Local ensemble assimilation scheme with global constraints and conservation.

Alexander Barth, Yajing Yan, Martin Canter, Aida Alvera-Azcárate and Jean-Marie Beckers

Sangoma, ULg, April 2014

GHER
GeoHydrodynamics and Environment Research

SANGOMA

Need for covariance localization

- In ensemble assimilation schemes, the model error covariance $\mathbf{P}$ is represented by an **ensemble of model states** $\mathbf{x}^{(k)}, k = 1, \ldots, N$ ($\langle\rangle$ is the ensemble average).

$$\mathbf{P} = \langle(\mathbf{x} - \langle\mathbf{x}\rangle)(\mathbf{x} - \langle\mathbf{x}\rangle)^T\rangle = \mathbf{X}\mathbf{X}^T$$

- As N increases **convergence is relatively slow** $(N^{-1/2}) \rightarrow$ sampling error.
- This sampling error leads to **unrealistic long-range correlations**.
- Covariance localization suppresses these long-range correlations based on the horizontal distance based on a **specified length-scale**.

Issues with covariance localization

- **Ad-hoc approach** introduced by analogy to optimal interpolation (where it is done for computational reasons).
- It can **filter out realistic long-range correlations** (for example introduced through error in atmospheric fields).
- Spurious correlations between **weakly related model variables close-by** might not be filtered (in particular in coupled models).
- Which length-scale to use when **multiple processes with different scales** are present?
- Generally result in non-conservative assimilation scheme (even if the underlying model is conservative)

Localization

- One can distinguish different localization approaches (*e.g* Nerger et al., 2012):
  - **covariance localization**: every single observation point is assimilated sequentially and the correction are filtered by a localization function. (less suited for parallel processing and the domain localization).
  - **domain localization**: the state vector is decomposed into sub-domains (e.g. single grid points or vertical columns) where the assimilation is performed independently. Such algorithm are easily applied to parallel computers.
- The conservation requires a coupling of a model grid points which is filtered-out by the localization.
- Similar difficulty: non-local observation operator (*e.g* Campbell et al., 2010).

Local ensemble assimilation scheme with global constraints and conservation

- Global assimilation scheme have no problem in respecting linear conservation.
- Non-linear constraints can sometimes be transformed into linear constraints by a careful transformation model variable. Example:
  - For sea ice concentration c_i and sea ice height h_i, then :

$$\int_{\Omega} c_i h_i \, dx = \text{const} \tag{1}$$

    in the absence of ice melting and ice formation.
  - This conservation property is non-linear if a state vector including c_i and h_i .
  - ... but becomes linear if the state vector includes $c_i h_i$ and c_i (or h_i).

Method

- We propose a local assimilation scheme which is local and can satisfy global conservation properties and non-local observation operators.
- In essence:
  - Based on **covariance localization**
  - Localize ensemble covariance matrix (by using an element-wise matrix product).
  - Modify this localized covariance matrix to so that the uncertainty of the total amount of the conserved quantity is zero.
  - Algorithm should work on an **ensemble as input** (model forecast) and produce an **ensemble as outputs** (analysis).
  - Avoid the formation of huge matrices.
  - One should recover the original Kalman filter analysis if the covariance does not have spurious long-range correlation.
  - Parallel algorithm.

Variants of the scheme (overview)

- We have now a way to compute the analysis increment by applying the Kalman gain the difference between observation the model results

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}^f)$$

- We know how to update the mean of the ensemble, but how to update the whole ensemble?
- Different variants:
  - Using perturbed observations and apply the analysis update for every member individually
  - Avoid using perturbed observation, write-out formally $\mathbf{P}^a$ (the analysis error covariance) and project it on a suitable subspace (suffix $\mathbf{P}_c$).
  - Modify the previous approach such that the analysis ensemble tends to the forecast ensemble if the observation error covariance gets very large (suffix $\mathbf{S}\mathbf{S}^T$).

Variants of the scheme (the nasty details)

Ensemble covariance $\mathbf{P}$ $(n \times n)$ of a ensemble with N members can be written as:

$$\mathbf{P} = \mathbf{S}\mathbf{S}^T \tag{2}$$

where $\mathbf{S}$ is of size $n \times N - 1$ (scaled difference between ensemble member and ensemble mean).
Spurious long-range correlations are filtered by a function ρ with compact support:

$$\mathbf{P}' = \rho \circ \mathbf{P} \tag{3}$$

We want that the analysis increment satisfies an constrain (does not create heat or salt for example):

$$\mathbf{h}^T(\mathbf{x}^a - \mathbf{x}^f) = 0 \tag{4}$$

Normalize $\mathbf{h}^T$ so that $\mathbf{h}^T\mathbf{h} = 1$. Modify covariance for strong constraint:

$$\mathbf{P}_c = (\mathbf{I} - \mathbf{h}\mathbf{h}^T)\mathbf{P}'(\mathbf{I} - \mathbf{h}\mathbf{h}^T) \tag{5}$$

Thus the uncertainty in the conserved values $\mathbf{h}^T\mathbf{P}_c\mathbf{h}$ is zero.

Kalman gain based on modified covariance:

$$\mathbf{K} = \mathbf{P}_c\mathbf{H}^T \left(\mathbf{H}\mathbf{P}_c\mathbf{H}^T + \mathbf{R}\right)^{-1} \tag{6}$$

The product of $\mathbf{K}$ times a vector $\rightarrow$ **conjugate gradient algorithm** (forming explicit matrices) to solve:

$$\left(\mathbf{H}\mathbf{P}_c\mathbf{H}^T + \mathbf{R}\right)\mathbf{y} = \mathbf{y}^o - \mathbf{H}\mathbf{x}^f$$

Possible preconditioning: global scheme
The analysis (ensemble mean):

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}^f) = \mathbf{P}_c\mathbf{H}^T\mathbf{y} \tag{7}$$

Stochastic analysis scheme

Ensemble $\mathbf{x}^{a(k)}$ using perturbed observations:

$$\mathbf{x}^{a(k)} = \mathbf{x}^{f(k)} + \mathbf{K}(\mathbf{y}^{o(k)} - \mathbf{H}\mathbf{x}^{f(k)}) \tag{8}$$

This can already be used in with an realistic ocean model.

Deterministic analysis scheme

Can we get a formulation without perturbed observations?

$$\mathbf{x}^{a(k)} = (\mathbf{I} - \mathbf{KH})\mathbf{x}^{f\,(k)} + \mathbf{K}\mathbf{y}^{o(k)} \tag{9}$$

The covariance of this matrix is:

$$\mathbf{P}^a \;\; = \;\; (\mathbf{I} - \mathbf{KH})\mathbf{P}^f(\mathbf{I} - \mathbf{KH})^T + \mathbf{KRK}^T \tag{10}$$

If we take $\mathbf{P}^f = \mathbf{SS}^T$, thus the unfiltered covariance, we get:

$$\mathbf{P}^a \;\; = \;\; \left[\; (\mathbf{I} - \mathbf{KH})\mathbf{S} \quad \mathbf{KR}^{1/2} \;\right] [\text{idem}]^T \tag{11}$$

In general the rank of $\mathbf{P}^a$ increases.
We can try to project the term due to uncertain observations on the error space.
Onto which basis can be project $\mathbf{P}^a$? A good choice seem to be

$$\mathbf{S}' = (\mathbf{I} - \mathbf{KH})\mathbf{S} \tag{12}$$

as $\mathbf{h}^T\mathbf{S}' = 0$ if $\mathbf{h}^T\mathbf{S} = 0$.

eProjection operator:

$$\mathbf{S}'\left(\mathbf{S}'^T\mathbf{S}'\right)^{-1}\mathbf{S}'^T \tag{13}$$

Project the covariance matrix $\mathbf{P}^a$ onto the subspace defined by $\mathbf{S}$

$$\mathbf{P}^a = \mathbf{S}'\left(\mathbf{S}'^T\mathbf{S}'\right)^{-1}\mathbf{P}^a_{S'}\left(\mathbf{S}'^T\mathbf{S}'\right)^{-1}\mathbf{S}'^T + \text{contrib. in perp. space to be neglected} \tag{14}$$

$$\mathbf{P}^a_{S'} = \mathbf{S}'^T\mathbf{P}^a\mathbf{S}' \;\; = \;\; (\mathbf{S}'^T - \mathbf{S}'^T\mathbf{K}\mathbf{H})\mathbf{P}_c(\mathbf{S}'^T - \mathbf{S}'^T\mathbf{K}\mathbf{H})^T + \mathbf{S}'^T\mathbf{K}\mathbf{R}\mathbf{K}^T\mathbf{S}' \tag{15}$$

We need to compute the product $\mathbf{K}^T\mathbf{S}'$ efficiently:

$$\mathbf{K}^T\mathbf{S}' = \left(\mathbf{H}\mathbf{P}_c\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{P}_c\mathbf{S}' \tag{16}$$

Finally:

$$\mathbf{S}^a = \mathbf{S}'\left(\mathbf{S}'^T\mathbf{S}'\right)^{-1}(\mathbf{P}^a_{S'})^{1/2} \tag{17}$$

where $(\mathbf{P}^a_{S'})^{1/2}$ is the principal square root of $\mathbf{P}^a_{S'}$ (which is unique).
This requires to solve $2(N-1)$ systems of the size $m \times m$ for the error modes and 1 system for the ensemble mean (sounds reasonable). The N systems are independent and can be distributed on a parallel machine.

Problem: unwanted rotation

Even if $\mathbf{R}$ is very large (goes to infinity), the analysis ensemble is different from the forecast ensemble (mean and covariance are however unchanged) because

$$\mathbf{P}^a_{S'} \rightarrow \mathbf{S}^T \mathbf{P}_c \mathbf{S} \tag{18}$$

and the principal square root of this matrix introduces an unwanted rotation. However, we want that this tends to the following:

$$\mathbf{P}^a_{S'} \rightarrow \mathbf{S}^T \mathbf{S} \mathbf{S}^T \mathbf{S} \tag{19}$$

so that the principal square root of $\mathbf{P}^a_{S'}$ tends to $\mathbf{S}^T\mathbf{S}$ (because it is unique) and $\mathbf{S}^a$ will tends to $\mathbf{S}$. This can be achieved by mortifying (15), so that this equations reads:

$$\begin{aligned}
\mathbf{S}'^T \mathbf{P}^a \mathbf{S}' &= (\mathbf{S}'^T - \mathbf{S}'^T \mathbf{K}\mathbf{H}) \mathbf{S}\mathbf{S}^T (\mathbf{S}'^T - \mathbf{S}'^T \mathbf{K}\mathbf{H})^T + \mathbf{S}'^T \mathbf{K}\mathbf{R}\mathbf{K}^T \mathbf{S}' && (20)\\
&= \mathbf{S}'^T (\mathbf{I} - \mathbf{K}\mathbf{H}) \mathbf{S}\mathbf{S}^T (\mathbf{I} - \mathbf{K}\mathbf{H})^T \mathbf{S}' + \mathbf{S}'^T \mathbf{K}\mathbf{R}\mathbf{K}^T \mathbf{S}' && (21)\\
&= \mathbf{S}'^T \mathbf{S}' \mathbf{S}'^T \mathbf{S}' + \mathbf{S}'^T \mathbf{K}\mathbf{R}\mathbf{K}^T \mathbf{S}' && (22)
\end{aligned}$$

As $\mathbf{R} \rightarrow \infty$ we have:

$$\mathbf{S}^a \rightarrow \mathbf{S}' \left(\mathbf{S}'^T \mathbf{S}'\right)^{-1} \mathbf{S}^T \mathbf{S} = \mathbf{S} \tag{23}$$

Test case

Kuramoto-Sivashinsky equation

- Equations:

$$\partial_t v = -\partial_x^2 v - \partial_x^4 v - v\partial_x v \tag{24}$$

- Periodic domain: $L = 32\pi$ with 128 grid points
- Time-step: $\Delta t = 1/4$
- ETDRK4 (Exponential Time Differencing fourth-order Runge-Kutta)
- Conservation:

$$\frac{d}{dt}\int_0^L v \, dx = 0 \tag{25}$$

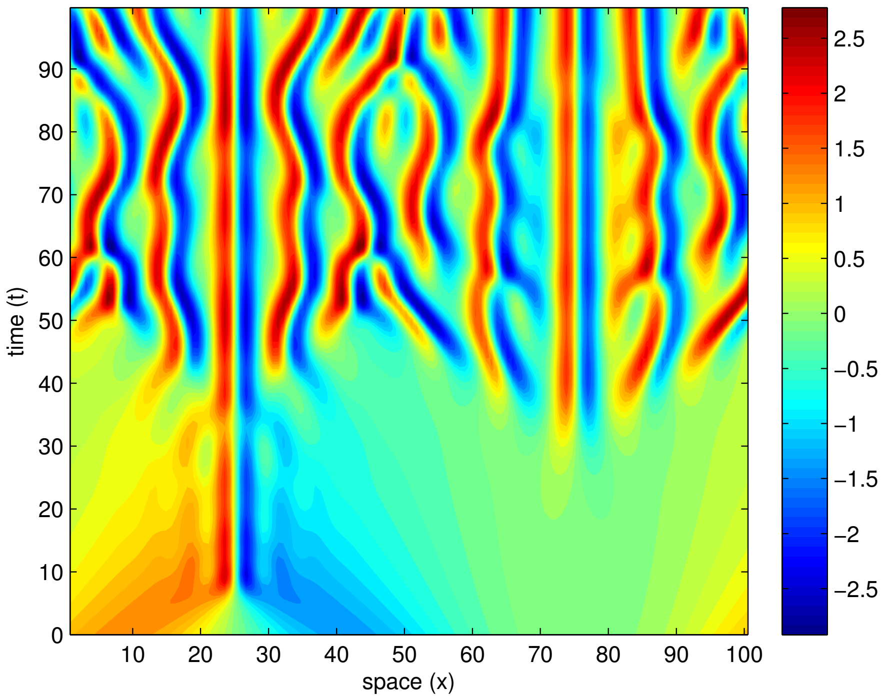


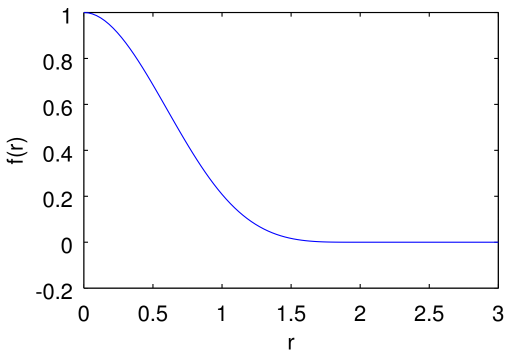
Figure 1: Solution of the KS equation (without assimilation)

Localization function

- Localization function from Gaspari and Cohn (1999): compactly supported 5th-order piecewise rational function:

$$f(r) = \begin{cases} -\frac{1}{4}r^5 + \frac{1}{2}r^4 + \frac{5}{8}r^3 - \frac{5}{3}r^2 + 1, & \text{if } r \leq 1 \\ \frac{1}{12}r^5 - \frac{1}{2}r^4 + \frac{5}{8}r^3 + \frac{5}{3}r^2 - 5r + 4 - \frac{2}{3r}, & \text{if } r \leq 2 \\ 0. & \text{if } r > 2 \end{cases}$$

- where r is the distance scaled by a given length-scale L.



Assimilation test cases:

- CL: standard covariance localization: observations are assimilated sequentially and the correction is multiplied with a localization function.
- CL-adj: The same as CL, but after the analysis the budget is corrected with an adjustment step.
- LEnKF-pert: Localized EnKF using perturbed observations without conservation constraint.
- CLEnKF-pert: Localized EnKF using perturbed observations with conservation constraint.
- LEnKF-Pc: Localized EnKF variant "$\mathbf{P}_c$" without conservation constraint.
- CLEnKF-Pc: Localized EnKF variant "$\mathbf{P}_c$" with conservation constraint.
- LEnKF-$\mathbf{SS}^T$: Localized EnKF variant "$\mathbf{SS}^T$" without conservation constraint.
- CLEnKF-$\mathbf{SS}^T$: Localized EnKF variant "$\mathbf{SS}^T$" with conservation constraint.

Assimilation setup

- Classical twin experiment.
- Every 8th grid point is observed (with an error variance of 0.1) at every 10 model time steps.
- The model with assimilation for 1000 time steps.
- The experiment is repeated 800 times and RMS errors relative to the true solution are averaged.
- Using different localization length-scale and inflation factors.

Results

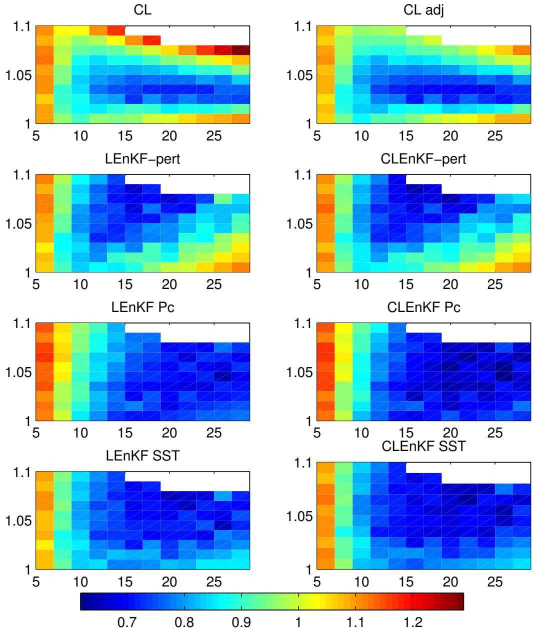


Figure 2: RMS error between the model run with assimilation and true solution for different schemes and localization length-scale (x-axis) and inflation factors (y-axis)

Optimal parameters

	RMS	L	inflation
CL	0.71733	21	1.03
CL adj	0.69039	23	1.03
LEnKF-pert	0.66798	19	1.07
CLEnKF-pert	0.64005	15	1.08
LEnKF P_c	0.63993	25	1.04
CLEnKF P_c	**0.61864**	25	1.05
LEnKF SS^T	0.65308	19	1.07
CLEnKF SS^T	0.63871	21	1.06

Table 1: Lowest RMS for different assimilation schemes and corresponding parameters

Minimal model for sea ice and salinity with conservation

We look for minimal model for sea ice and salinity where the amount of "freshwater" (or salt) is conserved. In this system, the integral of a function f (of the model parameter) over a closed domain remains constant over time:

$$\frac{d}{dt}\int_{\Omega} f dx = 0 \tag{26}$$

Based on Bert's ideas, the velocity (v) for salinity (S) is provided using the Kuramoto-Sivashinsky equation:

$$\partial_t v = -\partial_x^2 v - \partial_x^4 v - v\partial_x v - g\partial_x h \tag{27}$$

The flow v is not "incompressible" as it varies with x. Thus we use also the variable h, representing the height of the mixed layer:

$$\begin{aligned} \partial_t(hS) + \partial_x(vhS) &= \kappa\partial_x^2(hS) + \mu\mathcal{F} &(28)\\ \partial_t c + \partial_x((v_c + v)c) &= \mathcal{F} &(29) \end{aligned}$$

where v_c is the velocity of the sea ice (constant) and h is governed by:

$$\partial_t h + \partial_x (hv) = 0 \tag{30}$$

For a periodic domain Ω, salinity fluxes and ice fluxes cancel after integration over the whole domain and one obtains:

$$\frac{d}{dt}\int_\Omega (hS - \mu c)\, dx = 0 \tag{31}$$

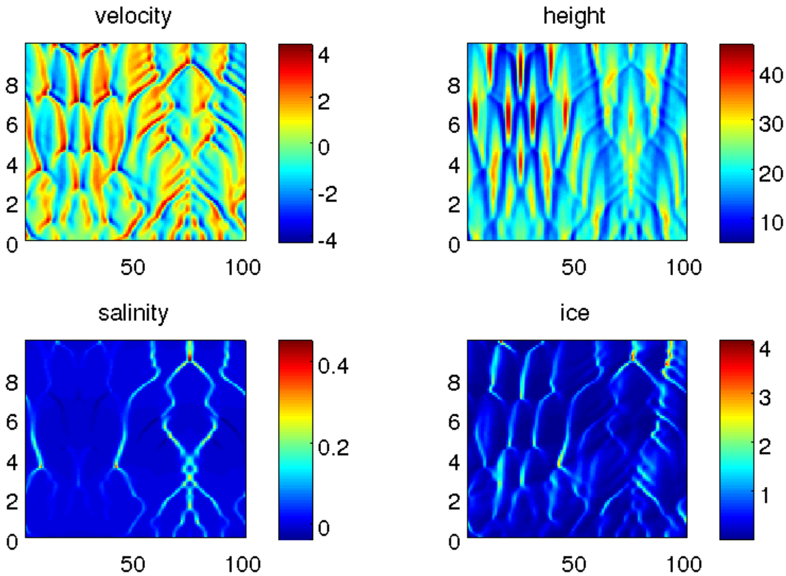


Figure 3: Sample model results

Results

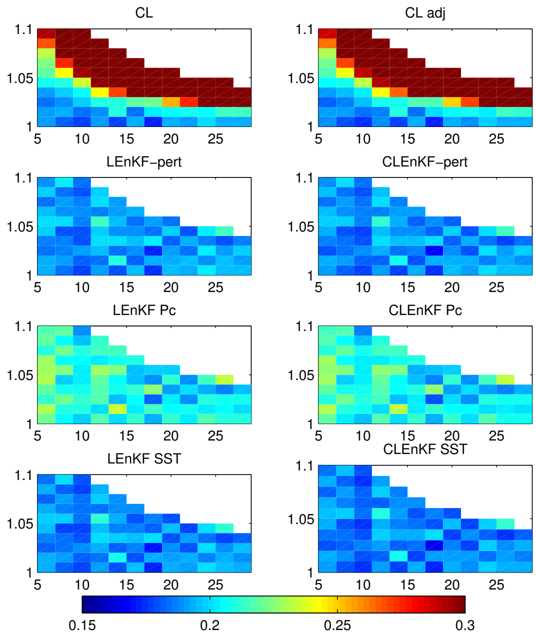


Figure 4: RMS error between the model run with assimilation and true solution for different schemes and localization length-scale (x-axis) and inflation factors (y-axis)

	RMS	L	inflation
CL	0.17423	17	1.00
CL adj	0.17374	17	1.00
LEnKF-pert	0.17108	17	1.02
CLEnKF-pert	0.16803	17	1.02
LEnKF $\mathbf{P}_c$	0.17895	17	1.02
CLEnKF $\mathbf{P}_c$	0.17670	17	1.02
LEnKF $\mathbf{SS}^T$	0.16846	17	1.02
CLEnKF $\mathbf{SS}^T$	**0.16545**	17	1.02

Table 2: Lowest RMS for different assimilation schemes and corresponding parameters (for salinity)

- Conservation gives a slight improvement
- Method CLEnKF SS^T is better than standard CL (covariance localization)

Analysis with sub-optimal gain

- The analysis error covariance for any gain matrix $\mathbf{K}'$ can be written as:

$$
\begin{aligned}
\mathbf{P}^a &= \mathbf{P}^f - \underbrace{\mathbf{KHP}^f}_{\text{error reduction}} \\
&\quad + \underbrace{(\mathbf{K}-\mathbf{K}')(\mathbf{HPH}^T+\mathbf{R})(\mathbf{K}-\mathbf{K}')^T}_{\text{error increase}}
\end{aligned}
$$

- The matrix $\mathbf{K}$ is the optimal Kalman gain and the matrix $\mathbf{K}'$ is the estimated Kalman gain.
- The 2nd term is always positive defined and represents an error reduction.

$$\mathbf{P}_r = \mathbf{KHP}^f$$

- The 3rd term is always negative defined and represents an error increase (unless the $\mathbf{K}'$ is the optimal Kalman gain).

$$\mathbf{P}_i = (\mathbf{K}-\mathbf{K}')(\mathbf{HPH}^T+\mathbf{R})(\mathbf{K}-\mathbf{K}')^T \tag{32}$$

- In the following we will explain how the matrices $\mathbf{K}$, $\mathbf{K}'$, $\mathbf{P}_r$ and $\mathbf{P}_i$ are estimated.

Approach

The method similar to bootstrapping in statistics:

1. An ensemble of e.g. $N = 100$ members is created by perturbed initial conditions, boundary conditions,...

2. The ensemble is **split in 2 sub-ensembles** of 50 members (at random)

$$
\begin{aligned}
{\mathbf{X}'_{i,j}}^{(m)} &= \mathbf{X}_{i,p(j)^{(m)}} & 1 \leq j \leq N/2 \\
{\mathbf{X}''_{i,j-N/2}}^{(m)} &= \mathbf{X}_{i,p(j)^{(m)}} & N/2 < j \leq N
\end{aligned}
$$

   where $p(j)^{(m)}$ is the m-th realization of a permutation vector.

3. The **observations are assimilated with the 2 sub-ensembles** of 50 members by calculating the Kalman gain ($\mathbf{K}'^{(m)}$ and $\mathbf{K}''^{(m)}$).

$$
\begin{aligned}
\Delta\mathbf{x}'^{(m)} &= \mathbf{K}'^{(m)} \left(\mathbf{y}^o - \mathbf{H}{\mathbf{x}'^f}^{(m)} \right) \\
\Delta\mathbf{x}''^{(m)} &= \mathbf{K}''^{(m)} \left(\mathbf{y}^o - \mathbf{H}{\mathbf{x}''^f}^{(m)} \right)
\end{aligned}
$$

We do not have access to the optimal Kalman gain (obtained if $N \to \infty$), but the Kalman gain error $\mathbf{K} - \mathbf{K}'$ of equation (32) can be approximated in average by $\mathbf{K}'^{(m)} - \mathbf{K}''^{(m)}$.

4. The difference between the two analyses increment is computed.

$$\delta \mathbf{x}^{(m)} = \Delta \mathbf{x}'^{(m)} - \Delta \mathbf{x}''^{(m)}$$

   The covariance of this difference is thus $\mathbf{P}_i$.

5. Step 2. to 4. are **repeated** with other 2 sub-ensembles.

6. We check where the analyses are consistent in all tests (variance of the increment).

$$\mathbf{z} = \frac{1}{m_{\mathsf{max}}} \sum_{m=1}^{m_{\mathsf{max}}} \delta \mathbf{x}^{(m)^2} \sim \mathrm{diag}(\mathbf{P}_i)$$

7. The localization function is built based on the **consistency of the analysis** using $\mathbf{z}$ **and expected error reduction**,

$$\mathbf{f}_i = \exp\left(\frac{\mathbf{z}_i^2}{\mathbf{P}_{rii}^{\,2}} \frac{\mathbf{P}_{ii}^f}{\mathbf{P}_{rii}} \right)$$

All terms of this equation are **filtered spatially** to ensure smooth variations in space.

8. The analysis increment is multiplied element-by-element with this localization function

Notes

- The approach is pessimistic since the statistical fluctuations are based on an ensemble of half the size.
- The approach is optimistic since all realizations of the increment fluctuations use the same ensemble.

Idealized tests

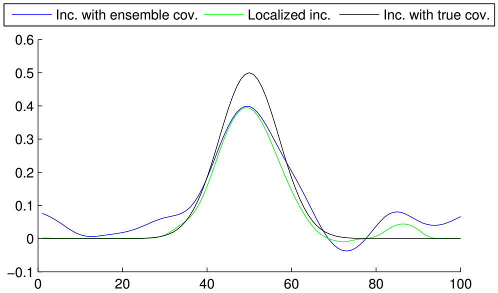
- 1 dimensional domain with an error covariance of $\mathbf{P}^f$ given by:

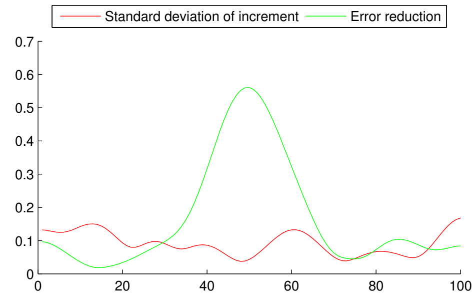
$$\mathbf{P}^f_{ij} = \exp(-(i-j)^2/L^2)$$

  with $L = 10$

- Ensemble of 100 members are drawn from a Gaussian distribution with this covariance

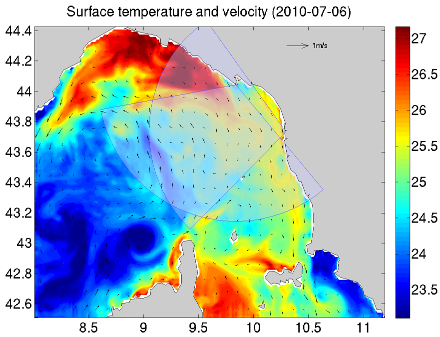
- Observations in the middle of the domain is assimilated ($y^o = R = 1$) using the ensemble error covariance and analytical error covariance





Model

- ROMS nested (off-line) in the Mediterranean Ocean Forecasting System
- 1/60 degree resolution and 32 vertical levels
- Atmospheric forcings come from the limited-area model COSMO (hourly at 2.8 km resolution)
- Currents: Western & Eastern Corsican Current, Northern Current, inertial oscillations, mesoscale currents



Model error covariance

- Estimated by ensemble simulation (with 100 members) where the uncertain aspects of the model are perturbed
- Perturbed zonal and meridional wind forcing
- Perturbed boundary conditions (elevation, velocity, temperature and salinity)
- Perturbed momentum equation (ε)

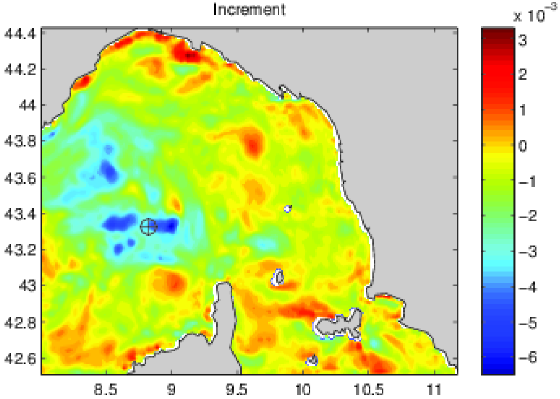
$$\begin{aligned}\frac{d\mathbf{u}}{dt} + \mathbf{\Omega} \wedge \mathbf{u} &= -\frac{1}{\rho_0}\nabla_h p + \frac{1}{\rho_0}\nabla \cdot \mathbf{F}^{\mathbf{u}} \\ &\quad + \nabla_h \wedge \ \varepsilon\, \mathbf{e}_z\end{aligned}$$

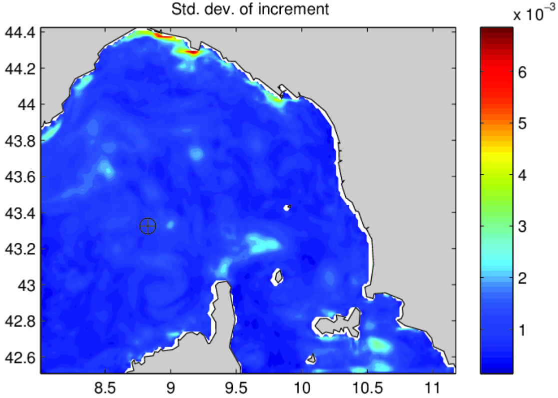
  - where $\nabla_h = \mathbf{e}_x \frac{\partial}{\partial x} + \mathbf{e}_y \frac{\partial}{\partial y}$
  - does not create horizontal convergence or divergence (linked to barotropic waves)
  - can create mesoscale flow structures (absent or misplaced)

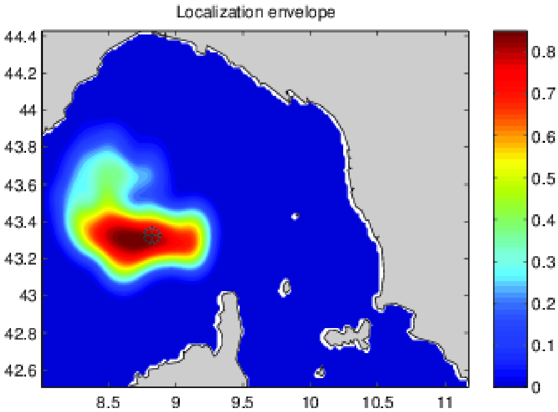
Test in realistic domain

- We assume an observation of the u-velocity (at the location of the marker) of 0.1 m/s
- Observational error covariance is $\mathbf{R} = (0.1 \text{ m/s})^2$

Observations in the interior of the model domain

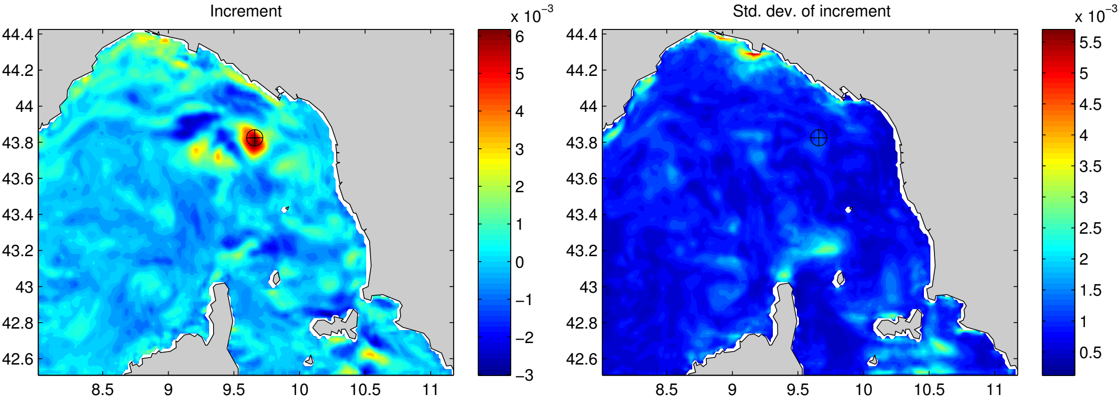


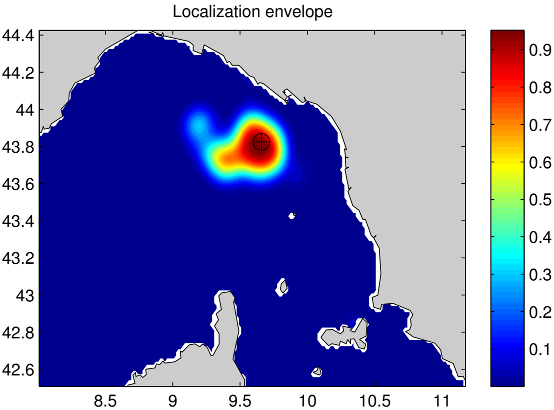




- Observation located at 8.8250 W and 43.3250 N
- Significant spurious long-range correlation, especially with parts of the domain having a large error variance
- The localization function naturally selects corrections near the location of the observations.

Observations in a highly variable area





- Covariance is more localized at this location, but spurious corrections are still present

Conclusions

- New assimilation scheme which is formulated globally (i.e. for the whole state vector)
    - where spurious long-range correlations can be filtered out
    - global conservation properties can be enforced and non-local observation operators can be used
- Test with Kuramoto-Sivashinsky show benefit of this approach compared to the traditional covariance localization scheme where observations are assimilated sequentially (even with an ad-hoc step enforcing conservation)
- But better results are obtained with the new assimilation variants (except Pc variant for KS-sea-ice)
- Even without conservation the new the new schemes produce a lower RMS errors.
- A method similar to bootstrapping can be used to estimate the uncertainty of the analysis increment
- The error increase can be compared to the expected error reduction to formulate a field that can be used as a localization envelope.

References

W.F. Campbell, C. H. Bishop, and D. Hodyss. Vertical covariance localization for satellite radiances in ensemble Kalman filters. *Monthly Weather Review*, 138:282–290, 2010.

Gregory Gaspari and Stephen E. Cohn. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125 (554):723–757, 1999. ISSN 1477-870X. doi: 10.1002/qj.49712555417.

T. M Hamill, J. S. Whitaker, and C. Snyder. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, 129:2776–2790, 2001.

I. Hoteit, D.-T. Pham, and J. Blum. A simplified reduced order Kalman filtering and application to altimetric data assimilation in Tropical Pacific. *Journal of Marine Systems*, 36:101–127, 2002.

P. L. Houtekamer and H. L. Mitchell. A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Reviev*, 129:123–137, 2001.

L. Nerger, T. Janjić, J. Schröter, and W. Hiller. A regulated localization scheme for ensemble-based kalman filters. *Quarterly Journal of the Royal Meteorological Society*, 138(664):802–812, 2012. ISSN 1477-870X. doi: 10.1002/qj.945.