EnKF Workshop 2021

Efficient Ensemble Based Methods for Sequential Data Assimilation

Elias D. Nino-Ruiz

Applied Math and Computer Science Laboratory (AML-CS)
Department of Computer Science
Universidad del Norte
BAQ 080001, Colombia

Outline I

Applied Math and Computer Science Lab

- The Applied Math and Computer Science Laboratory (AML-CS) at Universidad del Norte in Barranquilla, Colombia.
- The group was founded in Apr 2017. Our research focus on the next topics:
  1. Data Assimilation Methods.
  2. Inverse Problems and Parameter Estimation.
  3. Combinatorial Optimization.
  4. Numerical Optimization.
  5. Bayesian Inference.
- Current students: Juan Rodriguez, Omar Mejia, Randy Consuegra, Andres Movilla, and Felipe Acevedo.

Data Assimilation [BS12] I

- We want to estimate $\mathrm{x}^* \in \mathbb{R}^{n \times 1}$. $n \sim \mathcal{O}\left(10^8\right)$.
- Imperfect numerical model:

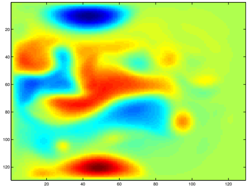
$$\mathrm{x}_{\mathrm{next}} = \mathcal{M}_{t_{\mathrm{current}} \rightarrow t_{\mathrm{next}}}\left(\mathrm{x}_{\mathrm{current}}\right),$$

  where $\mathrm{x} \in \mathbb{R}^{n \times 1}$.
- Noisy observations:

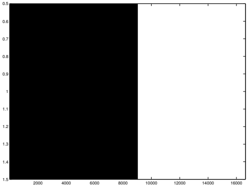
$$\mathrm{y} = \mathcal{H}\left(\mathrm{x}\right) + \boldsymbol{\epsilon} \in \mathbb{R}^{m \times 1},$$

  where $\mathcal{H} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\boldsymbol{\epsilon} \sim \mathcal{N}\left(0_m, \mathrm{R}\right)$.
- Prior estimate $\mathrm{x}^b \in \mathbb{R}^{n \times 1}$ with errors following $\mathcal{N}\left(0, \mathrm{B}\right)$.

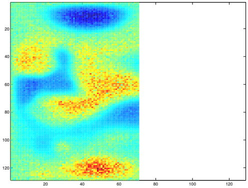
Data Assimilation [BS12] II



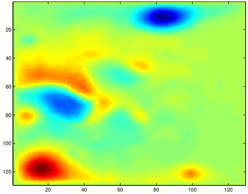
(a) x^*



(b) H



(c) $y = H \cdot x^* + \epsilon$



(d) x^b

Data Assimilation [BS12] III

- By Bayes' Theorem we know that:

$$\mathcal{P}\left(\mathrm{x}|\mathrm{y}\right) \propto \mathcal{P}\left(\mathrm{x}\right) \cdot \mathcal{L}\left(\mathrm{x}|\mathrm{y}\right)$$

where

$$\begin{aligned}
\mathcal{P}\left(\mathrm{x}\right) &\propto \exp\left(-\frac{1}{2} \cdot \left\|\mathrm{x} - \mathrm{x}^b\right\|_{\mathrm{B}^{-1}}^2\right) \\
\mathcal{L}\left(\mathrm{x}|\mathrm{y}\right) &\propto \exp\left(-\frac{1}{2} \cdot \left\|\mathrm{y} - \mathrm{H} \cdot \mathrm{x}\right\|_{\mathrm{R}^{-1}}^2\right)
\end{aligned}$$

and therefore,

$$\mathrm{x}^a = \arg\max_{\mathrm{x}} \mathcal{P}\left(\mathrm{x}|\mathrm{y}\right) \,,$$

Data Assimilation [BS12] IV

- It can be easily shown that:

$$
\begin{aligned}
x^a &= x^b + A \cdot H^T \cdot R^{-1} \cdot d = A \cdot \left[B^{-1} \cdot x^b + H^T \cdot R^{-1} \cdot y\right] \\
&= x^b + B \cdot H^T \cdot \left[R + H \cdot B \cdot H^T\right]^{-1} \cdot d
\end{aligned}
$$

where $A = \left[B^{-1} + H^T \cdot R^{-1} \cdot H\right]^{-1} \in \mathbb{R}^{n \times n}$, and $d = y - H \cdot x^b \in \mathbb{R}^{m \times 1}$.

- Posterior distribution:

$$
x \sim \mathcal{N}\left(x^a, A\right) .
$$

- How do we estimate x^b and B?.

Ensemble Based Methods

- We can make use of an ensemble of model realizations:

$$\mathrm{X}^b = \left[\mathrm{x}^{b[1]},\, \mathrm{x}^{b[2]},\, \ldots,\, \mathrm{x}^{b[N]}\right] \in \mathbb{R}^{n\times N}$$

- Empirical moments of the ensemble:

$$\begin{aligned} \mathrm{x}^b \approx \overline{\mathrm{x}}^b &= \frac{1}{N} \cdot \mathrm{X}^b \cdot \mathbf{1}_N \in \mathbb{R}^{n\times 1}\,, \\ \mathrm{B} \approx \mathrm{P}^b &= \frac{1}{N-1} \cdot \boldsymbol{\delta}\mathrm{X} \cdot \boldsymbol{\delta}\mathrm{X}^T \in \mathbb{R}^{n\times n}\,, \end{aligned}$$

and $\boldsymbol{\delta}\mathrm{X} = \mathrm{X}^b - \overline{\mathrm{x}}^b \cdot \mathbf{1}_N^T \in \mathbb{R}^{n\times N}$.

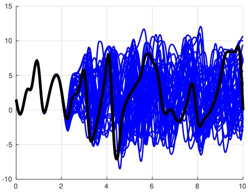
The Lorenz 96 Model - Toy Model I

- The Lorenz 96 model:

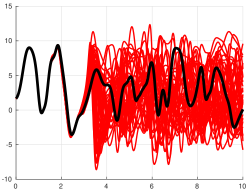
$$\frac{dx_j}{dt} = \begin{cases} (x_2 - x_{n-1}) \cdot x_n - x_1 + F & \text{for } i = 1, \\ (x_{i+1} - x_{i-2}) \cdot x_{i-1} - x_i + F & \text{for } 2 \leq i \leq n-1, \\ (x_1 - x_{n-2}) \cdot x_{n-1} - x_n + F & \text{for } i = n, \end{cases}$$

  where x_i stands for the i-th model component, for $1 \leq i \leq n$, usually $n = 40$.
- Each model component stands for a particle that fluctuates in the atmosphere.
- Exhibits chaotic behavior when the external force F is set to 8.

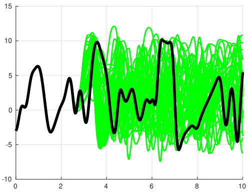
The Lorenz 96 Model - Toy Model II



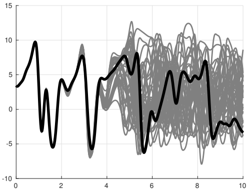
(e) x_5



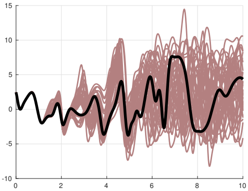
(f) x_{10}



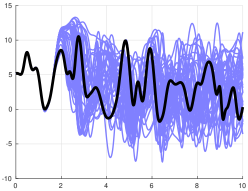
(g) x_{20}



(h) x_{30}

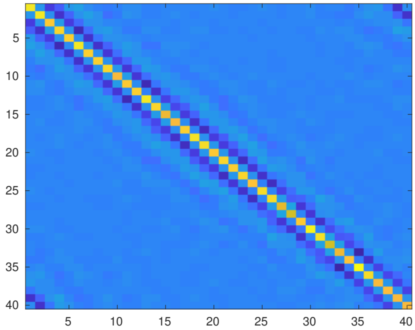


(i) x_{35}

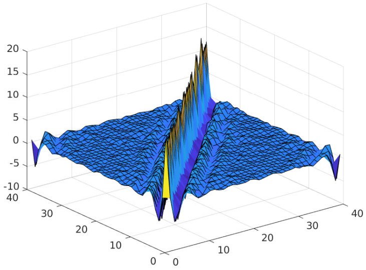


(j) x_{40}

Estimation of B via $N = 10^5$.



(a) Structure



(b) Surf

Figure: Estimation of B via $N = 10^5$.

The Stochastic Ensemble Kalman Filter [Eve03, Eve06] I

- Sequential Monte Carlo method for parameter and state estimation.
- Analysis ensemble (posterior ensemble):

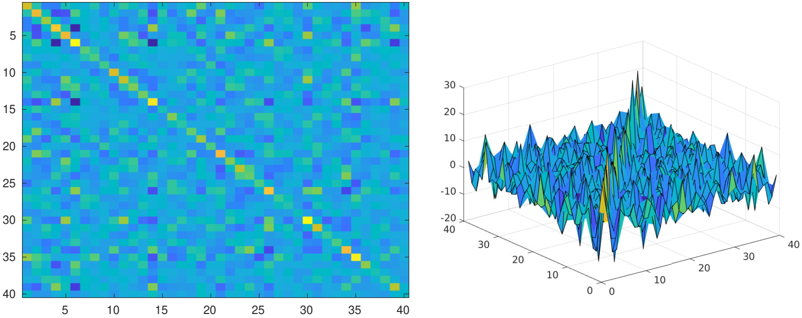
$$\begin{aligned} \mathbf{X}^a &= \mathbf{X}^b + \mathbf{P}^b \cdot \mathbf{H}^T \cdot \left[\mathbf{R} + \mathbf{H} \cdot \mathbf{P}^b \cdot \mathbf{H}\right]^{-1} \cdot \Delta\mathbf{Y} \\ \mathbf{X}^a &= \mathbf{X}^b + \mathbf{P}^a \cdot \mathbf{H}^T \cdot \mathbf{R}^{-1} \cdot \Delta\mathbf{Y} \in \mathbb{R}^{n \times N}, \\ \mathbf{X}^a &= \mathbf{P}^a \cdot \left[\mathbf{H}^T \cdot \mathbf{R}^{-1} \cdot \mathbf{Y}^s + \left[\mathbf{P}^b\right]^{-1} \cdot \mathbf{X}^b\right] \in \mathbb{R}^{n \times N}, \end{aligned}$$

where $\mathbf{P}^a = \left[\mathbf{H}^T \cdot \mathbf{R}^{-1} \cdot \mathbf{H} + \left[\mathbf{P}^b\right]^{-1}\right]^{-1} \in \mathbb{R}^{n \times n}$, and the e-th column of $\Delta\mathbf{Y} \in \mathbb{R}^{m \times N}$ and $\mathbf{Y}^s \in \mathbb{R}^{n \times N}$ are:

$$\mathbf{d}^{[e]} = \mathbf{y} + \boldsymbol{\epsilon}^{[e]} - \mathcal{H}\left(\mathbf{x}^{b[e]}\right) \in \mathbb{R}^{m \times 1}, \text{ and } \mathbf{y}^{s[e]} = \mathbf{y} + \boldsymbol{\epsilon}^{[e]},$$

respectively, for $1 \le e \le N$, and $\boldsymbol{\epsilon}^{[e]} \sim \mathcal{N}(\mathbf{0}_m, \mathbf{R})$.

Estimation of B via $N = 10$



(a) Structure (b) Surf

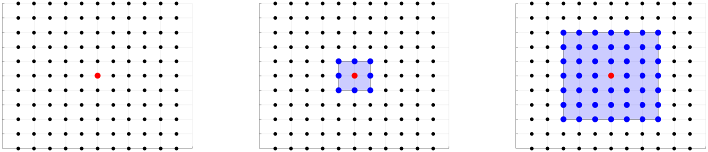
Figure: Estimation of B via $N = 10$.

Localization Methods

- Avoid the impact of spurious correlations.
- Increase the rank of $\mathbf{P}^b$.
- Covariance Matrix Localization. (Precision Localization) [NRSD15, NRSD17, NR17, NRSD18].
- Spatial Domain Localization [OHS$^+$04].
- Observation Localization [AND07, AND09].

Precision Matrix Localization I

- Component-wise products are prohibitive in high-dimensional spaces.
- *When two model components are conditional independent, their corresponding entry in the precision covariance matrix is zero*.



(a) $\delta = 0$ (b) $\delta = 1$ (c) $\delta = 3$

Precision Matrix Localization II

- Modified Cholesky Decomposition [BL+08]:

$$\widehat{\mathrm{B}}^{-1} = \mathrm{T}^T \cdot \mathrm{D}^{-1} \cdot \mathrm{T}$$

where the non-zero elements from $\mathrm{T} \in \mathbb{R}^{n \times n}$ are given by fitting models of the form:

$$\mathrm{x}^{[i]} = \sum_{q \in P(i,\delta)} \mathrm{x}^{[q]} \cdot \{-\mathrm{T}\}_{i,q} + \boldsymbol{\epsilon}^{[i]} \in \mathbb{R}^{N \times 1}, \text{ for } 1 \leq i \leq n\,,$$

and $\{\mathrm{D}\}_{i,i} = \mathrm{var}\left(\boldsymbol{\epsilon}^{[i]}\right)$.

1	5	9	13
2	6	10	14
3	7	11	15
4	8	12	16

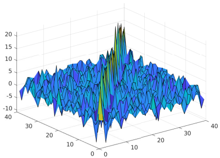
(a) $N(6,1)$

1	5	9	13
2	6	10	14
3	7	11	15
4	8	12	16

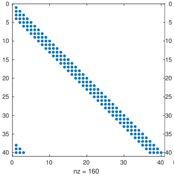
(b) $P(6,1)$

Precision Matrix Localization III

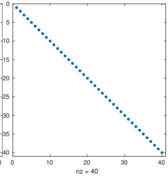
- Sparse estimates:



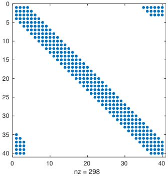
(a) P^b



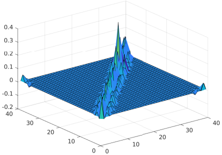
(b) T



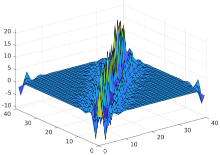
(c) D



(d) $\widehat{B}^{-1}$ Str



(e) $\widehat{B}^{-1}$



(f) $\widehat{B}$

Efficient Implementations of Ensemble Based Methods via a Modified Cholesky Decomposition

1. Nino-Ruiz, E. D. (2021). A data-driven localization method for ensemble-based data assimilation. Journal of Computational Science, 51, 101328.
2. Nino-Ruiz, E. D., Guzman-Reyes, L. G., & Beltran-Arrieta, R. (2020). An adjoint-free four-dimensional variational data assimilation method via a modified Cholesky decomposition and an iterative Woodbury matrix formula. Nonlinear Dynamics, 99(3), 2441-2457.

A Data-Driven Localization Method for Ensemble Based Data Assimilation I

- We want to estimate δ_i^*, for $1 \leq i \leq n$,
- this can be done by analyzing the error dynamics of model components, and
- errors are typically driven by model dynamics which develops error correlations in model components [MHR+19].
- Key assumptions:
  1. variances of model components can tell us how much information can be spread onto their local neighbourhood (domain), and
  2. covariances of model components can tell us how fast error correlations should decay.
- The first step is to consider a statistical model which can link radius lengths δ to empirical variances and covariances (i.e., from X^b).

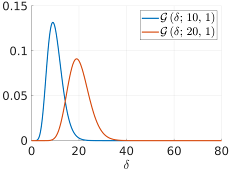
A Data-Driven Localization Method for Ensemble Based Data Assimilation II

- This can be done by employing a Gamma probability density function (pdf) with parameters $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$:

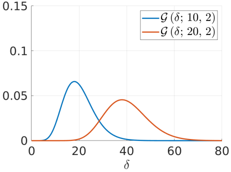
$$\mathcal{G}\left(\delta;\, \alpha,\, \beta\right) \propto \delta^{\alpha-1} \cdot \exp\left(-\delta/\beta\right)\,.$$

- We let the parameters α and β to influence the radius size and the decorrelation decay, respectively.

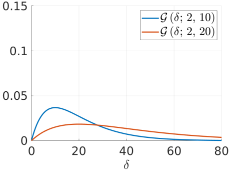
A Data-Driven Localization Method for Ensemble Based Data Assimilation III



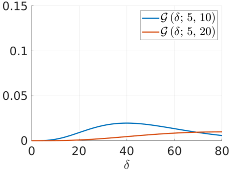
(a) $\alpha \in \{10, 20\}$, and $\beta = 1$



(b) $\alpha \in \{10, 20\}$, and $\beta = 2$



(c) $\beta \in \{10, 20\}$, and $\alpha = 2$



(d) $\beta \in \{10, 20\}$, and $\alpha = 5$

A Data-Driven Localization Method for Ensemble Based Data Assimilation IV

For each model component i,

$$\mathcal{P}\left(\delta_i|\mathbf{X}^b\right) \propto \mathcal{P}\left(\delta_i\right) \cdot \mathcal{L}\left(\delta_i|\mathbf{X}^b\right) \,,$$

where hyper-parameters read $\tilde{\alpha}$ and $\tilde{\beta}$ for all model components:

$$\mathcal{P}\left(\delta_i\right) = \mathcal{G}\left(\delta_i;\, \tilde{\alpha},\, \tilde{\beta}\right) \,,$$

We seek the radius length δ_i^* which maximizes the posterior probability:

$$\delta_i^* = \arg\max_{\delta_i} \mathcal{P}\left(\delta_i|\mathbf{X}^b\right) \,,$$

A Data-Driven Localization Method for Ensemble Based Data Assimilation V

look at the posterior moments:

$$\begin{aligned}
\mathcal{P}\left(\delta_i | \mathsf{X}^b\right) &\propto \delta_i^{\widetilde{\alpha}-1} \cdot \exp\left(-\delta_i / \widetilde{\beta}\right) \cdot \delta_i^{\widehat{\alpha}_i - 1} \cdot \exp\left(-\delta_i / \widehat{\beta}_i\right) \\
&\propto \delta_i^{\widetilde{\alpha}+\widehat{\alpha}_i-1-1} \cdot \exp\left(-\delta_i \cdot \left(\frac{\widetilde{\beta}+\widehat{\beta}_i}{\widetilde{\beta}\cdot\widehat{\beta}_i}\right)^{-1}\right) \\
&\propto \delta_i^{\widetilde{\alpha}+\widehat{\alpha}_i-1-1} \cdot \exp\left(-\delta_i \cdot \left(\frac{\widetilde{\beta}\cdot\widehat{\beta}_i}{\widetilde{\beta}+\widehat{\beta}_i}\right)\right) \, ,
\end{aligned}$$

therefore,

$$\mathcal{P}\left(\delta_i | \mathsf{X}^b\right) = \mathcal{G}\left(\delta_i ;\, \alpha_i^* , \beta_i^*\right) \, ,$$

A Data-Driven Localization Method for Ensemble Based Data Assimilation VI

where $\alpha_i^* = \widetilde{\alpha} + \widehat{\alpha}_i - 1$, and

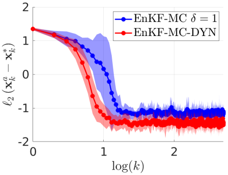
$$\beta_i^* = \frac{\widetilde{\beta} \cdot \widehat{\beta}_i}{\widetilde{\beta} + \widehat{\beta}_i} \,.$$

Thus, the solution of (1) reads:

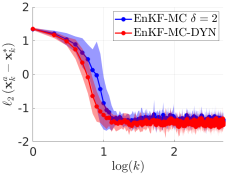
$$\mathbb{E}\left(\delta_i | \mathsf{X}^b\right) = \alpha_i^* \cdot \beta_i^* = \delta_i^* \,.$$

Experimental Results with the Lorenz-96 Model

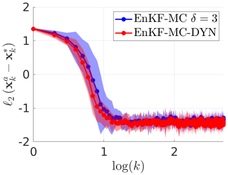
- Ensemble size 40.
- The number of observations reads 50%.
- Standard deviations of observation errors 0.01.
- Number of assimilation steps 500.



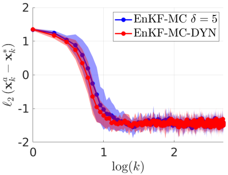
(a) $\delta = 1$



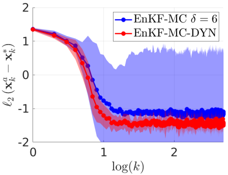
(b) $\delta = 2$



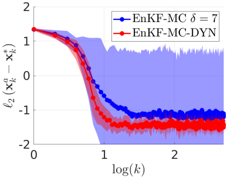
(c) $\delta = 3$



(d) $\delta = 5$

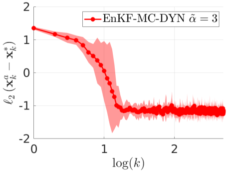


(e) $\delta = 6$

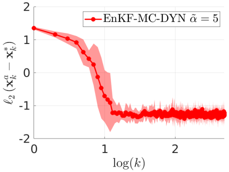


(f) $\delta = 7$

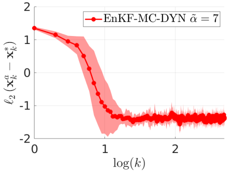
Figure: Mean and standard deviations, in log scale, of ℓ_2-norm of errors across 10 runs. 50% of model components are observed.



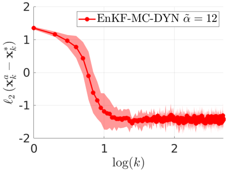
(a) $\widehat{\delta} = 3$



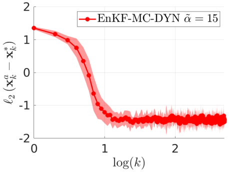
(b) $\widehat{\delta} = 5$



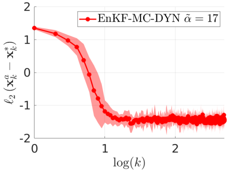
(c) $\widehat{\delta} = 7$



(d) $\widehat{\delta} = 12$

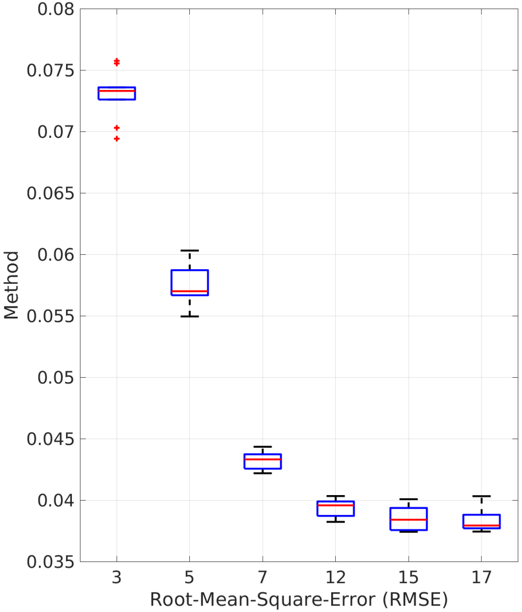


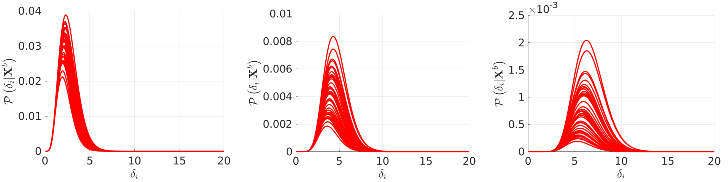
(e) $\widehat{\delta} = 15$



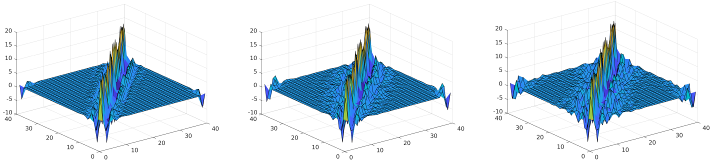
(f) $\widehat{\delta} = 17$

Figure: Different hyper-parameters $\widetilde{\alpha}$ are tried during experiments. The number of observations reads 50% of model components.


0.08
0.075
0.07
0.065
0.06
0.055
0.05
0.045
0.04
0.035
Method
3
5
7
12
15
17
Root-Mean-Square-Error (RMSE)



(a) $\widehat{\delta} = 7$ (b) $\widehat{\delta} = 12$ (c) $\widehat{\delta} = 17$



(d) $\widehat{\delta} = 7$ (e) $\widehat{\delta} = 12$ (f) $\widehat{\delta} = 17$

Figure: Posterior estimates of radius lengths and its corresponding covariance matrix estimations.

An adjoint-free 4D-Var method via a modified Cholesky decomposition and an iterative Woodbury matrix formula I

Recall the 4D-Var cost funcion:

$$\mathcal{J}(\mathrm{x}) = \frac{1}{2}\left\|\mathrm{x}-\mathrm{x}_0^b\right\|^2_{\mathrm{B}_0^{-1}} + \frac{1}{2}\cdot\sum_{k=0}^{M}\left\|\mathrm{y}_k-\mathcal{H}_k\left(\mathrm{x}\right)\right\|^{-1}_{\mathrm{R}_k^{-1}}\,,$$

this requires adjoint computation. Consider the following:

$$\mathrm{x}_k = \overline{\mathrm{x}}_k^b + \boldsymbol{\delta}\mathrm{X}_k\cdot\mathrm{w}\,,$$

where $\mathrm{w}\in\mathbb{R}^{N\times 1}$, we want:

$$\mathrm{x}_k-\overline{\mathrm{x}}_k^b \in \mathrm{range}\left\{\boldsymbol{\delta}\mathrm{X}_k\right\} \approx \mathrm{range}\left\{\mathrm{B}_k^{1/2}\right\}\,,$$

An adjoint-free 4D-Var method via a modified Cholesky decomposition and an iterative Woodbury matrix formula II

we get,

$$\begin{aligned}\mathcal{J}(\mathrm{x}_0) &= \mathcal{J}\left(\overline{\mathrm{x}}_0^b + \boldsymbol{\delta}\mathrm{X}_0 \cdot \mathrm{w}\right) \\ &= \mathcal{J}_{\text{ens}}(\mathrm{w}) = \frac{(N-1)}{2} \cdot \|\mathrm{w}\|^2 + \frac{1}{2} \cdot \sum_{k=0}^{M} \|\mathrm{d}_k - \mathrm{Q}_k \cdot \mathrm{w}\|^2_{\mathrm{R}_k^{-1}} \,,\end{aligned}$$

where $\mathrm{P}_k^b = \frac{1}{N-1} \cdot \boldsymbol{\delta}\mathrm{X}_k \cdot \boldsymbol{\delta}\mathrm{X}_k^T$, and $\mathrm{d}_k = \mathrm{y}_k - \mathrm{H}_k \cdot \overline{\mathrm{x}}_k^b$therefore,

$$\mathrm{w}^* = \arg\min_{\mathrm{w}} \mathcal{J}_{\text{ens}}(\mathrm{w})\,.$$

but, we continue working onto the ensemble space: impact of sampling errors.

An adjoint-free 4D-Var method via a modified Cholesky decomposition and an iterative Woodbury matrix formula III

Consider:

$$\widehat{\mathrm{B}}_k^{-1/2} = \widehat{\mathrm{T}}_k^T \cdot \widehat{\mathrm{D}}_k^{-1/2} \in \mathbb{R}^{n \times n}\,,$$

new space (more degrees of freedom):

$$\mathrm{x}_k = \overline{\mathrm{x}}_k^b + \widehat{\mathrm{B}}_k^{1/2} \cdot \boldsymbol{\alpha} \in \mathbb{R}^{n \times 1}\,.$$

Prior weights:

$$\boldsymbol{\alpha}^b \sim \mathcal{N}\left(0,\, \mathrm{I}\right)\,.$$

An adjoint-free 4D-Var method via a modified Cholesky decomposition and an iterative Woodbury matrix formula IV

The 4D-Var cost function:

$$\begin{aligned}\mathcal{J}\left(\mathrm{x}_0\right) &= \mathcal{J}\left(\overline{\mathrm{x}}_0^b + \widehat{\mathrm{B}}_0^{1/2}\cdot\boldsymbol{\alpha}\right) = \widehat{\mathcal{J}}\left(\boldsymbol{\alpha}\right)\\ &= \frac{1}{2}\cdot\left\|\boldsymbol{\alpha}\right\|^2 + \frac{1}{2}\cdot\sum_{k=0}^{M}\left\|\widehat{\mathrm{d}}_k - \widehat{\mathrm{Q}}_k\cdot\boldsymbol{\alpha}\right\|_{\mathrm{R}_k^{-1}}^2 ,\end{aligned}$$

where $\widehat{\mathrm{d}}_k = \mathrm{y}_k - \mathrm{H}_k\cdot\overline{\mathrm{x}}_k^b \in \mathbb{R}^{m\times 1}$, and $\widehat{\mathrm{Q}}_k = \mathrm{H}_k\cdot\widehat{\mathrm{B}}_k^{1/2} \in \mathbb{R}^{m\times n}$.
Problem to solve:

$$\boldsymbol{\alpha}^a = \arg\min_{\boldsymbol{\alpha}} \widehat{\mathcal{J}}\left(\boldsymbol{\alpha}\right) ,$$

An adjoint-free 4D-Var method via a modified Cholesky decomposition and an iterative Woodbury matrix formula V

note that:

$$\nabla_{\boldsymbol{\alpha}} \widehat{\mathcal{J}}(\boldsymbol{\alpha}) = \left[\mathbf{I} + \sum_{k=0}^{M} \widehat{\mathbf{Q}}_k^T \cdot \mathbf{R}_k^{-1} \cdot \widehat{\mathbf{Q}}_k \right] \cdot \boldsymbol{\alpha} - \sum_{k=0}^{M} \widehat{\mathbf{Q}}_k^T \cdot \mathbf{R}_k^{-1} \cdot \widehat{\mathbf{d}}_k \,,$$

whose root reads:

$$\boldsymbol{\alpha}^a = \left[\mathbf{I} + \sum_{k=0}^{M} \widehat{\mathbf{Q}}_k^T \cdot \mathbf{R}_k^{-1} \cdot \widehat{\mathbf{Q}}_k \right]^{-1} \cdot \left[\sum_{k=0}^{M} \widehat{\mathbf{Q}}_k^T \cdot \mathbf{R}_k^{-1} \cdot \widehat{\mathbf{d}}_k \right] ,$$

estimate of the initial analysis state:

$$\overline{\mathbf{x}}_0^a = \overline{\mathbf{x}}_0^b + \widehat{\mathbf{B}}_0^{1/2} \cdot \boldsymbol{\alpha}^a \,,$$

An adjoint-free 4D-Var method via a modified Cholesky decomposition and an iterative Woodbury matrix formula VI

the posterior ensemble onto the control space:

$$\boldsymbol{\alpha}^{a[e]} \sim \mathcal{N}\left(\boldsymbol{\alpha}^{a},\, \left[\mathbf{I} + \sum_{k=0}^{M} \widehat{\mathbf{Q}}_k^T \cdot \mathbf{R}_k^{-1} \cdot \widehat{\mathbf{Q}}_k\right]^{-1}\right),\ \text{for } 1 \leq e \leq N\,,$$

with corresponding analysis members in the model space:

$$\mathbf{x}_0^{a[e]} = \overline{\mathbf{x}}^{a} + \widehat{\mathbf{B}}_0^{1/2} \cdot \boldsymbol{\alpha}^{a[e]}\,.$$

Matrix-free Implementation I

Consider the sequence of matrices:

$$\mathrm{G}^{(q)} = \mathrm{G}^{(q-1)} + \widehat{\mathrm{Q}}_{q-1}^{T} \cdot \mathrm{R}_{q-1}^{-1} \cdot \widehat{\mathrm{Q}}_{q-1} \,,$$

for $1 \leq q \leq M+1$, with $\mathrm{G}^{(0)} = \mathrm{I}$, the posterior weights:

$$\left[\mathrm{G}^{(M+1)}\right] \cdot \boldsymbol{\alpha}^{a} = \boldsymbol{\pi} \,,$$

where $\boldsymbol{\pi} = \sum_{k=0}^{M} \widehat{\mathrm{Q}}_{k}^{T} \cdot \mathrm{R}_{k}^{-1} \cdot \widehat{\mathrm{d}}_{k} \in \mathbb{R}^{n \times 1}$. Hence,:

$$\left[\mathrm{G}^{(M)} + \widehat{\mathrm{Q}}_{M}^{T} \cdot \mathrm{R}_{M}^{-1} \cdot \widehat{\mathrm{Q}}_{M}\right] \cdot \boldsymbol{\alpha}^{a} = \boldsymbol{\pi} \,,$$

by using the Woodbury matrix identity,

$$\boldsymbol{\alpha}^{a} = \mathrm{z}^{(M+1)} = \mathrm{z}^{(M)} - \mathrm{Z}^{(M)} \cdot \left[\mathrm{R}_{M} + \widehat{\mathrm{Q}}_{M} \cdot \mathrm{Z}^{(M)}\right]^{-1} \cdot \widehat{\mathrm{Q}}_{M} \cdot \mathrm{z}^{(M)} \,,$$

Matrix-free Implementation II

where $\mathrm{G}^{(M)} \cdot \mathrm{z}^{(M)} = \boldsymbol{\pi}$ and $\mathrm{G}^{(M)} \cdot \mathrm{Z}^{(M)} = \widehat{\mathrm{Q}}_M^T$. By noting that $\mathrm{G}^{(M)} = \mathrm{G}^{(M-1)} + \widehat{\mathrm{Q}}_{M-1}^T \cdot \mathrm{R}_{M-1}^{-1} \cdot \widehat{\mathrm{Q}}_{M-1}$, subsequent linear systems, in general:

$$
\begin{aligned}
\mathrm{z}^{(q)} &= \mathrm{z}^{(q-1)} \\
&- \mathrm{Z}^{(q-1,q-1)} \cdot \left[\mathrm{R}_{q-1} + \widehat{\mathrm{Q}}_{q-1} \cdot \mathrm{Z}^{(q-1,q-1)}\right]^{-1} \cdot \widehat{\mathrm{Q}}_{q-1} \cdot \mathrm{z}^{(q-1)}\,, \\
\mathrm{Z}^{(q,w)} &= \mathrm{Z}^{(q-1,w)} \\
&- \mathrm{Z}^{(q-1,q-1)} \cdot \left[\mathrm{R}_{q-1} + \widehat{\mathrm{Q}}_{q-1} \cdot \mathrm{Z}^{(q-1,q-1)}\right]^{-1} \cdot \widehat{\mathrm{Q}}_{q-1} \cdot \mathrm{Z}^{(q-1,w)}\,,
\end{aligned}
$$

where $\mathrm{G}^{(q-1)} \cdot \mathrm{z}^{(q-1)} = \boldsymbol{\pi}$ and $\mathrm{G}^{(q)} \cdot \mathrm{Z}^{(q-1,w)} = \widehat{\mathrm{Q}}_w^T$, for $0 \le w \le 1 \le q \le M$.

Recurrences converge to (trivial) linear systems of the form:

$$
\mathrm{G}^{(0)} \cdot \mathrm{z}^{(0)} = \boldsymbol{\pi}, \text{ and } \mathrm{G}^{(0)} \cdot \mathrm{Z}^{(0,w)} = \widehat{\mathrm{Q}}_w^T\,.
$$

Matrix-free Implementation III

Set $z^{(0)} \leftarrow \boldsymbol{\pi}$

for $w \leftarrow 0 \rightarrow M$ **do**

  Set $Z^{(0,w)} \leftarrow \widehat{Q}_w^T$

end for

for $q \leftarrow 1 \rightarrow M+1$ **do**

  $Z^{\mathrm{piv}} \leftarrow Z^{(q-1,q-1)}$ ▷ Pivot element

  for $w \leftarrow q \rightarrow M$ **do**

    $Z^{(q,w)} \leftarrow$

$$Z^{(q-1,w)} - Z^{\mathrm{piv}} \cdot \left[R_{q-1} + \widehat{Q}_{q-1} \cdot Z^{\mathrm{piv}} \right]^{-1} \cdot \widehat{Q}_{q-1} \cdot Z^{(q-1,w)}$$

  end for

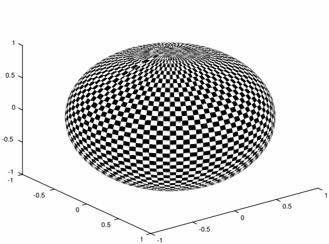
  $z^{(q)} \leftarrow z^{(q-1)} - Z^{\mathrm{piv}} \cdot \left[R_{q-1} + \widehat{Q}_{q-1} \cdot Z^{\mathrm{piv}} \right]^{-1} \cdot \widehat{Q}_{q-1} \cdot z^{(q-1)}$

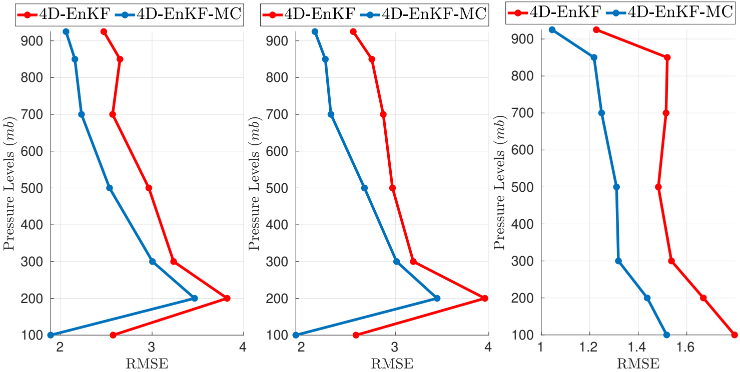
end for

return $z^{(M+1)}$ ▷ $z^{(M+1)}$ equals $\boldsymbol{\alpha}^a$

Experimental results with the SPEEDY model

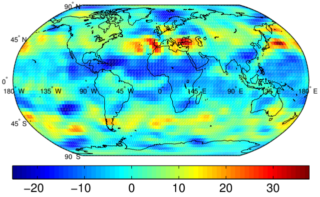
- Atmospheric General Circulation Model (AT-GCM).
- 5 model variables (u, v, T, ρ, H).
- 8 numerical layers.
- $T30 - 48 \times 96$
- $p = 60\%$, $N = 40$
- Observation operator:



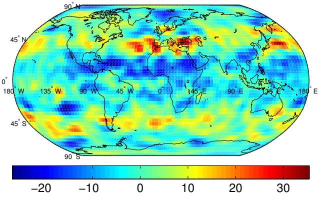


(a) x (b) v (c) T

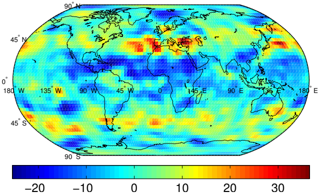
Figure: RMSE across pressure levels for $N = 10$ and $p = 60\%$. The errors per layer are shown for the initial analysis state and the model variables u, v, and T.



(a) x_0^*



(b) $\overline{x}^a_{0\{\mathrm{EnKF}\}}$



(c) $\overline{x}^a_{0\{\mathrm{EnKF-MC}\}}$

Figure: Snapshots of the Zonal Wind Component u at 200 mb level for the compared filter implementations and the reference solution. The snapshots are taken at the initial condition.

Conclusions

- Efficient implementations of EnKF methods can be obtained via a modified Cholesky decomposition.
- Special structure of Cholesky factors allow for efficient and practical implementations.
- Computational costs can be linearly bounded regarding model resolutions.
- The precision estimator presented here can be exploited in pre-conditioning.

EnKF-MC Publications

1. Elias D. Nino-Ruiz, Adrian Sandu, and Xinwei Deng. "An Ensemble Kalman Filter Implementation Based on Modified Cholesky Decomposition for Inverse Covariance Matrix Estimation", SIAM Journal on Scientific Computing 40:2, A867-A886 (2018).
2. Elias D. Nino-Ruiz, Adrian Sandu, and Xinwei Deng. "A parallel implementation of the ensemble Kalman filter based on modified Cholesky decomposition", Journal of Computational Science, Elsevier, (2017).

Bibliography I

[AND07] JEFFREY L. ANDERSON. An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus A*, 59(2):210–224, 2007.

[AND09] JEFFREY L. ANDERSON. Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus A*, 61(1):72–83, 2009.

[BL+08] Peter J Bickel, Elizaveta Levina, et al. Regularized estimation of large covariance matrices. *The Annals of Statistics*, 36(1):199–227, 2008.

[BS12] M. Bocquet and P. Sakov. Combining Inflation-free and Iterative Ensemble Kalman Filters for Strongly Nonlinear Systems. *Nonlinear Processes in Geophysics*, 19(3):383–399, 2012.

[Eve03] Geir Evensen. The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367, 2003.

[Eve06] Geir Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[MHR+19] Sammy Metref, Alexis Hannart, Juan Ruiz, Marc Bocquet, Alberto Carrassi, and Michael Ghil. Estimating model evidence using ensemble-based data assimilation with localization–the model selection problem. *Quarterly Journal of the Royal Meteorological Society*, 145(721):1571–1588, 2019.

[NR17] Elias D Nino-Ruiz. A matrix-free posterior ensemble kalman filter implementation based on a modified cholesky decomposition. *Atmosphere*, 8(7):125, 2017.

[NRSD15] Elias D. Nino-Ruiz, Adrian Sandu, and Xinwei Deng. A parallel ensemble kalman filter implementation based on modified cholesky decomposition. In *Proceedings of the 6th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems*, ScalA '15, pages 4:1–4:8, New York, NY, USA, 2015. ACM.

[NRSD17] Elias D Nino-Ruiz, Adrian Sandu, and Xinwei Deng. A parallel implementation of the ensemble kalman filter based on modified cholesky decomposition. *Journal of Computational Science*, 2017.

[NRSD18] Elias D Nino-Ruiz, Adrian Sandu, and Xinwei Deng. An ensemble kalman filter implementation based on modified cholesky decomposition for inverse covariance matrix estimation. *SIAM Journal on Scientific Computing*, 40(2):A867–A886, 2018.

[OHS+04] Edward Ott, Brian R. Hunt, Istvan Szunyogh, Aleksey V. Zimin, Eric J. Kostelich, Matteo Corazza, Eugenia Kalnay, D. J. Patil, and James A. Yorke. A local ensemble kalman filter for atmospheric data assimilation. *Tellus A*, 56(5):415–428, 2004.