Objective localization of ensemble covariances: theory and applications

Yann Michel[1], B. Ménétrier[2] and T. Montmerle[1]

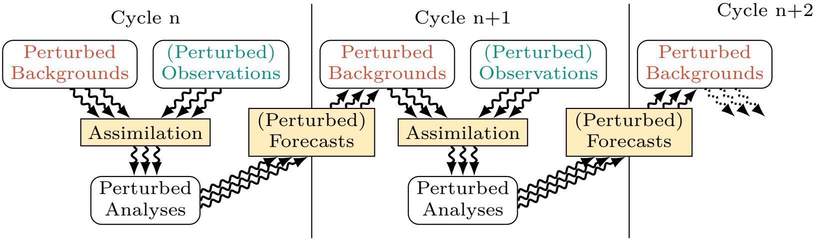
(1) Météo-France & CNRS, Toulouse, France
(2) NCAR, Boulder, CO, USA

25 Feb. 2015

4th International Symposium on Data Assimilation.
RIKEN AICS, Kobe, Japan.

Introduction: Ensemble data assimilation

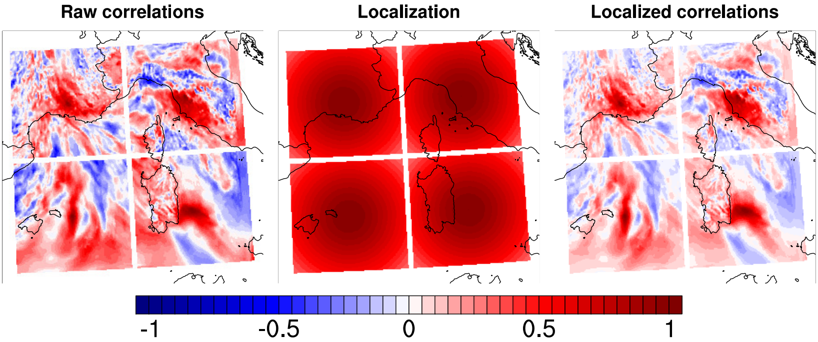
- Early days variationnal data assimilation considered modelled **B** estimated from NMC method.
- EnKF uses Monte Carlo sampling (ensembles) to better estimate covariances.
- We can also use ensembles in variationnal data assimilation, but there is sampling noise.



Introduction: Localization

Covariance localization through a Schur product:

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

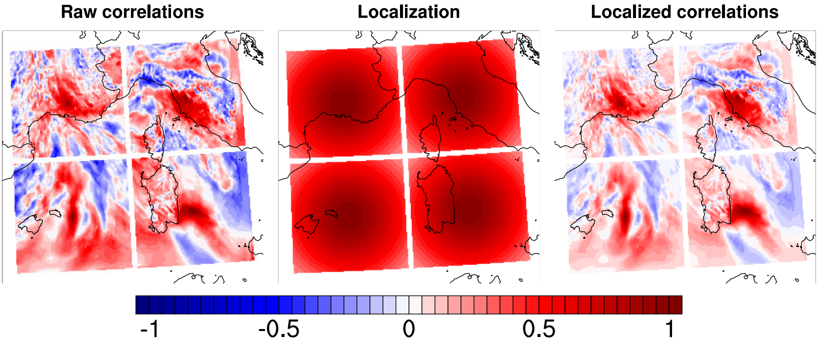


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

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

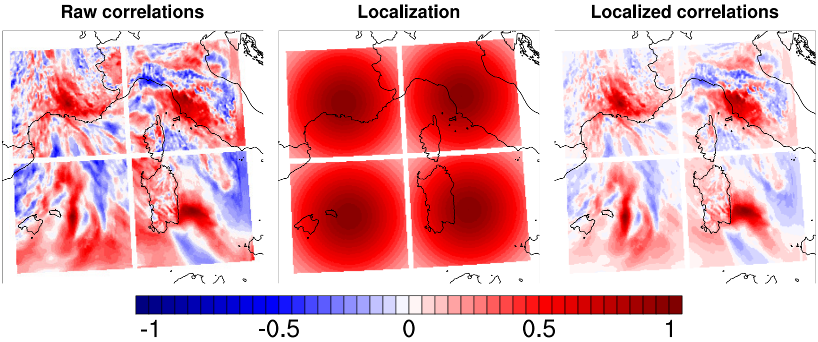


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

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

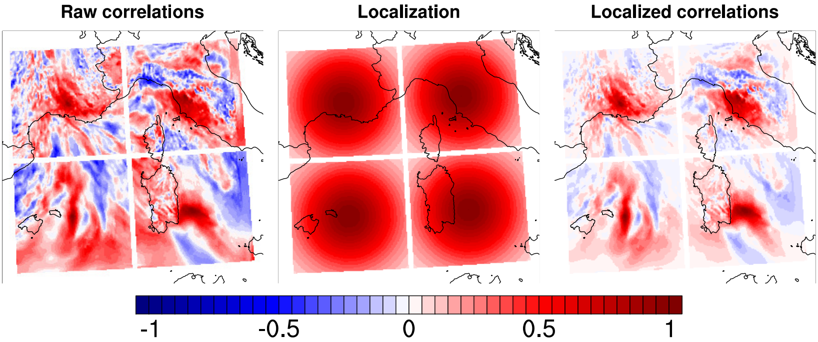


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

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

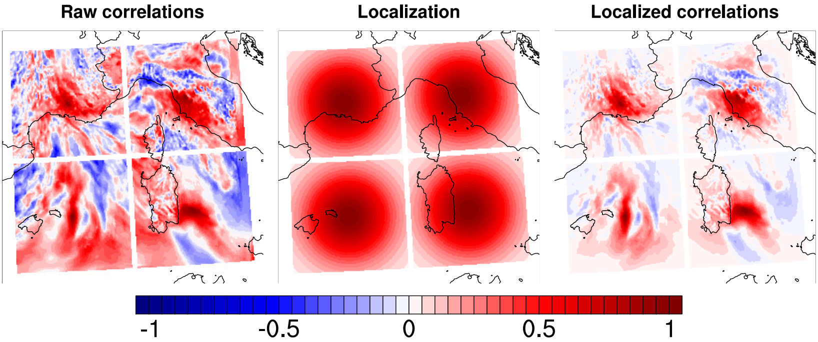


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

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

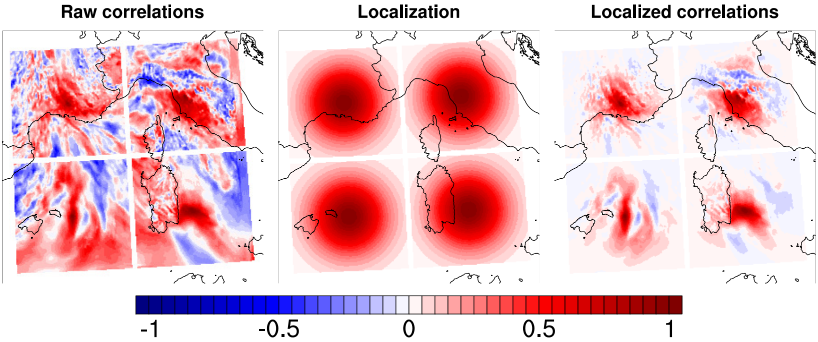


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

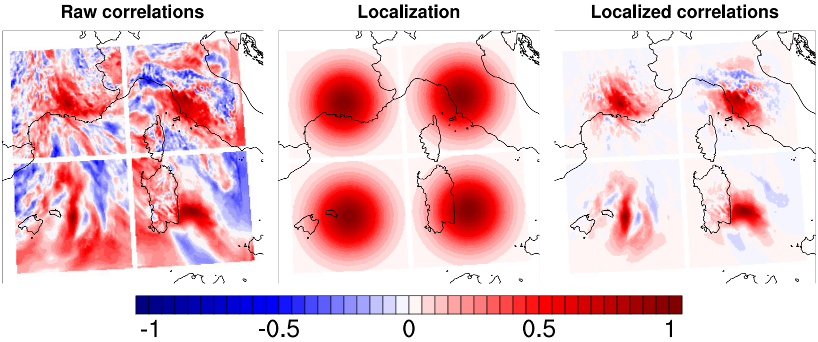
$$\widehat{\mathbf{B}} = \mathbf{L} \circ \widetilde{\mathbf{B}} \quad \Leftrightarrow \quad \widehat{B}_{ij} = L_{ij}\widetilde{B}_{ij}$$



Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Covariance localization through a Schur product:

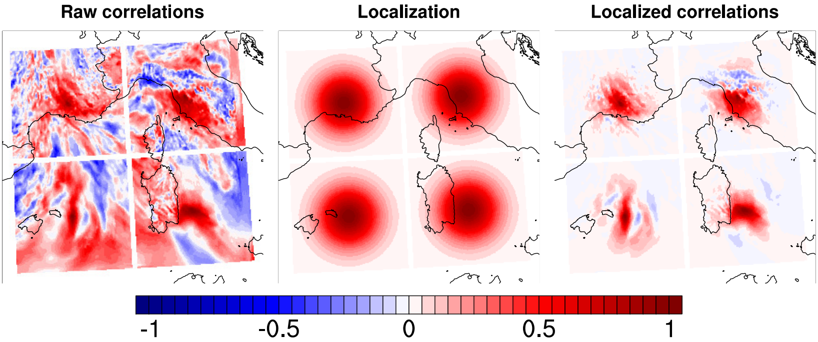
$$\widehat{\mathbf{B}} = \mathbf{L} \circ \widetilde{\mathbf{B}} \quad \Leftrightarrow \quad \widehat{B}_{ij} = L_{ij} \widetilde{B}_{ij}$$



Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Covariance localization through a Schur product:

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

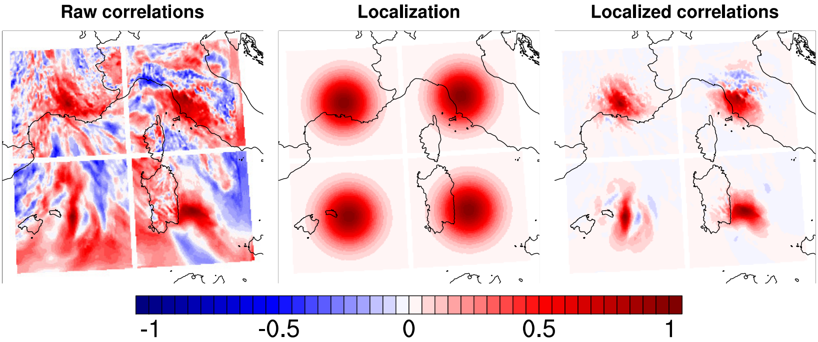


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

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

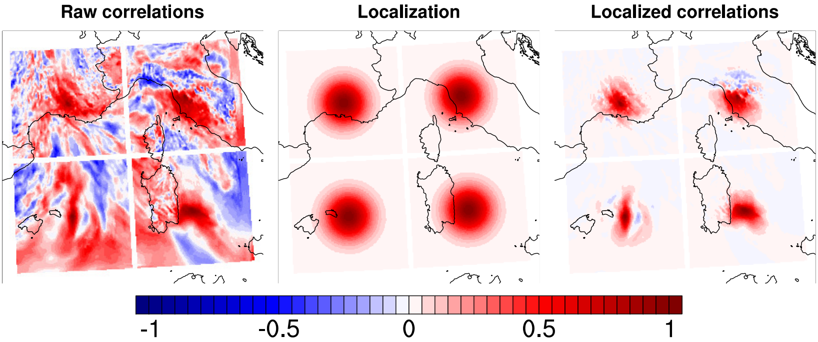


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

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

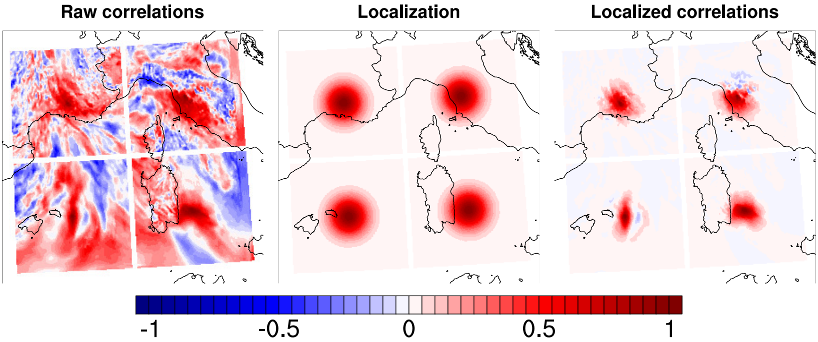


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

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

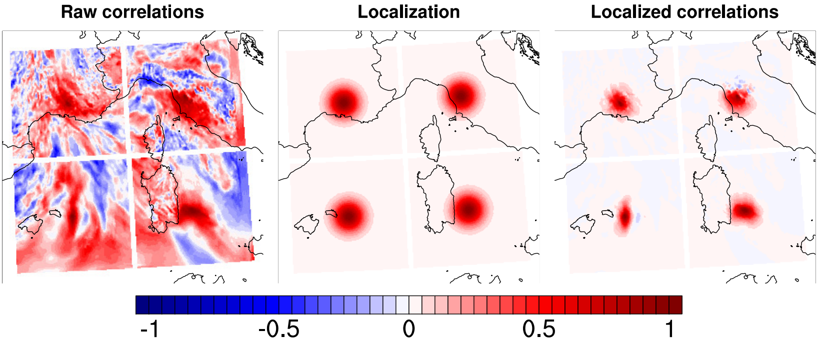


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

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

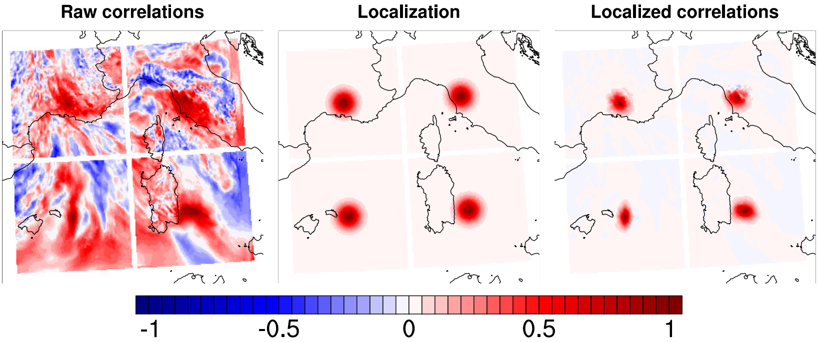


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

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

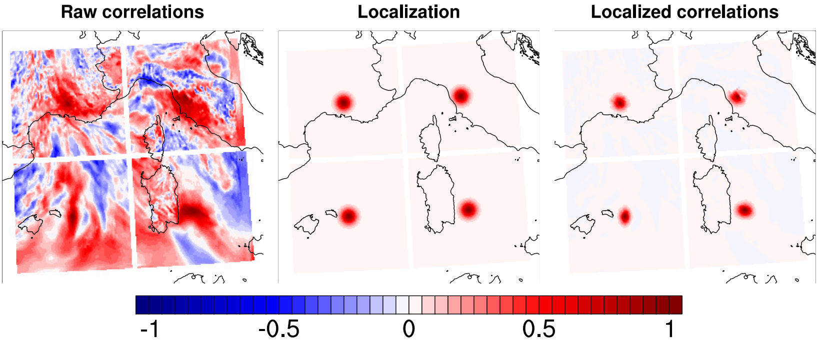


Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction: Localization

Covariance localization through a Schur product:

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



Sample correlation, localization and localized correlation
for specic humidity in the boundary layer (AROME-France - 30 members)

Introduction

Aim of this study:

- Development of a general theory for the linear filtering of background error sample covariances.
- (Application to the spatial filtering of variances).
- Application to the spatial localization of covariances.
- Illustration with the global ARPEGE model and the convective scale AROME model.

Main results:

- Dependency on the ensemble size;
- Dependency on the variable;
- Vertical dependency of horizontal and vertical localizations.

Outline

Outline

Theory: how to get started (I)

- The sample covariance matrix $\widetilde{\mathbf{B}}$ is estimated using standard unbiased formulae:
$$\widetilde{\mathbf{B}} = \frac{1}{N-1}\mathbf{X}\mathbf{X}^{\mathrm{T}}$$
where $\mathbf{X}$ is the matrix of perturbations.
- Asymptotic convergence with ensemble size
$$\widetilde{\mathbf{B}}^{\star} = \lim_{N\to\infty} \widetilde{\mathbf{B}}$$
- Sampling error is
$$\widetilde{\mathbf{B}}^{e} = \widetilde{\mathbf{B}} - \widetilde{\mathbf{B}}^{\star}$$
- Estimate using a linear filter is
$$\widehat{\mathbf{B}} = \mathbf{F}\widetilde{\mathbf{B}} + \mathbf{f}$$
- Filtering error is
$$\widehat{\mathbf{B}}^{e} = \widehat{\mathbf{B}} - \widetilde{\mathbf{B}}^{\star}$$

Theory: how to get started (II)

From our statistical knowledge

- Unbiased sampling error $\mathbb{E}[\widetilde{\mathbf{B}}^e] = 0$
- Covariance of sampling error has known expression

$$\begin{aligned}\mathrm{Cov}(\widetilde{B}^e_{ij}, \widetilde{B}^e_{kl}) &= \frac{1}{N}\widetilde{\Xi}^\star_{ijkl} - \frac{1}{N}\widetilde{B}^\star_{ij}\widetilde{B}^\star_{kl} + \frac{1}{N(N-1)}(\widetilde{B}^\star_{ik}\widetilde{B}^\star_{jl} + \widetilde{B}^\star_{il}\widetilde{B}^\star_{jk}) \\ &= \frac{1}{N-1}(\widetilde{B}^\star_{ik}\widetilde{B}^\star_{jl} + \widetilde{B}^\star_{il}\widetilde{B}^\star_{jk})\end{aligned}$$

(under Gaussianity)

From the linear filtering theory

- Optimal filter has unbiased error $\mathbb{E}[\widehat{\mathbf{B}}^e] = 0$
- Orthogonality relationship

$$\mathbb{E}[\widehat{B}^e_{ij}\widetilde{B}_{kl}] = 0$$

Theory: end result

$$L_{ij} = \frac{(N-1)^2}{N(N-3)} - \frac{N}{(N-2)(N-3)} \frac{\mathbb{E}[\widetilde{\Xi}_{ijij}]}{\mathbb{E}[\widetilde{B}_{ij}^2]} + \frac{N-1}{N(N-2)(N-3)} \frac{\mathbb{E}[\widetilde{v_i}\widetilde{v_j}]}{\mathbb{E}[\widetilde{B}_{ij}^2]}$$

$$\approx \frac{(N-1)}{(N+1)(N-2)} \left(N - 1 - \frac{\mathbb{E}[\widetilde{v_i}\widetilde{v_j}]}{\mathbb{E}[\widetilde{B}_{ij}^2]} \right)$$

(under Gaussianity)

$$\approx \frac{(N-1)}{(N+1)(N-2)} \left(N - 1 - \frac{1}{\mathbb{E}[\widetilde{C}_{ij}^2]} \right)$$

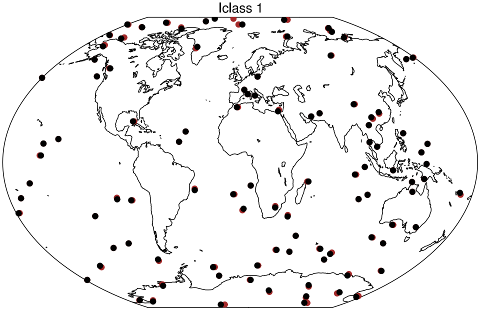
(at small separation distance)

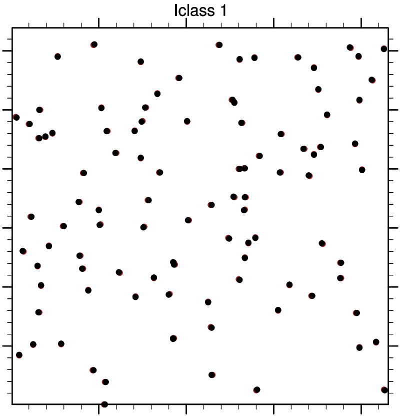
and where:

- N is ensemble size,
- $\widetilde{\Xi}_{ijij}$ is the sample fourth order moment,
- $\widetilde{v}_j$ is the sample variance at j,
- $\widetilde{B}_{ij}^2$ is the sample covariance squared,
- $\widetilde{C}_{ij}^2$ is the sample correlation squared.

Ergodic assumptions

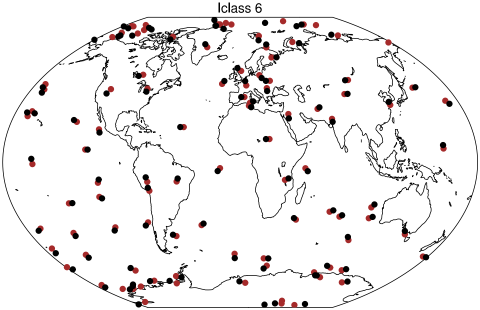
- $\mathbb{E}$ is replaced by averaging over sampling points.
- Sampling is randomly distributed over the domain.
- Binning is isotropic and based on distance.

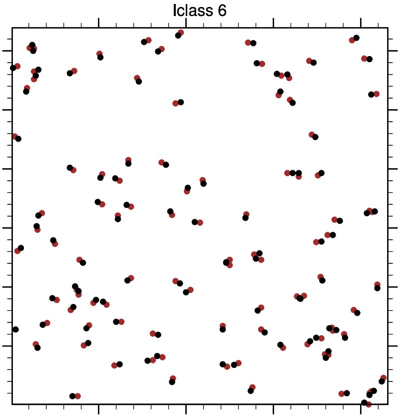




Ergodic assumptions

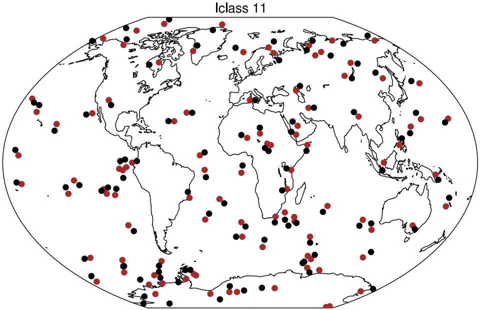
- $\mathbb{E}$ is replaced by averaging over sampling points.
- Sampling is randomly distributed over the domain.
- Binning is isotropic and based on distance.

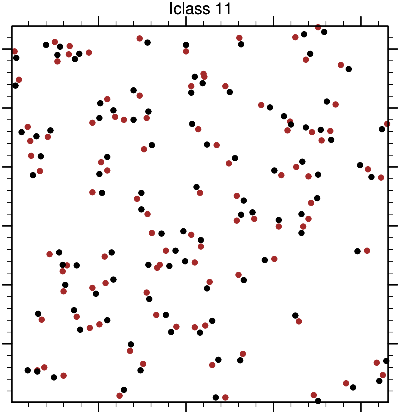




Ergodic assumptions

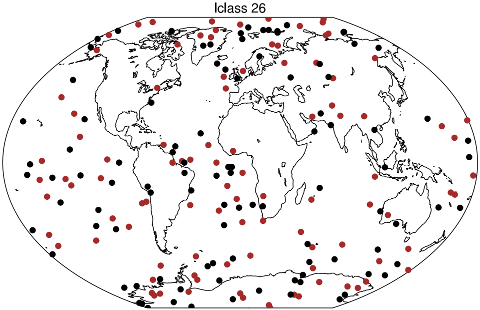
- $\mathbb{E}$ is replaced by averaging over sampling points.
- Sampling is randomly distributed over the domain.
- Binning is isotropic and based on distance.

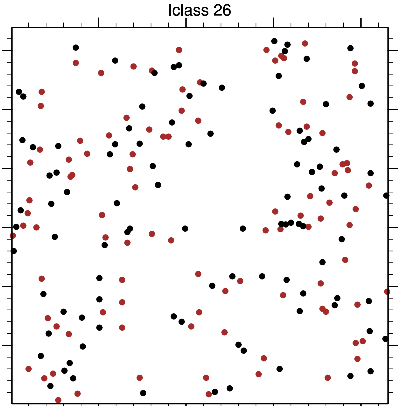




Ergodic assumptions

- $\mathbb{E}$ is replaced by averaging over sampling points.
- Sampling is randomly distributed over the domain.
- Binning is isotropic and based on distance.

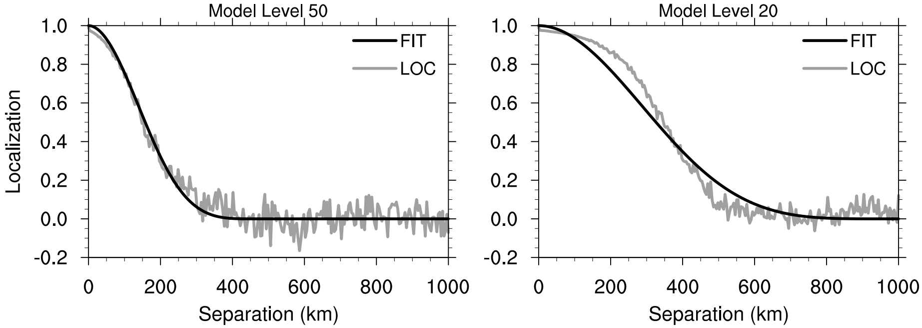




Outline

Results: horizontal localizations

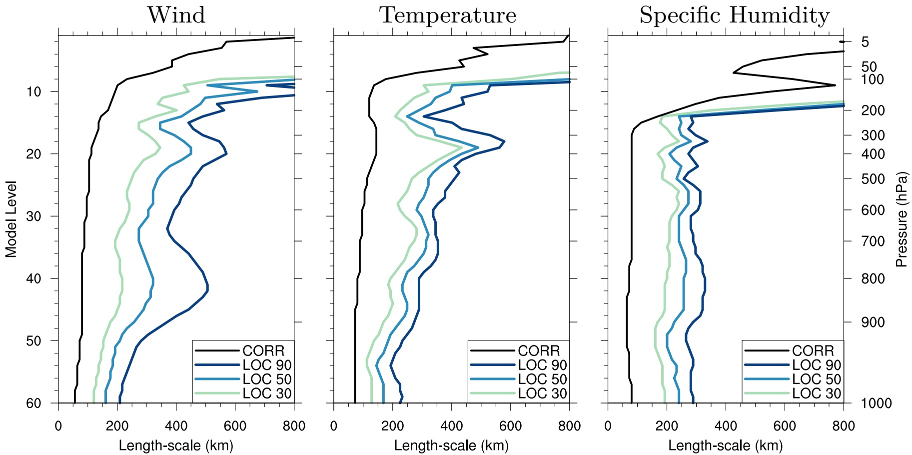
for the convective scale AROME model (I)



Raw horizontal localizations are well approximated by Gaspari and Cohn's compactly supported function.

Results: horizontal localizations length-scales

for the convective scale AROME model (II)

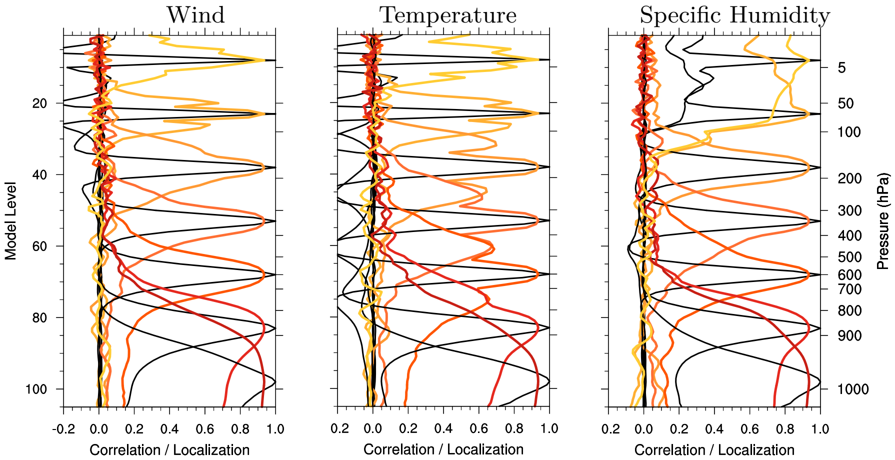


Localization length-scale increases with ensemble size. Local maxima at the top of boundary layer and at the tropopause.

Outline

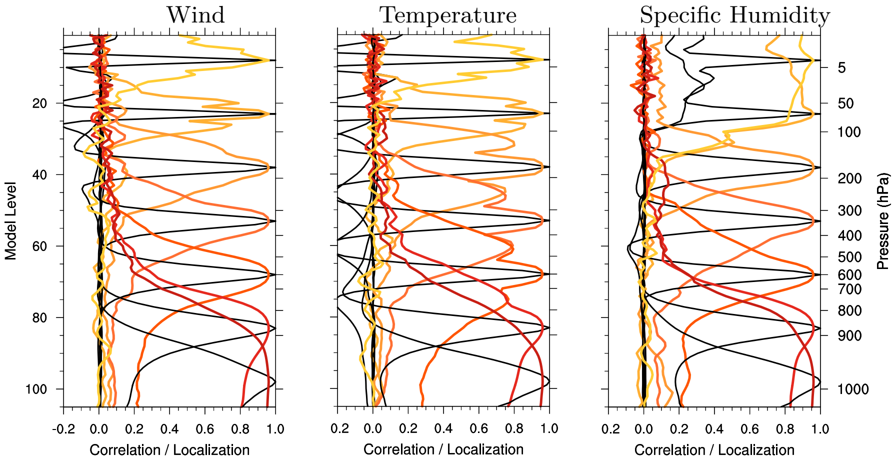
Results: vertical localizations

Comparison between variables for the global ARPEGE model (ens. size=30)



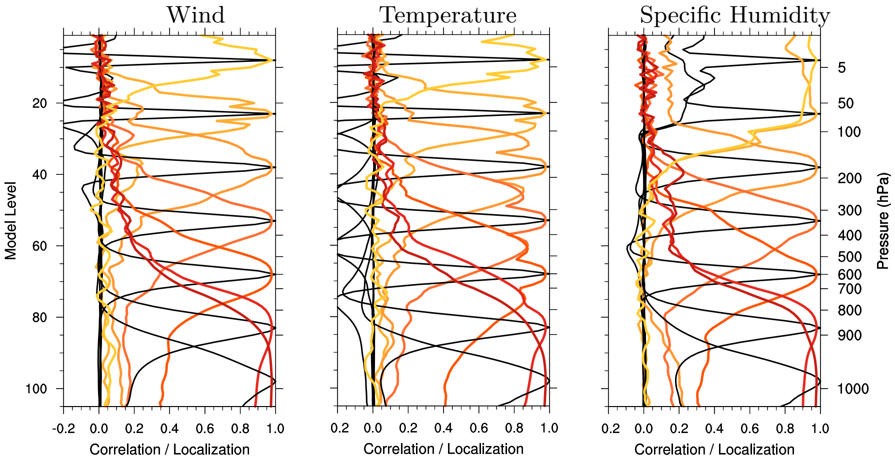
Results: vertical localizations

Comparison between variables for the global ARPEGE model (ens. size=50)



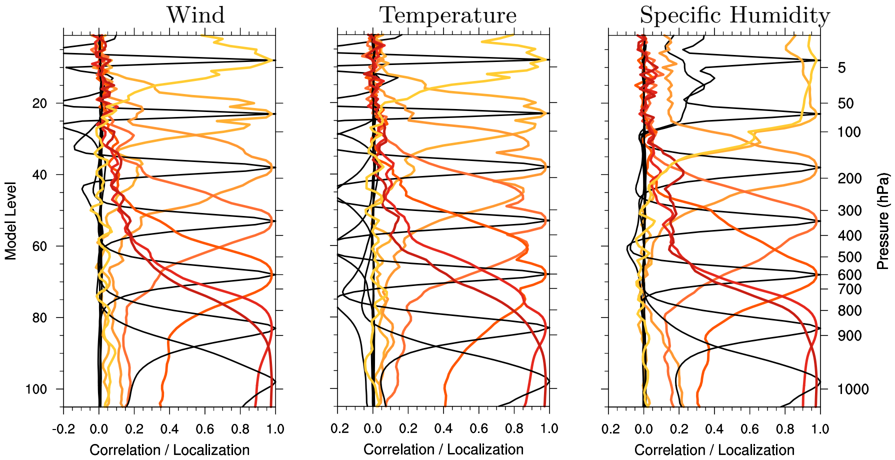
Results: vertical localizations

Comparison between variables for the global ARPEGE model (ens. size=90)



Results: vertical localizations

Comparison between variables for the global ARPEGE model (ens. size=90)



Mostly similar between variables except when there are negative correlations (*e.g.* for T) - as localization follows *squared* correlation.

Conclusion and perspectives

Theory

- Localization is a linear filter
- Optimal (in the RMSE sense) filtering gives orthogonality relationships
- Sampling noise on covariances has known statistical structure

⇒ "Optimal" localization of background error covariances, depending only on quantities estimated over the ensemble.

Ménétrier, B., Montmerle, T., Michel, Y. and Berre, L. (2014)
Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part I: Optimality Criteria and Application to Variances Filtering and Covariances Localization.
Monthly Weather Review.

Conclusion and perspectives

Application

- Localization length-scale increases with ensemble size.
- There are differences between variables: lobes on the vertical for T, larger length-scales on the horizontal at the tropopause (T, U) and at the top of the boundary layer (U).

$\Rightarrow$ We will test these localizations in the 3D/4D-EnVar schemes being developped for AROME model.

Ménétrier, B., Montmerle, T., Michel, Y. and Berre, L. (2014)
Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part II: Application to a Convective-Scale NWP Model.
Monthly Weather Review.