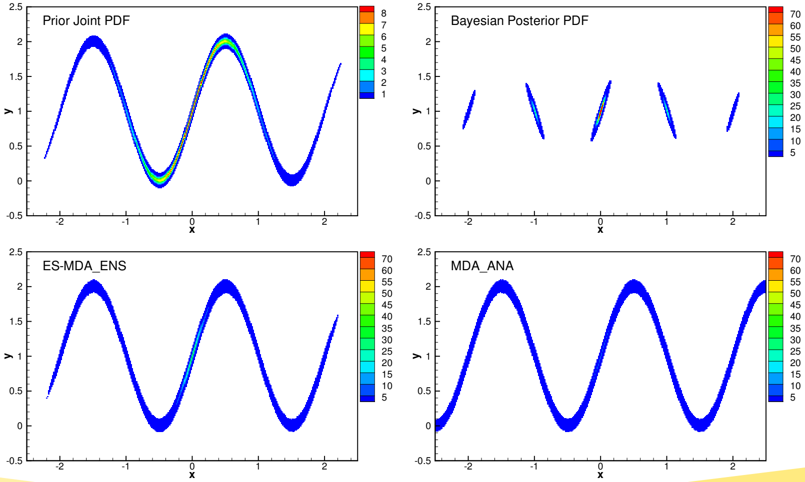


Case 3: IES



Prior Joint PDF
Bayesian Posterior PDF
IES_ENS
IES_ANA
x
y
8
7
6
5
4
3
2
1
70
60
55
50
45
40
35
30
25
20
15
10
5
2.5
2
1.5
0.5
0
-0.5
-2
-1

Case 3: ESMDA



Model errors in iterative smoothers

Including model errors

Model

$$\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{q})$$

Bayes

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

Marginal pdf

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

Ensemble of cost functions

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

ES including model errors

Prior

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

Update

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

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

Indirect update

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

or direct update

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

ESMDA including model errors

Cost function for $\boldsymbol{z}_j^{\mathrm{T}} = \left(\boldsymbol{x}_j^{\mathrm{T}}, \boldsymbol{q}_j^{\mathrm{T}}\right)$ at step i

$$\begin{aligned}\mathcal{J}(\boldsymbol{z}_{i+1}) &= \left(\boldsymbol{z}_{i+1} - \boldsymbol{z}_i\right)^{\mathrm{T}} \left(\boldsymbol{C}_{zz}^i\right)^{-1} \left(\boldsymbol{z}_{i+1} - \boldsymbol{z}_i\right) \\ &\quad + \left(\boldsymbol{g}(\boldsymbol{z}_{i+1}) - \boldsymbol{d} - \sqrt{\alpha_i}\boldsymbol{e}_i\right)^{\mathrm{T}} \left(\alpha_i \boldsymbol{C}_{dd}\right)^{-1} \left(\boldsymbol{g}(\boldsymbol{z}_{i+1}) - \boldsymbol{d} - \sqrt{\alpha_i}\boldsymbol{e}_i\right)\end{aligned}$$

where we must have

$$\sum_{i=1}^{N_{\text{mda}}} \frac{1}{\alpha_i} = 1.$$

$$\begin{aligned}\boldsymbol{x}_{i+1} &= \boldsymbol{x}_i + \overline{\boldsymbol{C}}_{xy}^i \left(\overline{\boldsymbol{C}}_{yy}^i + \alpha_i \boldsymbol{C}_{dd}\right)^{-1} \left(\boldsymbol{d} + \sqrt{\alpha_i}\boldsymbol{e}_i - \boldsymbol{y}_i)\right) \\ \boldsymbol{q}_{i+1} &= \boldsymbol{q}_i + \overline{\boldsymbol{C}}_{qy}^i \left(\overline{\boldsymbol{C}}_{yy}^i + \alpha_i \boldsymbol{C}_{dd}\right)^{-1} \left(\boldsymbol{d} + \sqrt{\alpha_i}\boldsymbol{e}_i - \boldsymbol{y}_i)\right) \\ \boldsymbol{y}_{i+1} &= \boldsymbol{g}(\boldsymbol{x}_{i+1}, \boldsymbol{q}_{i+1})\end{aligned}$$

IES including model errors

Cost function for $\boldsymbol{z}_j^{\mathrm{T}} = \left(\boldsymbol{x}_j^{\mathrm{T}}, \boldsymbol{q}_j^{\mathrm{T}}\right)$

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

Iterative solution

$$\begin{aligned} \boldsymbol{z}_{j,i+1} &= \boldsymbol{z}_{j,i} + \overline{\boldsymbol{C}}_{zz}^{i} \overline{\boldsymbol{C}}_{zz}^{-1} (\boldsymbol{z}_{j,i} - \boldsymbol{z}_j^{\mathrm{f}}) \\ &\quad - \overline{\boldsymbol{C}}_{zy}^{i} \left(\overline{\boldsymbol{C}}_{yy}^{i} + \boldsymbol{C}_{dd}\right)^{-1} \left(\overline{\boldsymbol{C}}_{yz}^{i} \overline{\boldsymbol{C}}_{zz}^{-1} (\boldsymbol{z}_{j,i} - \boldsymbol{z}_j^{\mathrm{f}}) - \left(\boldsymbol{g}(\boldsymbol{z}_{j,i}) - \boldsymbol{d}_j\right)\right) \\ \boldsymbol{y}_{j,i+1} &= \boldsymbol{g}(\boldsymbol{z}_{j,i}) \end{aligned}$$

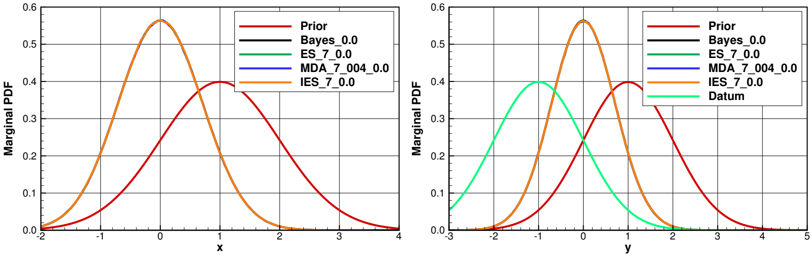
Scalar example

Model

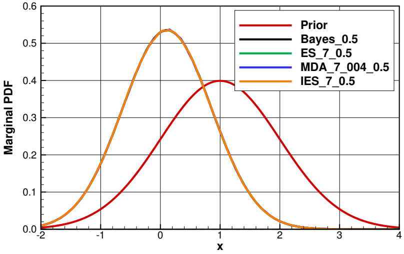
$$y = x(1 + \beta x^2) + \sqrt{C_{qq}} q$$

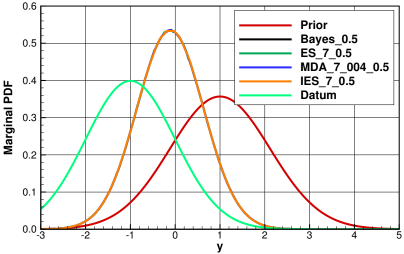
- Linear case: $\beta = 0$.
- Nonlinear case: $\beta = 0.2$.
- Prior ensemble for x: $N(1, 1)$.
- Likelihood for measurement of y: $N(-1, 1)$.

Linear model without model errors

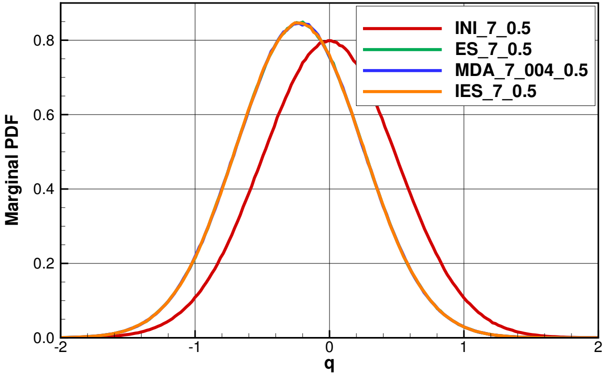


Linear model including model errors

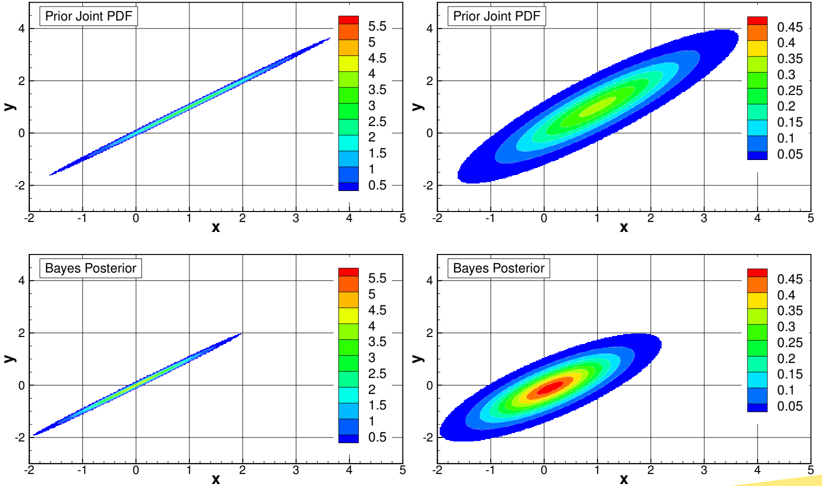




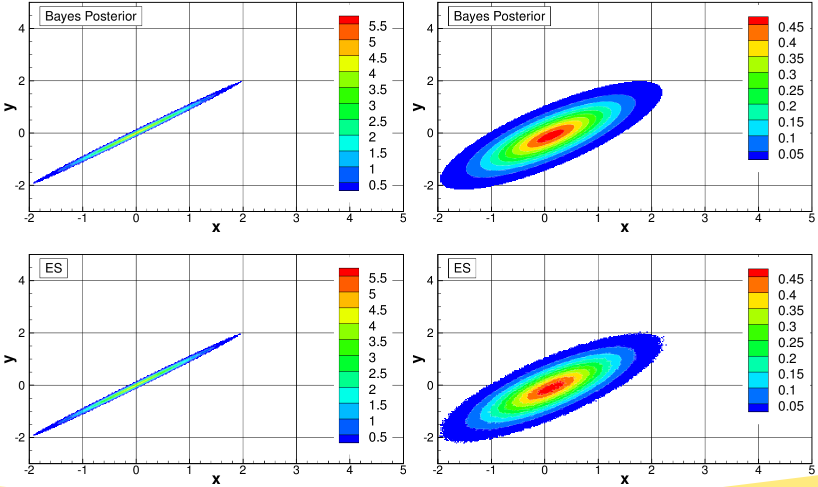
Pdfs for model error with linear model



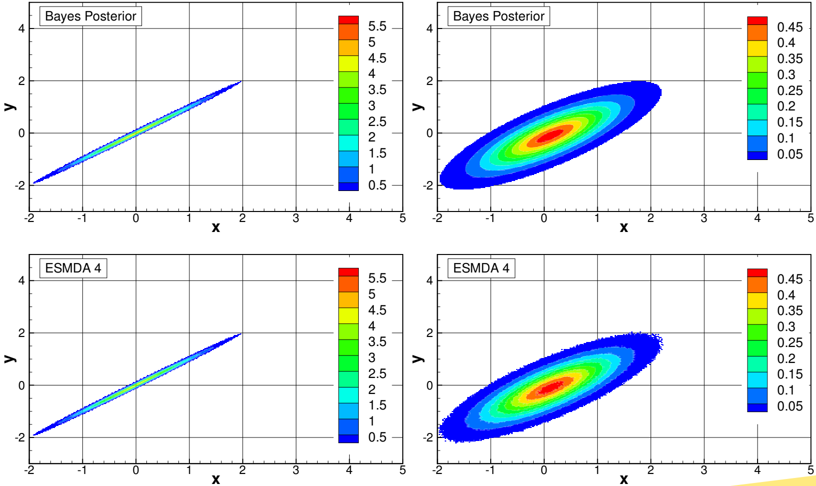
Joint pdfs linear model



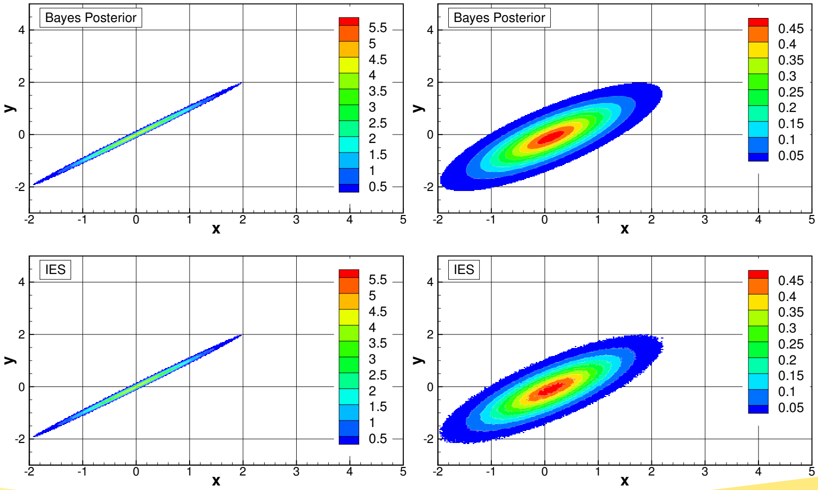
Joint pdfs linear model



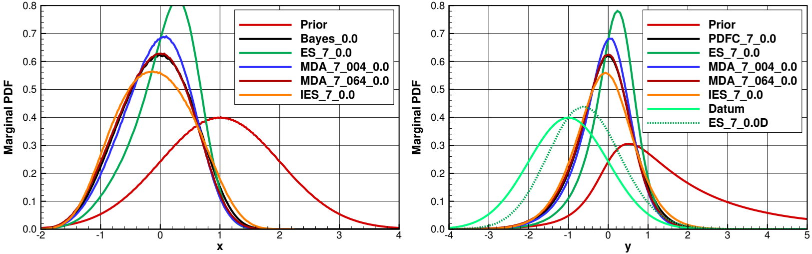
Joint pdfs linear model



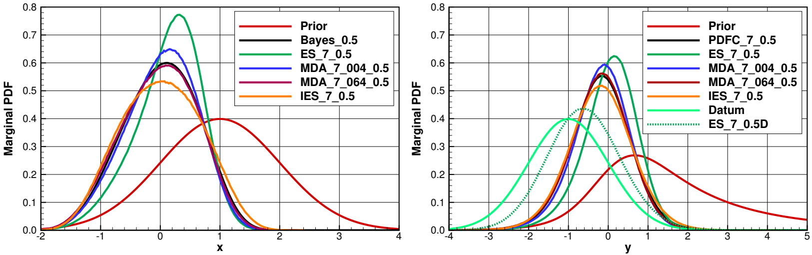
Joint pdfs linear model



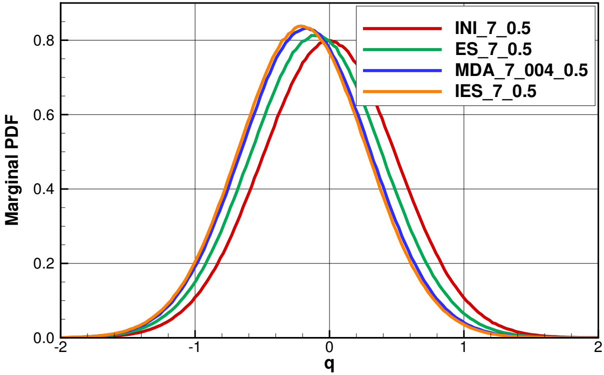
Nonlinear model without model errors



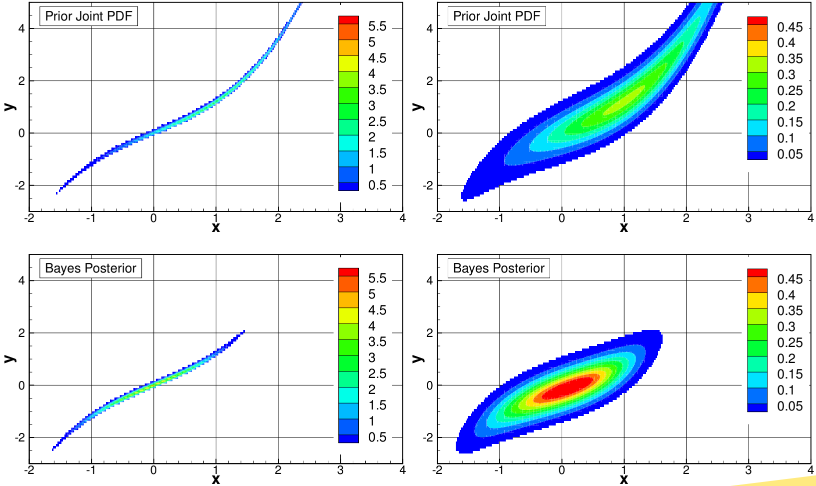
Nonlinear model including model errors



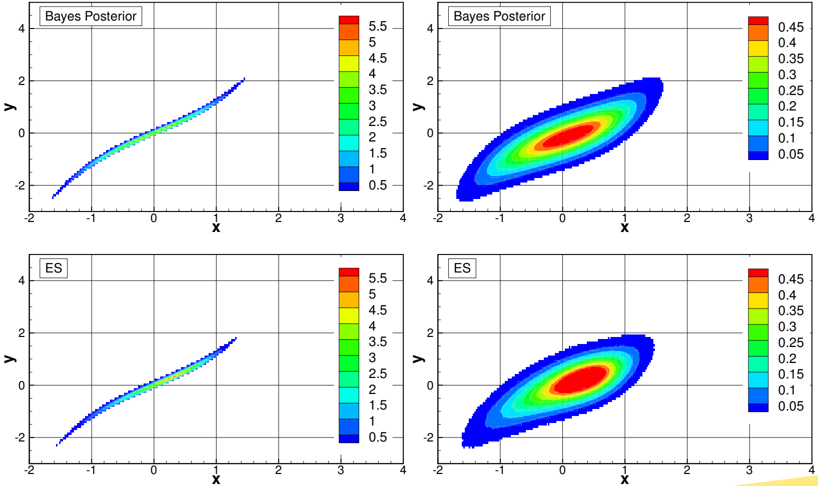
Pdfs for model error with nonlinear model



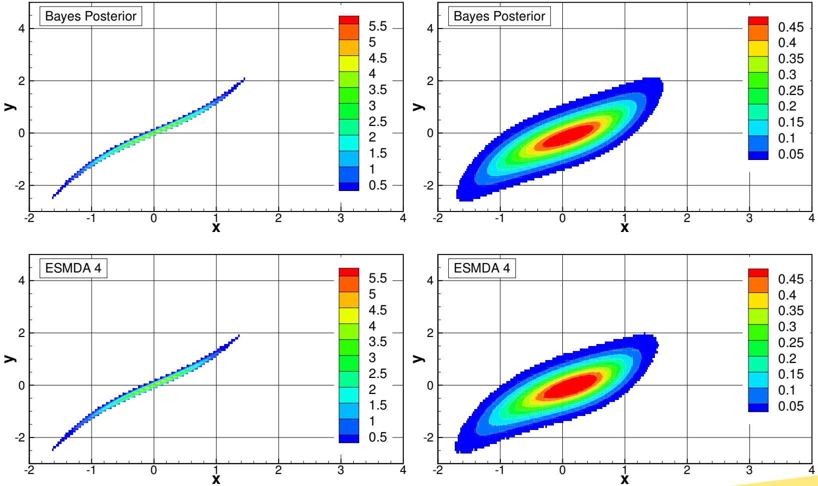
Joint pdfs nonlinear model



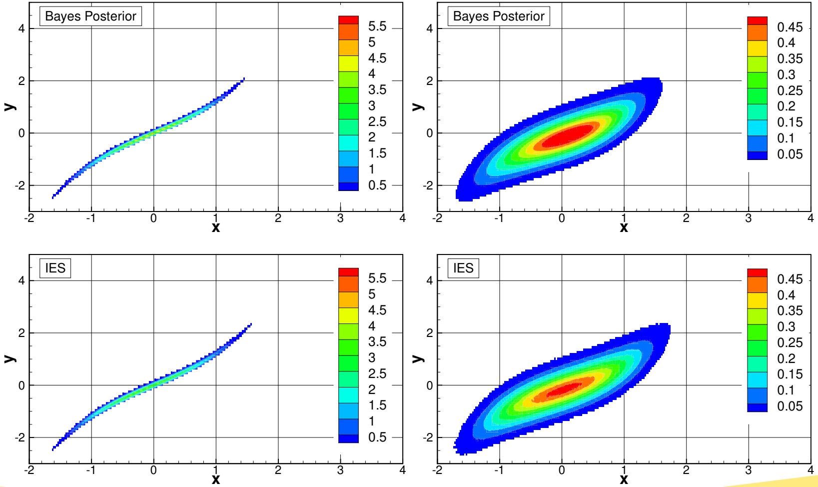
Joint pdfs nonlinear model



Joint pdfs nonlinear model



Joint pdfs nonlinear model



Acknowledgements

- RCN DIGIRES project with sponsors Equinor, Aker BP, Eni, Petrobras, VNG, and DEA, for funding this work.

Y. Chen 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.

A. A. Emerick and A. C. Reynolds. Ensemble smoother with multiple data assimilation. *Computers and Geosciences*, 55:3–15, 2013.

G. Evensen. Analysis of iterative ensemble smoothers for solving inverse problems. *Computat Geosci*, 22(3):885–908, 2018a. doi:10.1007/s10596-018-9731-y. URL `https://doi.org/10.1007/s10596-018-9731-y`.

G. Evensen. Accounting for model errors in iterative ensemble smoothers, 2018b. URL `https://arxiv.org/abs/1806.00237`.

P. K. Kitanidis. Quasi-linear geostatistical therory for inversing. *Water Resources Research*, 31(10): 2411–2419, 1995.

D. S. Oliver, N. He, and A. C. Reynolds. Conditioning permeability fields to pressure data. ECMOR – 5th European Conference on the Mathematics of Oil Recovery, 1996.

J.-A. Skjervheim, G. Evensen, J. Hove, and J. Vabø. An ensemble smoother for assisted history matching. SPE 141929, 2011.

P. J. van Leeuwen and G. Evensen. Data assimilation and inverse methods in terms of a probabilistic formulation. *Mon. Weather Rev.*, 124:2898–2913, 1996.

Summary

IRIS

- We noted the duality of the DA and inverse problem.
- Iterations improve the estimate for models with modest nonlinearity.
- Iterative methods allows for inclusion of general model errors.
- Results published in Evensen (2018a,b).