Ensemble background-error variances: objective filtering and impact studies

L. Raynaud, L. Berre and G. Desroziers
CNRM/GAME, Météo-France/CNRS, Toulouse, France

8th International Workshop on Adjoint Model Applications in Dynamic Meteorology,

18-22 May 2009, Tannersville, PA, USA
Key element of any DA schemes: background-error covariance matrix \( B \)

**ENDA** provides a suitable framework to estimate \( B \)

- Simulation of the estimation errors along analyses and forecasts
- Documentation of error covariances:
  - over a long period \( \Rightarrow \) “climatological error”
  - for a particular date \( \Rightarrow \) “error of the day”

\( (\text{Evensen, 1997}; \text{Fisher, 2004}; \text{Berre et al., 2007}) \)

(From Ehrendorfer 2006)
Only small size ensembles \((10 \rightarrow \mathcal{O}(10^2))\) are affordable \(\Rightarrow\) detrimental sampling noise for the estimation of \(\mathbf{B}\):

- noisy variance fields \((\text{Berre et al., 2007; Raynaud et al., 2008})\)
- spurious non-zero correlations at long distances \((\text{Houtekamer and Mitchell, 1998; Buehner and Charron, 2007; Pannekoucke et al., 2007})\)
Only small size ensembles \((10 \to \mathcal{O}(10^2))\) are affordable \(\implies\) detrimental sampling noise for the estimation of \(\mathbf{B}\):

- **noisy variance fields** *(Berre et al., 2007; Raynaud et al., 2008)*

- spurious non-zero correlations at long distances *(Houtekamer and Mitchell, 1998; Buehner and Charron, 2007; Pannekoucke et al., 2007)*
Only small size ensembles (10 $\rightarrow$ $\mathcal{O}(10^2)$) are affordable $\implies$ detrimental sampling noise for the estimation of $\mathbf{B}$:

- **noisy variance fields** (Berre et al., 2007; Raynaud et al., 2008)
- spurious non-zero correlations at long distances (Houtekamer and Mitchell, 1998; Buehner and Charron, 2007; Pannekoucke et al., 2007)

Aims of the presentation:

- Introduce an objective filtering method for ensemble-based **variances**
Only small size ensembles ($10 \rightarrow \mathcal{O}(10^2)$) are affordable \implies detrimental sampling noise for the estimation of $B$:

- **noisy variance fields** \cite{Berre2007,Raynaud2008}
- spurious non-zero correlations at long distances \cite{Houtekamer1998,Buehner2007,Pannekoucke2007}

Aims of the presentation:

- Introduce an objective filtering method for ensemble-based variances
- Present an application of the filter to a real NWP ensemble
1.1. Ensemble Data Assimilation

- Only small size ensembles ($10 \rightarrow \mathcal{O}(10^2)$) are affordable
  $\implies$ detrimental sampling noise for the estimation of $\mathbf{B}$:
  - **noisy variance fields** \cite{Berre2007, Raynaud2008}
  - spurious non-zero correlations at long distances
    \cite{Houtekamer1998, Buehner2007, Pannekoucke2007}

Aims of the presentation:

- Introduce an objective filtering method for ensemble-based variances
- Present an application of the filter to a real NWP ensemble
- Estimate the impact of this filter on forecast scores
Only small size ensembles ($10 \rightarrow O(10^2)$) are affordable
\implies detrimental sampling noise for the estimation of $B$:

- **noisy variance fields** (*Berre et al., 2007; Raynaud et al., 2008*)
- spurious non-zero correlations at long distances
  (*Houtekamer and Mitchell, 1998; Buehner and Charron, 2007; Pannekoucke et al., 2007*)

Aims of the presentation:

- Introduce an objective filtering method for ensemble-based variances
- Present an application of the filter to a real NWP ensemble
- Estimate the impact of this filter on forecast scores
- Give some more general results about the benefits of using errors of “the day”
Spatial structure of sampling noise

(Fisher and Courtier 1995 (Fig 6), Raynaud et al., 2008)

True variance field
\( \mathbf{V}^* \sim \text{large scale} \)
\[ \Downarrow \]
Sampling noise
\[ \mathbf{V}^e = \mathbf{\tilde{V}}(\mathcal{N}) - \mathbf{V}^* \sim \text{large scale too?} \]
\[ \rightarrow \text{depend on } L_{\epsilon^b} \]

\[ \Rightarrow \text{Close link between the spatial structures of sampling noise and background-error} \]

(a) \( N = 50, L_{\epsilon^b} = 200 \text{km} \)

(b) \( N = 50, L_{\epsilon^b} = 1000 \text{km} \)
Notations
- $\tilde{B}$ : the estimated $B$ matrix,
- $\tilde{B}^* = E[\tilde{B}]$ : the noise-free estimated $B$ matrix,
- $V^e = \hat{B} - \tilde{B}^*$ : the sampling noise or random error component.
2.1. Empirical insight

- $\tilde{B}$: the estimated $B$ matrix,
- $\tilde{B}^* = E[\tilde{B}]$: the noise-free estimated $B$ matrix,
- $V^e = \tilde{B} - \tilde{B}^*$: the sampling noise or random error component.

2.2. Analytical results

Analytically, it can be shown that the noise covariance matrix is

$$E[V^e V^{eT}] = \frac{2}{N-1} \tilde{B}^* \circ \tilde{B}^*,$$

where $\circ$ stands for the Hadamard product:

- spatial structures of sampling noise and background-error are directly related,
- the relative error of the variance estimation, $\frac{E[(V^e)^2]}{(\tilde{V}^*)^2} = \frac{2}{N-1}$, is inversely proportional to the ensemble size $N$. 
Verification of the analytical formula \((N = 6 \text{ and } N_{exp} = 1000)\)

\[ \Rightarrow \text{Very good agreement between empirical and analytical results.} \]

Following Daley (1991), it can be shown that the noise length-scale is

\[ L_{Ve} = \frac{L_{\epsilon b}}{\sqrt{2}}. \]

\[ \Rightarrow \text{The sampling noise } V^e \text{ is smaller scale than the bkg-error field.} \]
3.1. Formulation

Objective filtering of sampling noise (Raynaud et al., 2009)

Notations
Let’s $S$ be the spherical spectral transform, we define the spectral fields:

$$
\tilde{S} = S(\tilde{V}) \quad \tilde{S}^* = S(\tilde{V}^*) \quad S^e = S(V^e)
$$

Formulation
An objective filter $\rho$, such that $\tilde{S}^*(n,m) \sim \rho(n)\tilde{S}(n,m)$, is defined by (Berre et al. (2007))

$$
\rho = \frac{1}{1 + \frac{P(S^e)}{P(\tilde{S}^*)}}
$$

where $P(\cdot)$ is the power spectrum.
3.1. Formulation

Objective filtering of sampling noise (Raynaud et al., 2009)

Notations
Let’s $S$ be the spherical spectral transform, we define the spectral fields:

$$\tilde{S} = S(\tilde{V}) \quad \tilde{S}^* = S(\tilde{V}^*) \quad S^e = S(V^e)$$

Formulation
An objective filter $\rho$, such that $\tilde{S}^*(n,m) \sim \rho(n)\tilde{S}(n,m)$, is defined by (Berre et al. (2007))

$$\rho = \frac{1}{1 + \frac{P(S^e)}{P(S^*)}}, \text{ where } P(\ . \ ) \text{ is the power spectrum.}$$

- $\rho$ is a simple function of the noise/signal ratio,
Objective filtering of sampling noise (Raynaud et al., 2009)

**Notations**
Let’s $S$ be the spherical spectral transform, we define the spectral fields:

$$
\tilde{S} = S(\tilde{V}) \quad \tilde{S}^* = S(\tilde{V}^*) \quad S^e = S(V^e)
$$

**Formulation**
An objective filter $\rho$, such that $\tilde{S}^*(n, m) \sim \rho(n)\tilde{S}(n, m)$, is defined by (Berre et al. (2007))

$$
\rho = \frac{1}{1 + \frac{P(S^e)}{P(S^*)}}, \text{ where } P(\cdot) \text{ is the power spectrum.}
$$

- $\rho$ is a simple function of the **noise/signal ratio**, 
- it can be estimated with the help of the $E[V^eV^{eT}]$ formula.
Application to the Arpège model $\sigma^b$ (Raynaud et al., 2009)

Experimental setup

- Météo-France Arpège operational model

- Ensemble of 6 independent 3D-Fgat assimilation experiments (Berre et al., 2007, operational since July 2008):
  - explicit perturbation of observations
  - implicit perturbation of background
  - perfect model framework
1. Introduction

2. Spatial structure of sampling noise

3. Design of an objective filter

4. Application to a NWP context

5. Impact studies

6. Conclusions and perspectives

4.1. Experimental setup

4.2. Results

![VO 500 hPa](image1)

![VO surface](image2)
- The truncation of the filter depends on the vertical level: it tends to decrease with altitude.
- The truncation of the filter depends on the vertical level: it tends to decrease with altitude.
- The truncation of the filter depends on the vertical level: it tends to decrease with altitude.

- Filter values are close to 1 in the largest scales, since these components are well-sampled spatially.
1. Introduction
2. Spatial structure of sampling noise
3. Design of an objective filter
4. Application to a NWP context
5. Impact studies
6. Conclusions and perspectives

4.1. Experimental setup
4.2. Results

Raw std

Filtered std

VO surface

VO 500hPa
Does spatial filtering of variances have an impact in the (very) end? *(Raynaud et al., 2009)*

2 impact studies from 15/02/08 to 20/03/08, using:

- vorticity variances “of the day”,
  - either raw
  - or objectively filtered
- climatological variances for other variables.
So, the response is **YES**! Spatial filtering has a positive impact.
Impact of errors “of the day” on an extreme weather event: case of the French storm of 10 February 2009

- 48h-forecasts using:
  - climatological variances
  - variances “of the day” (including VO,D,T,Ps,Q)

- Analysis valid on 10/02/09 at 00 UTC
About the filtering of variances

- Close link between spatial structures of background-error and sampling noise
- Objective filter based on noise-to-signal ratio
- In a NWP context, this filter is robust and nearly cost-free
- Filtered variance maps accurately reflect the underlying flow
About the filtering of variances

- Close link between spatial structures of background-error and sampling noise
- Objective filter based on noise-to-signal ratio
- In a NWP context, this filter is robust and nearly cost-free
- Filtered variance maps accurately reflect the underlying flow

About impact

- Filtered variances improve the background fit to observations and provide more accurate forecasts than raw variances
- The use of a complete set of variances “of the day” results in better forecasts, especially in cases of intense weather events
 Perspectives

- Validation and tuning of the filtered variances (*Desroziers et al.*, 2005).
- Use of such filtered flow-dependent variances in the operational Arpège B matrix.
- Ultimate goal: combined use of filtered flow-dependent variances and correlations (*Pannekoucke et al.*, 2007).
Perspectives

- Validation and tuning of the filtered variances (Desroziers et al., 2005).
- Use of such filtered flow-dependent variances in the operational Arpège B matrix.
- Ultimate goal: combined use of filtered flow-dependent variances and correlations (Pannekoucke et al., 2007).

Thank you for your attention!
Details on the calculation of the objective filter \( \rho \)

- **Estimation of the noise spectrum** \( P(S^e) \)
  
  \[
  E[V^e V^{eT}] = \frac{2}{N-1} \tilde{B}^* \circ \tilde{B}^* \rightarrow \text{needs to estimate } \tilde{B}^* = E[\tilde{B}]
  \]

  With *Ergodic + homogeneous* hypotheses:

  \[
  E[\tilde{B}_{j.}] \approx \frac{1}{N_t N_i} \sum_{t=1,i=1}^{N_t,N_i} \tilde{B}_{i.}(t),
  \]

  - \( \tilde{B}_{i.} \) is the local spatial covariance at gridpoint \( i \),
  - \( N_t \) is the number of dates in the time average,
  - \( N_i \) is the number of gridpoints over the globe.

  In the isotropic case, \( E[\tilde{B}_{j.}] = \overline{C} \) (1D) and:

  \[
  P(S^e) = \frac{2}{N-1} L(\overline{C}^2)
  \]

- **Estimation of the noise-free variance spectrum** \( P(\tilde{S}^*) \)

  \[
  P(\tilde{S}^*) = P(\tilde{S}) - P(S^e)
  \]