Ensemble background-error variances: objective filtering and impact studies

L. Raynaud, L. Berre and G. Desroziers

CNRM/GAME, Météo-France/CNRS, Toulouse, France

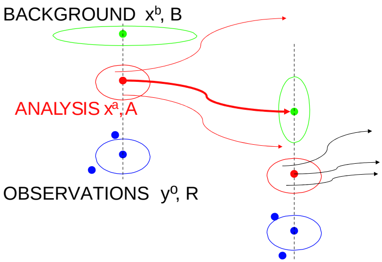
8th International Workshop on Adjoint Model Applications in Dynamic Meteorology,

18-22 May 2009, Tannersville, PA, USA

- Key element of any DA schemes : background-error covariance matrix $\mathbf{B}$
- ENDA provides a suitable framework to estimate $\mathbf{B}$

  - 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"**

(Evensen, 1997 ; Fisher, 2004 ; Berre et al., 2007)



(From Ehrendorfer 2006)

- Only small size ensembles ($10 \rightarrow \mathcal{O}(10^2)$) are affordable
  $\Longrightarrow$ 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
  $\Longrightarrow$ 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
  $\Longrightarrow$ 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
  $\Longrightarrow$ 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

- Present an application of the filter to a real NWP ensemble

- Only small size ensembles ($10 \rightarrow \mathcal{O}(10^2)$) are affordable
  $\Longrightarrow$ 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

- Present an application of the filter to a real NWP ensemble

- Estimate the impact of this filter on forecast scores

- Only small size ensembles ($10 \rightarrow \mathcal{O}(10^2)$) are affordable
  $\Longrightarrow$ 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)

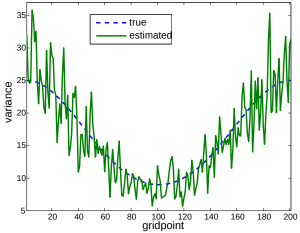
- Aims of the presentation :

- Introduce an objective filtering method for ensemble-based variances

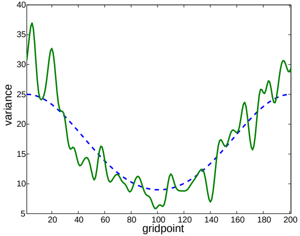
- Present an application of the filter to a real NWP ensemble

- Estimate the impact of this filter on forecast scores

- Give some more general results about the benefits of using errors of "the day"



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



(b) $N = 50, L_{\epsilon^b} = 1000\text{km}$

Spatial structure of sampling noise
(Fisher and Courtier 1995 (Fig 6), Raynaud et al., 2008)

True variance field

$\mathbf{V}^\star \sim$ large scale

$\Downarrow$

Sampling noise

$\mathbf{V}^e = \tilde{\mathbf{V}}(N) - \mathbf{V}^\star \sim$ large scale too ?

$\rightarrow$ depend on L_{ϵ^b}

$\Longrightarrow$ **Close link between the spatial structures of sampling noise and background-error**

- **Notations**
  -$\tilde{\mathbf{B}}$: the estimated $\mathbf{B}$ matrix,

  -$\tilde{\mathbf{B}}^\star = E[\tilde{\mathbf{B}}]$: the noise-free estimated $\mathbf{B}$ matrix,

  -$\mathbf{V}^e = \tilde{\mathbf{B}} - \tilde{\mathbf{B}}^\star$: the sampling noise or *random error component*.

- **Notations**
  -$\tilde{\mathbf{B}}$: the estimated $\mathbf{B}$ matrix,

  -$\tilde{\mathbf{B}}^\star = E[\tilde{\mathbf{B}}]$: the noise-free estimated $\mathbf{B}$ matrix,

  -$\mathbf{V}^e = \tilde{\mathbf{B}} - \tilde{\mathbf{B}}^\star$: the sampling noise or *random error component.*

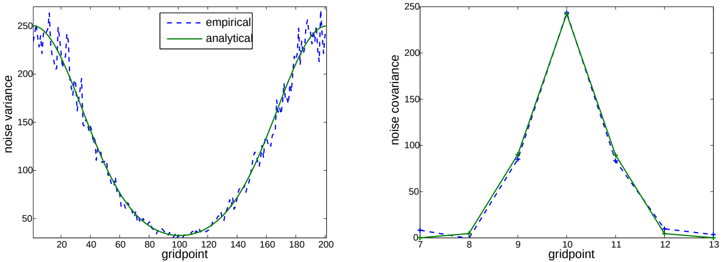
- Analytically, it can be shown that the noise covariance matrix is

$$\boxed{E[\mathbf{V}^e \mathbf{V}^{e^T}] = \frac{2}{N-1} \tilde{\mathbf{B}}^\star \circ \tilde{\mathbf{B}}^\star},$$

  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, $\frac{E[(\mathbf{V}^e)^2]}{(\tilde{\mathbf{V}}^\star)^2} = \frac{2}{N-1}$, is inversely proportional to the ensemble size N.

- Verification of the analytical formula ($N = 6$ and $N_{exp} = 1000$)



$\Longrightarrow$ Very good agreement between empirical and analytical results.

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

$$\boxed{L_{V^e} = \frac{L_{\epsilon^b}}{\sqrt{2}}.}$$

$\Longrightarrow$ The sampling noise $\mathbf{V}^e$ is **smaller scale** than the bkg-error field.

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

Notations

Let's $\mathcal{S}$ be the spherical spectral transform, we define the spectral fields :

$$\tilde{\mathbf{S}} = \mathcal{S}(\tilde{\mathbf{V}}) \qquad \tilde{\mathbf{S}}^\star = \mathcal{S}(\tilde{\mathbf{V}}^\star) \qquad \mathbf{S}^e = \mathcal{S}(\mathbf{V}^e)$$

Formulation

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

$$\boxed{\boldsymbol{\rho} = \frac{1}{1 + \frac{\mathbf{P}(\mathbf{S}^e)}{\mathbf{P}(\tilde{\mathbf{S}}^\star)}}, \text{ where } \mathbf{P}(\ .\ ) \text{ is the power spectrum.}}$$

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

Notations

Let's $\mathcal{S}$ be the spherical spectral transform, we define the spectral fields :

$$\tilde{\mathbf{S}} = \mathcal{S}(\tilde{\mathbf{V}}) \quad \tilde{\mathbf{S}}^\star = \mathcal{S}(\tilde{\mathbf{V}}^\star) \quad \mathbf{S}^e = \mathcal{S}(\mathbf{V}^e)$$

Formulation

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

$$\boxed{\boldsymbol{\rho} = \frac{1}{1 + \frac{\mathbf{P}(\mathbf{S}^e)}{\mathbf{P}(\tilde{\mathbf{S}}^\star)}}, \ \text{where } \mathbf{P}(\ .\ ) \text{ is the power spectrum.}}$$

- $\boldsymbol{\rho}$ is a simple function of the **noise/signal ratio**,

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

Notations

Let's $\mathcal{S}$ be the spherical spectral transform, we define the spectral fields :

$$\tilde{\mathbf{S}} = \mathcal{S}(\tilde{\mathbf{V}}) \quad \tilde{\mathbf{S}}^\star = \mathcal{S}(\tilde{\mathbf{V}}^\star) \quad \mathbf{S}^e = \mathcal{S}(\mathbf{V}^e)$$

Formulation

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

$$\boxed{\boldsymbol{\rho} = \frac{1}{1 + \frac{\mathbf{P}(\mathbf{S}^e)}{\mathbf{P}(\tilde{\mathbf{S}}^\star)}}, \ \text{where } \mathbf{P}(\ .\ ) \text{ is the power spectrum.}}$$

- $\boldsymbol{\rho}$ is a simple function of the **noise/signal ratio**,
- it can be estimated with the help of the $E[\mathbf{V}^e\mathbf{V}^{e^T}]$ formula.

Application to the Arpège model $\boldsymbol{\sigma}^b$ *(Raynaud et al., 2009)*

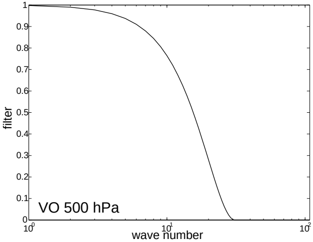
Experimental setup

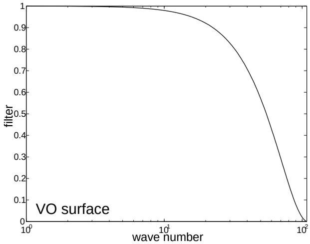
- Météo-France Arpège operational model
- Ensemble of 6 independent 3D-Fgat assimilation experiments *(Berre et al., 2007, operational since July 2008)* :

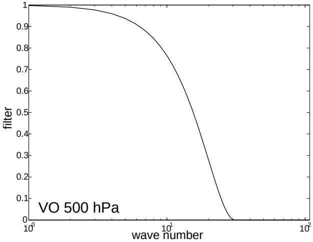
  - explicit perturbation of observations

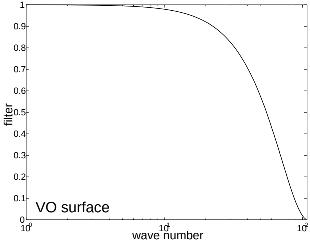
  - implicit perturbation of background

  - perfect model framework

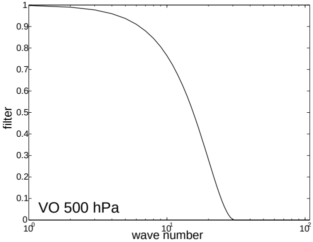

VO 500 hPa
filter
wave number

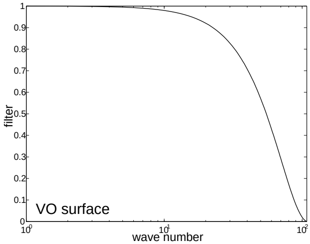

VO surface
filter
wave number



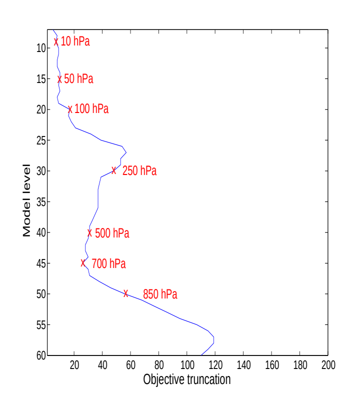


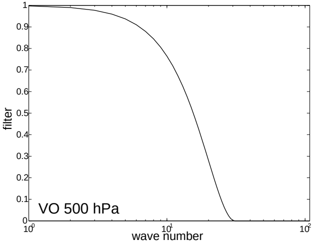
- The truncation of the filter depends on the vertical level : it tends to decrease with altitude.

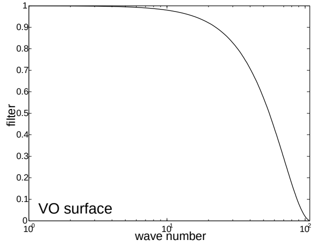




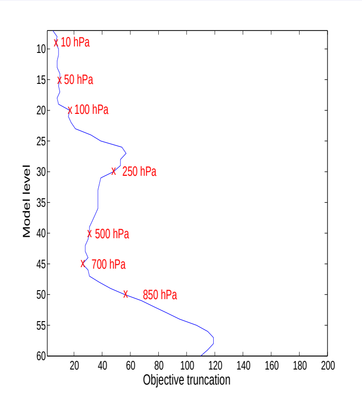
- The truncation of the filter depends on the vertical level : it tends to decrease with altitude.







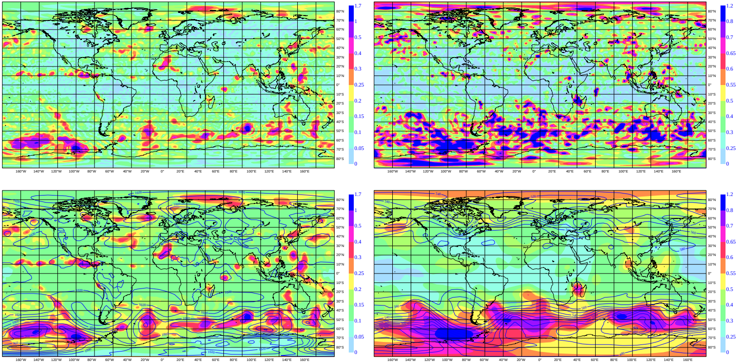
- The truncation of the filter depends on the vertical level : it tends to decrease with altitude.



- Filter values are close to 1 in the largest scales, since these components are well-sampled spatially.

Raw std

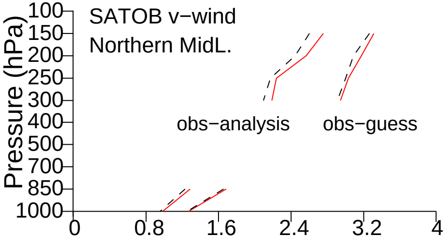
Filtered std

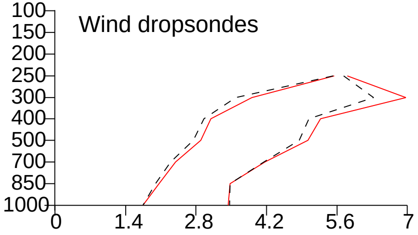


VO surface

VO 500hPa

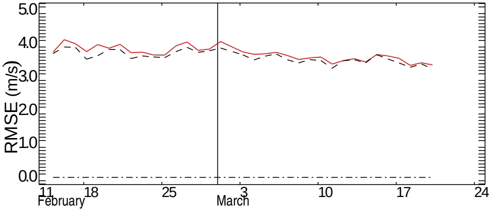
Does spatial filtering of variances have an impact in the (very) end ?
(Raynaud et al., 2009)



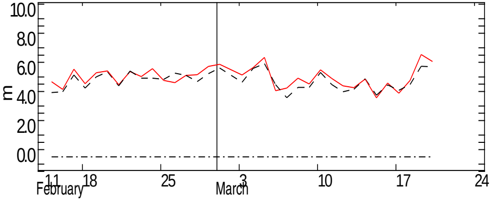


2 impact studies from 15/02/08 to 20/03/08, using :

- vorticity variances “of the day”,
  - either raw
  - or objectively filtered
- climatological variances for other variables.



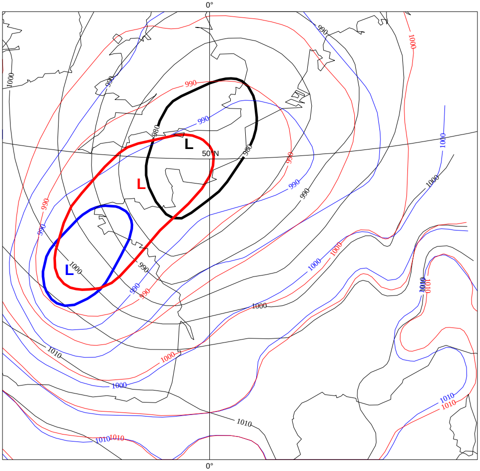
(a) 6h-forecast of 250 hPa wind in Tropics



(b) 6h-forecast of 850 hPa geopotential over Euratl

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 $\boldsymbol{\rho}$

- **Estimation of the noise spectrum $\mathbf{P}(\mathbf{S}^e)$**

  $E[\mathbf{V}^e \mathbf{V}^{e^T}] = \frac{2}{N-1} \tilde{\mathbf{B}}^\star \circ \tilde{\mathbf{B}}^\star \rightarrow$ needs to estimate $\tilde{\mathbf{B}}^\star = E[\tilde{\mathbf{B}}]$

  With *Ergodic + homogeneous* hypotheses :

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

  - $\tilde{\mathbf{B}}_{i.}$ 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, $E[\tilde{\mathbf{B}}_{j.}] = \overline{C}$ (1D) and :

$$\boxed{\mathbf{P}(\mathbf{S}^e) = \frac{2}{N-1} L(\overline{C}^2)}$$

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

$$\boxed{\mathbf{P}(\tilde{\mathbf{S}}^\star) = \mathbf{P}(\tilde{\mathbf{S}}) - \mathbf{P}(\mathbf{S}^e)}$$