Relation between Five Data Assimilation Methods for a Simplistic Weakly Non-linear Problem

Trond Mannseth

Uni Research CIPR

Methods

Methods

Randomized Maximum Likelihood (RML)
(Simultaneous estimation, Iterative updates)

Methods

Randomized Maximum Likelihood (RML)
(Simultaneous estimation, Iterative updates)

Ensemble Smoother (ES)
(Sim. est., A single update)

Methods

Randomized Maximum Likelihood (RML)
(Simultaneous estimation, Iterative updates)

Ensemble Smoother (ES)
(Sim. est., A single update)

Ensemble Smoother with Multiple Data Assimilations (ESMDA)
(Sim. est., Multiple upd. with inflated data covariance)

Methods

Randomized Maximum Likelihood (RML)
(Simultaneous estimation, Iterative updates)

Ensemble Smoother (ES)
(Sim. est., A single update)

Ensemble Smoother with Multiple Data Assimilations (ESMDA)
(Sim. est., Multiple upd. with inflated data covariance)

Half-iteration Ensemble Kalman Filter (HIEnKF)
(Sequential est., A single upd.)

Methods

Randomized Maximum Likelihood (RML)
(Simultaneous estimation, Iterative updates)

Ensemble Smoother (ES)
(Sim. est., A single update)

Ensemble Smoother with Multiple Data Assimilations (ESMDA)
(Sim. est., Multiple upd. with inflated data covariance)

Half-iteration Ensemble Kalman Filter (HIEnKF)
(Sequential est., A single upd.)

Half-iteration Ensemble Kalman Filter with MDA (HIEnKFMDA)
(Seq. est., Multiple upd. with inflated data covariance)

Motivation

Motivation

Methods are equivalent for gauss-linear problems

Motivation

Methods are equivalent for gauss-linear problems

Methods behave differently for non-linear problems

Motivation

Methods are equivalent for gauss-linear problems

Methods behave differently for non-linear problems

RML is iterative, ES is purely non-iterative, while ESMDA, HIEnKF and HIEnKFMDA 'lie somewhere in between'

Motivation

Methods are equivalent for gauss-linear problems

Methods behave differently for non-linear problems

RML is iterative, ES is purely non-iterative, while ESMDA, HIEnKF and HIEnKFMDA 'lie somewhere in between'

Systematic differences for weakly non-linear problems?

Investigation

Investigation

Compare methods on simplistic, weakly non-linear parameter estimation problem

Investigation

Compare methods on simplistic, weakly non-linear parameter estimation problem

Focus on differences in data handling – remove other differences

Investigation

Compare methods on simplistic, weakly non-linear parameter estimation problem

Focus on differences in data handling – remove other differences

Asymptotic calculations (additional assumptions) to first order in non-linearity strength

Investigation

Compare methods on simplistic, weakly non-linear parameter estimation problem

Focus on differences in data handling – remove other differences

Asymptotic calculations (additional assumptions) to first order in non-linearity strength

Numerical calculations with full methods and relaxed assumptions

Simplistic Weakly Non-linear Parameter Estimation Problem

Simplistic Weakly Non-linear Parameter Estimation Problem

Estimate x from d, where

Simplistic Weakly Non-linear Parameter Estimation Problem

Estimate x from d, where

$$d = (d_1 \ldots d_D)^T,$$

Simplistic Weakly Non-linear Parameter Estimation Problem

Estimate x from d, where

$$d = (d_1 \ldots d_D)^T,$$

$$d_i = y_i\left(x_{ref}\right) + \epsilon_i; \quad \epsilon_i \sim N(0, \sigma_i^2),$$

Simplistic Weakly Non-linear Parameter Estimation Problem

Estimate x from d, where

$$d = (d_1 \dots d_D)^T,$$

$$d_i = y_i\,(x_{ref}) + \epsilon_i; \quad \epsilon_i \sim N(0, \sigma_i^2),$$

$$y_i(x) = \sum_{m=1}^{M} c_{im} x_m^{1+n_{im}}; \quad |n_{im}| \text{ `not too big'}$$

Focus on Diff. in Data Handling – Remove Other Diff.

Focus on Diff. in Data Handling – Remove Other Diff.

Equip ES (and ESMDA, HIEnKF, HIEnKFMDA) with local gains

Focus on Diff. in Data Handling – Remove Other Diff.

Equip ES (and ESMDA, HIEnKF, HIEnKFMDA) with local gains

Kalman gain (global) $\quad K = C_{xy}\left(C_{yy} + C_d\right)^{-1}$

Focus on Diff. in Data Handling – Remove Other Diff.

Equip ES (and ESMDA, HIEnKF, HIEnKFMDA) with local gains

Kalman gain (global) $\quad K = C_{xy} \left(C_{yy} + C_d \right)^{-1}$

Replace C_{xy} by $C_x G_e^T$ and C_{yy} by $G_e C_x G_e^T$

Focus on Diff. in Data Handling – Remove Other Diff.

Equip ES (and ESMDA, HIEnKF, HIEnKFMDA) with local gains

Kalman gain (global) $K = C_{xy} \left(C_{yy} + C_d\right)^{-1}$

Replace C_{xy} by $C_x G_e^T$ and C_{yy} by $G_e C_x G_e^T$

$\rightarrow$ Local gains $K_e = C_x G_e^T \left(G_e C_x G_e^T + C_d\right)^{-1}$

Additional Assumptions facilitating Asymptotic Calculations

Additional Assumptions facilitating Asymptotic Calculations

I consider the updates of a single (arbitrary) ensemble member

Additional Assumptions facilitating Asymptotic Calculations

I consider the updates of a single (arbitrary) ensemble member

Assumptions

$$\text{Univariate } x \quad \rightarrow \quad y_i(x) = x^{1+n_i}$$

Additional Assumptions facilitating Asymptotic Calculations

I consider the updates of a single (arbitrary) ensemble member

Assumptions

$$\text{Univariate } x \quad \rightarrow \quad y_i(x) = x^{1+n_i}$$

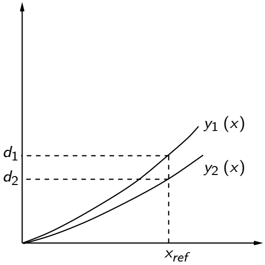
$$\text{Negligible data error} \quad \rightarrow \quad d_i = y_i\left(x_{ref}\right),$$

Additional Assumptions . . . (continued)

Additional Assumptions . . . (continued)

So far

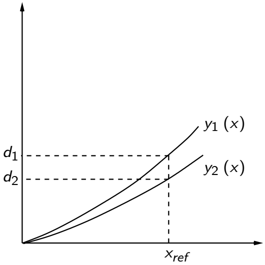
$|n_1| \wedge |n_2|$ 'not too big'



Additional Assumptions . . . (continued)

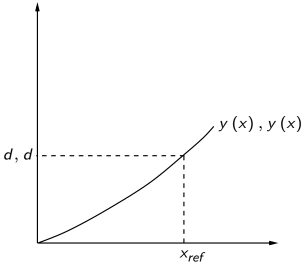
So far

$|n_1| \wedge |n_2|$ 'not too big'



I assume

$n_1 = n_2 = n, \quad |n| \ll 1$



Results from Asymptotic Calculations to $\mathcal{O}(n)$

Results from Asymptotic Calculations to $\mathcal{O}(n)$

$$x_{\mathrm{RML}} = d - (d \ln d)n + \mathcal{O}(n^2)$$

Results from Asymptotic Calculations to $\mathcal{O}(n)$

$$x_{\mathrm{RML}} = d - (d \ln d)n + \mathcal{O}(n^2)$$

Define

$$\Delta_{\mathrm{method}} = |x_{\mathrm{method}} - x_{\mathrm{RML}}|$$

Results from Asymptotic Calculations to $\mathcal{O}(n)$

$$x_{\mathrm{RML}} = d - (d \ln d)n + \mathcal{O}(n^2)$$

Define

$$\begin{aligned} \Delta_{\mathrm{method}} &= \left| x_{\mathrm{method}} - x_{\mathrm{RML}} \right| \\ Q &= \left| d(\ln d - \ln x_{\mathrm{prior}}) - (d - x_{\mathrm{prior}}) \right| \end{aligned}$$

Results from Asymptotic Calculations to $\mathcal{O}(n)$

$$x_{\mathrm{RML}} = d - (d \ln d)n + \mathcal{O}(n^2)$$

Define

$$\begin{aligned} \Delta_{\mathrm{method}} &= \left| x_{\mathrm{method}} - x_{\mathrm{RML}} \right| \\ Q &= \left| d(\ln d - \ln x_{\mathrm{prior}}) - (d - x_{\mathrm{prior}}) \right| \end{aligned}$$

$$\Delta_{\mathrm{ES}} = Qn + \mathcal{O}(n^2)$$

Results from Asymptotic Calculations to $\mathcal{O}(n)$

$$x_{\mathrm{RML}} = d - (d \ln d)n + \mathcal{O}(n^2)$$

Define

$$\begin{aligned} \Delta_{\mathrm{method}} &= \left| x_{\mathrm{method}} - x_{\mathrm{RML}} \right| \\ Q &= \left| d(\ln d - \ln x_{\mathrm{prior}}) - (d - x_{\mathrm{prior}}) \right| \end{aligned}$$

$$\begin{aligned} \Delta_{\mathrm{ES}} &= Qn + \mathcal{O}(n^2) \\ \Delta_{\mathrm{ESMDA}} &= A^{-1}Qn + \mathcal{O}(n^2) \end{aligned}$$

Results from Asymptotic Calculations to $\mathcal{O}(n)$

$$x_{\mathrm{RML}} = d - (d \ln d)n + \mathcal{O}(n^2)$$

Define

$$\begin{aligned} \Delta_{\mathrm{method}} &= \left| x_{\mathrm{method}} - x_{\mathrm{RML}} \right| \\ Q &= \left| d(\ln d - \ln x_{\mathrm{prior}}) - (d - x_{\mathrm{prior}}) \right| \end{aligned}$$

$$\begin{aligned} \Delta_{\mathrm{ES}} &= Qn + \mathcal{O}(n^2) \\ \Delta_{\mathrm{ESMDA}} &= A^{-1} Qn + \mathcal{O}(n^2) \\ \Delta_{\mathrm{HIEnKF}} &= D^{-1} Qn + \mathcal{O}(n^2) \end{aligned}$$

Results from Asymptotic Calculations to $\mathcal{O}(n)$

$$x_{\mathrm{RML}} = d - (d \ln d) n + \mathcal{O}(n^2)$$

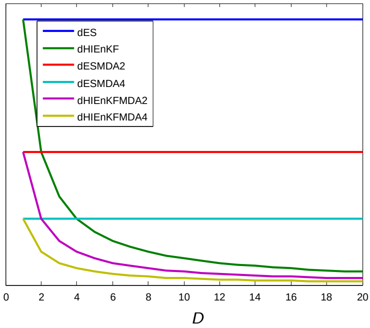
Define

$$\begin{aligned} \Delta_{\mathrm{method}} &= \left| x_{\mathrm{method}} - x_{\mathrm{RML}} \right| \\ Q &= \left| d(\ln d - \ln x_{\mathrm{prior}}) - (d - x_{\mathrm{prior}}) \right| \end{aligned}$$

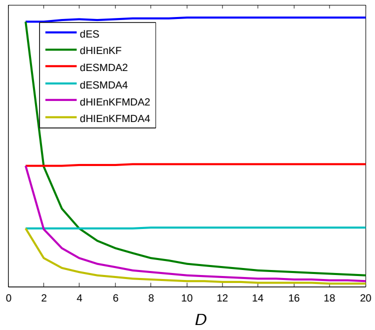
$$\begin{aligned} \Delta_{\mathrm{ES}} &= Qn + \mathcal{O}(n^2) \\ \Delta_{\mathrm{ESMDA}} &= A^{-1} Qn + \mathcal{O}(n^2) \\ \Delta_{\mathrm{HIEnKF}} &= D^{-1} Qn + \mathcal{O}(n^2) \\ \Delta_{\mathrm{HIEnKFMDA}} &= (AD)^{-1} Qn + \mathcal{O}(n^2) \end{aligned}$$

Results from Asymptotic Calculations to $\mathcal{O}(n)$

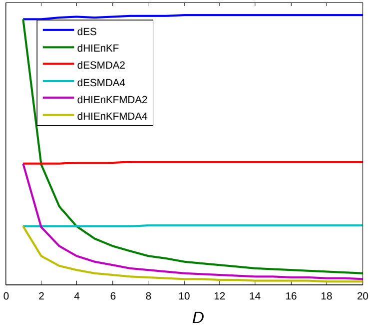
$$
\begin{aligned}
\Delta_{\mathrm{ES}} &\approx Qn \\
\Delta_{\mathrm{ESMDA}} &\approx A^{-1}Qn \\
\Delta_{\mathrm{HIEnKF}} &\approx D^{-1}Qn \\
\Delta_{\mathrm{HIEnKFMDA}} &\approx (AD)^{-1}Qn
\end{aligned}
$$



Numerical Results with Full Methods



Numerical Results with Full Methods



'Ranking' stable for $n \in [-0.5, 5]$

Num. Calc. with Full Methods and Relaxed Assumptions

$$d = (d_1 \ldots d_D)^T,$$

$$d_i = y_i\,(x_{ref}) + \epsilon_i; \quad \epsilon_i \sim N(0, \sigma_i^2),$$

$$y_i(x) = \sum_{m=1}^{M} c_{im} x_m^{1+n_{im}}; \quad \bar{n}_{im} = 0.4$$

Num. Calc. with Full Methods and Relaxed Assumptions

$$d = (d_1 \dots d_D)^T,$$

$$d_i = y_i\,(x_{ref}) + \epsilon_i; \quad \epsilon_i \sim N(0, \sigma_i^2),$$

$$y_i(x) = \sum_{m=1}^{M} c_{im} x_m^{1+n_{im}}; \quad \bar{n}_{im} = 0.4$$

$M = 1, 2, 5.$

Num. Calc. with Full Methods and Relaxed Assumptions

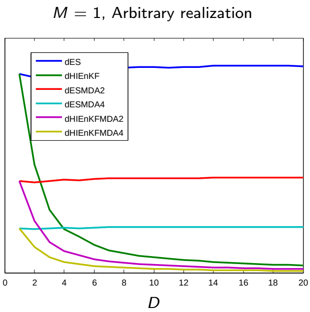
$$d = (d_1 \ldots d_D)^T,$$

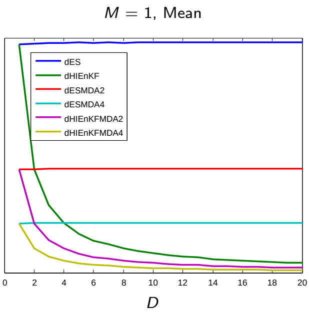
$$d_i = y_i\,(x_{ref}) + \epsilon_i; \quad \epsilon_i \sim N(0, \sigma_i^2),$$

$$y_i(x) = \sum_{m=1}^{M} c_{im} x_m^{1+n_{im}}; \quad \bar{n}_{im} = 0.4$$

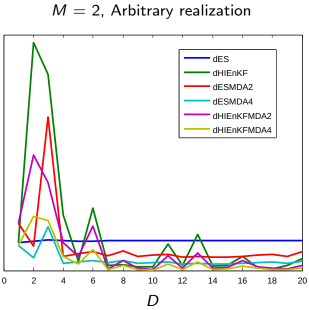
$M = 1, 2, 5.$ Draw 300 realizations of x_{ref}, x_{prior}, n_{im}, c_{im}

Num. Res. with Full Methods and Relaxed Assumptions

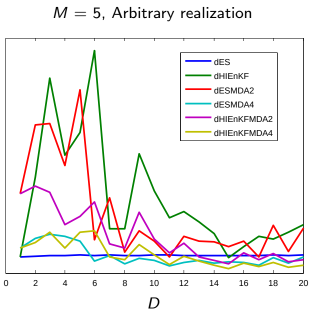


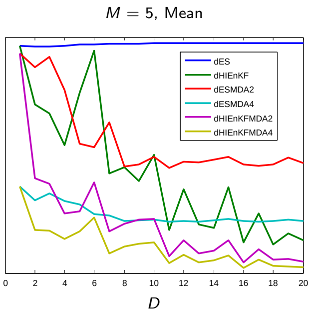


Num. Res. with Full Methods and Relaxed Assumptions



Num. Res. with Full Methods and Relaxed Assumptions





Summary

Summary

Compared five different ways to assimilate data on simplistic, weakly non-linear parameter estimation problem

Summary

Compared five different ways to assimilate data on simplistic, weakly non-linear parameter estimation problem

Asymptotic calculations to first order in non-linearity strength (relying on further simplifications) reveals nature of similarity with iterative methods for (local-gain) ESMDA, HIEnKF and HIEnKFMDA methods

Summary

Compared five different ways to assimilate data on simplistic, weakly non-linear parameter estimation problem

Asymptotic calculations to first order in non-linearity strength (relying on further simplifications) reveals nature of similarity with iterative methods for (local-gain) ESMDA, HIEnKF and HIEnKFMDA methods

Numerical results with full (local-gain) methods and relaxed assumptions support asymptotic calculations for low values of M