Use of ensemble assimilation to represent flow-dependent B in 3D/4D-Var

Loïk Berre
Météo-France (CNRM/GAME)

Thanks to Gérald Desroziers
1. Simulation of the error evolution

2. Illustration of flow-dependent features

3. Filtering of sampling noise in variances

4. Filtering of sampling noise in correlations

5. Validation results and issues
Ensemble assimilation (EnDA = EnVar, EnKF, ...) : simulation of the error evolution

\[ \varepsilon_b = M \varepsilon_a + \varepsilon_m \]

Explicit observation perturbations, and implicit (but effective) background perturbations.

(Houtekamer et al 1996; Fisher 2003; Ehrendorfer 2006; Berre et al 2006)
The analysis error equation
(e.g. Daley 1991, Berre et al 2006)

Analysis state : \[ x_a = (I-KH) x_b + K y \]

True state (with \( y^* = H x^* \)):
\[ x^* = (I-KH) x^* + K y^* \]

Analysis error : \[ e_a = (I-KH) e_b + K e_o \]

with \( e_a = x_a - x^* \)

This is true even if \( K \) is suboptimal. NL case ok too.
The analysis perturbation equation

Perturbed analysis: \[ x'_a = (I-KH)x'_b + Ky' \]

Unperturbed analysis: \[ x_a = (I-KH)x_b + Ky \]

Analysis perturbation: \[ \varepsilon_a = (I-KH)\varepsilon_b + K\varepsilon_o \]

with \[ \varepsilon_a = x'_a - x_a \]
Formal comparison with NMC method
(Bouttier 1994, Berre et al 2006)

Analysis error: \[ e_a = (I-KH) e_b + K e_o \]

Analysis perturbation (EnDA): \[ \varepsilon_a = (I-KH) \varepsilon_b + K \varepsilon_o \]

Analysis increment (NMC): \[ dx = -KH e_b + K e_o \]

with \( I-KH \sim \) high-pass filter, whereas \( KH \sim \) low-pass filter.

⇒ Sharper correlations in EnDA than in NMC
  (e.g. Belo Pereira and Berre 2006, Fisher 2003).

⇒ Better simulation of the analysis error equation in EnDA than in NMC.
Simulation of the error evolution : open issue(s)

Ex: reference 4D-Var and +/- high resolution model.

Possible approximations in the ensemble:

• reduce the horizontal resolution of the model.

• approximate the reference 4D-Var gain matrix $K$, either with 3D-Fgat, or with 4D-Var and fewer outer loops, or with ETKF/EnKF: by deriving $K$ from the ensemble « only ».

⇒ What is the best approach (for a given computation cost)?

EnKF-Var hybrid approaches: the « hybrid $K$ » is not accounted for in the analysis perturbation update.

• (Model error representation...)

1. Simulation of the error evolution

2. Illustration of flow-dependent features

3. Filtering of sampling noise in variances

4. Filtering of sampling noise in correlations

5. Validation results and issues
The operational MF ensemble Var assimilation

- 6 perturbed global members T359 L60 with 3D-Fgat (Arpege).

- Spatial filtering of error variances (see later), to further increase the sample size and robustness (~90%).

- Inflation of ensemble B (by $1.3^2$), as in the static approach, to represent model error contributions.

- The Arpege 4D-Var uses these « $\sigma$b’s of the day ». ⇒ operational since July 2008.

- Coupling with six LAM members during two seasons of two weeks, with both Aladin (10 km) and Arome (2.5 km).
LOCAL SIGMAB’s OF THE DAY
(connections with cyclones and troughs)
Connection between large sigmab and intense weather
(08/12/2006, 03-06UTC)

Ensemble spread:
large sigmab over France

Mean sea level pressure:
storm over France

NB: changes in sigmab’s are relatively localized.
Connection between large sigmab and intense weather (15/02/2008, 12 UTC)

Colours:
sigmab field

Purple isolines:
mean sea level pressure

Large sigmab near the tropical cyclone
Outline

1. Simulation of the error evolution
2. Illustration of flow-dependent features
3. Filtering of sampling noise in variances
4. Filtering of sampling noise in correlations
5. Validation results and issues
While the signal of interest is large scale, the sampling noise is rather small scale.

Explanation: \( \text{cor}(V^e[i], V^e[j]) = \text{cor}(\varepsilon^b[i], \varepsilon^b[j])^2 \)
General expectation: increasing the ensemble size reduces sampling noise, whereas the signal remains.

Experimental result: when increasing the ensemble size, small scale details tend to vanish, whereas the large scale part remains.

This indicates/confirms that the sampling noise is small scale, and that the signal of interest is large scale.
The noise contribution is relatively small in the large scales, and large in the small scales.
To minimize estimation errors, apply a filter $\rho$ accounting for amplitudes and structures of signal and noise:

$$V_b^* \sim \rho \; V_b$$

with

$$\rho = \frac{\text{signal}}{\text{signal}+\text{noise}}$$

$\Rightarrow \rho$ is a low-pass filter (as $K$ in data assimilation).
“OPTIMIZED” SPATIAL FILTERING
OF THE SIGMAB FIELD

« TRUE » SIGMAB’S

FILTERED SIGMAB’s (N = 6)

RAW SIGMAB’s (N = 6)

\[ \varepsilon_b = B^{1/2} \eta \]

\[ V_b \sim \rho V_b \]
Outline

1. Simulation of the error evolution

2. Illustration of flow-dependent features

3. Filtering of sampling noise in variances

4. Filtering of sampling noise in correlations

5. Validation results and issues
Implementation of ensemble-based B in 3D/4D-Var

Two possible implementation techniques:

1. additional $\alpha$-control variable,
   with Schur filtering of correlations.

2. $\sigma_b$'s in gridpoint space, and correlations in wavelet space,
   with spatial filtering of variances and correlations.

The two approaches differ (?):

- in the way of filtering sampling noise in ensemble-based covariances,
- balance operators are applied in technique n°2 (also possible in n°1 (?)).
Schur filtering of correlations (Hamill 2008)

Background-error correlations estimated from 25 members of a 200-member ensemble exhibit a large amount of structure that does not appear to have any physical meaning. Without correction, an observation at the dotted location would produce increments across the globe.

Proposed solution is element-wise multiplication of the ensemble estimates (a) with a smooth correlation function (c) to produce (d), which now resembles the large-ensemble estimate (b). This has been dubbed “covariance localization.”

from Hamill, Chapter 6 of “Predictability of Weather and Climate”
Spatial structure of raw correlation length-scale field

Sampling noise: artificial small scale variations.
=> Use spatial filtering techniques, e.g. wavelets. (Pannekoucke et al. 2007)

\[
L(\varepsilon_b) = \frac{1}{\sqrt{-2 \frac{d^2 \text{cor}/ds^2}{s=0}}}
\]

\[
\varepsilon_b = B^{1/2} \eta
\]

\[
N = 10
\]
Wavelet diagonal modelling of B
(Fisher 2003, Pannekoucke et al 2007)

Local spatial averages of correlation functions $\text{cor}(x,s)$:

$$\text{cor}_W(x,s) \sim \sum_{x'} \text{cor}(x',s) \Phi(x',s)$$

with scale-dependent weighting functions $\Phi$:

- increase of sample size thanks to spatial averaging,
- with main geographical variations thanks to local approach.
Wavelet filtering of correlation functions

Wavelet approach: sampling noise is reduced, leading to a lesser need of Schur localization.

\( N = 10 \)
\( L_s = 6000 \, \text{km} \)

(Pannekoucke et al. 2007)
Impact of wavelet filtering on analysis quality

Wavelet approach: sampling noise is reduced, and there is a lesser need of Schur localization. (Pannekoucke et al 2007)
Wavelet filtering of flow-dependent correlations

Synoptic situation (geopotential near 500 hPa)

Anisotropic wavelet based correlation functions (N = 12)

(Lindskog et al 2007, Deckmyn et al 2005)
Wavelet filtering of correlations « of the day »

Raw length-scales

Wavelet length-scales

(Fisher 2003, Pannekoucke et al 2007)

N = 6
1. Simulation of the error evolution

2. Illustration of flow-dependent features

3. Filtering of sampling noise in variances

4. Filtering of sampling noise in correlations

5. Validation results and issues
Innovation-based sigmab estimate (Desroziers et al 2005)

\[ \text{cov}(H \, dx, dy) \sim H B H^T \]

⇒ This can be calculated for a specific date, to examine flow-dependent features, but then the local sigmab is calculated from a single error realization ( \( N = 1 \) )!

⇒ Conversely, if we calculate local spatial averages of these sigmab's, the sample size will be increased, and comparison with ensemble can be considered.
Validation of ensemble sigmab’s « of the day » in HIRS 7 space (28/08/2006 00h) (Berre et al 2007)

```
Ensemble sigmab’s
```

```
« Observed » sigmab’s
cov( H dx , dy ) ~ H B Hᵀ
(Desroziers et al 2005)
```

=> model error estimation.
Estimates of $\sigma_b$ of the day, in HIRS 7 space

$cov(H \ dx, \ dy)$ of 4D-Var

Ensemble 3D-Fgat

Ensemble 4D-Var

(from Gibier, 2009)
REDUCTION OF NORTHERN AMERICA
AVERAGE GEOPOTENTIAL RMSE
WHEN USING SIGMAB’s OF THE DAY

NOV 2006 - JAN 2007 (3 months)

FEB - MARCH 2008 (1 month)

SEPT - OCT 2007 (1 month)

Forecast range (hours)

Height (hPa)
+24h 500 hPa WIND RMSE over EUROPE

(σb’s of the day versus static σb’s)

⇒ Reduction of RMSE peaks (intense weather systems)
Impact on a severe storm (10/02/2009):
36h forecasts versus verifying analysis

With static $\sigma_b$'s

With $\sigma_b$'s « of the day »

$\Rightarrow$ Positive impact on the depth of the low + gradient intensity
Impact of $\sigma_b$ « of the day »
on the forecast of cyclone *Jokwe*

(Montroty 2008)
Beneficial amplification of analysis increments

MSLP analysis increment with STATIC $\sigma_b$’s

$\sigma_b$’s OF THE DAY (850 hPa vorticity)

MSLP analysis increment with $\sigma_b$’s OF THE DAY

Vendredi 19 décembre 2008
Conclusions

- Ensemble assimilation allows analysis/background error cycling to be simulated, and flow-dependent covariances to be estimated.

- Ensemble-based covariances are affected by sampling noise, but optimized filtering techniques can be applied.

- Comparisons with innovation diagnostics:
  - for validation, and for estimation of model error covariances.

- Impact studies: positive impacts
  - on intense/severe weather events (mid-lat. storms, tropical cyclones).

- Open issues: optimization of error simulation, covariance estimation/filtering, and implementation techniques in 3D/4D-Var.
Belo Pereira, M. and L. Berre, 2006: 
The Use of an Ensemble Approach to Study the Background Error Covariances in a Global NWP Model. 

Berre, L., S.E., Stefanescu, and M. Belo Pereira, 2006: 
The representation of the analysis effect in three error simulation techniques. 
{\it Tellus}, {\bf 58A}, 196-209.

Berre, L., O. Pannekoucke, G. Desroziers, S.E. Stefanescu, B. Chapnik and L. Raynaud, 2007: 
A variational assimilation ensemble and the spatial filtering of its error covariances: increase of sample size by local spatial averaging. 
Proceedings of the ECMWF workshop on flow-dependent aspects of data assimilation, 11-13 June 2007, 151-168. (available on line at: 
http://www.ecmwf.int/publications/library/do/references/list/14092007 )

Bouttier, F., 1994: 
Sur la prévision de la qualité des prévisions météorologiques. 
PhD dissertation, Université Paul Sabatier, 240 pages.

Daley, R., 1991: 
Atmospheric Data Analysis. 
Cambridge University Press, 460 pages.

Deckmyn, A., and L. Berre, 2005: 
A wavelet approach to representing background error covariances in a limited area model. 
{\it Mon. Wea. Rev.}, {\bf 133}, 1279-1294.
References

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005:
Diagnosis of observation, background, and analysis error statistics in observation space,

Ehrendorfer, M., 2006:
Review of issues concerning Ensemble-Based data assimilation techniques.
Oral presentation at the Seventh Adjoint Workshop, Obergurgl, Austria.

Fisher, M., 2003:
Background error covariance modelling.
Proceedings of the ECMWF seminar on recent developments in data assimilation
for atmosphere and ocean, 45-63.

Fisher, M., and P. Courtier, 1995:
Estimating the covariance matrices of analysis and forecast error in variational data assimilation.

Hamill, T., 2008:
Chapter 6 of "Predictability of Weather and Climate".

Houtekamer, P., L. LeFaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996:

Isaksen et al 2007:
Use of analysis ensembles in estimating flow-dependent background error variance.
References

Lindskog et al 2007:
Background error variances in HIRLAM variational data assimilation.
Proceedings of the ECMWF workshop on flow-dependent aspects of data assimilation, 11-13 June 2007, 113-123.

Montroty, R., 2008:
Impact d’une assimilation de données à mésoéchelle sur la prévision cyclonique.
PhD dissertation, Université Paul Sabatier, 219 pages.

Pannekoucke, O, L. Berre and G. Desroziers, 2007:
Filtering properties of wavelets for the local background error correlations.

Raynaud L., L. Berre et G. Desroziers, 2008:
Spatial averaging of ensemble-based background error variances.

Raynaud L., L. Berre et G. Desroziers, 2009:
Objective filtering of ensemble-based background error variances.
Thank you
for your attention