Data assimilation aspects at Météo-France

C. Fischer
E. Arbogast, L. Berre, G. Desroziers, E. Wattrelot, T. Montmerle
Outline of talk

- Radar obs assimilation
- 4DEnVar
- Plans: obs, applications
Evolution of total number of assimilated observations in AROME-France

Nov'08 : Arome oper with Vr
Spring'10 : Refl oper
Autumn'10 : Improved assim of NoRain Refl
Spring'15 : 1.3km resol Arome
Autumn : Higher density radar obs
Use of foreign data; OPERA; plans

- Tested impact of foreign radars (Spain): beneficial
- Wish to implement OPERA data directly in AROME => requirements for ODIM-HDF5
- Both raw data and cleaned data are required
- Quality flags have to contain the identification of anomalies and the quantification of the quality for an optimal use in NWP

- Near future: improve obs operator (DPOL, X-band),
- Implement some foreign radars in e-suite (D, Be, NL),
- Increase number of radar data in screening in next e-suite,
- Assess importance of initializing precipitating hydrometeors (DPOL)
EnVar for ARPEGE or AROME: perturbations

Use an ensemble data assimilation (EDA) with $L$ members to compute background perturbations:

$$\delta \tilde{x}^b_p = \frac{1}{\sqrt{L-1}} (\tilde{x}^b_p - \langle \tilde{x}^b \rangle) \quad \tilde{B} = \frac{1}{L-1} \sum_{p=1}^{L} (\tilde{x}^b_p - \langle \tilde{x}^b \rangle) (\tilde{x}^b_p - \langle \tilde{x}^b \rangle)^T$$

- Explicit obs. perturb.
- Implicit bkgd perturb.

- Explicit obs. and LBCs perturb.
- Implicit bkgd perturb.

Fisher 2003; Kucukkaraca and Fisher (2006); Berre et al 2006
deterministic 4D-Var at Météo-France
- already uses an ensemble to evolve $C_0$ (auto-correlations).
- wavelet representation of $C_0^w$: provides smoothed correlations.
- covariances evolve implicitly with time: $B_k = M_k K_0^b \Sigma_0^b C_0^w \Sigma_0^{bT} K_0^{bT} M_k^T$.

4DEnVar using evolved $\delta x^b$ perturbations derived from an ensemble of fcts
- Ensemble of $L=150$ members.
- In the present tests, the 150 members also are used to generate the $B_0$ “climatological” matrix for 4D-Var.
Comparison 4D-Var / 4DEnVar “δx^b ens” (L = 150)
Temperature increment at $t_0$ and ~ 850 hPa
Comparison 4D-Var / 4DEnVar "\(\delta x^b_{\text{ens}}\)" (L = 150)
Specific humidity increment at \(t_0\) and ~ 850 hPa
Comparison 4D-Var / 4DEnVar “$\delta x^b \text{ ens}$” ($L = 150$)
Specific humidity increment at $t_0$ and $\sim 850 \text{ hPa}$
4D-Var / 4DEnVar 150 $\delta x^b \ « \text{ ens }$ hybrid

- TEMP q NH
- TEMP T NH
- TEMP V NH
- AMSUB NH

$\delta x^b \ « \text{ ens }$
Overview about EnVar

- 4DEnVar ARPEGE/AROME already possesses quite a few options:
  - Renormalize perturbations using a filtered $\sigma^b$ field.
  - $B$ hybrid.
  - Change of variable in order to allow a unique localization.
  - Advection of localization.
  - Advection of the so-called “climatological” $B$.
  - Optimizations to decrease the numerical cost of the localization (“spectral”).
  - Various possible formulations for the localization.
  - Reminder: all this is coded starting from the OOPS/C++ layer

- State-of-the-art results with 4DEnVar ARPEGE
  - Obstats scores to read with some caution.
  - Randomized $B$ tested in 4DEnVar against pure ensemble $B$: “ens” version seems better than “rand”.
  - Increments of 4D-Var and 4DEnVar still look quite different.
  - 4D aspects require more studies and optimization.
Strategy for further testing for 4DEnVar in ARPEGE

✓ Which size of ens is needed, which one is tractable on HPC?
✓ Ensemble generation methods.
✓ Test 4DEnVar in T149.
✓ Hybridization seems necessary and promising.
✓ Obtain a quasi-optimal and cheap localization approach.
✓ Use all available observations.
✓ Add VarBC.
✓ 4D aspects, external loops, initialization.
✓ Port VAR prototypes to OOPS-IFS based on CY42R3
✓ Documentation.
Plans for observations in ARPEGE and AROME

- assimilation of Lidar winds from ADM-AEOLUS (provided these data are made available by Eumetcast dissemination),

- assimilation of new scatterometer data (ScatSat),

- get prepared for using data from IRS/MTG,

- consider new satellites: China (FY3-C, FY3-D), ATMS and CrIS obs. from JPSS1 (USA, successor of Suomi-NPP),

- AROME:
  - start using radar data provided by OPERA (technical throughput enabling to monitor a few OPERA radars should be available by the end of 2016)
  - study the potential of Mode-S data (ASD-B format),
  - all-sky microwave radiances using a Bayesian inversion approach,
Plans for MF’s NWP applications in a general overview

- Migration to new BULL HPC Phase 2 (ongoing)
- Transfer to operations of Arome-EPS (2.5kmL90, 12 members, twice a day)
- Next E-suite: starting mid 2016 (CY42_op1 ?), includes a new convection scheme and Surfex in Arpège; operational switch beginning of 2017

- 2017-2018:
  - Arome-EDA,
  - Arome-EPS 4 times/day,
  - Arpège new resolution (about 5km over Western Europe),
  - very likely also an increase of the horizontal resolution of Arpège EDA and EPS,
  - GRIB2, etc.
End of the talk

✓ Obrigado pela sua atenção. Questões por favor.
Present configuration of the operational global assimilation in ARPEGE

- **deterministic 4D-Var:**
  - Time window of 6 h.
  - 2 external loops:
    - T1198 C2.2 (7.5 km min) L105 / T149 (~135 km), T399 (~50 km).
  - Jc-DFI, VarBC.
  - \( B_0^{1/2} = K_0^b \Sigma_0^b C_0^{1/2} \), \( C_0 \) in wavelet, \( K_0^b = \) spectral + NL balances.

- **ensemble assimilation:**
  - 25 perturbed 4D-Vars.
  - 1 external loop T479 C1.0 (40 km) / T149 C1.0.
  - Multiplicative inflation of perturbations applied to the 3h fcts.

- \( \Sigma_0^b \) filtered, with the last 25 perturbations and updated every 6 h.
- \( C_0 \) wavelet with the 6 x 25 last perturb. 3h (30 h), updated every 6 h.
From these $L$ sampled perturbations, a localized $B_e$ is computed

\[ B_e = \tilde{B} \circ C = X^b X^{bT} \circ C \]

\[ X^b = [\delta\tilde{x}_1^b, \ldots, \delta\tilde{x}_L^b] \]

with:

The localization matrix $C$ aims at reducing sampling noise by damping covariances:

\[ C = \begin{pmatrix} I_N \\ \vdots \\ I_N \end{pmatrix} \hspace{1cm} C(I_N \ldots I_N) = 1_N C 1_N^T \]

$I_N$ is a $N \times N$ identity matrix, $1_N$ is composed of $M \times (K+1)$ $I_N$ blocks, and $C$ is a $N \times N$ correlation matrix.