Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Adaptive Localization for Ensemble Methods in Data Assimilation

África Periáñez, Hendrik Reich, Roland Potthast

DWD, Germany

International Conference on Ensemble Methods in Geophysical Sciences

Table of contents

Our problem

- Understand the basic properties of localization in the ensemble Kalman filter scheme.
- Define an adaptive localization depending on the density of data, observation and background error.
- Decomposition of the error sources to determine its effect on the optimal localization length scale.
- Perspective of approximation theory and functional analysis.
- Addressed with numerical experimental results.

Basic Approach

Let H be the observation operator mapping the state φ onto the measurements f. Then we need to find φ by solving the equation

$$H(\varphi) = f, \tag{1}$$

where H^{-1} is unstable or unbounded. When we have some initial guess $\varphi^{(b)}$, we transform the equation into

$$H(\varphi - \varphi^{(b)}) = f - H\varphi^{(b)}, \tag{2}$$

with the incremental form

$$\varphi = \varphi^{(b)} + H^{-1}(f - H\varphi^{(b)}). \tag{3}$$

Least Squares

In order to find out φ we should minimize the functional

$$J(\varphi) := \left\|\varphi - \varphi^{(b)}\right\|^2 + \left\|f - H\varphi^{(b)}\right\|^2. \tag{4}$$

The normal equations are obtained from first order optimality conditions

$$\nabla_\varphi J = 0. \tag{5}$$

Usually, the relation between variables at different points is incorporated by using covariances/weighted norms:

$$J(\varphi) := \left\|\varphi - \varphi^{(b)}\right\|^2_{B^{-1}} + \left\|f - H\varphi^{(b)}\right\|^2_{R^{-1}}, \tag{6}$$

The update formula is now

$$\varphi^{(a)} = \varphi^{(b)} + BH^*(R + HBH^*)^{-1}(f - H\varphi^{(b)}) \tag{7}$$

Ensemble Kalman Filter

In the Kalman filter method B also evolves with the model dynamics M,

$$B_{k+1} = MB_k M^*. \tag{8}$$

Ensemble filter methods use reduced rank estimation techniques to approximate the classical filters to reduce the computational costs. Starting with an ensemble $\varphi_0^{(l)}, \;\; l = 1, ..., L$, this leads to ensemble members

$$\varphi_k^{(l)} = M[t_k - 1, t_k]\varphi_{k-1}^{(l)}, \;\; k = 1, 2, 3, ... \tag{9}$$

In the EnKF methods the background covariance matrix is represented by $B^{(ens)} := \frac{1}{L-1} Q_k Q_k^*$. The ensemble matrix Q_k is defined as

$$Q_k := \left(\varphi_k^{(1)} - \overline{\varphi}_k^{(b)}, ..., \varphi_k^{(L)} - \overline{\varphi}_k^{(b)} \right), \tag{10}$$

where $\overline{\varphi}^{(b)}$ denotes the mean $\frac{1}{L} \sum_{l=1}^{L} \varphi^{(l)}$.

Thus, we solve the update in a low-dimensional subspace

$$U^{(L)} := \mathrm{span}\{ \varphi_k^{(1)} - \overline{\varphi}_k^{(b)}, ..., \varphi_k^{(L)} - \overline{\varphi}_k^{(b)} \}. \tag{11}$$

The update formula now is

$$\varphi_k^{(a)} = \varphi_k^{(b)} + Q_k Q_k^* H^* (R + H Q_k Q_k^* H^*)^{-1} (f_k - H\varphi_k^{(b)}) \tag{12}$$

The updates of the EnKF are a linear combination of the columns of Q_k. We can therefore write

$$\varphi_k - \varphi_k^{(b)} = \sum_{l=1}^{L} \gamma_l \left(\varphi_k^{(l)} - \overline{\varphi}_k^{(b)} \right) = Q_k \gamma \tag{13}$$

The resulting expression to minimize is

$$J(\gamma) := \| Q_k \gamma \|^2_{B_k^{-1}} + \| f_k - H\varphi_k^{(b)} - H Q_k \gamma \|^2_{R^{-1}}. \tag{14}$$

Error analysis without background contribution

We can write the observational error like

$$\|f^{(\delta)} - f^{(true)}\| \leq \delta. \tag{15}$$

The analysis error is given by

$$E_k := \|\varphi^{(a)} - \varphi^{(true)}\|. \tag{16}$$

We now provide estimates for the analysis error in norms which fit the particular setup of the assimilation task.

$$J(Q) := \|Q_k\gamma\|^2_{B_k^{-1}} + \|f_k - H\varphi_k^{(b)} - HQ_k\gamma\|^2_{R^{-1}}, \tag{17}$$

Error analysis without background contribution

Lemma

Assume that H is injective, that we study true measurement data $f = H\varphi^{(true)}$ and consider the EnKF with data term only

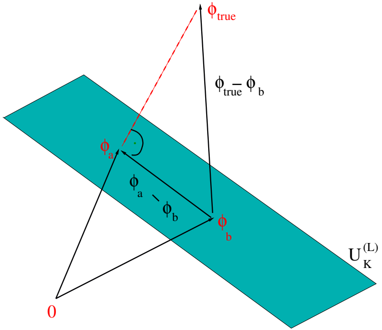
$$J^{(data)}(\gamma) = \|(f - H\varphi^{(b)}) - HQ_k\gamma\|^2_{R^{-1}} \tag{18}$$

Then, for the analysis $\varphi^{(a)}$ calculated by the EnKF the difference $\varphi^{(a)} - \varphi^{(b)}$ is the orthogonal projection of $\varphi^{(true)} - \varphi^{(b)}$ onto the ensemble space $U_k^{(L)}$ and the analysis error is given by

$$E_k = d_{H^* R^{-1} H}\Big(U_k^{(L)}, \varphi_k^{(true)} - \varphi^{(b)}\Big), \tag{19}$$

where the right-hand side denotes the distance between a point $\psi = \varphi_k^{(true)} - \varphi^{(b)}$ and the subspace $U^{(L)}$ with respect to the norm induced by the scalar product $< .,. >_{H^* R^{-1} H}$.

Illustration of Lemma



Error analysis with background term

Theorem

Assume that H is injective, that we study true measurement data $f = H\varphi^{(true)}$ and consider an assimilation step using the EnKF. Then, for the analysis error in the step k we have the analysis error estimate

$$\|\varphi_k^{(true)} - \varphi^{(b)}\|_{H^* R^{-1} H} \geq E_k \geq d_{H^* R^{-1} H}\Big(U_k^{(L)}, \varphi_k^{(true)} - \varphi_k^{(b)}\Big). \quad (20)$$

Localization

LETKF basic idea: Localization to D, leading to

$$Q_{k,loc} := \Big(\chi_D(\varphi_k^{(1)} - \overline{\varphi}_k^{(b)}), ..., \chi_D(\varphi_k^{(L)} - \overline{\varphi}_k^{(b)}) \Big) . \quad (21)$$

We now have

$$B = \frac{1}{L-1} Q_{k,loc} Q_{k,loc}^T \quad (22)$$

and

$$f_{k,loc} = \chi_D f_k \quad (23)$$

We now solve the equations in the locally low-dimensional subspace

$$U_k^{(L,D)} := \mathrm{span}\{ \chi_D(\varphi_k^{(1)} - \overline{\varphi}_k^{(b)}), ..., \chi_D(\varphi_k^{(L)} - \overline{\varphi}_k^{(b)}) \}. \quad (24)$$

Localization

Thus, in the previous error estimates we just have to replace

$$
\begin{aligned}
E_k &\rightarrow E_{k,loc} \\
(\varphi_k^{(true)} - \varphi^{(b)}) &\rightarrow \chi_D(\varphi_k^{(true)} - \varphi^{(b)}) \\
U_k^{(L)} &\rightarrow U_{k,loc}^{(L)}.
\end{aligned}
\tag{25}
$$

to get the local error estimates.

Localization

We define

$$\hat{E}^{(\rho)}(x_0) := \frac{1}{|B_\rho(x_0)|} \left\| \hat{\varphi}^{(a,\rho)} - \varphi^{(true)} \right\|_{H^* R^{-1} H, \rho}, \qquad (26)$$

where $|B_\rho(x_0)|$ denotes the volume or area of the ball $B_\rho(x_0)$ in two, three or more dimensions.

$\hat{\varphi}^{(a,\rho)}$ is the analysis obtained when the data is assumed to be perfect and we are performing the analysis without background term.

Localization. Convergence results.

Theorem

We study assimilation in the case where true data $\varphi^{(true)}$ *are used and* $\hat{\varphi}^{(a,\rho)}$ *is chosen such that* $\hat{\varphi}^{(a,\rho)} - \varphi^{(b)}$ *is the orthogonal projection of* $\varphi^{(true)} - \varphi^{(b)}$ *onto the ensemble space* $U^{(L)}$. *Assume that there is* c, $C > 0$ *such that for all* $x \in D$ *there is* $l \in \{1, ..., L\}$ *such that*

$$|\varphi^{(l)}(x) - \varphi^{(b)}(x)| \geq c, \tag{27}$$

and that the ensemble members are continuously differentiable on D *with*

$$|\nabla(\varphi^{(j)}(x) - \varphi^{(b)}(x))| \leq C, \ \ x \in D, \ \ j \in \{1, ..., L\}. \tag{28}$$

Further assume that $\varphi^{(true)} - \varphi^{(b)}$ *is continuously differentiable on* D. *Then, we have*

$$\sup_{x_0 \in D} \hat{E}^{\rho}(x_0) \leq \tilde{C}\rho \ \ \to 0, \ \ \rho \to 0 \tag{29}$$

with some constant $\tilde{C}$ *depending on* C, H *and* R.

Least Squares Analysis model

- One dimensional model without cycling
- Least square estimation to obtain the analysis (LSA) and the truth is given by a high-order function.
- The analysis is obtained using both all available observations and only a **local** subset.
- Estimation performed with and without background correction.
- Observations are generated from the truth with a specified observation error σ_{obs}.
- Analysis approximated by straight lines $a + bx$

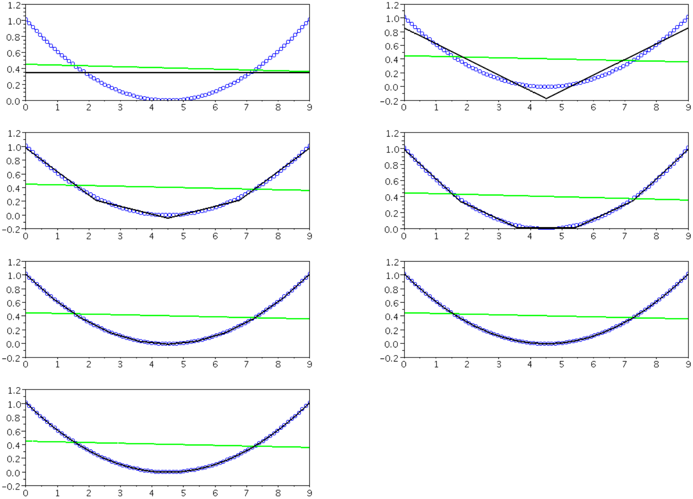


Fig.1: truth (blue line), observations (blue circles), background (green), no background LSA (red), background LSA (black) for $\sigma_{obs} = 0.0005$.

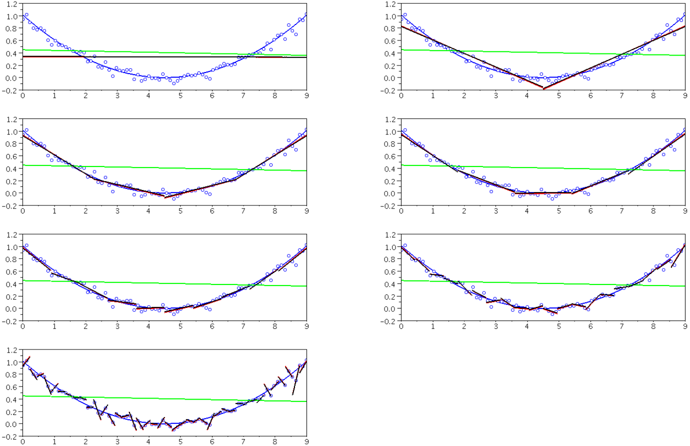


Fig.1: Truth (blue line), observations (blue circles), background (green), no background LSA (red) and background LSA (black) for $\sigma_{obs} = 0.05$ and different localization radii.

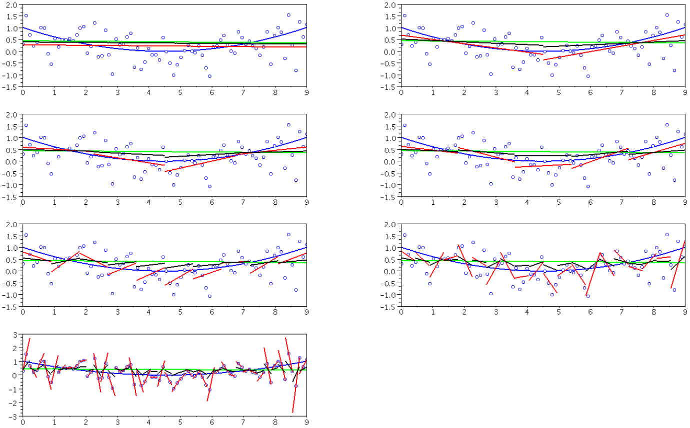


Fig.2: Truth (blue line), observations (blue circles), background (green), no background LSA (red) and background LSA (black) for $\sigma_{obs} = 0.5$ and different localization radii.

Remarks

- The optimal value of ρ_{loc} takes smaller values when σ_{obs} decreases.
- For large values of σ_{obs} the analysis without the background correction is worse than analysis considering the background.
- For small σ_{obs} the results with and without background influence become very similar.

Optimal localization radius

- Estimation of the optimal localization radius ρ_{loc} as a function of σ_{obs} and observation density d (without considering the background).
- **Approximation error** decreases with smaller localization radius.
- **Sampling error** decreases with a larger localization radius, as a larger number of observations gives a better statistical estimation. It is then $\sim 1/\sqrt{N_{obs}}$, with $N_{obs} = \int_V d(x)dV = 2d\rho_{loc}$.
- The sum of both errors leads to the ansatz:

$$\hat{e} \sim \rho_{loc}^2 \quad + \quad \frac{\sigma_{obs}}{\sqrt{2d\rho_{loc}}} \tag{30}$$

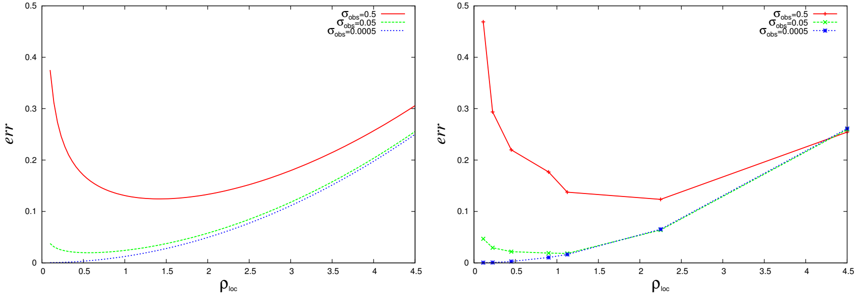
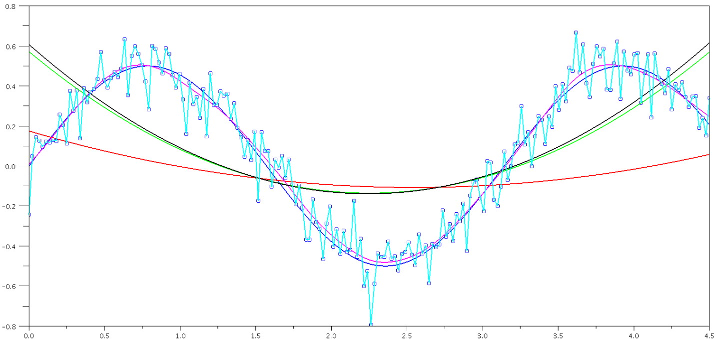


Fig.4: theoretical and numerical results for error as a function of ρ_{loc}, $\sigma_{obs} = [0.0005\,0.05\,0.5]$.

The optimal value of the localization radius ρ_{loc} takes smaller values when σ_{obs} decreases.

LETKF example

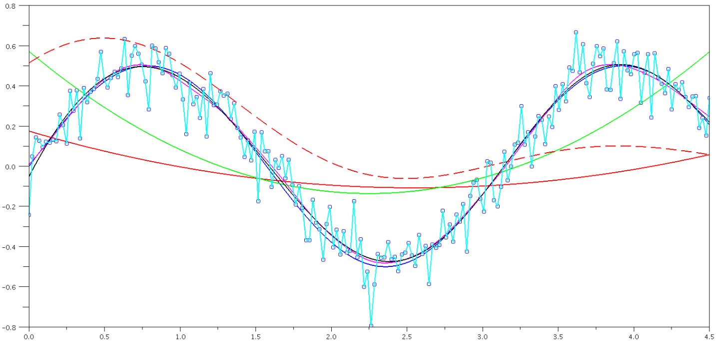
- Computation of the analysis with the LETKF algorithm, 3dVar and spline interpolation.
- The truth is a sin-type function.
- LETKF/3DVar uses either quadratic functions or sin / cos background functions.



o o o observations —— truth —— LSA/LETKF/3dVar background
—— LSA analysis —— LETKF —— 3dVar —— spline

Fig.4 $\sigma_{obs} = 0.1$, $N_{ens} = 10$, $N_P = 3$ in LETKF background ensemble

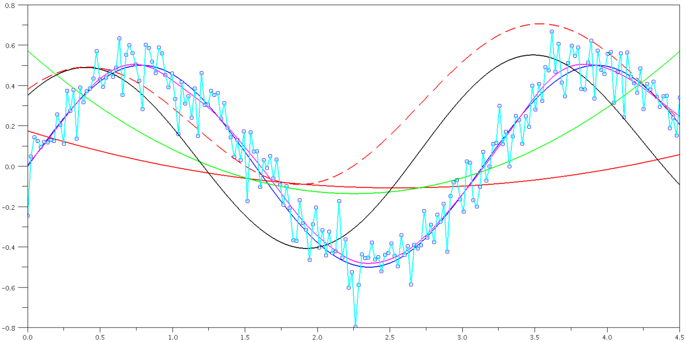
LETKF is similar to LSA; 3dVar analysis is the best



o o o observations —— truth —— LSA/3dVar background
- - - LETKF background —— LSA analysis —— LETKF —— 3dVar —— spline

Fig.5: $\sigma_{obs} = 0.1$, $N_{ens} = 10$, sin/cos base functions as LETKF background ensemble.

LETKF similar to 3dVar



o o o observations —— truth —— LSA/3dVar background
- - - LETKF background —— LSA analysis —— LETKF —— 3dVar —— spline

Fig.6: $\sigma_{obs} = 0.1$, sin/cos base functions as LETKF background ensemble, but

$$N_{ens} = 3$$

LETKF analysis is worse than 3dVar

Remarks

- When is LETKF similar to LSA (with background term)?
  - the order of the background function both LETKF and LSA, N_P, should be the same; additionally we need $(N_{ens} - 1) \geq N_P$
- LETKF cannot "fit" more than N_{ens} observations, but we have to distinguish two cases:
  - if N_P and N_{ens} are comparable to "order" of the truth (ensemble subspace is good approximation to truth $\rightarrow$ *approximation error* small), the error will decrease $\sim \frac{1}{\sqrt{N_{obs}}}$ even if $N_{obs} > (N_{ens} - 1)$ (*sampling error*)
  - if LETKF subspace is too small/not appropriate (model error?); *approximation error* dominates, additional obs do not have positive impact for $N_{obs} > (N_{ens} - 1)$

Adaptive horizontal localization

- Localization length scales depend on weather situation and observation density.
- Simple adaptive localization method: keep the number of *effective observations* fixed, vary localization radius (*effective observations*: sum of **observation weights**).
- Up to now only implemented in horizontal direction.
- A minimum/maximum radius is defined, as well a number of *effective observations* $N_{obs}^{eff} = \alpha(N_{ens} - 1)$, $\alpha \geq 1$
- The ideal number of effective observations depends on ensemble size.

Outlook / Conclusion

- 1d model: optimal localization length ρ_{loc} depends on σ_{obs}; this (first results) also seems to be the case for the L95-LETKF.
- 2-step analysis gives better results if two observation types and $\sigma_{obs}^1 >> \sigma_{obs}^2$.
- 1d model: for fixed ρ_{loc} in LETKF: $N_{obs} > (N_{ens} - 1)$ gives better results only if ensemble-subspace is appropriated.
- 2d model LETKF: similar results found
- "Classical" view (avoid spurious correlations) of localization in EnKF: up to which distance can we trust the correlations in the ensemble?
- Are both approaches connected? Do they lead to similar optimal localization radii? Should be tested (L95-LETKF).