Overview of localization techniques for ensemble based Kalman filter algorithms

T.Janjić

Why ensemble Kalman filter

- The Kalman filter is difficult to implement in realistic systems because of:
  - computational costs,
  - the nonlinearity of dynamics and
  - poorly characterized error sources.
- The ensemble Kalamn filter (EnKF) (Evensen 1994) uses ensembles (a sample) to calculate the uncertainty of the background and analysis error covariance.
- Ensembles are propagated with full nonlinear numerical model. This can be done over long time period, and results in flow dependent covariances.

Pros of ensemble Kalman filter

- Cross correlations are represented naturally
- Covariances are flow dependent
- Computationally algorithm is not expensive

Cons of ensemble Kalman filter

- Only small number of ensembles can be evolved due to complexity of the dynamical systems;
- Due to the small ensemble numbers covariances are not representing correctly uncertainty, in particular long-distance correlations.
- The analysis increment is restricted to the r dimensional subspace.

Outline localization

- What is localization?
- Two basic approaches for localization:
  - Covariance localization or direct forecast error localization (used in Houtekamer and Mitchell (1998, 2001))
  - Domain localization (used in Haugen and Evensen 2002; Brusdal et al. 2003; Ott et al. 2004; Hunt et al. 2007; Miyoshi and Yamane 2007)
- Effects of localization on each of the steps of ensemble Kalman filter algorithm
- Simple 1D experiment
- Localization and Balance
- Conclusion

Step 1: Analysis

$$\mathbf{w}_k^a = \mathbf{w}_k^b + \mathbf{K}_k(\mathbf{y}_k^o - \mathbf{H}_k\mathbf{w}_k^b),$$

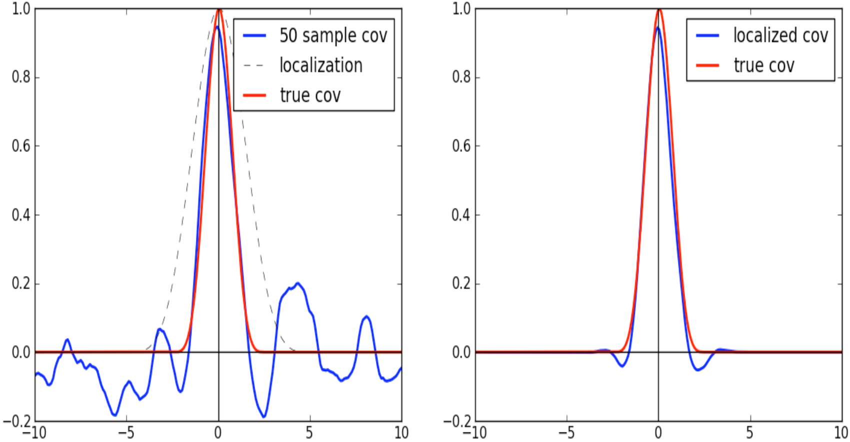
$\mathbf{K}_k$ is taken as

$$\mathbf{K}_k = \mathbf{B}_k^b\mathbf{H}_k^T(\mathbf{H}_k\mathbf{B}_k^b\mathbf{H}_k^T + \mathbf{R}_k)^{-1}$$

with $\mathbf{B}_k^b$ represented by:

$$\mathbf{B}_k^b = \frac{1}{r-1}\sum_{i=1}^{r}[\mathbf{w}_k^{b,i} - \mathbf{w}_k^b][\mathbf{w}_k^{b,i} - \mathbf{w}_k^b]^T.$$

What is localization?



Estimates of the covariance from small ensemble size will be noisy, especially signal to noise ratio is large when covariances are small (from Hamill and Whitaker 2009).

Covariance localization

Covariance localization: The ensemble derived forecast error covariance matrix is Schur multiplied with a stationary a priori chosen correlation matrix that is compactly supported.

Let $\mathbf{C}$ be a matrix of rank M that is used for the Schur product.

$$\mathbf{C} \circ \mathbf{B}_k^b$$

Let $\circ$ denotes the element-wise product (Schur product) where

$$[\mathbf{C} \circ \mathbf{B}_k^b]_{ij} = [\mathbf{C}]_{ij} [\mathbf{B}_k^b]_{ij}$$

Covariance localization

Schur product theorem:

If $\mathbf{A}$, $\mathbf{B}$ are positive semi-definite matrices, then $\mathbf{A} \circ \mathbf{B}$ is also positive semi-definite. If $\mathbf{A}$, $\mathbf{B}$ are positive definite matrices, then $\mathbf{A} \circ \mathbf{B}$ is also positive definite.

Covariance localization

Schur product theorem:

If $\mathbf{A}$, $\mathbf{B}$ are positive semi-definite matrices, then $\mathbf{A} \circ \mathbf{B}$ is also positive semi-definite. If $\mathbf{A}$, $\mathbf{B}$ are positive definite matrices, then $\mathbf{A} \circ \mathbf{B}$ is also positive definite.

$$min(diag(\mathbf{B}_k^b))\lambda_{min}(\mathbf{C}) \leq \lambda_{min}(\mathbf{B}_k^b \circ \mathbf{C}) \leq \lambda_{max}(\mathbf{B}_k^b \circ \mathbf{C}) \leq max(diag(\mathbf{B}_k^b))\lambda_{max}(\mathbf{C})$$

Covariance localization

Schur product theorem:

If $\mathbf{A}$, $\mathbf{B}$ are positive semi-definite matrices, then $\mathbf{A} \circ \mathbf{B}$ is also positive semi-definite. If $\mathbf{A}$, $\mathbf{B}$ are positive definite matrices, then $\mathbf{A} \circ \mathbf{B}$ is also positive definite.

$$min(diag(\mathbf{B}_k^b))\lambda_{min}(\mathbf{C}) \leq \lambda_{min}(\mathbf{B}_k^b \circ \mathbf{C}) \leq \lambda_{max}(\mathbf{B}_k^b \circ \mathbf{C}) \leq max(diag(\mathbf{B}_k^b))\lambda_{max}(\mathbf{C})$$

Let $\mathbf{C}$ be a matrix of rank M that is used for the Schur product. Let $\mathbf{v}_j$ represent eigenvectors of matrix $\mathbf{C}$ multiplied with the square root of the corresponding eigenvalue.

$$\mathbf{C} = \sum_{j=1}^{M} \mathbf{v}_j \mathbf{v}_j^T .$$

Covariance localization increases the rank of ensemble derived covariances

For any vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and $\mathbf{d}$:

$$(\mathbf{a} \circ \mathbf{c})(\mathbf{b} \circ \mathbf{d})^T = (\mathbf{a}\mathbf{b}^T) \circ (\mathbf{c}\mathbf{d}^T).$$

The localized error covariance $\mathbf{B}_k^b \circ \mathbf{C}$ can be represented as

$$\begin{aligned} \mathbf{B}_k^b \circ \mathbf{C} &= \frac{1}{r-1} \sum_{i=1}^{r} [\mathbf{w}_k^{b,i} - \mathbf{w}_k^b][\mathbf{w}_k^{b,i} - \mathbf{w}_k^b]^T \circ \sum_{j=1}^{M} \mathbf{v}_j \mathbf{v}_j^T \\ &= \frac{1}{r-1} \sum_{i,j=1}^{r,M} [\mathbf{w}_k^{b,i} - \mathbf{w}_k^b][\mathbf{w}_k^{b,i} - \mathbf{w}_k^b]^T \circ \mathbf{v}_j \mathbf{v}_j^T \\ &= = \sum_{i,j=1}^{r,M} \mathbf{u}_{i,j} \mathbf{u}_{i,j}^T \end{aligned}$$

with $\mathbf{u}_{i,j} = \frac{1}{\sqrt{r-1}}[\mathbf{w}^{f,i}(t_k) - \mathbf{w}_k^b] \circ \mathbf{v}_j$

This representation implies that instead of using ensemble members $\mathbf{w}^{f,i}$ for the calculation of the analysis error covariance, we can use the ensemble $\mathbf{u}_{i,j}$, and the same formulas as in original algorithms apply.

This representation implies that instead of using ensemble members $\mathbf{w}^{f,i}$ for the calculation of the analysis error covariance, we can use the ensemble $\mathbf{u}_{i,j}$, and the same formulas as in original algorithms apply.

- This procedure increases the cost of the algorithm for calculating the analysis error covariance matrix since now $\mathbf{W}_k^b = [\mathbf{u}_{1,1} \ldots \mathbf{u}_{r,M}]$.

This representation implies that instead of using ensemble members $\mathbf{w}^{f,i}$ for the calculation of the analysis error covariance, we can use the ensemble $\mathbf{u}_{i,j}$, and the same formulas as in original algorithms apply.

- This procedure increases the cost of the algorithm for calculating the analysis error covariance matrix since now $\mathbf{W}_k^b = [\mathbf{u}_{1,1} \dots \mathbf{u}_{r,M}]$.
- The increase in cost depends on the rank of the chosen localization matrix C.

This representation implies that instead of using ensemble members $\mathbf{w}^{f,i}$ for the calculation of the analysis error covariance, we can use the ensemble $\mathbf{u}_{i,j}$, and the same formulas as in original algorithms apply.

- This procedure increases the cost of the algorithm for calculating the analysis error covariance matrix since now $\mathbf{W}_k^b = [\mathbf{u}_{1,1} \dots \mathbf{u}_{r,M}]$.
- The increase in cost depends on the rank of the chosen localization matrix C.
- We are choosing **C** and therefore we can choose **C** with any rank M.

This representation implies that instead of using ensemble members $\mathbf{w}^{f,i}$ for the calculation of the analysis error covariance, we can use the ensemble $\mathbf{u}_{i,j}$, and the same formulas as in original algorithms apply.

- This procedure increases the cost of the algorithm for calculating the analysis error covariance matrix since now $\mathbf{W}_k^b = [\mathbf{u}_{1,1} \ldots \mathbf{u}_{r,M}]$.
- The increase in cost depends on the rank of the chosen localization matrix C.
- We are choosing **C** and therefore we can choose **C** with any rank M.
- This way we would still propagate only r ensemble members with the numerical model (the expensive part of the algorithm.)

This representation implies that instead of using ensemble members $\mathbf{w}^{f,i}$ for the calculation of the analysis error covariance, we can use the ensemble $\mathbf{u}_{i,j}$, and the same formulas as in original algorithms apply.

- This procedure increases the cost of the algorithm for calculating the analysis error covariance matrix since now $\mathbf{W}_k^b = [\mathbf{u}_{1,1} \dots \mathbf{u}_{r,M}]$.
- The increase in cost depends on the rank of the chosen localization matrix C.
- We are choosing **C** and therefore we can choose **C** with any rank M.
- This way we would still propagate only r ensemble members with the numerical model (the expensive part of the algorithm.)
- However for the calculation of the analysis error covariance matrix $\mathbf{B}_k^a$ with rank higher then r will be used.

This representation implies that instead of using ensemble members $\mathbf{w}^{f,i}$ for the calculation of the analysis error covariance, we can use the ensemble $\mathbf{u}_{i,j}$, and the same formulas as in original algorithms apply.

- This procedure increases the cost of the algorithm for calculating the analysis error covariance matrix since now $\mathbf{W}_k^b = [\mathbf{u}_{1,1} \dots \mathbf{u}_{r,M}]$.
- The increase in cost depends on the rank of the chosen localization matrix C.
- We are choosing **C** and therefore we can choose **C** with any rank M.
- This way we would still propagate only r ensemble members with the numerical model (the expensive part of the algorithm.)
- However for the calculation of the analysis error covariance matrix $\mathbf{B}_k^a$ with rank higher then r will be used.

$$\mathbf{w}_k^a = \mathbf{w}_k^b + \mathbf{K}_k(\mathbf{f}_k^o - \mathbf{H}_k\mathbf{w}_k^b)$$

$$\mathbf{w}_k^a - \mathbf{w}_k^b = \mathbf{W}_k^b\mathbf{T}_k{\mathbf{T}_k}^T(\mathbf{H}_k\mathbf{W}_k^b)^T\mathbf{R}_k^{-1}(\mathbf{f}_k^o - \mathbf{H}_k\mathbf{w}_k^b)$$

$$\mathbf{w}_k^a = \mathbf{w}_k^b + \mathbf{K}_k(\mathbf{f}_k^o - \mathbf{H}_k\mathbf{w}_k^b)$$

$$\mathbf{w}_k^a - \mathbf{w}_k^b = \mathbf{W}_k^b\mathbf{T}_k{\mathbf{T}_k}^T(\mathbf{H}_k\mathbf{W}_k^b)^T\mathbf{R}_k^{-1}(\mathbf{f}_k^o - \mathbf{H}_k\mathbf{w}_k^b)$$

- We have increased the subspace space where the solution is searched for.
- Matrices $\mathbf{W}_k^b$ is of size $n \times Mr$.

Step 2

In resampling step we can not produce an ensemble with size r if we use the analysis error covariance calculated from $\mathbf{W}_k^b = [\mathbf{u}_{1,1} \ldots \mathbf{u}_{r,M}]$.

$$\mathbf{w}_k^{a,i} = \mathbf{w}_k^a + \sqrt{r-1}[\mathbf{W}_k^b \mathbf{T}_k \mathbf{U}]_i$$

One approach to limit the ensemble to r ensemble members that we can propagate to time $k+1$ around the newly calculated analysis:

- use instead $\mathbf{W}_k^b = \frac{1}{\sqrt{r-1}}[\mathbf{w}_k^{b,1} - \mathbf{w}_k^b \ldots \mathbf{w}_k^{b,r} - \mathbf{w}_k^b]$

Example

The state vector **w** to be estimated will be taken as a realization of normally distributed random function $w(y) \sim \mathcal{N}(0, W(y_1, y_2))$ on the circle of radius $D/2\pi$, where the covariance $W(y_1, y_2)$ is either

$$W(y_1, y_2) = (1 + \frac{|y_1 - y_2|}{L}) e^{\frac{-|y_1 - y_2|}{L}}, \quad (1)$$

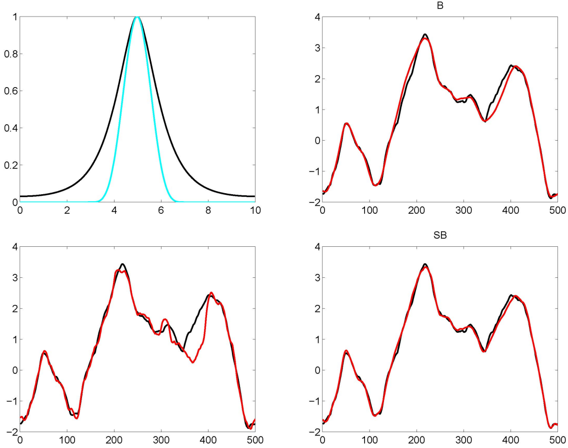
or

$$W(y_1, y_2) = e^{-\frac{|y_1 - y_2|}{L}}. \quad (2)$$

Here, $|y_1 - y_2|$ represents the chord length between the points y_1 and y_2 on the circle of radius $D/2\pi$.

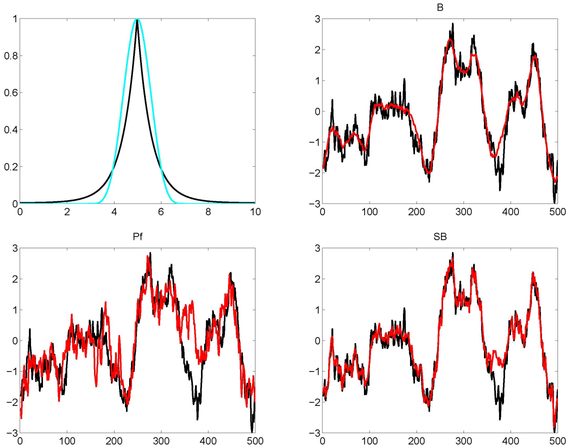
The observations are given as a vector of values of the realization at all grid points contaminated by normally distributed random noise with standard deviation of 0.05, the observations from two subdomains were removed.

Example $W(y_1, y_2) = (1 + \frac{|y_1 - y_2|}{L})e^{\frac{-|y_1-y_2|}{L}}$



Upper Left: True covariance (black) and approximate C covariance (blue). Upper Right: True state (black) and analysis (red) after one assimilation step with approximate B covariance. Lower Left: True state (black) and analysis (red) after one assimilation step with ensemble covariance from 30 ensemble members. Lower Right: True state (black) and analysis (red) after one assimilation step with localized ensemble covariance.

Example nonsmooth field $W(y_1, y_2) = e^{-\frac{|y_1 - y_2|}{L}}$



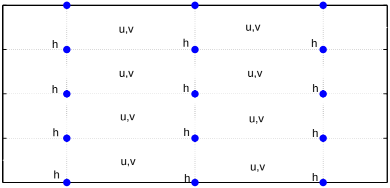
Upper Left: True covariance (black) and approximate B covariance (blue).
Upper Right: True state (black) and analysis (red) after one assimilation step with approximate B covariance.
Lower Left: True state (black) and analysis (red) after one assimilation step with ensemble covariance from 30 ensemble members.
Lower Right: True state (black) and analysis (red) after one assimilation step with localized ensemble covariance.

Covariance localization

- Distant correlation are removed
- Positive definite correlation matrix is introduced that increases the rank of forecast error covariance and this way
- increases the space where the solution can be searched for
- usually correlation function **C** is chosen full rank, positive definite, isotropic matrix, compactly supported. Usually 5th order polynomial correlation function (Gaspari and Cohn 1999).

Domain localization

Domain localization: Disjoint domains in the physical space are considered as domains on which the analysis is performed. Therefore, for each subdomain an analysis step is performed independently using observations not necessarily belonging only to that subdomain. Results of the local analysis steps are pasted together and then the global forecast step is performed.



Basic properties:

- The localized error covariance is calculated using

$$\mathbf{B}_k^{f,loc} = \sum_{i,j=1}^{r,L} \mathbf{u}_{i,j}\mathbf{u}_{i,j}^T \tag{3}$$

where $\mathbf{u}_{i,j} = \frac{1}{\sqrt{r}}[\mathbf{w}^{f,i}(t_k) - \mathbf{w}_k^b] \circ \mathbf{1}_{Dj}$ with $j = 1, \ldots, L$ and L is the number of subdomains. Here $\mathbf{1}_{Dj}$ is a vector whose elements are 1 if the corresponding point belongs to the domain Dj.

Domain localization

- $\mathbf{C}$ positive semidefinite, has block structure and is the sum of rank one matrices $\mathbf{1}_{Dj}\mathbf{1}_{Dj}^T$. The rank of matrix $\mathbf{C}$ corresponds to the number of subdomains.
- In case that $\text{rank}(\mathbf{C})\text{rank}(\mathbf{B}_k^f) < n$, the matrix $\mathbf{C} \circ \mathbf{B}_k^f$ is singular.
- In domain localization methods, the rank is not increased locally on each subdomain. Accordingly, it is possible to resample exactly on that subdomain in contrast to direct forecast error localization.
- Because the assimilations are performed independently in each local region, the smoothness of the analysis fields is of more concern in domain localization methods than with direct forecast error localization. In particular, two neighboring subdomains might produce strongly different analysis estimates when the assimilated observations have gaps, because distinct sets of observations are used for the analyses.

Ensemble based Kalman filters apply the observation operator directly on each ensemble member before localization is applied. The localization is usually performed on the matrices $\mathbf{H_k B}_k^b$ and $\mathbf{H_k B}_k^b \mathbf{H_k}^T$

$$\mathbf{H_k B}_k^b = \frac{1}{r} \sum_{i=1}^{r+1} [\mathbf{H_k}(\mathbf{x}^{f,i}(t_k)) - \mathbf{H_k}(\mathbf{x}_k^b)][\mathbf{x}^{f,i}(t_k) - \mathbf{x}_k^b]^T$$

$$\mathbf{H_k B}_k^b \mathbf{H_k}^T = \frac{1}{r} \sum_{i=1}^{r+1} [\mathbf{H_k}(\mathbf{x}^{f,i}(t_k)) - \mathbf{H_k}(\mathbf{x}_k^b)][\mathbf{H_k}(\mathbf{x}^{f,i}(t_k)) - \mathbf{H_k}(\mathbf{x}_k^b)]^T.$$

Once these matrices are calculated, they are Schur multiplied with the matrices $\mathbf{H_k C}$ and $\mathbf{H_k C H_k}^T$, respectively.

For the domain localization methods, different analysis results can be obtained depending on the treatment of the observations.

If all the observations in the full domain are used for the analysis in each disjoint subdomain, the algorithm without localization is recovered. This follows from

$$\frac{1}{r}\sum_{i=1}^{r+1}\sum_{j=1}^{L}[\mathbf{H_k}(\mathbf{x}^{f,i}(t_k)) - \mathbf{H_k}(\mathbf{x}_k^b)][\mathbf{x}^{f,i}(t_k)\circ \mathbf{1}_{D_j} - \mathbf{x}_k^b \circ \mathbf{1}_{D_j}]^T =$$

$$\frac{1}{r}\sum_{i=1}^{r+1}[\mathbf{H_k}(\mathbf{x}^{f,i}(t_k)) - \mathbf{H_k}(\mathbf{x}_k^b)][\mathbf{x}^{f,i}(t_k) - \mathbf{x}_k^b]^T = \mathbf{H_k}\mathbf{B}_k^b \; .$$

If, on the other hand, we restrict observations to the local analysis subdomains the covariance matrix is given by (3).

Why is domain localization used?

- As for OI, one of the major advantages of using domain localization is computational. The updates on the smaller domains can be done independently, and therefore in parallel.
- In certain algorithms this is more natural way of localizing. Examples of such methods are the ensemble transform Kalman filter ETKF and the singular evolutive interpolated Kalman filter SEIK.

Why is domain localization used?

- In these algorithms, the forecast error covariance matrix is never explicitly calculated. Therefore, direct forecast localization as in Houtekamer and Mitchell (1998, 2001) is not immediately possible.

- In these methods an ensemble resampling is used that ensures that the ensemble statistics represent exactly the analysis state and error covariance matrix.

- Ways of including full rank, positive definite and isotropic matrix in domain localized algorithms were developed. Two methods will be presented Method SD+Loc and Method SD+ObsLoc introduced by Hunt et al. 2007.

Method SD+Loc

Let $\mathbf{1}_{Dmj}$ be a vector that has a value of 1 if the observation belongs to the domain Dm otherwise has a value of 0, and let $Dj \subseteq Dmj$.

$$\frac{1}{r}\sum_{i=1}^{r+1}\sum_{j=1}^{L}[\mathbf{H_k}\mathbf{x}^{f,i}(t_k)\circ\mathbf{1}_{Dmj}-\mathbf{H_k}\mathbf{x}_k^b\circ\mathbf{1}_{Dmj}][\mathbf{x}^{f,i}(t_k)\circ\mathbf{1}_{Dj}-\mathbf{x}_k^b\circ\mathbf{1}_{Dj}]^T$$

$$=\sum_{j=1}^{L}(\mathbf{1}_{Dmj}\mathbf{1}_{Dj}^T)\circ\mathbf{H_k}\mathbf{B}_k^b$$

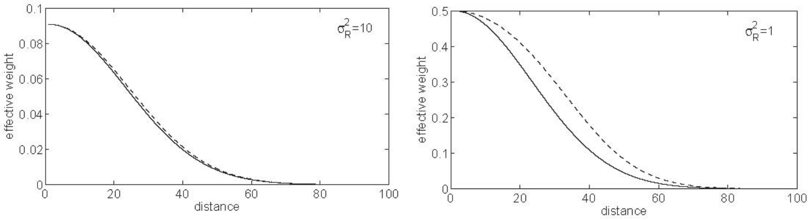
where matrix $\sum_{j=1}^{L}\mathbf{1}_{Dmj}\mathbf{1}_{Dj}^T$ has entries of zeros and ones since the domains Dj are disjoint.

Method (SD+Loc): An modification to this algorithm is to use for each subdomain $(\mathbf{1}_{Dmj}\mathbf{1}_{Dj}^T)\circ\mathbf{H_k}\mathbf{B}_k^b\circ\mathbf{H_k}\mathbf{C}$ and $\mathbf{1}_{Dmj}\mathbf{1}_{Dmj}^T\circ\mathbf{H_k}\mathbf{B}_k^b\mathbf{H_k}^T\circ\mathbf{H_k}\mathbf{C}\mathbf{H_k}^T$.
Resampling done on subdomains with local $\mathbf{B}_k^a$.

Observational error localization: Method (SD+ObLoc)

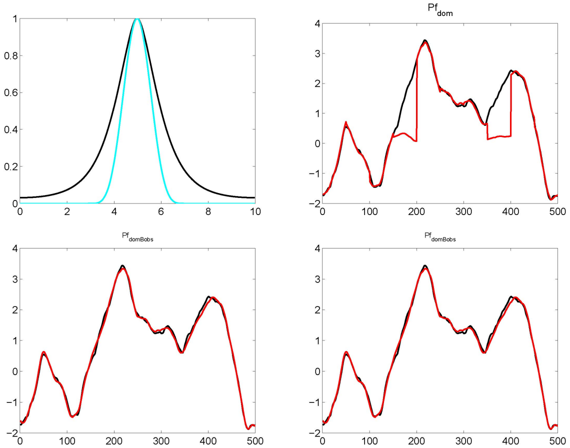
The observation localization method modifies the observational error covariance matrix **R**.

Let us consider a single observation example, in observation error localization method, the observation error σ^2_{obs} is modified to $\sigma^2_{obs}/weight_d$ where $weight_d$ can be calculated using any of the correlation functions.



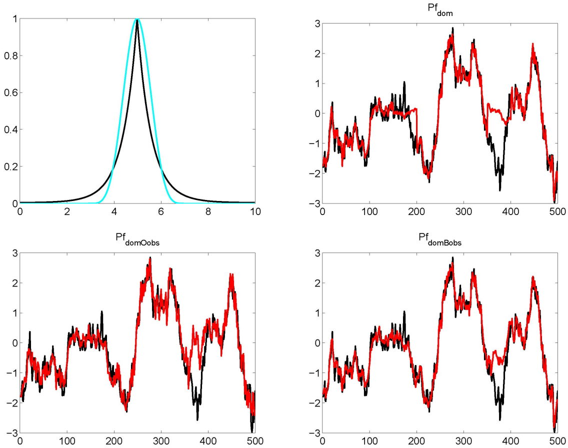
Resampling step includes modification of **R**.

Example domain localization cont.



Upper Left: True covariance (black) and approximate B covariance (blue). Upper Right: True state (black) and analysis (red) after one assimilation step with domain localized covariance. Lower Left: True state (black) and analysis (red) after one assimilation step with domain localized with overlapping observations. Lower Right: True state (black) and analysis (red) after one assimilation step with localized ensemble covariance with overlapping observations and B.

Example domain localization cont.



Upper Left: True covariance (black) and approximate B covariance (blue). Upper Right: True state (black) and analysis (red) after one assimilation step with domain localized covariance. Lower Left: True state (black) and analysis (red) after one assimilation step with domain localized with overlapping observations. Lower Right: True state (black) and analysis (red) after one assimilation step with localized ensemble covariance with overlapping observations and B.

Model Lorenz40

- $dX_i dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F$
- Lorenz40 model is governed by 40 coupled ordinary differential equations in domain with cyclic boundary conditions.
- The state vector dimension is 40.
- The observations are given as a vector of values contaminated by uncorrelated normally distributed random noise with standard deviation of 1.
- The observations are assimilated at every time step.
- After a spin-up period of 1000 time steps, assimilation is performed for another 50 000 time steps.
- A 10-member ensemble is used.

Domain Localization Methods

Method (SD+):Let $\mathbf{1}_{Dmj}$ be a vector that has a value of 1 if the observation belongs to the domain Dm otherwise has a value of 0, and let $Dj \subseteq Dmj$. where matrix $\sum_{j=1}^{L} \mathbf{1}_{Dmj} \mathbf{1}_{Dj}^{T}$. Use for each subdomain $(\mathbf{1}_{Dmj}\mathbf{1}_{Dj}^{T}) \circ \mathbf{H_k} \mathbf{B}_k^b$ and $\mathbf{1}_{Dmj}\mathbf{1}_{Dmj}^{T} \circ \mathbf{H_k} \mathbf{B}_k^b \mathbf{H_k}^T$.

Method (SD+Loc): Use for each subdomain $(\mathbf{1}_{Dmj}\mathbf{1}_{Dj}^{T}) \circ \mathbf{H_k} \mathbf{B}_k^b \circ \mathbf{H_k} \mathbf{C}$ and $\mathbf{1}_{Dmj}\mathbf{1}_{Dmj}^{T} \circ \mathbf{H_k} \mathbf{B}_k^b \mathbf{H_k}^T \circ \mathbf{H_k} \mathbf{C} \mathbf{H_k}^T$.

Method (SD+ObsLoc): Its implementation requires for each observation a weight that depends on the distance of the observation from the analysis location (Penduff et al. 2002; Hunt et al. 2007; Nerger and Gregg 2007).

Domain Localization Methods

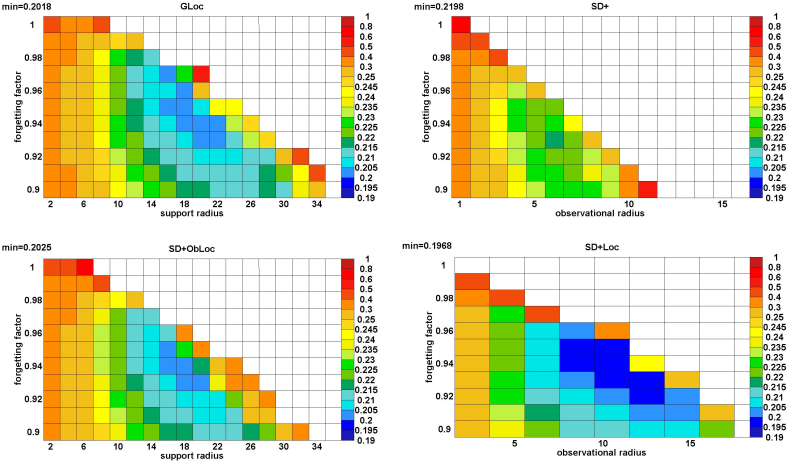
Method (SD+):Let $\mathbf{1}_{Dmj}$ be a vector that has a value of 1 if the observation belongs to the domain Dm otherwise has a value of 0, and let $Dj \subseteq Dmj$. where matrix $\sum_{j=1}^{L} \mathbf{1}_{Dmj}\mathbf{1}_{Dj}^{T}$. Use for each subdomain $(\mathbf{1}_{Dmj}\mathbf{1}_{Dj}^{T}) \circ \mathbf{H_k}\mathbf{B}_k^b$ and $\mathbf{1}_{Dmj}\mathbf{1}_{Dmj}^{T} \circ \mathbf{H_k}\mathbf{B}_k^b\mathbf{H_k}^T$.

Method (SD+Loc): Use for each subdomain $(\mathbf{1}_{Dmj}\mathbf{1}_{Dj}^{T}) \circ \mathbf{H_k}\mathbf{B}_k^b \circ \mathbf{H_k}\mathbf{C}$ and $\mathbf{1}_{Dmj}\mathbf{1}_{Dmj}^{T} \circ \mathbf{H_k}\mathbf{B}_k^b\mathbf{H_k}^T \circ \mathbf{H_k}\mathbf{C}\mathbf{H_k}^T$.

Method (SD+ObsLoc): Its implementation requires for each observation a weight that depends on the distance of the observation from the analysis location (Penduff et al. 2002; Hunt et al. 2007; Nerger and Gregg 2007).

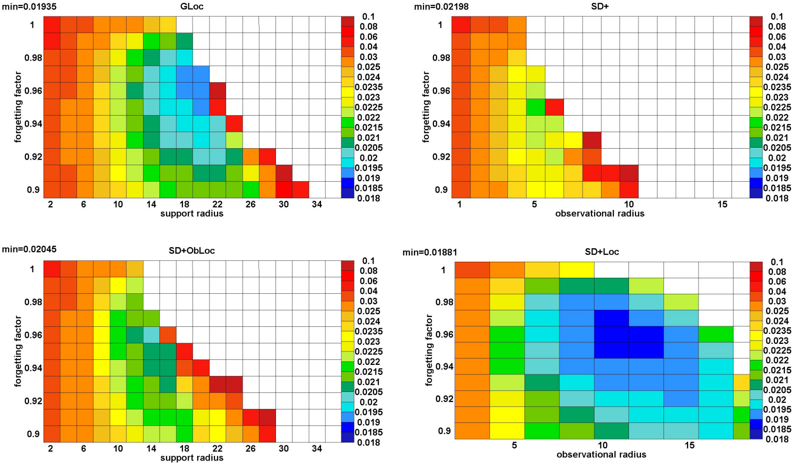
Method (GLocEn): An ensemble square root filter as in Whitaker and Hamill 2002 is applied. Covariance localization is applied.

L40 results: $\sigma_{obs} = 1$



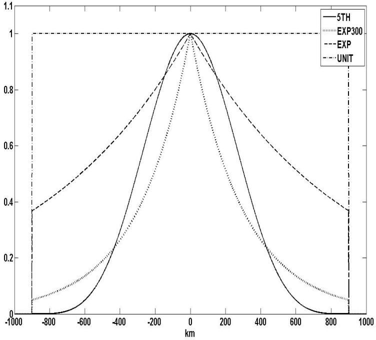
RMS error for different covariance localization techniques.

L40 results: $\sigma_{obs} = 0.1$



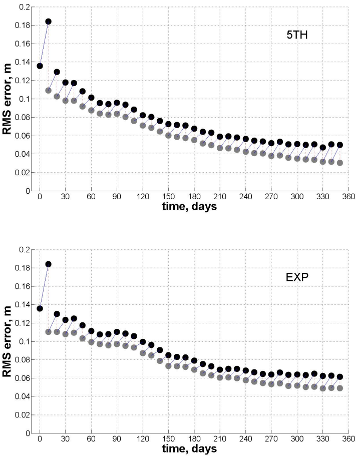
RMS error for different covariance localization techniques.

Goal: To study the filtering behavior when different correlation functions for the weighting of observations are applied.



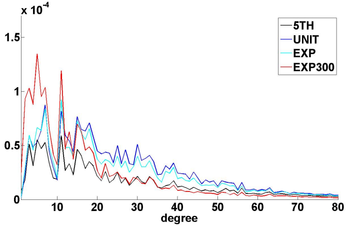
Correlation function used for localization.

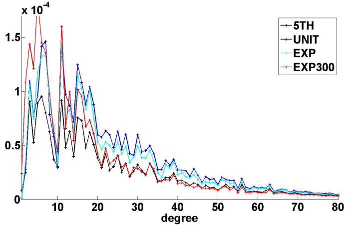
RMS errors



RMS error for different covariance localisation techniques.

Spectral properties of the errors

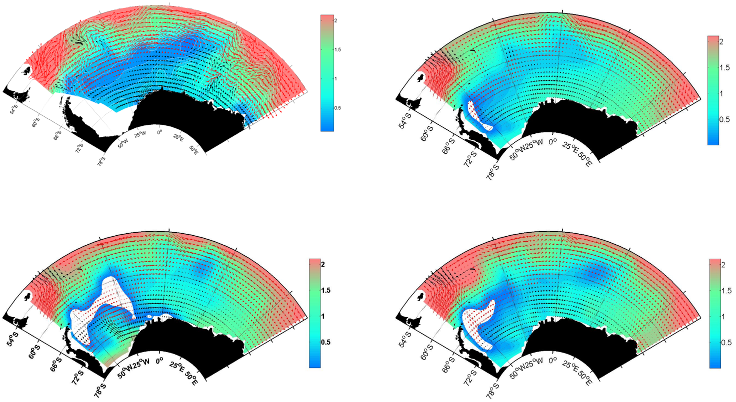




Logarithm of the spectral difference between analysis and the data (left) and forecast and the data (right) depending on spherical harmonic degree.

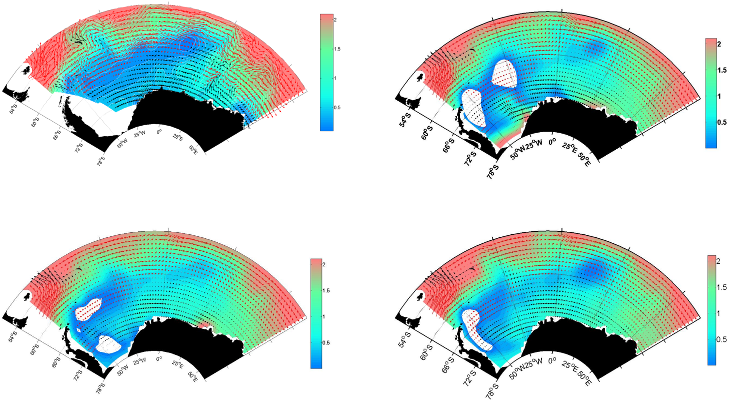
$$\epsilon_\ell^{oi} = \sum_m \left(T_{\ell m}^o - T_{\ell m}^i \right)^2$$

The Weddell gyre flow



In-situ temperature at 800 m depth. Composite from the ARGO data (1999 to 2010) (upper left). Model only (upper right). As result of assimilation of geodetic DOT filtered up to 241 km and UNIT (lower left) and 5TH lower right.

The Weddell gyre flow



In-situ temperature at 800 m depth. Composite from the ARGO data (1999 to 2010) (upper left). As result of assimilation of geodetic DOT filtered up to 241 km EXP (upper right) and EXP300 (lower left) and of geodetic DOT filtered up to 121 km lower right.

Domain localization conclusion

- The domain localization technique has been investigated here and compared to direct forecast error localization on simple example and L40 model.
- It was shown that domain localization is equivalent to direct forecast error localization with a Schur product matrix that has a block structure and is not isotropic.
- The rank of the matrix corresponding to the domain localization depends on the number of subdomains that are used in the assimilation. This matrix is positive semidefinite.
- Inclusion of positive definate matrix either through method SD+Loc or SD+ObsLoc is beneficial for domain localization methods.

Localization and balance

Assume we have two variables h and v defined at the model grid points, i.e. $\mathbf{h}$ and $\mathbf{v}$:

$$\mathbf{P_k^b} \equiv \begin{bmatrix} cov(\mathbf{h},\mathbf{h}) & cov(\mathbf{h},\mathbf{v}) \\ cov(\mathbf{v},\mathbf{h}) & cov(\mathbf{v},\mathbf{v}) \end{bmatrix}$$

Let us assume that we want to apply direct forecast error localization with diagonal matrix then

$$\mathbf{P_k^b} \circ \mathbf{I} \equiv \begin{bmatrix} cov(h_1,h_1) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots \\ 0 & 0 & cov(h_n,h_n) & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0\cdots 0 & cov(v_n,v_n) \end{bmatrix}$$

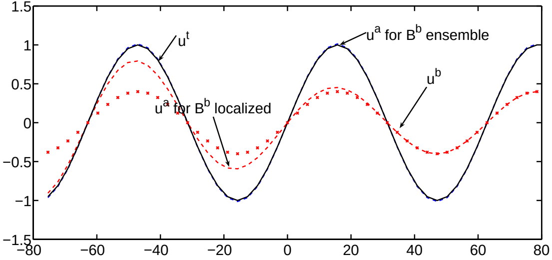
Example 2: How good are our unobserved variables?

$$\frac{\partial h}{\partial t} + c\frac{\partial h}{\partial x} = 0$$
$$u(x,t) = \frac{\partial h}{\partial x}$$
$$h(x,0) = sin(x)$$

Solution is given by $h^t(x,t) = sin(x - ct), u^t(x,t) = cos(x - ct)$.

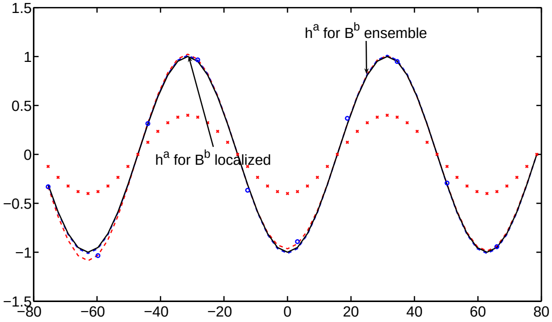
- We observe only h as in Example 1.
- Our $\mathbf{w}_k = [\begin{matrix} \mathbf{h} \\ \mathbf{u} \end{matrix}]$
- Field u should be corrected through the background error covariance!

Example 2: How good are our unobserved variables?



- Diagonal covariance matrix does not correct *u* since *u* is not observed.
- Field *u* was corrected through the ensemble background error covariance using the cross correlations between variables u and h as given by model dynamics!
- Once we localize the covariances we will loose the cross correlations specified by the model.

Example 2: How good are our unobserved variables?



- In the field which is observed we can fit the data better by localizing.
- RMS error calculations against data that is assimilated. In this case: RMS $= 0.0083$ for localized $\mathbf{B}_k^b$, RMS $= 0.0132$ for 5 ensemble members.

Localization and Balance

- By applying localization we destroy correlation given by numerical model between two fields.
- In Greybush et al. MWR 2011 methods (SD+ObsLoc) and direct forecast error localization with $\exp \frac{-d(i,j)^2}{2L^2}$ were compared in order to investigate the effects on geostrophic balance
- It was shown that the observational error localization preserves better balance then the direct covariance error localization.

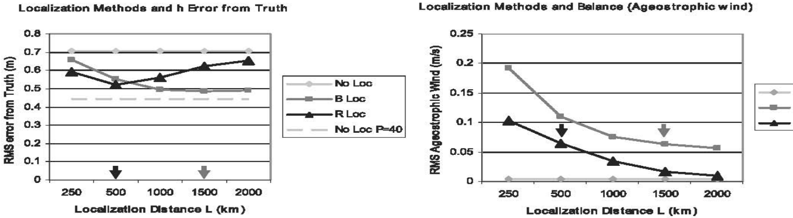


FIG. 3. (left) RMS error of the analysis from the truth for height (m) and (right) RMS ageostrophic wind (m using no localization, B localization, and R localization for five ensemble members and a variety of locali distances L. For comparison, an analysis with no localization and 40 ensemble members is also plotted. Arrows optimum values of L.

Other ways of localization

- Kepert (2009) analysed the impact of localisation on an analytical covariance model that contained certain exact balances and showed that the Schur-product localisation had severe impact on those balances.
- it was shown that applying localization to geopotential, streamfunction and velocity potential produces better balanced forecast fields than localization applied on geopotential and zonal and meridional velocities.
- This was shown using identical-twin experiments with a global spectral shallow-water model and no separate initialization step.

Conclusion

- Localization is necessary for application of ensemble Kalman filter algorithms for large scale probelms.
- Several localization techniques are in use.
- Localization removes spurious long range correlations, but it also introduces an ad hoc procedure that requires tuning in ensemble Kalman filter methods.
- Localization is topic of active research especially concerning the effect of localization on balance.
- Proper ways of performing multivariate localization are still not fully understood.
- Proper localization scales depend on the properties of dynamical system and observations.