Ensemble of Data Assimilations methods for the initialization of EPS

Laure RAYNAUD

Météo-France

ECMWF Annual Seminar
Reading, 12 September 2017

METEO FRANCE

Introduction

- Estimating the uncertainty in the initial conditions is a key aspect for both data assimilation and ensemble prediction
- Accurate estimates of the uncertainty in analysis and background states :
  - help better use the observations in DA
    $\Rightarrow$ better analyses and forecasts
  - provide guidance for the design of appropriate initial perturbations for EPS
    $\Rightarrow$ better ensemble forecasts

Introduction

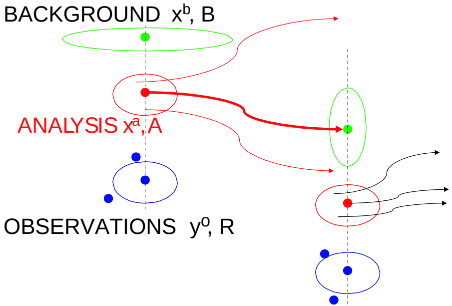
- The data assimilation problem can be described from a Bayesian perspective : $p(x|y^o) = \frac{p(y^o|x) \times p(x)}{p(y^o)}$
- In the last decade, Monte-Carlo methods became operationally feasible approaches to converge to the solution of the Bayesian filtering, e.g., :
  - Ensemble Kalman Filters (N. Bowler's talk)
  - Ensemble of Data Assimilations (EDA, this talk !)
    $\Rightarrow$ ensemble of cycled 3D or 4D-Vars with perturbed observations
    $\sim$ variational counterpart of the EnKF

- **Why running an EDA ?**
  - It provides an ensemble of background states from which the background-error statistics (**B**-matrix) can be estimated
  - It provides an ensemble of analyses that can be used to initialize ensemble predictions.

- **Aim of this talk**
  - Present the principle and configurations of some operational/research EDA systems
  - Present the impact of coupling EPS to EDA, based on experiences at Météo-France and MetOffice.

Plan

1 - EDA : how does it work ?



▷ Simulate the error cycling through evolution of observation perturbations (drawn from $\mathcal{N}(\mathbf{0}, \mathbf{R})$) and model perturbations.

Evolution of true errors ($\mathbf{e}^a = \mathbf{x}^a - \mathbf{x}^t$, $\mathbf{e}^b = \mathbf{x}^b - \mathbf{x}^t$)

$\mathbf{e}^a = (\mathbf{I} - \mathbf{KH})\mathbf{e}^b + \mathbf{K}\mathbf{e}^o$

$\mathbf{e}^b = M\mathbf{e}^a + \mathbf{e}^m$

Evolution of ensemble perturbations

($\boldsymbol{\epsilon}^a = \mathbf{x}^a_{pert} - \mathbf{x}^a_{ctrl}$, $\boldsymbol{\epsilon}^b = \mathbf{x}^b_{pert} - \mathbf{x}^b_{ctrl}$)

$\boldsymbol{\epsilon}^a = (\mathbf{I} - \mathbf{KH})\boldsymbol{\epsilon}^b + \mathbf{K}\boldsymbol{\epsilon}^o$

$\boldsymbol{\epsilon}^b = M\boldsymbol{\epsilon}^a + \boldsymbol{\epsilon}^m$

▷ The evolution of ensemble perturbations is the same as the evolution of exact errors !

1 - Some issues

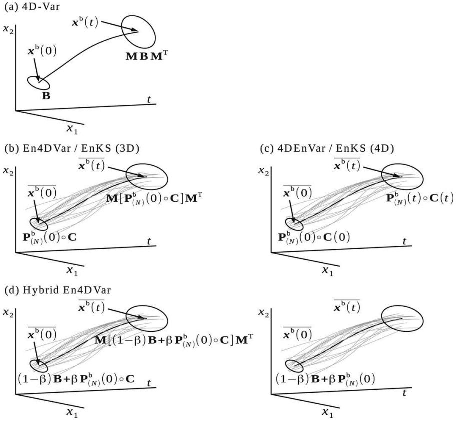
- Natural way to develop an ensemble DA when already running a variational DA scheme (reduced maintenance costs)
- Requires accurate estimates of observation error and model error statistics (**R** and **Q** matrix)
- EDA implementations come with some post-processing tools such as variance filtering and covariance localization to reduce the impact of sampling noise (due to the use of small-size ensembles)
- Consistency with the deterministic 3D/4D-Var schemes and EPS, e.g., choice of horizontal resolution.

1 - Practical applications

▷ Current implementations of EDA systems apply to :

- 3D/4D-Var
- En3D/4D-Var, e.g., Météo-France, ECMWF (operational)
- 3D/4D-EnVar, e.g., UKMO (in test)

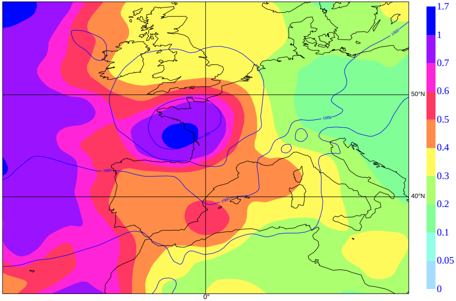
$$J(\underline{\delta \mathbf{x}}) = \frac{1}{2}\underline{\delta \mathbf{x}}^T \underline{\mathbf{P}}^{-1} \underline{\delta \mathbf{x}} + \frac{1}{2}(\underline{\mathbf{y}} - \underline{\mathbf{y}}^o)^T \underline{\mathbf{R}}^{-1} (\underline{\mathbf{y}} - \underline{\mathbf{y}}^o)$$



▷ 4D-Var and En4D-Var use linear model $\mathbf{P} = \underline{\mathbf{M}}\mathbf{P}\underline{\mathbf{M}}^T$

▷ 4D-EnVar uses an ensemble of 4D perturbation trajectories $\underline{\mathbf{P}} = \underline{\mathbf{X}\mathbf{X}}^T$, $\underline{\mathbf{X}} = [\underline{\mathbf{x}}'_1, \underline{\mathbf{x}}'_2, \ldots, \underline{\mathbf{x}}'_N]/\sqrt{N-1}$

Is the EDA a good indicator of analysis uncertainty ?



▷ Connection between large errors and intense weather
(Xynthia storm, 28/02/2010, 03 UTC, vorticity standard deviations from an ensemble of Arpège 4D-Vars)

Plan

2 - Arpège EDA

▷ Global EDA operational at Météo-France since 2008, based on the Arpège 6h 4D-Var scheme (Berre *et al.*, 2007) :

- 25 members with 4D-Var, T479 (40 km) L105, minim T149
- Perturbations of 4D-Var analyses : obs perturbs. (drawn from $\mathbf{R}$) and background perturbs (cycling of analysis perturbs and model perturbs)
- Model error accounted for with a multiplicative inflation (cycled) of forecast perturbations, based on innovation estimates (N. Bowler's talk).

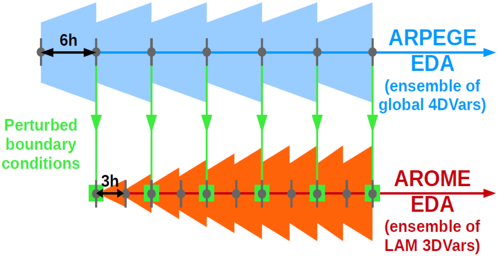
▷ Used to :

- Provide a flow-dependent **B**-matrix to the Arpège 4D-Var
- Provide perturbed initial states to Arpège EPS

2 - Arome EDA

▷ LAM EDA currently being developed at Météo-France, based on the convective-scale Arome 3D-Var scheme :

- 25 members with 3 hourly 3D-Var at 3.8km spatial resolution
- Perturbed observations, perturbed SST, inflation ("spread-skill") scheme
- Planned for operations in 2018

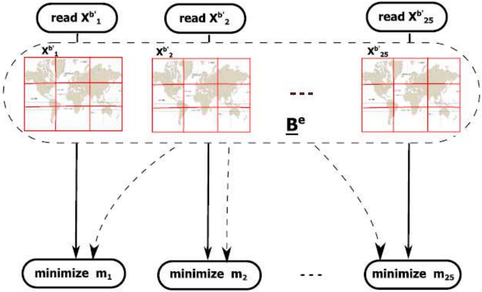


(Courtesy Y. Michel)

2 - Ensemble of 4D-EnVars at UKMO

▷ Early implementations of the En-4DEnVar system at Met Office, in order to replace the operational ETKF (Bowler *et al.*, 2017)

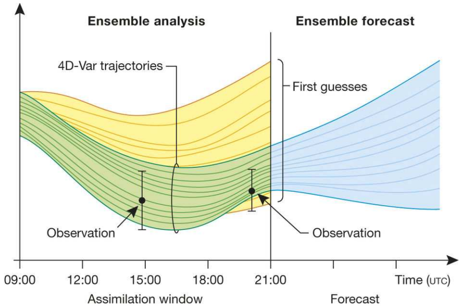
- Tests mainly with 44 members (operational expectation ∼ 100 members) on a 800×600 grid
- Perturbations to SST, soil moisture and temperature
- Relaxation-to-prior-perturbations and spread
- Model error simulation includes additive inflation, SKEB and random parameters
- Self-exclusion to avoid inbreeding



Plan

Why coupling EPS to EDA ?

- EDA provides a flow-dependent estimate of the initial uncertainty
- EDA provides consistent initial perturbations at all scales resolved by the model (providing EDA and EPS resolutions are close ...)



How to couple EPS to EDA ?

- Direct use of perturbed analysis (or background) states to initialize EPS
- Recentre analysis (or background) perturbations around a higher-resolution analysis
- Combine centered EDA perturbations with other perturbation methods (e.g., singular vectors, breeding modes)

Reference

S. Lang, M. Bonavita and M. Leutbecher, 2015 : On the impact of re-centring initial conditions for ensemble forecasts, *Q. J. R. Meteorol. Soc.*, 141, 2571-2581.

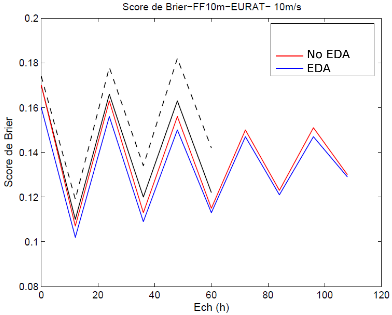
3 - Arpège-EPS

- 34 perturbed members + control run
- Running at : 06UTC (90h range) and 18UTC (108h range)
- Forecasts resolution : T798C2.4L90 (≈10km over Europe, 60km on the opposite side of the globe)
- Initial conditions : combination of Arpège EDA perturbed states (centred on the higher resolution deterministic analysis) with singular vectors
- Model error accounted for with the multiphysics approach, considered to provide a valuable flow-dependent sampling of the uncertainty in the physical parametrizations :
  - 10 different physical parametrization sets, including the Arpège deterministic physical package
  - different schemes for turbulence, shallow convection, deep convection and for the computation of oceanic fluxes.

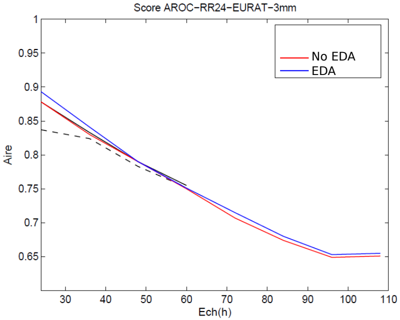
Reference

L. Descamps, C. Labadie, A. Joly, E. Bazile, P. Arbogast and P. Cébron, 2015 : PEARP, the Météo-France short-range ensemble prediction system, *Q. J. R. Meteorol. Soc.*, 141, 1671-1685.

3 - Initializing Arpège-EPS with Arpège-EDA



(a) Brier FF10m



(b) ROCA rr24

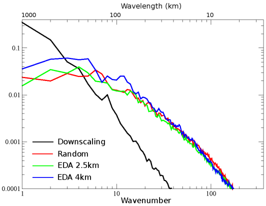
3 - AROME-EPS

- Based on the non-hydrostatic convective-scale Arome-France model with a 2.5km horizontal resolution
- 12 perturbed members
- Running at 09UTC and 21UTC up to 45h
- Initial perturbations and lateral boundary conditions provided by selected runs of the Arpège EPS (through a clustering technique)
  ⇒ these ICPs should be replaced by AROME-EDA by 2018
- Initial perturbations are centred around the high-resolution deterministic analysis (at 1.3km)
- Random perturbations added to some surface variables (including SST, soil temperature and humidity)
- Model error represented with stochastic physics (SPPT scheme).

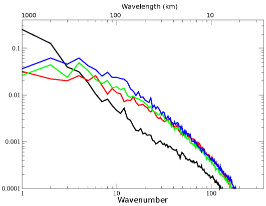
Reference

F. Bouttier, O. Nuissier, B. Vié and L.Raynaud, 2012 : Impact of stochastic physics in a convection-permitting ensemble, *Monthly Weather Review*, 140, 3706-3721.

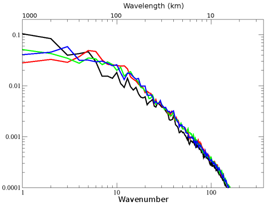
3 - Initializing AROME-EPS with AROME-EDA



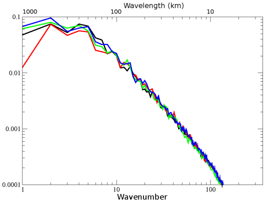
(a) 0h



(b) 1h



(c) 6h



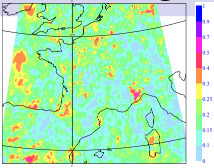
(d) 12h

- EDA perturbations are much more smaller scale than downscaled Arpège EPS perturbations
- Downscaled perturbations show a very large growth rate for smaller scales during the first hours, but it requires about 12h to complete a reasonable spectrum at all scales.

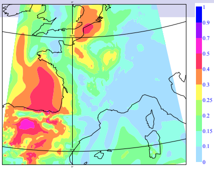
Reference

L. Raynaud and F. Bouttier, 2016 : Comparison of initial perturbation methods for ensemble prediction at convective scale, *QJRMS*, 142, 854-866.

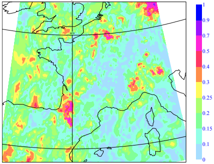
3 - Initializing AROME-EPS with AROME-EDA



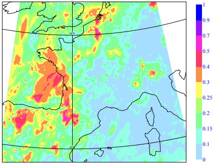
(a) 0h - EDA 2.5km



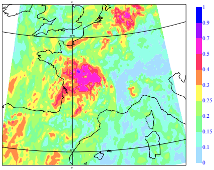
(b) 0h - DOWN



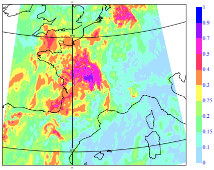
(c) 3h



(d) 3h



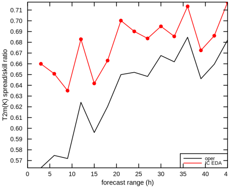
(e) 9h



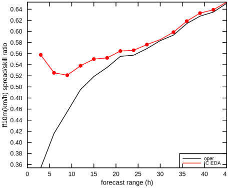
(f) 9h

- In agreement with the spectra, the ensemble spread converges towards that of the EDA-based EPS only from 9h $\Rightarrow$ there should be an impact of the EDA initialization at very short ranges.

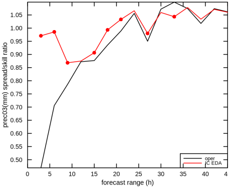
▷ Scores computed over the period 1-28 February 2017.



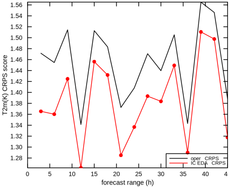
(a) T2m sp/sk



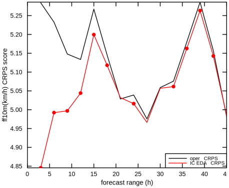
(b) 10m-wind sp/sk



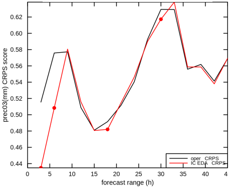
(c) rr3h sp/sk



(d) T2m CRPS



(e) 10m-wind CRPS

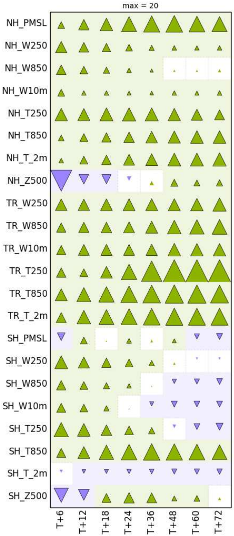


(f) rr3h CRPS

⇒ Large impact of EDA, especially at short ranges.

3 - Initializing MetUM EPS with En-4DEnVar

▷ En-4DEnVar vs ETKF (44 members each, Bowler *et al.*, 2017)



4 - Conclusions and future works

- Ensemble of variational data assimilation methods are straightforward to implement upon existing 3D/4D-Var schemes
- They provide perturbed states that can be used to initialize EPS, both at global and regional scales
- Operational at Météo-France and ECMWF
- Next step is the development of ensembles of 4DEnVars competitive with existing EDA and/or EnKF.

4 - Conclusions and future works

▷ Future works/open questions (non-exhaustive)

- What is the impact of EDA resolution and size on EPS ?
- Until now EDA size $\leq$ EPS size but this may change in the next years with the development of 4D-EnVar schemes
  $\Rightarrow$ Is there a clever way to select the best EDA members to initialize EPS ?
- Recentring of EDA perturbations on a deterministic high-resolution analysis has proved beneficial in several systems, but how does this affect initial balances and short-range EPS skill ?
- Delay between EDA and EPS productions, what is the impact ?
- Initialize EPS with EDA only (without SV for instance) : dream or close reality ?