Optimized localization and hybridization to filter ensemble-based covariances

Benjamin Ménétrier (CNRM, Météo-France / CNRS)
Thomas Auligné (NCAR / JCSDA / U of Md / ESSIC)

Summary

- **Localization and hybridization** improve the accuracy of **ensemble-based covariances**.
- Localization functions and hybridization weights can be **jointly and objectively optimized**.
- The proposed method uses the **ensemble members only** and is **affordable for high-dimensional systems**.
- It has been tested on various **atmospheric and oceanographic models** (ARPEGE/AROME, GFS, MPAS, WRF, NEMO).
- Localization and hybridization diagnostics can be used for both **EnVar algorithms and sequential filters** (e.g. EnKF).

Theory

- Covariance matrix sampled from N members: $\widetilde{\mathbf{B}}$
- Asymptotic value for $N \rightarrow \infty$: $\widetilde{\mathbf{B}}^\star$
- 4th order centered moment sampled from N members: $\widetilde{\boldsymbol{\Xi}}$
- General sampling theory (non-Gaussian):

$$\mathbb{E}\left[\widetilde{B}^{\star 2}_{ij}\right] = P(N)\, \mathbb{E}\left[\widetilde{B}^2_{ij}\right] + Q(N)\, \mathbb{E}\left[\widetilde{B}_{ii}\widetilde{B}_{jj}\right] + R(N)\, \mathbb{E}\left[\widetilde{\Xi}_{iijj}\right]$$

where P, Q and R are known fractions of polynomials.

- Localized covariance matrix: $\widehat{\mathbf{B}} = \mathbf{L} \circ \widetilde{\mathbf{B}}$
- Optimal localization matrix $\mathbf{L}$ minimizes $\mathbb{E}\left[\|\widehat{\mathbf{B}} - \widetilde{\mathbf{B}}^\star\|^2\right]$:

$$L_{ij} = \frac{\mathbb{E}\left[\widetilde{B}^{\star 2}_{ij}\right]}{\mathbb{E}\left[\widetilde{B}^2_{ij}\right]}$$

- Static covariance matrix for hybridization: $\overline{\mathbf{B}}$
- Localized/hybridized covariance matrix: $\widehat{\mathbf{B}}^h = \mathbf{L}^h \circ \widetilde{\mathbf{B}} + \beta^{c2}\, \overline{\mathbf{B}}$
- Optimal $\mathbf{L}^h$ and static weight β^c minimize $\mathbb{E}\left[\|\widehat{\mathbf{B}}^h - \widetilde{\mathbf{B}}^\star\|^2\right]$:

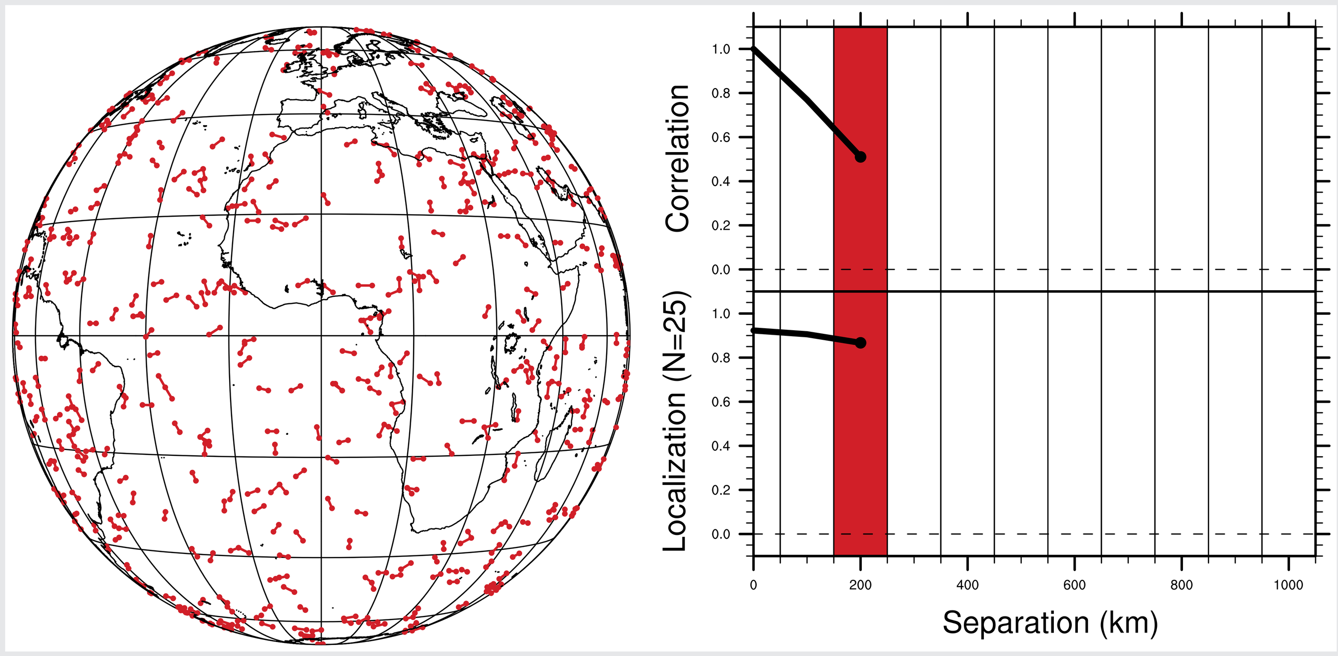
$$\beta^{c2} = \frac{\sum_{ij} \left(1 - L_{ij}\right) \mathbb{E}\left[\widetilde{B}_{ij}\right] \overline{B}_{ij}}{\sum_{ij} \frac{\mathrm{Var}\left[\widetilde{B}_{ij}\right]}{\mathbb{E}\left[\widetilde{B}^2_{ij}\right]} \overline{B}^2_{ij}} \quad \text{and} \quad L^h_{ij} = L_{ij} - \frac{\mathbb{E}\left[\widetilde{B}_{ij}\right]}{\mathbb{E}\left[\widetilde{B}^2_{ij}\right]} \beta^{c2}\, \overline{B}_{ij}$$

Ménétrier *et al.* (2015), MWR, 143, 1622-1643.
Ménétrier and Auligné (2015), MWR, 143, 3931-3947.

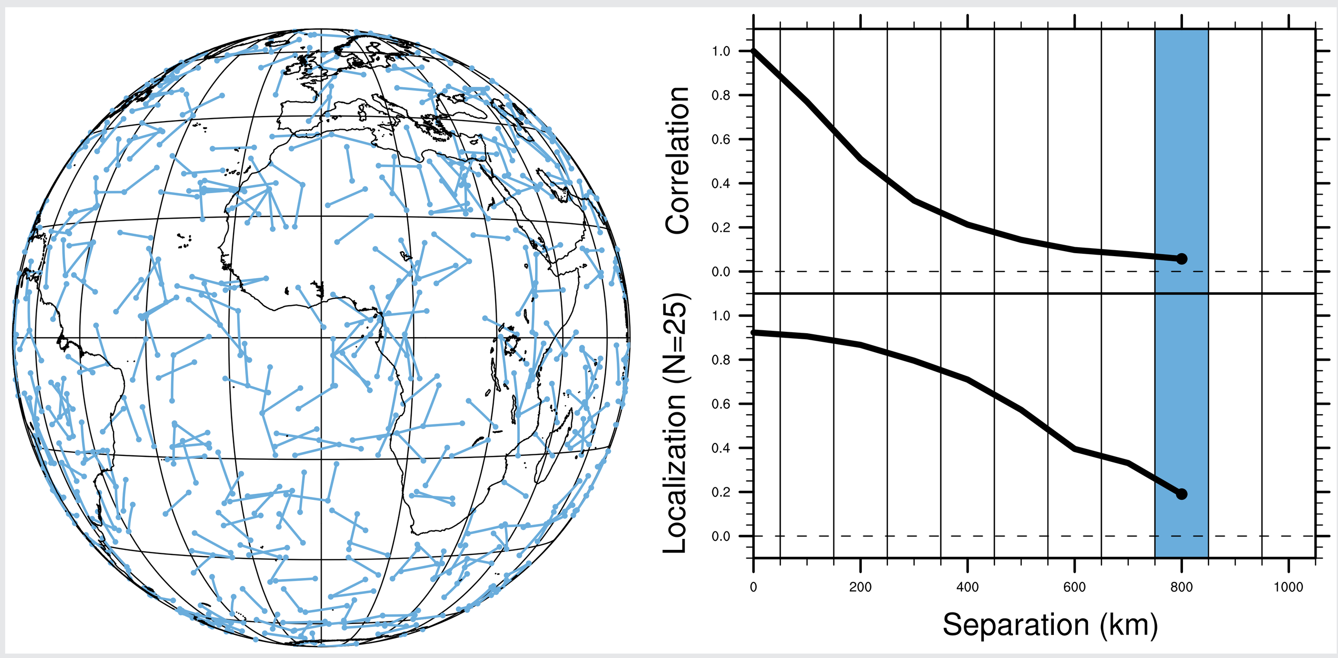
Implementation

- Expectations $\mathbb{E}[\cdot]$ estimated via an **ergodicity assumption**.
- For instance, **spatial and angular** ergodicity: quantites are sampled with couples of points for each separation class.

150 - 250 km class



750 - 850 km class

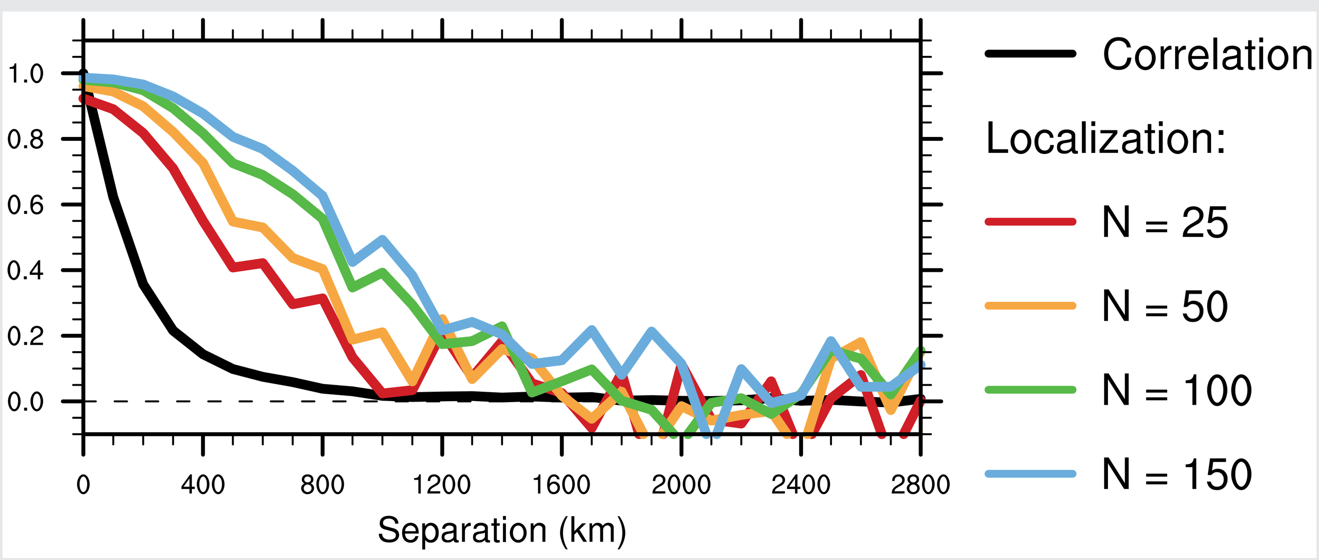


ARPEGE, mid-troposphere temperature, 25-member EDA

- Available for both **horizontal and vertical** localizations.
- **Multivariate diagnostics** capability.
- **Heterogeneous diagnostics** with geographical masks.
- **Generic computational core**, independent from the model grid structure (adding a new model is very simple).
- **Low computational cost** (a few minutes on a desktop PC).

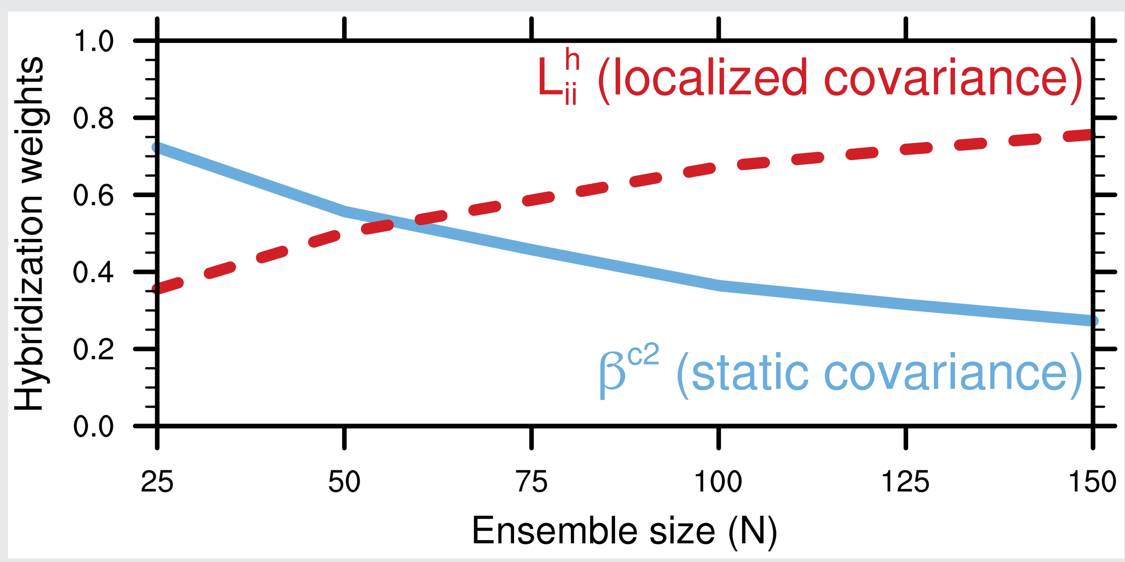
Results for the ARPEGE EDA, mid-troposphere temperature

Localization only



- Localization length-scale and amplitude increase with increasing N (less sampling noise to filter).
- Localization top is flatter than the correlation top.

Localization and hybridization



- If N increases: more weight on $\mathbf{L}^h \circ \widetilde{\mathbf{B}}$, less on $\overline{\mathbf{B}}$.
- $L^h_{ii} + \beta^{c2}$ can be different from 1 depending on $\widetilde{B}_{ii}$ and $\overline{B}_{ii}$ (if $\overline{B}_{ii}$ were twice larger, β^{c2} would be twice smaller).

An open-source and generic code is available at: **https://opensource.cnrm-game-meteo.fr/projects/hybrid_diag**

Contact: **benjamin.menetrier@meteo.fr**