Spatial filtering of assimilation ensemble statistics: increase of sample size by local spatial averaging

Olivier Pannekoucke, Loïk Berre, Gerald Desroziers
Simona Stefanescu, Bernard Chapnik and Laure Raynaud

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
Introduction and Plan

Importance of sample size:
- amplitude of sampling noise in the estimated covariances.
- the ensemble size influences the ensemble cost.

1 - Increase of sample size by local spatial averaging

2 - Spatial filtering of standard deviations

3 - Spatial filtering of correlations
1 - Increase of sample size by local spatial averaging
Strategies for ensemble size and cost reduction
(assimilation ensemble at Météo France)

- Experiments: a small number of members (e.g. 3 to 10) already provides a lot of information!

- Use ergodic properties: increase sample size by spatial averaging.

- The full assimilation system may be approximated in the error simulation (e.g. resolution or/and 4D-Var vs 3D-Fgat).

- Six global members T359 L46 with 3D-Fgat are running in nearly real time (Arpège).
Increase of statistical sample size through spatial sampling/averaging

Basic idea: MULTIPLY(!) the statistical sample size by a number Ng of gridpoint samples.

- An "ideal" case: local homogeneity of covariances and (relatively) short correlation length-scales.

- Another way to justify spatial filtering: sampling noise ~ small scale.

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
2 - Spatial filtering of standard deviations
Illustration in a simulated framework

True $\sigma_b$

RAW 6-member estimated $\sigma_b$

FILTERED 6-member estimated $\sigma_b$

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
The spatial structure of signal and noise from two independent 3-member ensembles

Common features ~ signal
Differences ~ sampling noise
The sampling noise is relatively small scale, which justifies the application of spatial filtering.
Can we design an objective and optimal filter?

\[ \rho = \text{cor}(S_1, S_2) \]

⇒ The two σb fields are more coherent in the large scales than in the small scales.

⇒ It can be shown that this spectral correlation can be used as an objective and optimal filtering coefficient!
Spatial filtering of standard deviations

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
Validation with innovation diagnostics
( for one specific day and for HIRS-7 )

Ensemble sigmab's

"Observed"
sigmab's

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

(Desroziers et al 2005)

Workshop on Flow-dependent
Impact of sigmab's of the day in the Arpège 4D-Var (~over two months ; versus “climatological” sigmab's)

Positive cases?

=> 7 to 9 major peaks

Europe wind RMS 500hPa
+24h / TEMPs
3 - Spatial filtering of correlations
Two extreme approaches in correlation modeling:

- Ens KF: local correlation functions are calculated for each gridpoint. => heterogeneity, but it needs a large ensemble.

- Var: a global average of correlation functions, via a spectral diagonal approach. => large (spatial) sample, but homogeneity.

=> An attractive compromise is to use wavelets, to calculate a local average of correlations.
Diagnosis of length-scale

\[ L_D = \sqrt{-\frac{1}{\Delta \rho(0)}}. \]

\[ L_{B&B} = \sqrt{\frac{\sigma(\varepsilon_b(x))^2}{\sigma(\partial_x \varepsilon_b(x))^2 - (\partial_x \sigma(\varepsilon_b(x)))^2}} \]

\[ L_{Pb} = \frac{|\delta x|}{\sqrt{2 \left(1 - \rho(\delta x)\right)}} \]

\[ L_{Gb} = \frac{|\delta x|}{\sqrt{-2 \ln \rho(\delta x)}} \]
The sampling noise is ~ small scale, and spatially filtered by wavelets.
Correlation length-scales of the day
(from a 6-member assimilation ensemble; Pannekoucke et al 2007)
The noise contribution is relatively small in the large scales, and large in the small scales.
Length-scales of the day

20/01/2005 12UTC (2/12)

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
Length-scales of the day

20/01/2005 18UTC (3/12)
Length-scales of the day

21/01/2005 00UTC  (4/12)
Length-scales of the day

21/01/2005 06UTC  (5/12)
Length-scales of the day

21/01/2005 12UTC (6/12)

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
Length-scales of the day

21/01/2005 18UTC (7/12)

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
Length-scales of the day

22/01/2005 00UTC (8/12)

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
Length-scales of the day

22/01/2005 06UTC  ( 9/12)

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
Length-scales of the day

22/01/2005 12UTC (10/12)

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
Length-scales of the day

22/01/2005 18UTC (11/12)

Workshop on Flow-dependent aspects of data assimilation 11 - 13 June 2007
Length-scales of the day

23/01/2005 00UTC (12/12)
Conclusions and perspectives

- Using an assimilation ensemble and spatial filtering is a promising way to obtain flow-dependent covariances (+ balances) (in a more realistic way than with a simplified background state dependence)

- The spatial filtering is justified by the small scale structure of sampling noise, and it can be optimized objectively.

- The local spatial averaging allows the sample size to be much increased, the ensemble size being MULTIPLIED by a 2D spatial sample size.
Conclusions and perspectives

- The spatial filtering is costless: it may help to make the ensemble size and cost reasonable.

- First impact experiments and comparisons with innovation diagnostics are encouraging. => operational in 2007-2008 ?

- Applications for assimilation diagnostics and ensemble prediction too.