Localization for ensemble DA: objective diagnostic and efficient application

Benjamin Menetrier, IRIT, Toulouse, France
(benjamin.menetrier@irit.fr)

Workshop on Climate Prediction in the Atlantic-Arctic sector - 7th June 2019

Introduction

- Background error covariance modelling is a key aspect of variational DA systems.
- Ensembles of perturbed backgrounds can be used to sample the background error covariance.
- For computational cost reasons, the ensemble size is limited.
- Localization is required to remove the sampling noise.
- In variational DA systems, localization is applied in model space (not in observation space).
- Optimal localization can be diagnosed from the ensemble.
- In practice, the localization matrix itself is not required, only its smoothing effect when applied on a state vector.

Outline

Outline

Sample covariance

An ensemble of N forecasts $\{\mathbf{x}_p^b\}$ is used to estimate the sample covariance matrix $\widetilde{\mathbf{B}}$:

$$\widetilde{\mathbf{B}} = \frac{1}{N-1} \sum_{p=1}^{N} \delta\mathbf{x}_p^b \delta\mathbf{x}_p^{b\mathrm{T}} \tag{1}$$

where $\delta\mathbf{x}_p^b$ is the p$^{\mathrm{th}}$ ensemble perturbation:

$$\delta\mathbf{x}_p^b = \mathbf{x}_p^b - \langle\mathbf{x}^b\rangle \quad \text{and} \quad \langle\mathbf{x}^b\rangle = \frac{1}{N}\sum_{p=1}^{N} \mathbf{x}_p^b \tag{2}$$

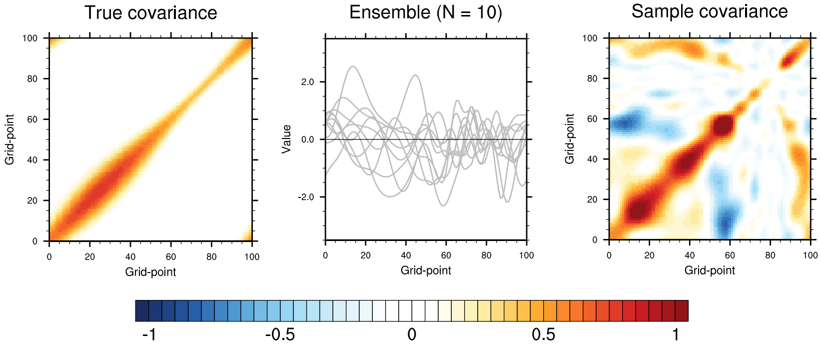
Asymptotic sample covariance: $\mathbf{B} = \lim_{N\rightarrow\infty} \widetilde{\mathbf{B}}$

Since the ensemble size $N < \infty$, $\widetilde{\mathbf{B}}$ is affected by sampling noise:

$$\widetilde{\mathbf{B}}^e = \widetilde{\mathbf{B}} - \mathbf{B} \tag{3}$$

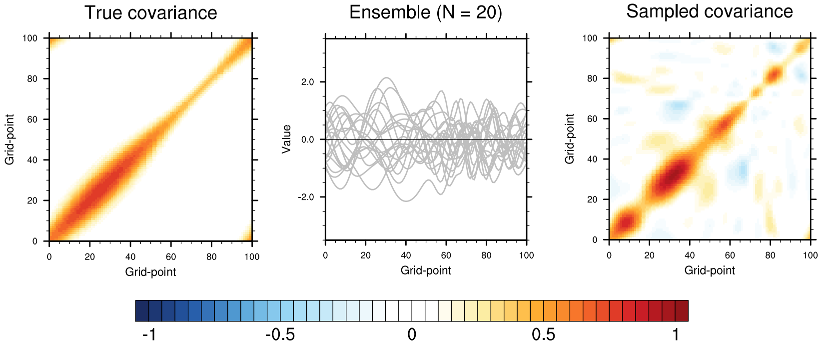
Sample covariance

Sampling noise strongly depends on the ensemble size:



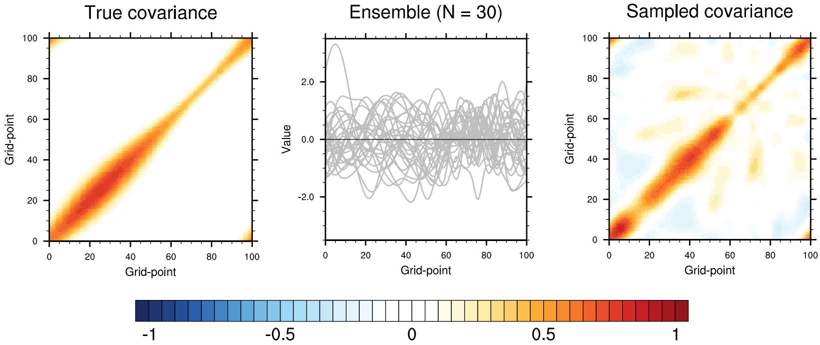
Sample covariance

Sampling noise strongly depends on the ensemble size:



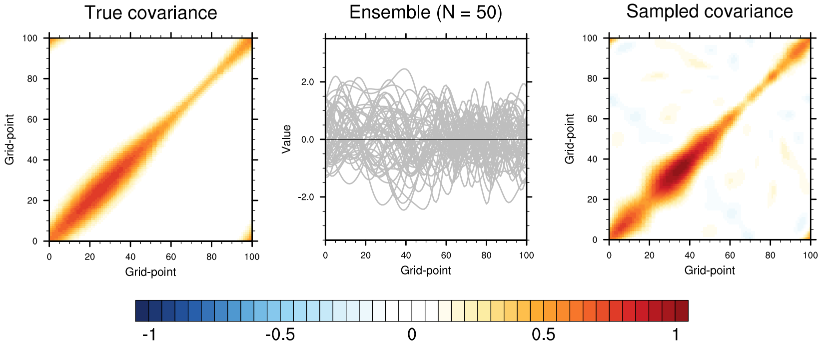
Sample covariance

Sampling noise strongly depends on the ensemble size:



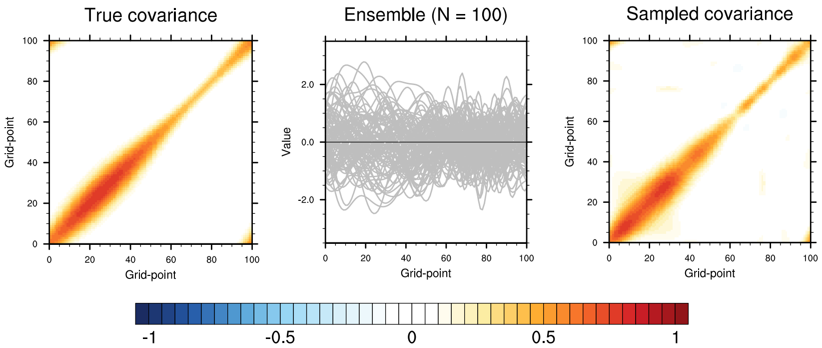
Sample covariance

Sampling noise strongly depends on the ensemble size:



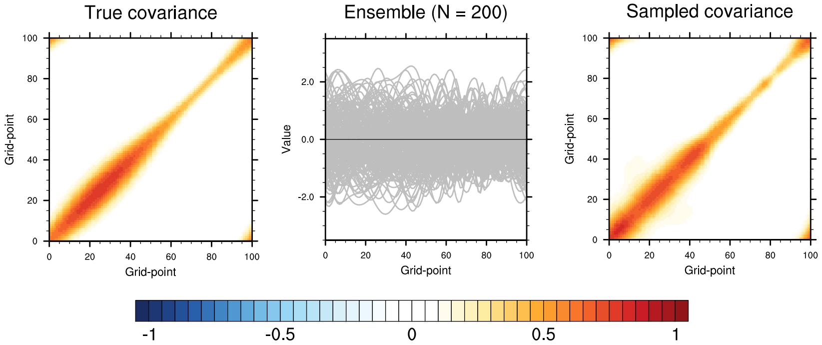
Sample covariance

Sampling noise strongly depends on the ensemble size:



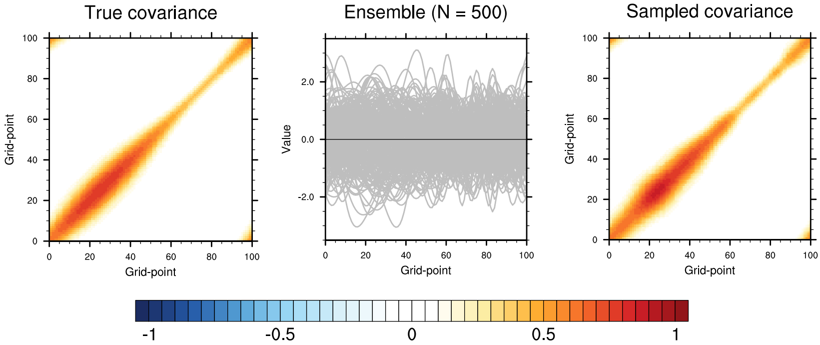
Sample covariance

Sampling noise strongly depends on the ensemble size:



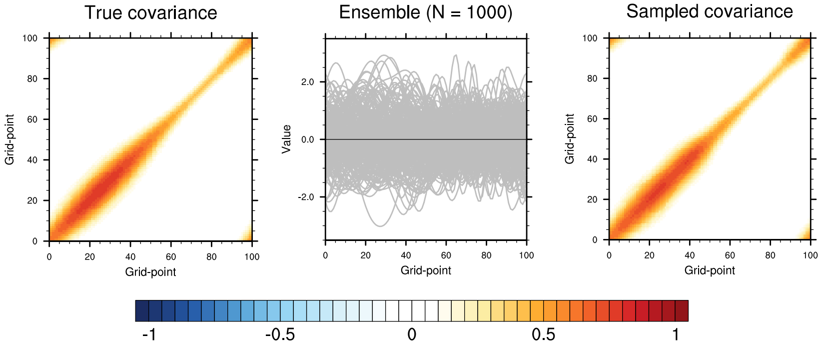
Sample covariance

Sampling noise strongly depends on the ensemble size:



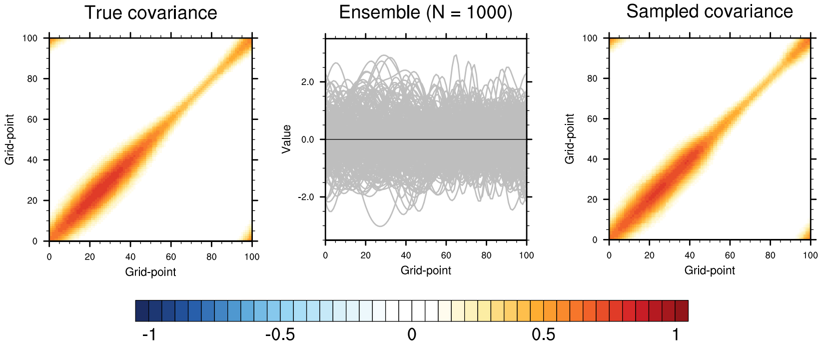
Sample covariance

Sampling noise strongly depends on the ensemble size:



Sample covariance

Sampling noise strongly depends on the ensemble size:



Solution: using a huge ensemble (really, really huge).

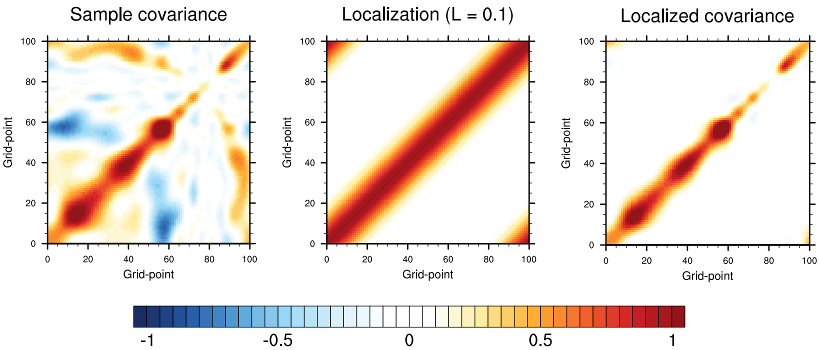
Outline

Localized covariance

Sampling noise on $\widetilde{\mathbf{B}}$ can be damped via a Schur product (element-by-element) with a localization matrix $\mathbf{L}$:

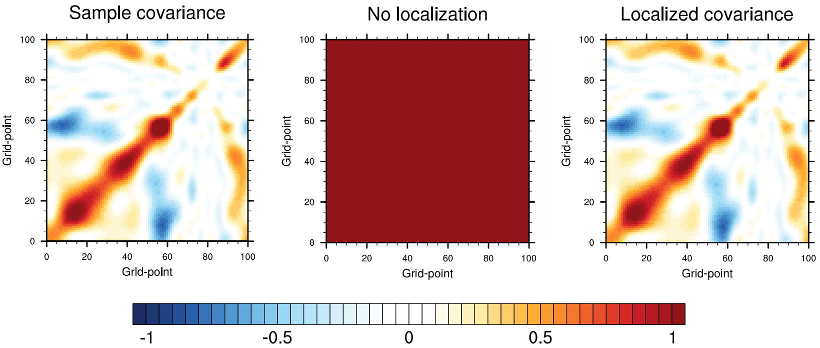
$$\widehat{\mathbf{B}} = \mathbf{L} \circ \widetilde{\mathbf{B}} \quad \Leftrightarrow \quad \widehat{B}_{ij} = L_{ij} \widetilde{B}_{ij} \tag{4}$$

In practice, $\mathbf{L}$ damps the long-distance correlations that are small and more affected by sampling noise (hence the “localization”).



Localization: what is the optimal length-scale?

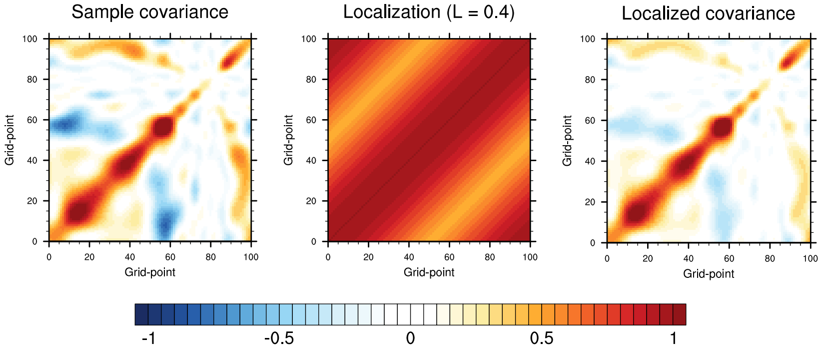
The localization length-scale is critical to remove the sampling noise while keeping the relevant covariance signal:



No impact

Localization: what is the optimal length-scale?

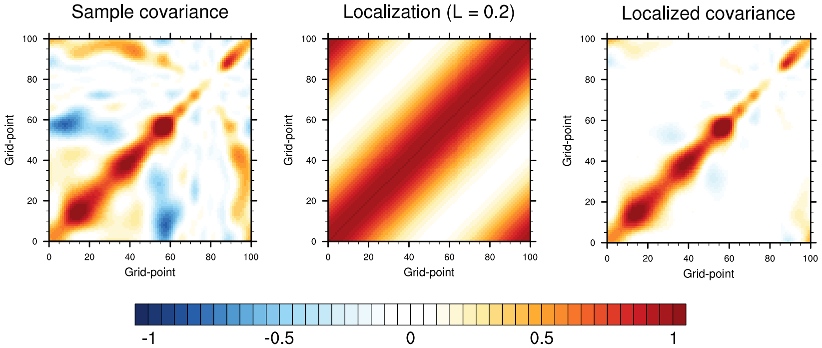
The localization length-scale is critical to remove the sampling noise while keeping the relevant covariance signal:



Start reducing the sampling noise...

Localization: what is the optimal length-scale?

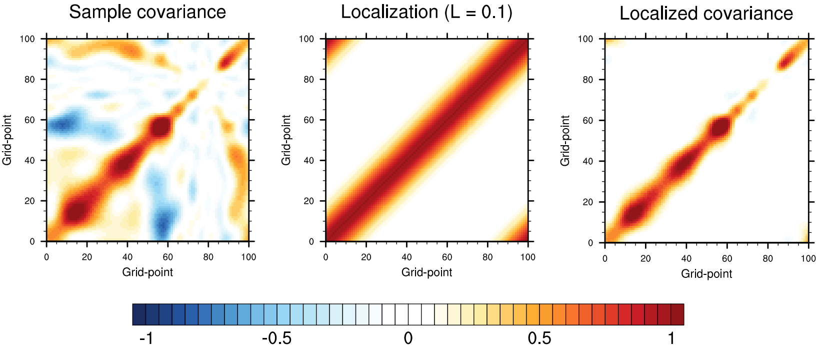
The localization length-scale is critical to remove the sampling noise while keeping the relevant covariance signal:



Less and less sampling noise...

Localization: what is the optimal length-scale?

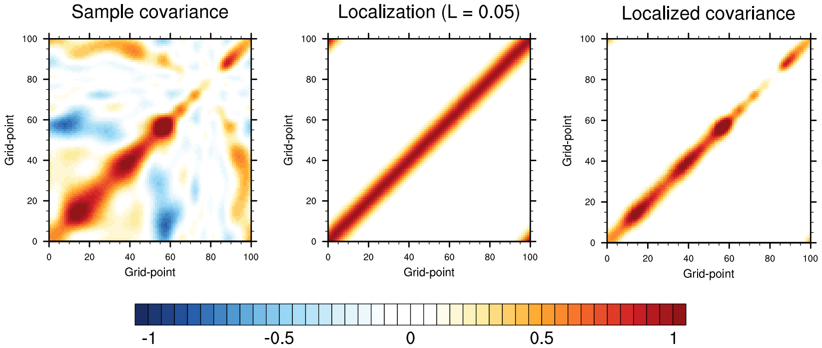
The localization length-scale is critical to remove the sampling noise while keeping the relevant covariance signal:



Good ! Almost no sampling noise anymore...

Localization: what is the optimal length-scale?

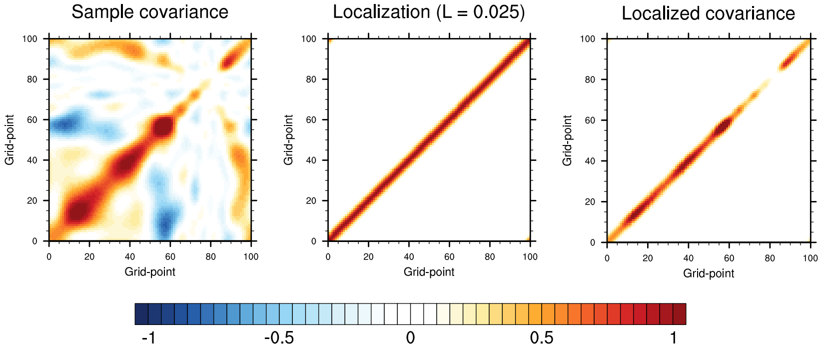
The localization length-scale is critical to remove the sampling noise while keeping the relevant covariance signal:



Well, we are loosing some signal now...

Localization: what is the optimal length-scale?

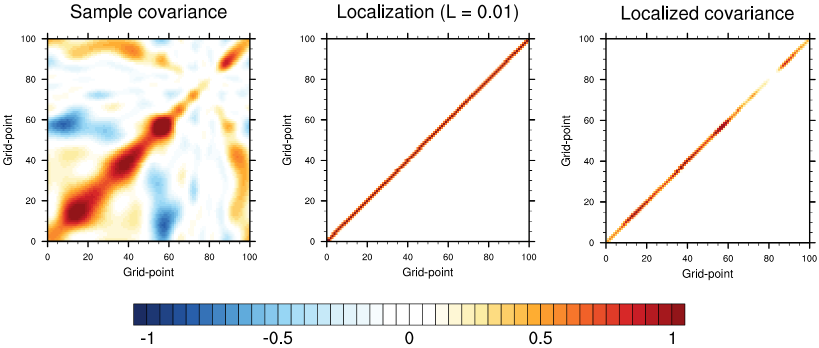
The localization length-scale is critical to remove the sampling noise while keeping the relevant covariance signal:



Hey, stop loosing signal !

Localization: what is the optimal length-scale?

The localization length-scale is critical to remove the sampling noise while keeping the relevant covariance signal:



No more signal !

Outline

How to optimize localization ?

Existing methods are empirical and costly (e.g. OSSE, brute-force optimization). We need a new method that:

- uses only ensemble members,
- is affordable for high-dimensional systems.

Objectives:

- Express $\mathbf{L}$ minimizing the error $\mathbb{E}\left[\|\mathbf{L} \circ \widetilde{\mathbf{B}} - \mathbf{B}\|^2\right]$

  $\rightarrow$ Linear filtering theory: $L_{ij} = \dfrac{\mathbb{E}\left[B_{ij}^2\right]}{\mathbb{E}\left[\widetilde{B}_{ij}^2\right]}$

- Express statistics on asymptotic quantities (unknown) with expected sample quantities (knowable).

  $\rightarrow$ Centered moments sampling theory (non-Gaussian case).

How to optimize localization?

Optimal localization :

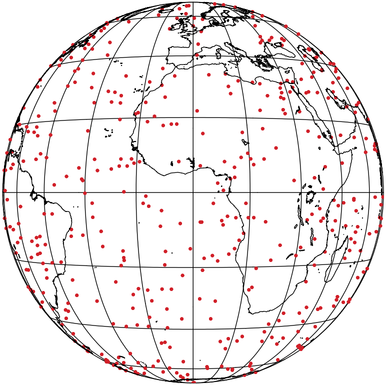
$$L_{ij} = \frac{(N-1)^2}{N(N-3)} + \frac{N-1}{N(N-2)(N-3)} \frac{\mathbb{E}\left[\widetilde{B}_{ii}\widetilde{B}_{jj}\right]}{\mathbb{E}\left[\widetilde{B}_{ij}^2\right]} - \frac{N}{(N-2)(N-3)} \frac{\mathbb{E}\left[\widetilde{\Xi}_{ijij}\right]}{\mathbb{E}\left[\widetilde{B}_{ij}^2\right]} \quad (5)$$

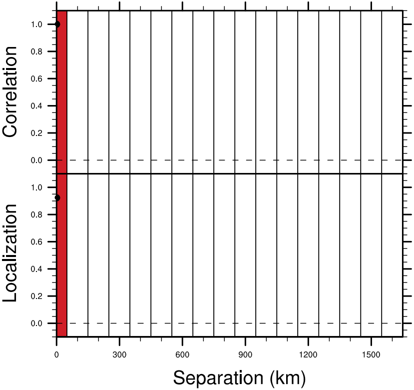
where $\widetilde{\boldsymbol{\Xi}}$ is the sampled fourth-order centered moment.

- Equation (5) involves only **sampled** quantities (with a limited ensemble size), not asymptotic ones.
- Extension available for hybrid weights diagnostic (Ménétrier and Auligné, 2015).
- Expectations $\mathbb{E}[\cdot]$ have to be estimated in practice.

Practical application

Spatial ergodicity assumption to estimate expectations $\mathbb{E}[\,\cdot\,]$:

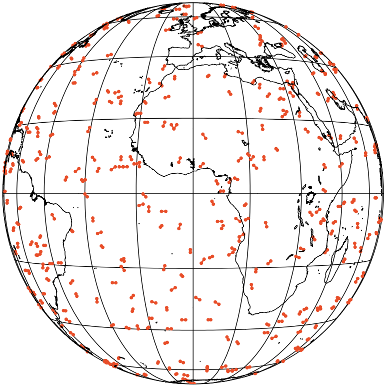


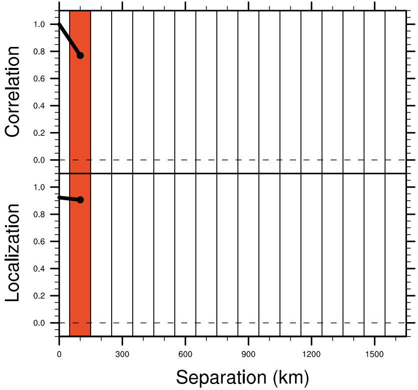


Estimation of horizontal correlation and localization
ARPEGE model, ensemble size N=25, mid-troposphere temperature

Practical application

Spatial ergodicity assumption to estimate expectations $\mathbb{E}[\,\cdot\,]$:

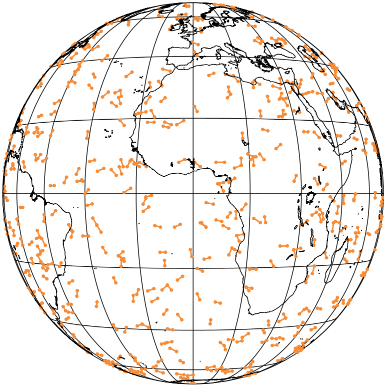


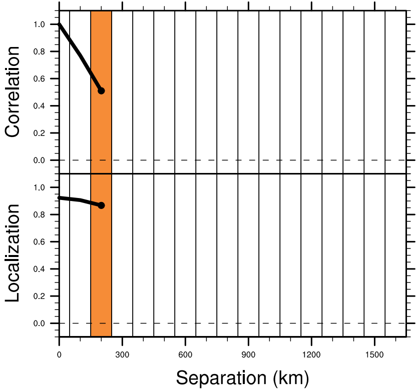


Estimation of horizontal correlation and localization
ARPEGE model, ensemble size N=25, mid-troposphere temperature

Practical application

Spatial ergodicity assumption to estimate expectations $\mathbb{E}[\,\cdot\,]$:

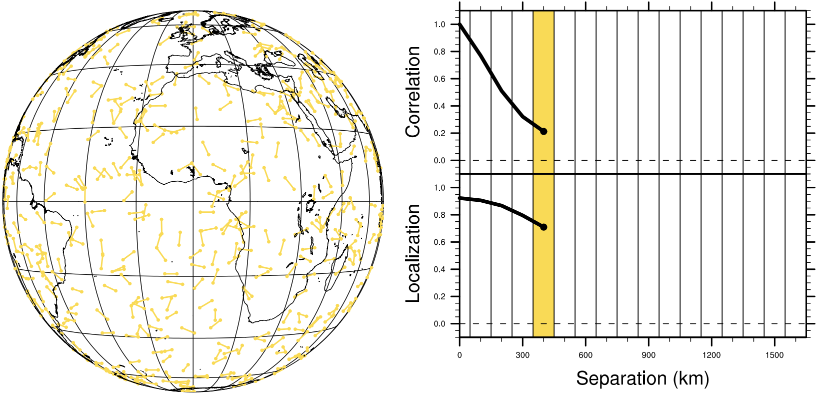




Estimation of horizontal correlation and localization
ARPEGE model, ensemble size N=25, mid-troposphere temperature

Practical application

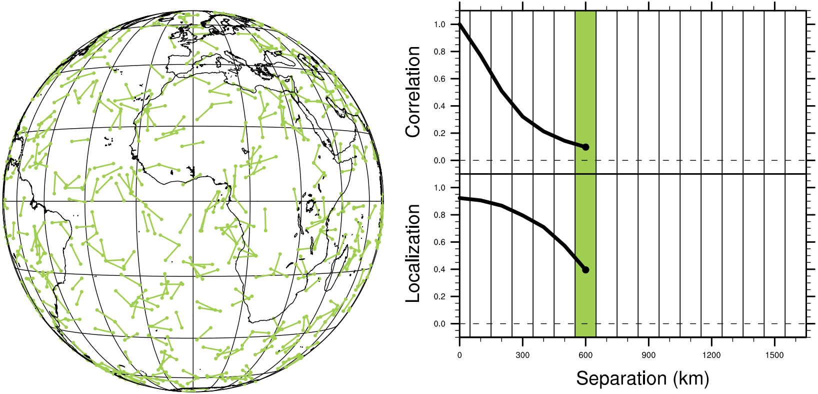
Spatial ergodicity assumption to estimate expectations $\mathbb{E}[\,\cdot\,]$:



Estimation of horizontal correlation and localization
ARPEGE model, ensemble size N=25, mid-troposphere temperature

Practical application

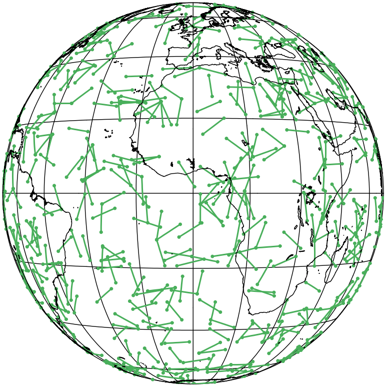
Spatial ergodicity assumption to estimate expectations $\mathbb{E}[\,\cdot\,]$:

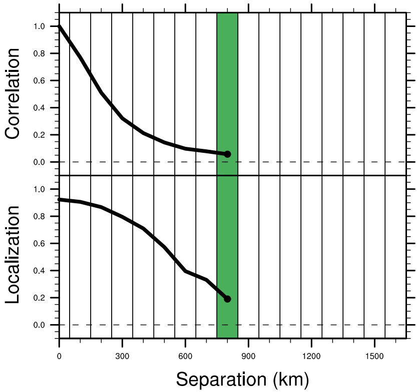


Estimation of horizontal correlation and localization
ARPEGE model, ensemble size N=25, mid-troposphere temperature

Practical application

Spatial ergodicity assumption to estimate expectations $\mathbb{E}[\,\cdot\,]$:

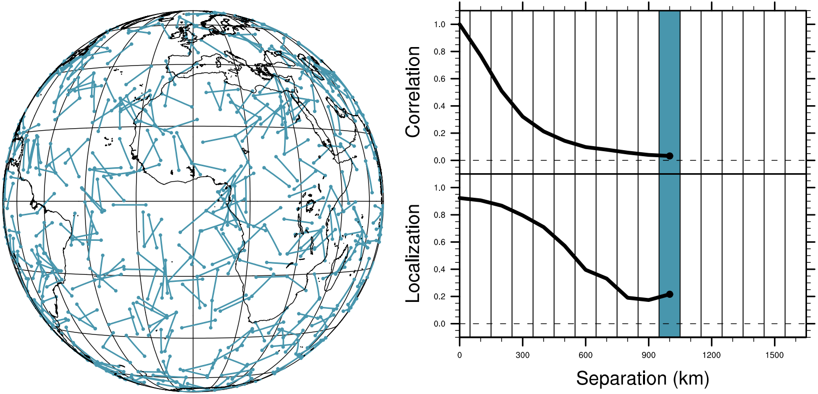




Estimation of horizontal correlation and localization
ARPEGE model, ensemble size N=25, mid-troposphere temperature

Practical application

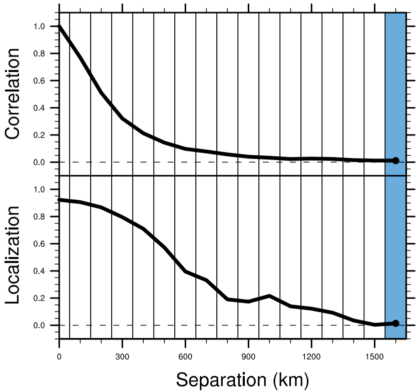
Spatial ergodicity assumption to estimate expectations $\mathbb{E}[\,\cdot\,]$:



Estimation of horizontal correlation and localization
ARPEGE model, ensemble size N=25, mid-troposphere temperature

Practical application

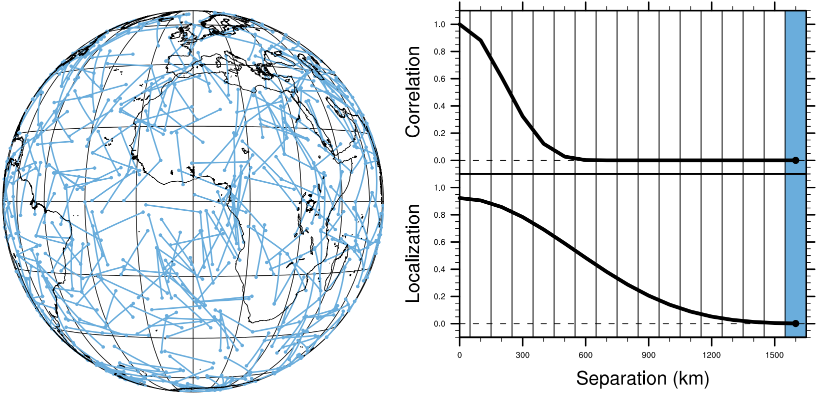
Spatial ergodicity assumption to estimate expectations $\mathbb{E}[\,\cdot\,]$:



Estimation of horizontal correlation and localization
ARPEGE model, ensemble size N=25, mid-troposphere temperature

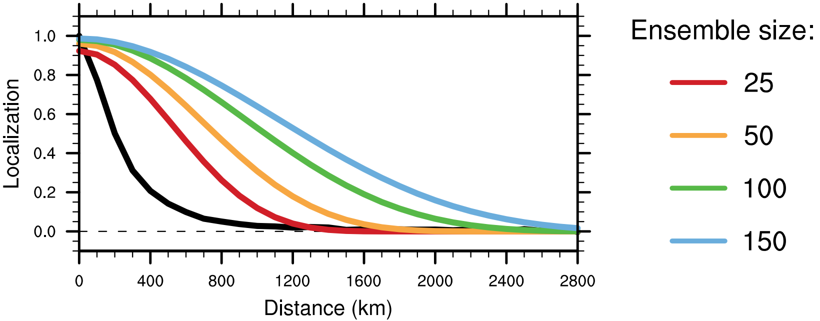
Practical application

Spatial ergodicity assumption to estimate expectations $\mathbb{E}[\,\cdot\,]$:



Fit of horizontal correlation and localization
ARPEGE model, ensemble size N=25, mid-troposphere temperature

Ensemble size sensitivity

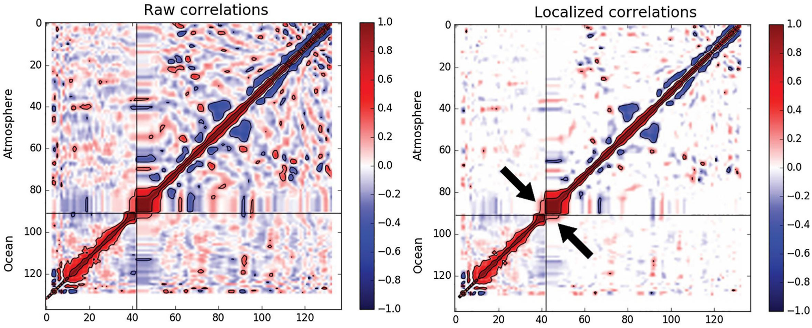


Correlation (black) et localization (colors) for various ensemble sizes

Localization length-scale increases as the ensemble size increases
(less sampling noise to remove)

Example with the coupled system CERA-20C

Temperature coupled background error on August 21, 2005
for the location (130° W, 0° N) from Laloyaux *et al.* (2018)



Vertical correlation matrix: (a) raw and (b) localized

Coupled localization diagnostic seems possible, but it still needs more work and refinements.

Outline

Explicit convolution

In variational methods, the localization matrix itself is not required, only its smoothing effect when applied to a state vector.

Standard methods:

- Spectral/wavelet transforms $\rightarrow$ regular grid required
- Recursive filters $\rightarrow$ regular grid required + normalization issue
- Explicit/implicit diffusion $\rightarrow$ potentially high cost + normalization issue

Advantages of an explicit convolution **C** :

- Work on any grid type
- Exact normalization ($C_{ii} = 1$)

Drawback: the computational cost scales as $O(n^2)$, where n is the size of the model grid...

Explicit convolution

To limit the computational cost, we approximate $\mathbf{C}$ on a subgrid (subset of n^s points of the model grid):

$$\mathbf{C} \approx \mathbf{S}\mathbf{C}^s\mathbf{S}^\mathrm{T} \tag{6}$$

where

- $\mathbf{S}$ is an interpolation from the subgrid to the model grid
- $\mathbf{C}^s$ is a convolution matrix on the subgrid

If $n^s \ll n$, then the total cost scales as $O(n)$ (interpolation cost).

Issues with this approach:

- If the subgrid density is too coarse compared to the convolution length-scale, the convolution is distorded.
- Normalization breaks down because of the interpolation: even if $\mathbf{C}^s$ is normalized, $\mathbf{S}\mathbf{C}^s\mathbf{S}^\mathrm{T}$ is not.

Explicit convolution

The **NICAS** method (Normalized Interpolated Convolution from an Adaptive Subgrid) is given by:

$$\widetilde{\mathbf{C}} = \mathbf{N}\mathbf{S}\mathbf{C}^s\mathbf{S}^{\mathrm{T}}\mathbf{N}^{\mathrm{T}} \tag{7}$$

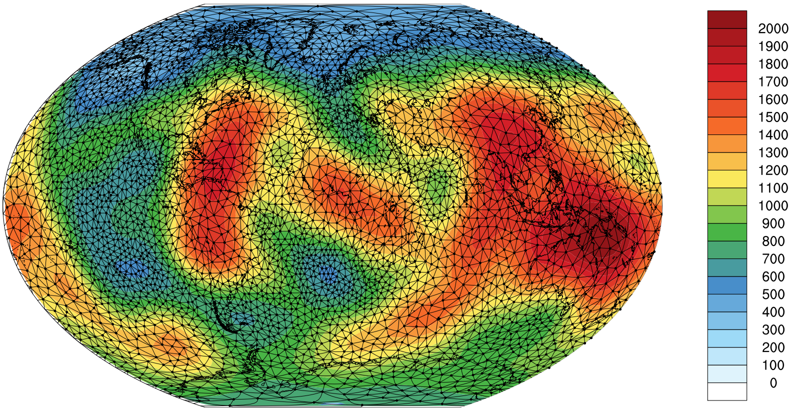
where $\mathbf{N}$ is a diagonal normalization matrix.

Features:

- The subgrid is locally adapted to the convolution length-scale.
- The convolution function is the Gaspari and Cohn (1999) function, modified to use heterogeneous length-scales, or even anisotropic local tensors.
- Communications are local, on the subgrid only.
- Hybrid MPI-OpenMP parallelization is enabled.

Illustrations

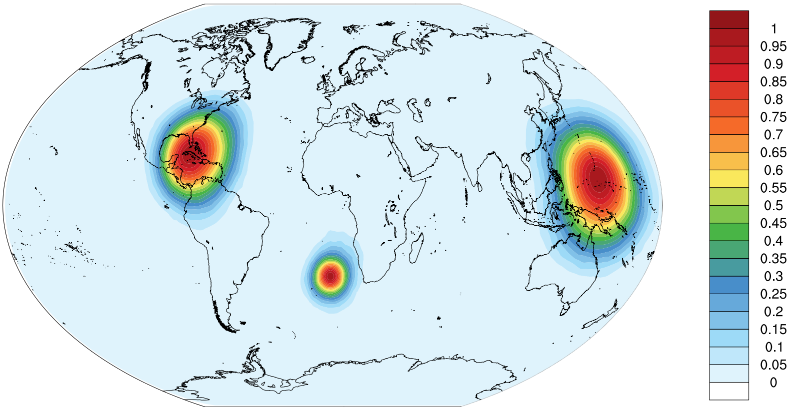
Heterogenous convolution length-scale → heterogenous subgrid:



A fast trial-and-error algorithm using a K-D tree ensures that the horizontal subsampling is well distributed.

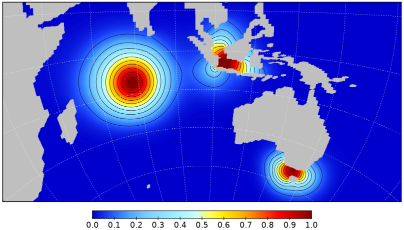
Illustrations

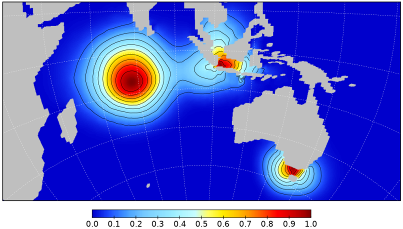
Convolution with a heterogeneous length-scale



Illustrations

Complex boundaries can be taken into account for both interpolation and convolution steps:

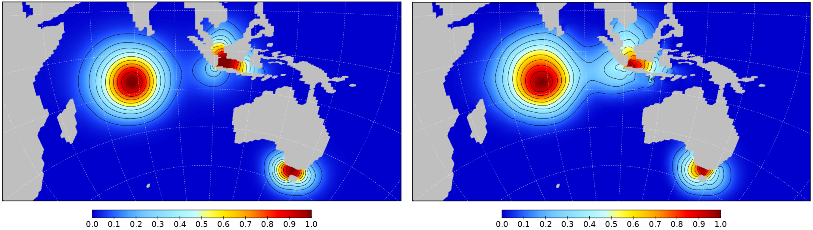




Implicit diffusion (left) and **NICAS** (right) on the ORCA grid.

Illustrations

Complex boundaries can be taken into account for both interpolation and convolution steps:



Implicit diffusion (left) and **NICAS** (right) on the ORCA grid.

Since NICAS can deal with any kind of grid, it should also work for coupled systems. Again, more work is needed to confirm this hope.

Outline

Conclusions

- Localization is required to remove the sampling noise for small ensembles, very large ensembles being unaffordable.
- In variational DA systems, localization is applied in model space (not in observation space).
- Optimal localization can be diagnosed from the ensemble.
- Localization application can be performed efficiently on any grid with the NICAS smoother.
- For coupled DA systems, ensemble-variational methods (EnVar) could be a powerful class of algorithms, but it requires a fully coupled sample covariance localization.
- Tests are underway to apply our recent methods to coupled systems, in order to build such a fully coupled localization.

The BUMP software

- An open-source library, the Background error on Unstructured Mesh Package (BUMP) is available at:

https://github.com/benjaminmenetrier/bump-standalone

- BUMP is also interfaced within OOPS, a generic DA system developed at ECMWF and at the JCSDA (JEDI project):

https://www.jcsda.org/jcsda-project-jedi

The BUMP software

- An open-source library, the Background error on Unstructured Mesh Package (BUMP) is available at:

`https://github.com/benjaminmenetrier/bump-standalone`

- BUMP is also interfaced within OOPS, a generic DA system developed at ECMWF and at the JCSDA (JEDI project):

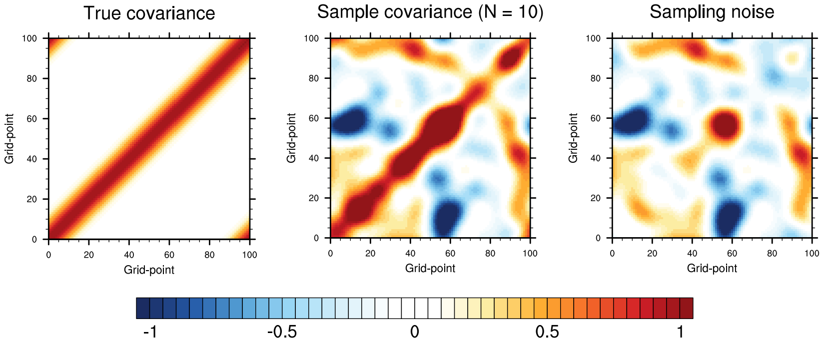
`https://www.jcsda.org/jcsda-project-jedi`

Thank you for your attention!

(and for the invitation)

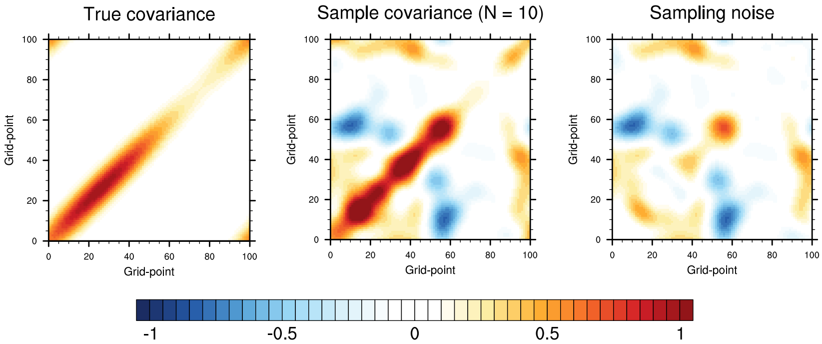
Sampling noise properties

Homogeneous variance / length-scale



Sampling noise properties

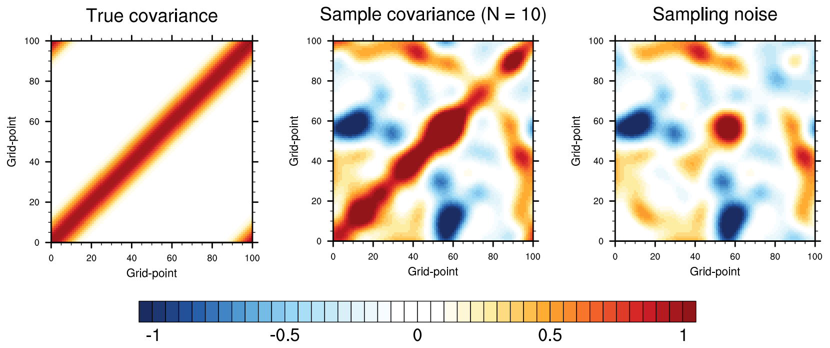
Heterogeneous variance / homogeneous length-scale



Sampling noise amplitude related to the asymptotic variance

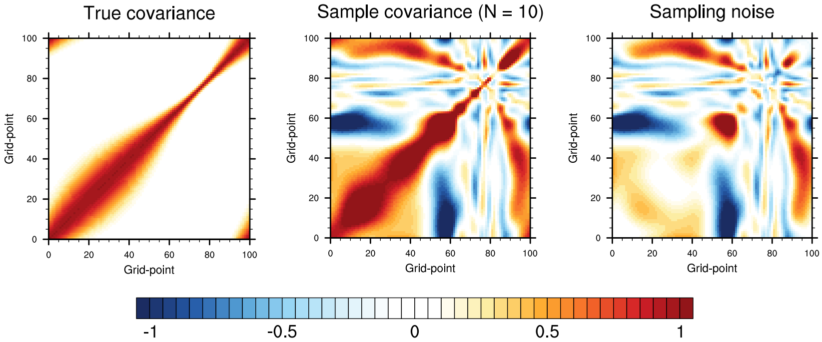
Sampling noise properties

Homogeneous variance / length-scale



Sampling noise properties

Homogeneous variance / heterogeneous length-scale



Sampling noise length-scale related to the asymptotic length-scale