A new method to diagnose optimal localization functions:

- **Localization** improves the accuracy of ensemble-based covariances and can be objectively optimized.
- The proposed method uses the ensemble members only and is affordable for high-dimensional systems.
- Localization diagnostics can be used for both EnVar algorithms and sequential filters (e.g., EnKF).
- **Hybridization weights** can be jointly optimized with the localization for hybrid-EnVar applications.

The method has been used with data from various atmospheric and oceanographic models (ARPEGE/AROME, GFS, GEOS, IFS, MPAS, WRF, NEMO), and is now tested with data from the global coupled reanalysis of the 20th century CERA-20C run at ECMWF and described in Laloyaux et al. (2016), QJRMS, 142, 65-78.

### Theory

- Covariance matrix sampled from N members: \( \hat{B} \)
- Asymptotic value for \( N \to \infty \): \( \hat{B} \)
- 4th order centered moment sampled from \( N \) members: \( \Xi \)
- General sampling theory (non-Gaussian):
  \[
  E \left[ \hat{B}_{ij}^2 \right] = P(N) E \left[ \hat{B}_{ij}^2 \right] + Q(N) E \left[ \hat{B}_i \hat{B}_j \right] + R(N) E \left[ \Xi_{ij} \right]
  \]
- Localized covariance matrix: \( \hat{B} = L \odot \hat{B} \)
- Optimal localization matrix \( L \) minimizes \( E \left[ \| \hat{B} - B^* \|_2^2 \right] : \)
  \[
  \frac{\partial E}{\partial L_{ij}} = 0 \quad \iff \quad L_{ij} = \frac{E \left[ \hat{B}_{ij}^2 \right]}{E \left[ \hat{B}_i^2 \right]} \quad \text{computed from (1)}
  \]

Ménétrier et al. (2015), MWR, 143, 1622-1643.

### Implementation

General implementation of the method:

- Expectations \( E[\cdot] \) estimated via an ergodicity assumption.
- For horizontal localization, spatial and angular ergodicity: quantities are averaged horizontally using couples of points for each separation class.
- For vertical localization, horizontal spatial ergodicity: quantities are averaged horizontally.
- Multivariate diagnostics capability.
- Local diagnostics capability with robust local averages over nearest neighbors (trimmed averages).
- Generic computational core, independent from the model grid structure (adding a new model is very simple).
- Low computational cost (a few minutes on a single proc.).

Technical specificities for coupled systems:

- Oceanic and atmospheric data are interpolated on the same horizontal grid (here the ORCA grid used by NEMO).

### Results

**Global diagnostics:**

- Correlation between oceanic and atmospheric errors is variable-dependent: significant for temperature but weak between salinity and temperature.
- For temperature, a signal could be used for data assimilation at the ocean-atmosphere interface.

**Local diagnostics:**

- Correlation between oceanic and atmospheric errors shows a significant spatial variability.
- Local diagnostics are required for a proper flow-dependent localization between coupled systems.

**Conclusions:**

- The proposed method to diagnose localization functions provides a rich flow-dependent information.
- This preliminary study shows that it can be successfully applied to coupled systems.
- Interested users are encouraged to contact us for applications on other coupled systems.