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.
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:

\[ \hat{B} = L \circ \tilde{B} \quad \Leftrightarrow \quad \hat{B}_{ij} = L_{ij} \tilde{B}_{ij} \]

Sample correlation, localization and localized correlation
for specific humidity in the boundary layer (AROME-France - 30 members)
Introduction: Localization

Covariance localization through a Schur product:

\[ \hat{B} = L \circ \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij} \]

Sample correlation, localization and localized correlation for specific humidity in the boundary layer (AROME-France - 30 members)
Covariance localization through a Schur product:

\[ \hat{B} = L \odot \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij} \]

Sample correlation, localization and localized correlation for specific humidity in the boundary layer (AROME-France - 30 members)
Covariance localization through a Schur product: \[
\hat{B} = L \circ \tilde{B} \quad \Leftrightarrow \quad \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}
\]

Sample correlation, localization and localized correlation for specific humidity in the boundary layer (AROME-France - 30 members)
Covariance localization through a Schur product:

\[ \hat{\mathbf{B}} = \mathbf{L} \circ \tilde{\mathbf{B}} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij} \]

Sample correlation, localization and localized correlation for specific humidity in the boundary layer (AROME-France - 30 members)
Covariance localization through a Schur product:

\[ \hat{B} = L \circ \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij} \]

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

Y. Michel et al., ISDA 2015
Introduction: Localization

Covariance localization through a Schur product:

\[
\hat{B} = L \circ \tilde{B} \quad \Leftrightarrow \quad \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}
\]

Sample correlation, localization and localized correlation for specific humidity in the boundary layer (AROME-France - 30 members)
Introduction: Localization

Covariance localization through a Schur product:

\[ \hat{B} = L \circ \tilde{B} \quad \iff \quad \hat{B}_{ij} = L_{ij} \tilde{B}_{ij} \]

Sample correlation, localization and localized correlation for specific humidity in the boundary layer (AROME-France - 30 members)
Introduction: Localization

Covariance localization through a Schur product:

\[ \hat{B} = L \circ \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij} \]

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

Y. Michel et al., ISDA 2015
Covariance localization through a Schur product:

\[
\hat{B} = L \circ \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}
\]

Sample correlation, localization and localized correlation for specific humidity in the boundary layer (AROME-France - 30 members)
Covariance localization through a Schur product:

\[
\hat{B} = L \circ \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}
\]

Sample correlation, localization and localized correlation for specific humidity in the boundary layer (AROME-France - 30 members)
Covariance localization through a Schur product:

\[
\hat{B} = L \circ \tilde{B} \quad \Leftrightarrow \quad \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}
\]

Sample correlation, localization and localized correlation for specific humidity in the boundary layer (AROME-France - 30 members)
Introduction: Localization

Covariance localization through a Schur product:

\[ \hat{B} = L \circ \tilde{B} \iff \hat{B}_{ij} = L_{ij} \tilde{B}_{ij} \]

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

Y. Michel et al., ISDA 2015
Introduction: Localization

Covariance localization through a Schur product:

\[
\hat{B} = L \circ \tilde{B} \quad \Leftrightarrow \quad \hat{B}_{ij} = L_{ij} \tilde{B}_{ij}
\]

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

Y. Michel et al., ISDA 2015
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

1. Introduction
2. Theory
3. Horizontal localization
4. Vertical localization
5. Conclusion and perspectives
Outline

1. Introduction
2. Theory
3. Horizontal localization
4. Vertical localization
5. Conclusion and perspectives

Y. Michel et. al., ISDA 2015

Objective localization
The sample covariance matrix $\tilde{\mathbf{B}}$ is estimated using standard unbiased formulae:

$$\tilde{\mathbf{B}} = \frac{1}{N - 1} \mathbf{X} \mathbf{X}^T$$

where $\mathbf{X}$ is the matrix of perturbations.

Asymptotic convergence with ensemble size

$$\tilde{\mathbf{B}}^* = \lim_{N \to \infty} \tilde{\mathbf{B}}$$

Sampling error is

$$\tilde{\mathbf{B}}^e = \tilde{\mathbf{B}} - \tilde{\mathbf{B}}^*$$

Estimate using a linear filter is

$$\hat{\mathbf{B}} = F\tilde{\mathbf{B}} + f$$

Filtering error is

$$\hat{\mathbf{B}}^e = \hat{\mathbf{B}} - \tilde{\mathbf{B}}^*$$
From our statistical knowledge

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

$$\text{Cov}(\tilde{B}^e_{ij}, \tilde{B}^e_{kl}) = \frac{1}{N} \Xi^*_{ijkl} - \frac{1}{N} \tilde{B}^*_{ij} \tilde{B}^*_{kl} + \frac{1}{N(N - 1)} (\tilde{B}^*_{ik} \tilde{B}^*_{jl} + \tilde{B}^*_{il} \tilde{B}^*_{jk})$$

$$= \frac{1}{N - 1} (\tilde{B}^*_{ik} \tilde{B}^*_{jl} + \tilde{B}^*_{il} \tilde{B}^*_{jk})$$

(under Gaussianity)

From the linear filtering theory

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

$$\mathbb{E}[\hat{B}^e_{ij} \hat{B}^e_{kl}] = 0$$
Theory: end result

\[ L_{ij} = \frac{(N-1)^2}{N(N-3)} - \frac{N}{(N-2)(N-3)} \frac{\mathbb{E}[\tilde{\Xi}_{ijij}]}{\mathbb{E}[B_{ij}^2]} + \frac{N-1}{N(N-2)(N-3)} \frac{\mathbb{E}[\tilde{v}_i\tilde{v}_j]}{\mathbb{E}[B_{ij}^2]} \]

\[ \approx \frac{(N-1)}{(N+1)(N-2)} \left( N - 1 - \frac{\mathbb{E}[\tilde{v}_i\tilde{v}_j]}{\mathbb{E}[B_{ij}^2]} \right) \]

(under Gaussianity)

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

(at small separation distance)

and where:

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

- \( E \) is replaced by averaging over sampling points.
- Sampling is randomly distributed over the domain.
- Binning is isotropic and based on distance.
Ergodic assumptions

- $E$ 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

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

1 Introduction
2 Theory
3 Horizontal localization
4 Vertical localization
5 Conclusion and perspectives
Results: horizontal localizations
for the convective scale AROME model (I)

Raw horizontal localizations are well approximated by Gaspari and Cohn’s compactly supported function.
Localization length-scale increases with ensemble size. Local maxima at the top of boundary layer and at the tropopause.
Outline

1 Introduction
2 Theory
3 Horizontal localization
4 Vertical localization
5 Conclusion and perspectives
Results: vertical localizations

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

Y. Michel et al., ISDA 2015

Objective localization
Results: vertical localizations
Comparison between variables for the global ARPEGE model (ens. size=50)

Wind

Temperature

Specific Humidity

Objective localization
Results: vertical localizations
Comparison between variables for the global ARPEGE model (ens. size=90)

Wind

Temperature

Specific Humidity

Y. Michel et al., ISDA 2015
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.

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 developed for AROME model.