The new Arome-France E-suite : data assimilation aspects and general performances

Maud Martet, P. Brousseau, V. Vogt, G. Roux, E. Arbogast, L. Berre
Météo-France, DESR/CNRM/GMAP

5th ACCORD ASW, 1st April 2025, Zalakaros Hungary.

3DEnVar since October 2024

- 3D-Var / 3DEnVar cost function :

3D-Var : $$J(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^T\mathbf{B}^{-1}(\delta\mathbf{x}) + \frac{1}{2}(\mathbf{d}-\mathbf{H}\delta\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{d}-\mathbf{H}\delta\mathbf{x})$$

$$\epsilon_l^b = \frac{1}{\sqrt{N_e - 1}}(\widetilde{\mathbf{x}}_l^b - \langle\widetilde{\mathbf{x}}^b\rangle)$$

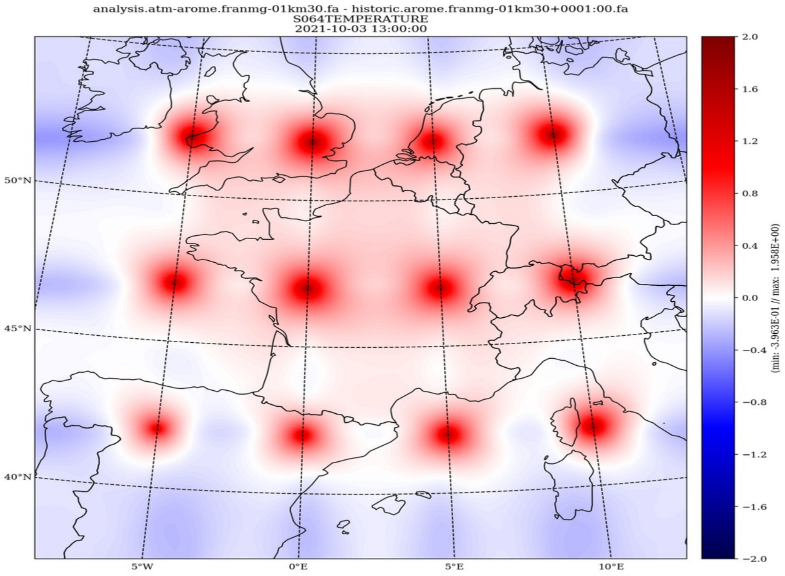
$$\mathbf{X}^b = [\epsilon_1^b, \ldots, \epsilon_{N_e}^b]$$

$$\mathbf{B} = \overline{\mathbf{B}} = \mathbf{K}\mathbf{B}_u\mathbf{K}^T$$

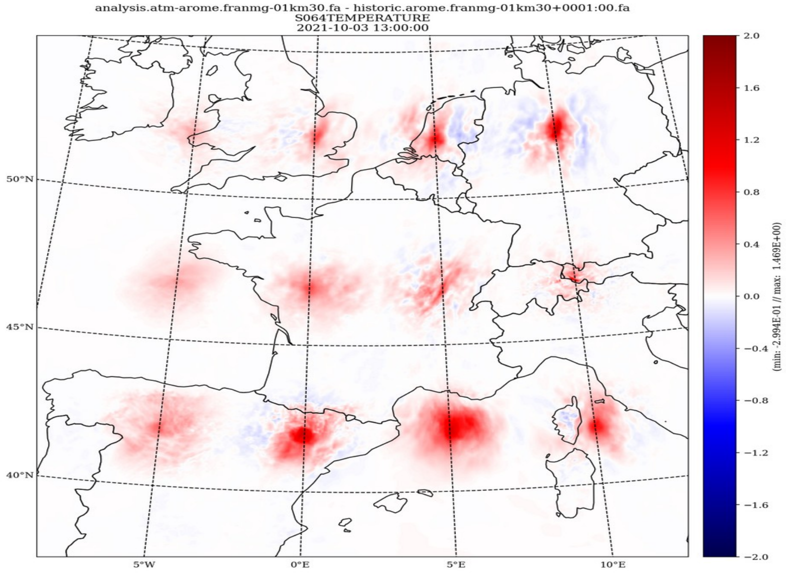
$$\mathbf{B}_u = \begin{pmatrix} \mathbf{C}_\zeta & 0 & 0 & 0 \\ 0 & \mathbf{C}_{\eta_u} & 0 & 0 \\ 0 & 0 & \mathbf{C}_{(T,P_s)_u} & 0 \\ 0 & 0 & 0 & \mathbf{C}_{q_u} \end{pmatrix}$$

$$\mathbf{K} = \begin{pmatrix} I & 0 & 0 & 0 \\ \mathcal{MH} & I & 0 & 0 \\ \mathcal{NH} & \mathcal{P} & I & 0 \\ \mathcal{QH} & \mathcal{R} & \mathcal{S} & I \end{pmatrix}.$$

$$\mathbf{B} = \widetilde{\mathbf{B}}_e = \mathbf{C} \circ \mathbf{X}^b\mathbf{X}^{bT}$$



3D-Var



3DEnVar

3DEnVar since October 2024 : localization

- 3D-Var / 3DEnVar cost function :

3D-Var :

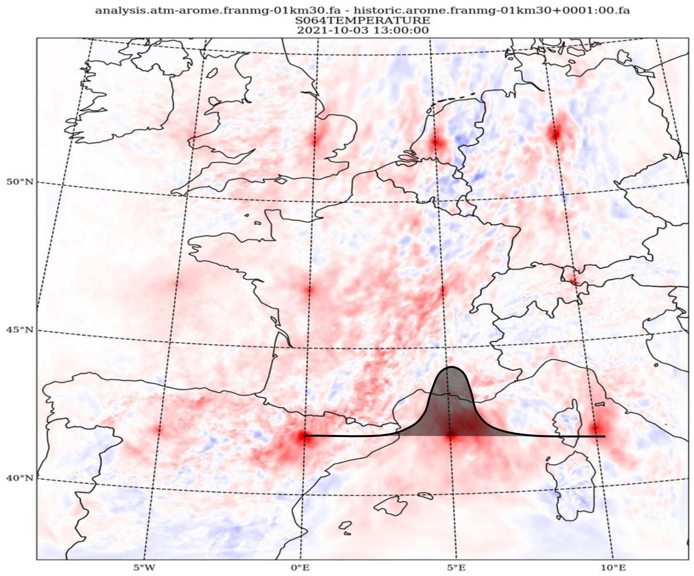
$$J(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^T\mathbf{B}^{-1}(\delta\mathbf{x}) + \frac{1}{2}(\mathbf{d}-\mathbf{H}\delta\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{d}-\mathbf{H}\delta\mathbf{x})$$

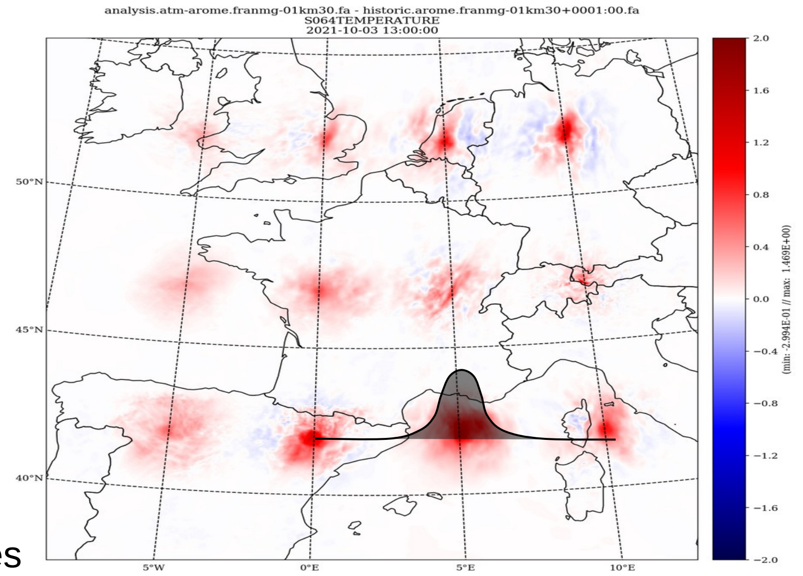
$$\epsilon_l^b = \frac{1}{\sqrt{N_e - 1}}(\widetilde{\mathbf{x}}_l^b - \langle\widetilde{\mathbf{x}}^b\rangle)$$

$$\mathbf{X}^b = [\epsilon_1^b, \ldots, \epsilon_{N_e}^b]$$

$$\mathbf{B} = \widetilde{\mathbf{B}}_e = \mathbf{C} \circ \mathbf{X}^b\mathbf{X}^{bT}$$

Need to localize to filter out the sampling noise





3DEnVar

Currently, localization length :
- Horizontally : varying from 25 km at lowest level to 150 km at highest ones
- vertically : 0.3 log(hPa)

3DEnVar since October 2024 : localization

- 3D-Var / 3DEnVar cost function :

3D-Var :

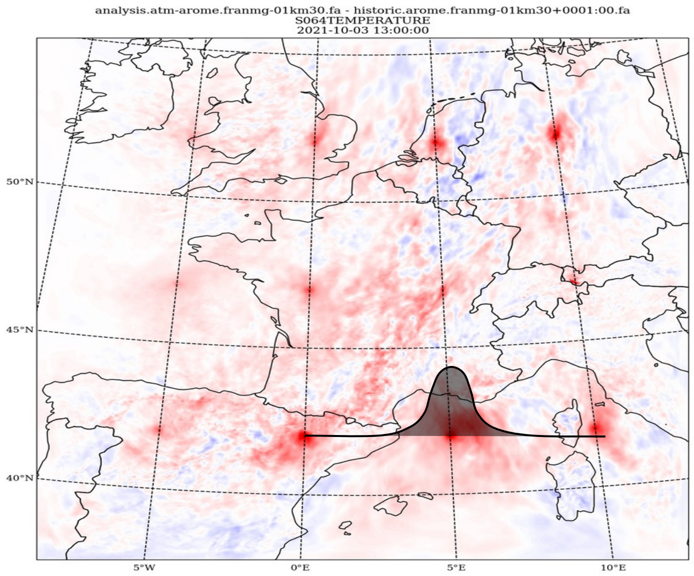
$$J(\delta \mathbf{x}) = \frac{1}{2}(\delta \mathbf{x})^T \mathbf{B}^{-1}(\delta \mathbf{x}) + \frac{1}{2}(\mathbf{d} - \mathbf{H}\delta \mathbf{x})^T \mathbf{R}^{-1}(\mathbf{d} - \mathbf{H}\delta \mathbf{x})$$

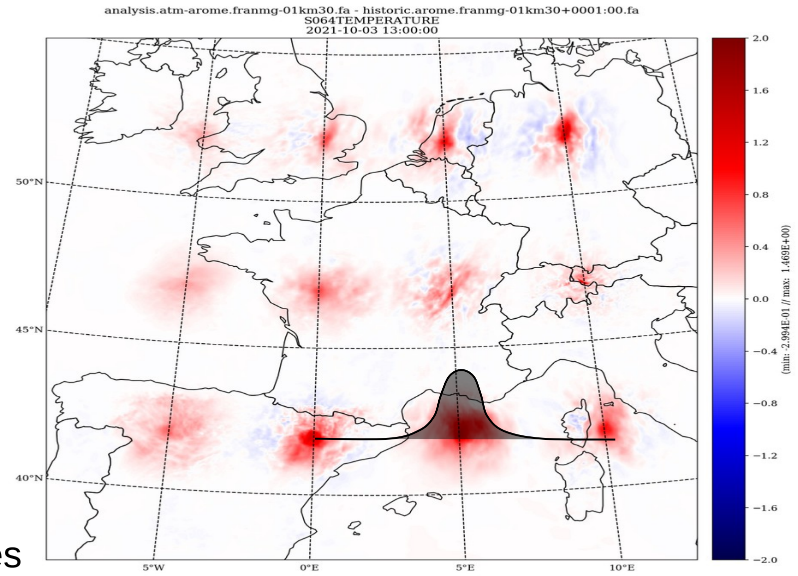
$$\epsilon_l^b = \frac{1}{\sqrt{N_e - 1}}(\widetilde{\mathbf{x}}_l^b - \langle \widetilde{\mathbf{x}}^b \rangle)$$

$$\mathbf{X}^b = [\epsilon_1^b, \ldots, \epsilon_{N_e}^b]$$

$$\mathbf{B} = \widetilde{\mathbf{B}}_e = \mathbf{C} \circ \mathbf{X}^b \mathbf{X}^{bT}$$

Need to localize to filter out the sampling noise





3DEnVar

Currently, localization length :

- Horizontally : varying from 25 km at lowest level to 150 km at highest ones
- vertically : 0.3 log(hPa)

3DEnVar since October 2024 : localization

- 3D-Var / 3DEnVar cost function :

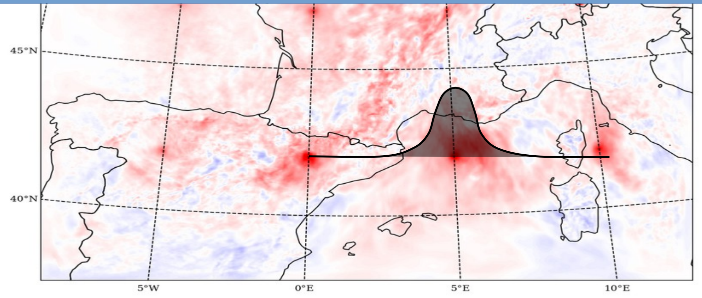
3D-Var : $J(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^T\mathbf{B}^{-1}(\delta\mathbf{x}) + \frac{1}{2}(\mathbf{d}-\mathbf{H}\delta\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{d}-\mathbf{H}\delta\mathbf{x})$

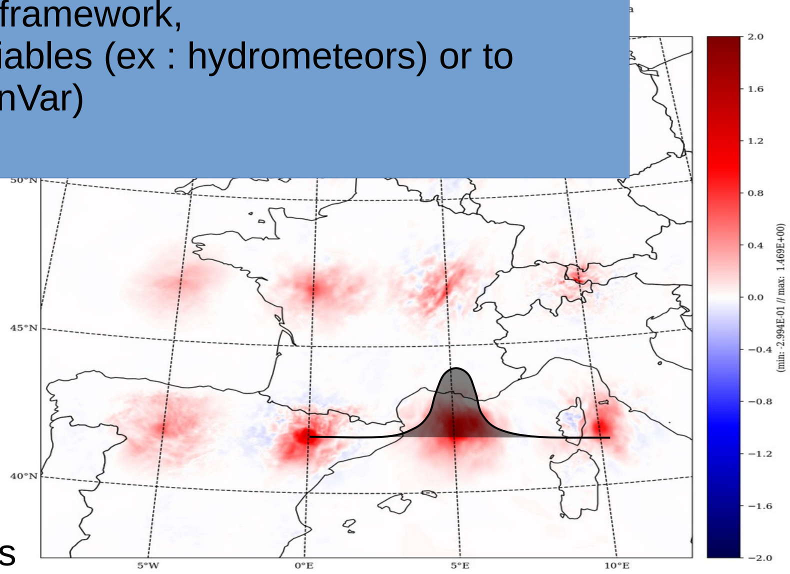
$$\epsilon_l^b = \frac{1}{\sqrt{N_e - 1}}(\widetilde{\mathbf{x}}_l^b - \langle\widetilde{\mathbf{x}}^b\rangle)$$

$$\mathbf{X}^b = [\epsilon_1^b, \ldots, \epsilon_{N_e}^b]$$

Need to locali
filter out the
sampling nois

This new kind of DA scheme allows numerous new developments,
facilitated by the use of the OOPS framework,
such as the extension of the control variable to new variables (ex : hydrometeors) or to the temporal dimension (4DEnVar)





3DEnVar

Currently, localization length :
- Horizontally : varying from 25 km at lowest level to 150 km at highest ones
- vertically : 0.3 log(hPa)

Outlines

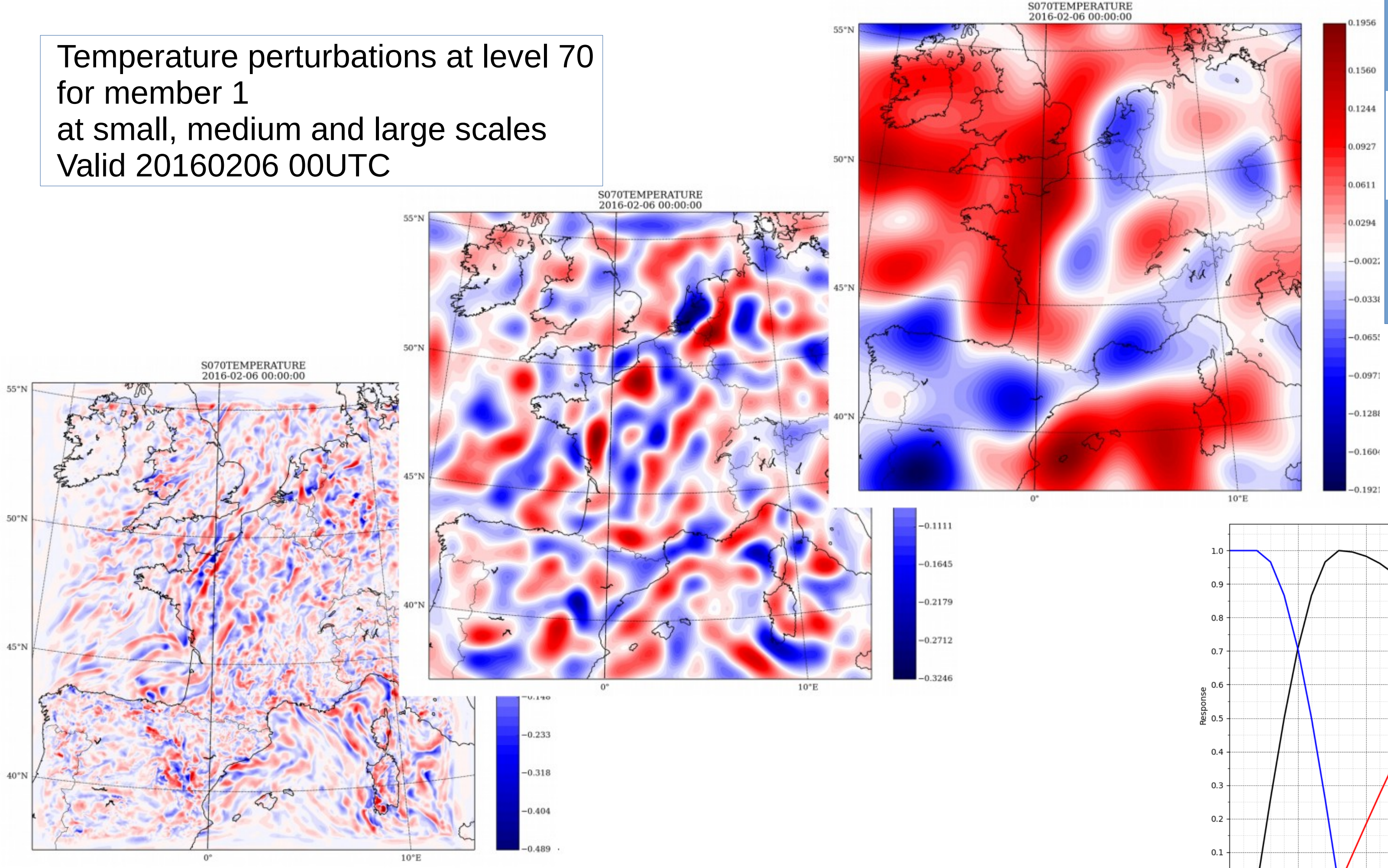
- Descriptions of the main components :
  - Scale Dependant Localisation
  - Direct assimilation of radar reflectivities
  - 4DEnVar

- General performances

- Links between EDA, DA and EPS

Scale Dependent Localization (1)

Temperature perturbations at level 70
for member 1
at small, medium and large scales
Valid 20160206 00UTC

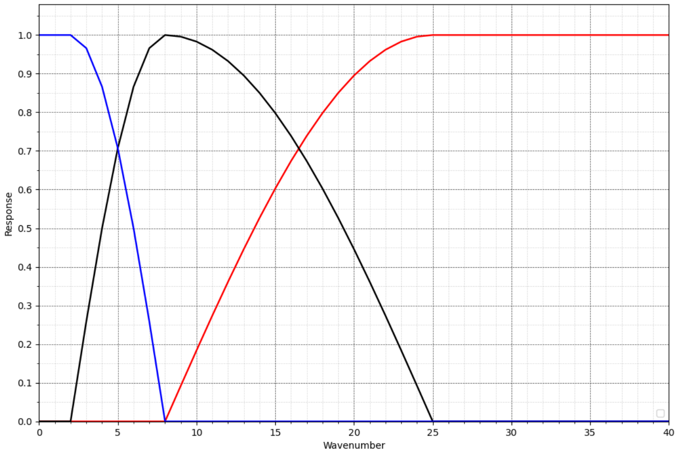


Analysis increment computed from control vector ($B^{1/2}$ preconditioning)

$$\Delta \mathbf{x}_o = \sum_k \sum_j \mathbf{e}_{k,j} \circ \left(\mathbf{L}_j^{1/2} \xi_k \right)$$

$e_{k,j}$ is scale j of normalized member k perturbation

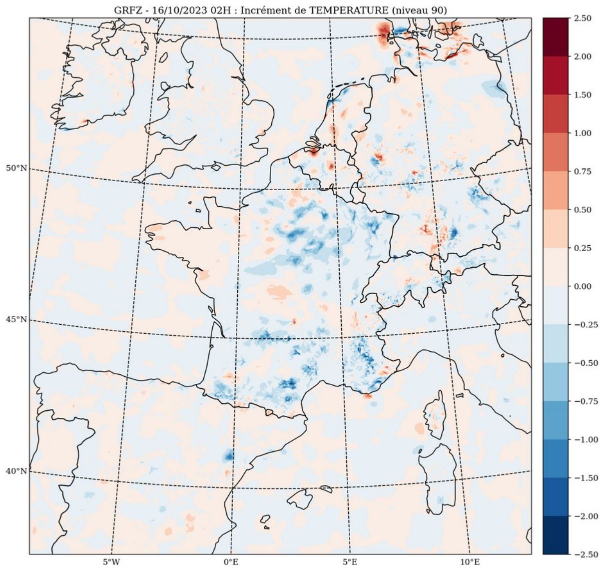
Filters response functions



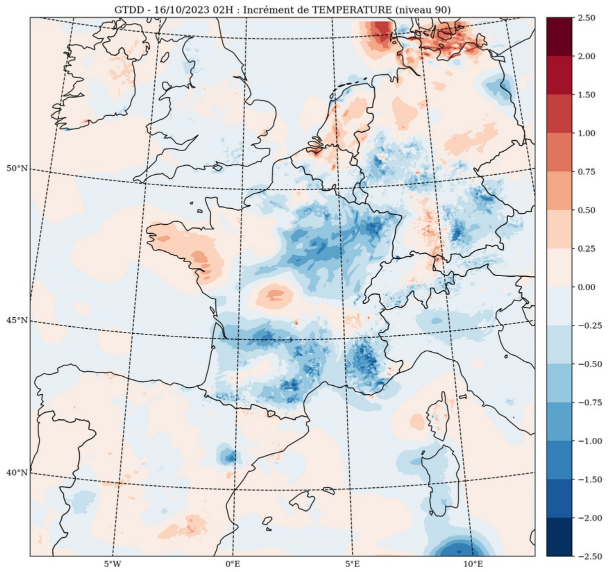
https://opensource.umr-cnrm.fr/attachments/download/5759/vogtv_sdl_vf.pdf

Scale Dependent Localization (2)

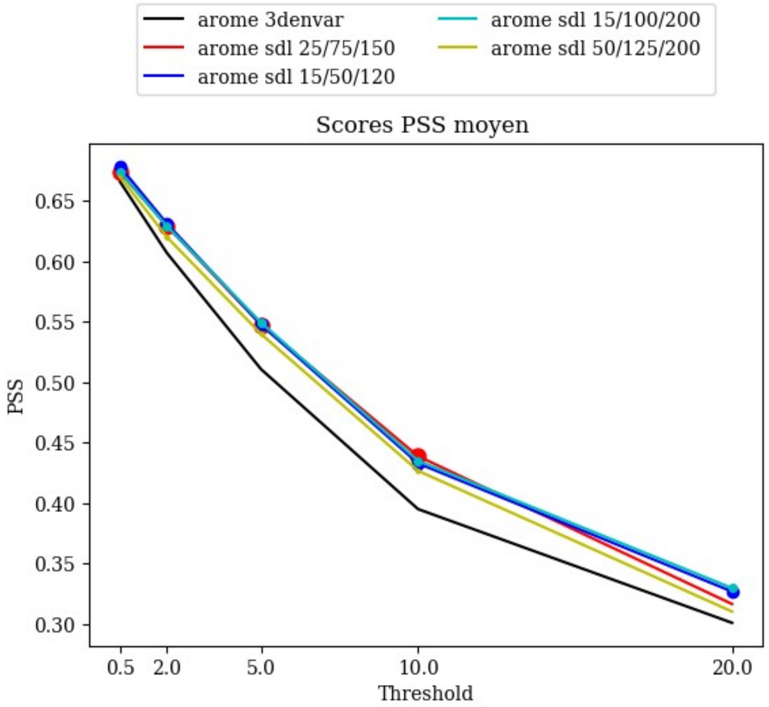
Temperature increment at level 90



REF 3DEnVar
height dependent localization
bottom 25km – top 150km



3DEnVar + SDL
band dependent localization
25/75/150km



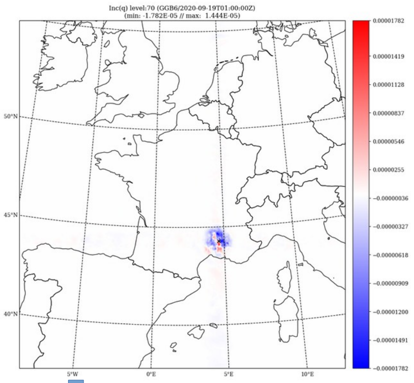
Pierce skill score computed on 12-hour accumulations from 0 and 12h runtimes
Cycled experiments
for several triplets of localization
28/08/2022-31/10/2022

=> 3DEnVar + SDL is used in the arome EDA (AEARO) E-suite instead of 3D-Var
=> SDL compensates for too severe localization of large-scale fields near the surface
=> SDL reduces the need for localization retuning (hydrometeors, 4D, …)

Direct assimilation of reflectivities (1)

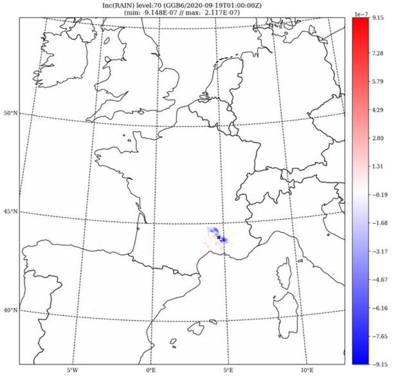
Bayesian inversion

Specific humidity level 70



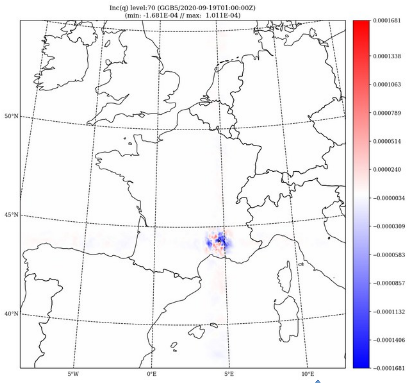
Cross-covariances

Rain content level 70



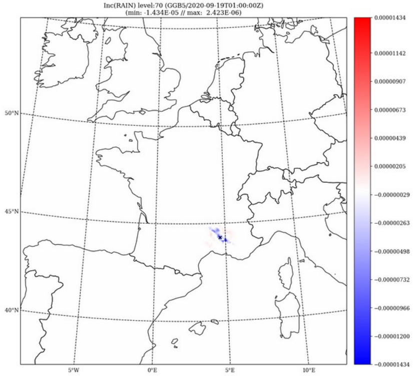
Reflectivity assimilation

Specific humidity level 70



Cross-covariances

Rain content level 70



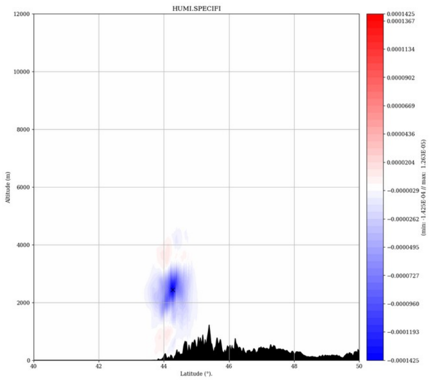
https://opensource.umr-cnrm.fr/attachments/5616

Hydrometeors in the control variable of a 3DEnVar assimilation system with OOPS allows to directly assimilate reflectivity provided you have **TL and AD** versions of observation operator.

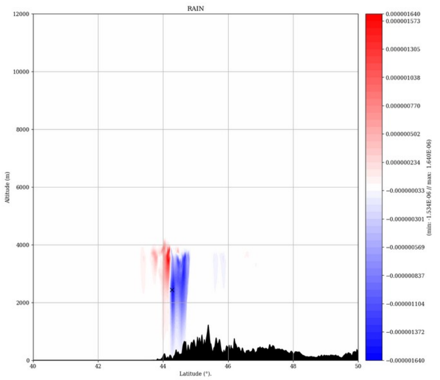
Assimilation of one observation with two different methods :

- pseudo-observation of humidity calculated by Bayesian inversion
- direct assimilation of reflectivity

Reflectivity assimilation



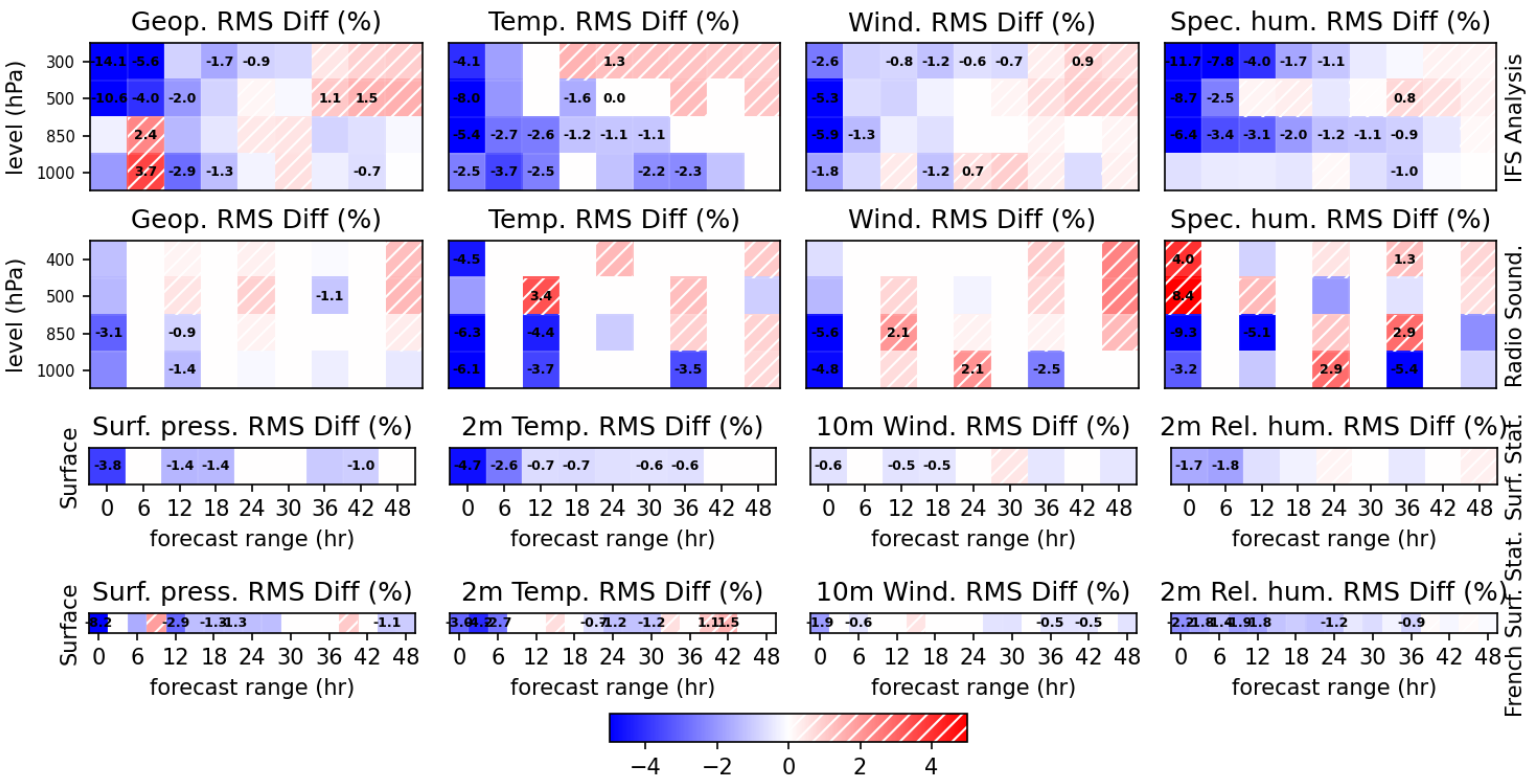
Specific humidity



Rain content

Direct assimilation of reflectivities (2)

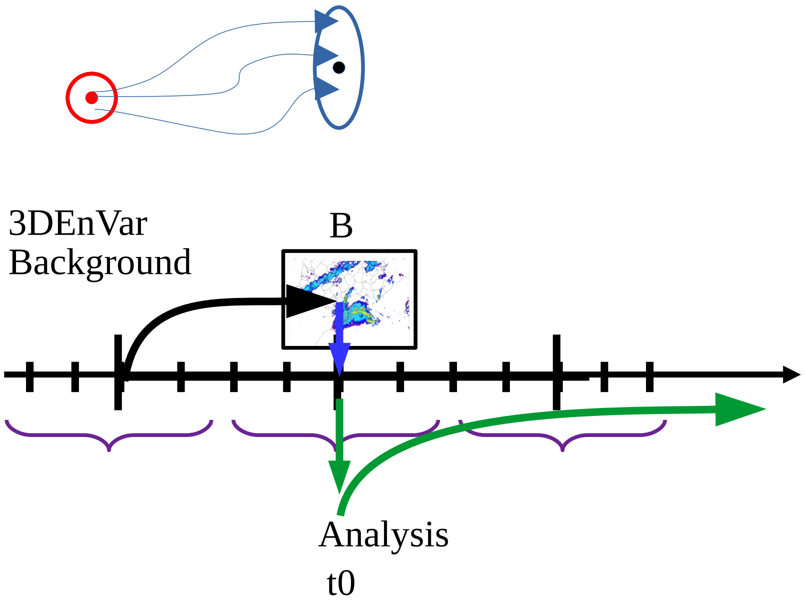
Relative RMSE differences between oper and dble against different references (IFS analysis, Radiosounding and surface station measurements) for different parameters depending on the forecast range :
- blue boxes : assim z better than 1D bayesian method
- numerical values in boxes : the difference is statistically significant at 95 % level of confidence using a bootstrap test



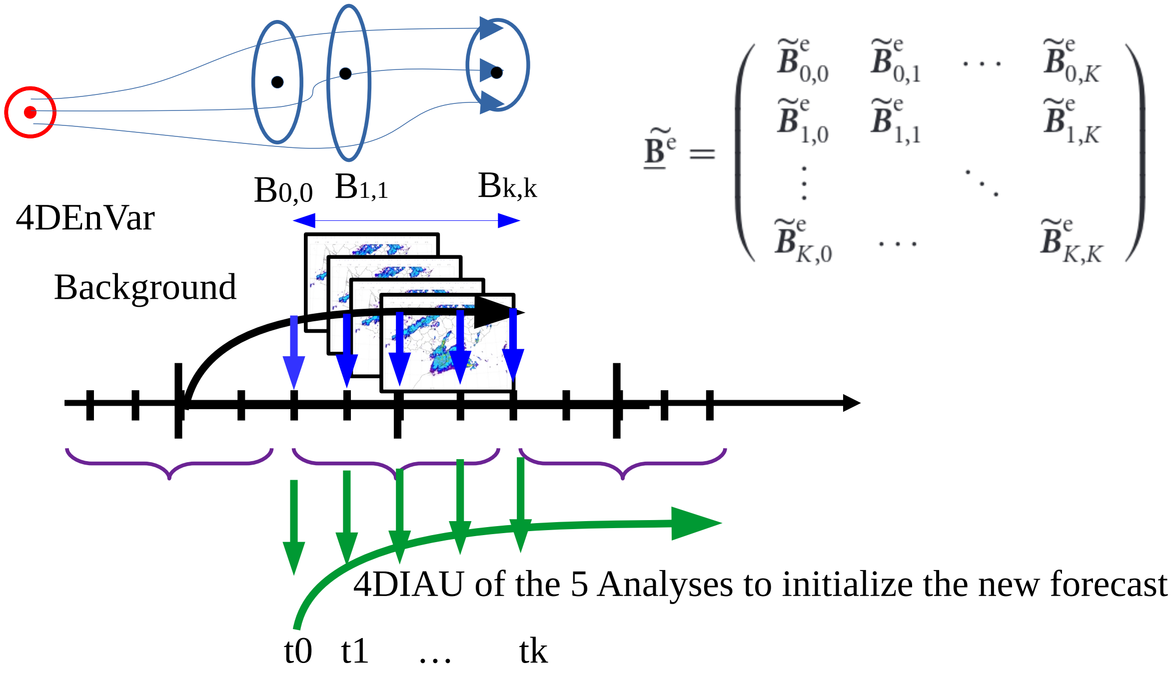
4DEnVar : minimization of a 4D state without any TL/AD forecast models

https://www.accord-nwp.org/IMG/pdf/brousseau_arome4denvar_asw2024.pdf

3D Control Variable
Cv=(U,V,T,q*,Ps)

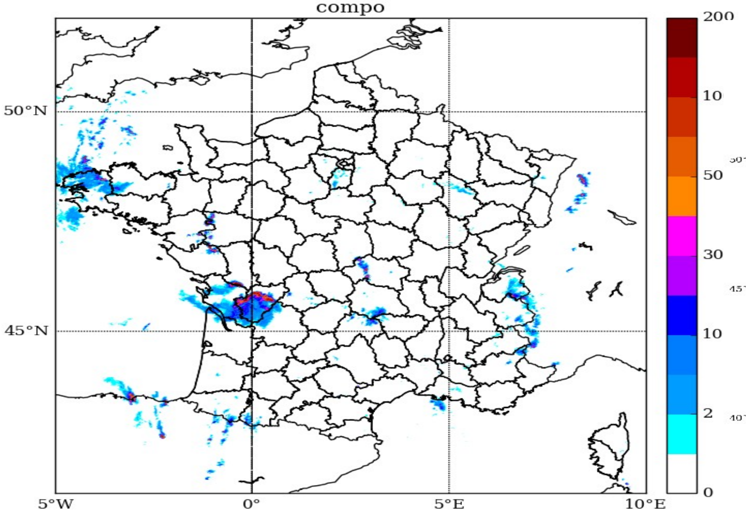


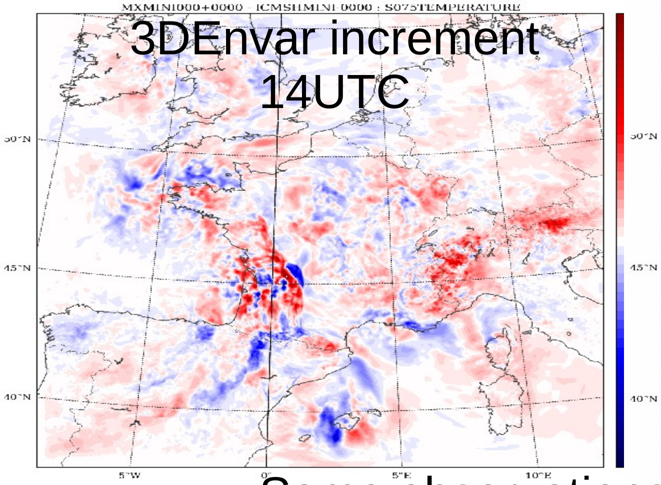
4D: Cv=(U0,V0,T0,q*0,Ps0, U1,V1,T1,q*1,Ps1, …,Uk,Vk,Tk,q*k,Psk)

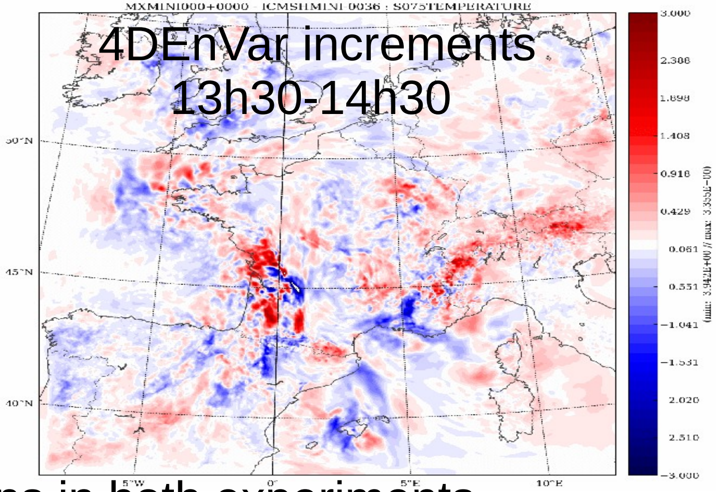


$$\underline{\widetilde{\mathbf{B}}}^e = \begin{pmatrix} \widetilde{B}^e_{0,0} & \widetilde{B}^e_{0,1} & \cdots & \widetilde{B}^e_{0,K} \\ \widetilde{B}^e_{1,0} & \widetilde{B}^e_{1,1} & & \widetilde{B}^e_{1,K} \\ \vdots & & \ddots & \\ \widetilde{B}^e_{K,0} & \cdots & & \widetilde{B}^e_{K,K} \end{pmatrix}$$

4D Temperature increment at 850 hPa







Same observations in both experiments

Retained configuration :

- 1h Cycle with 5 timeslots : 3*15min + 2*7min

- assimilating all operational observations with 15 min radar, Ground-based GNSS, SEVIRI radiances (geostationary platforms) and surface stations

- localization : same as in 3DEnvar

- Perturbations every 15 minutes from the same perturbed forecasts from Arome-France EDA (4h30-5h30-00h15, 5h30-6h30 and 6h30-7h30 depending on the assimilation time) 3.2km Hydrostatic

Others modifications (observations and model)

Observations (V. Pourret, C. Payan) :

- Modes-S data assimilation optimisation :
  - 4km Adaptative Weighting (AW) to avoid overfitting over Mode-S high density locations (triggered after QC and thinning)
  - New whitelist based on 2024 operational monitoring with 11007 selected aircraft over 14235 (~77 %)

- Revision of AMV quality control and error scheme :
  - Use of a prescribed observation error scheme that takes into account uncertainty in data altitude as a function of vertical wind shear, developed and applied at the MetOffice, then at CEP and EMC. This scheme applies to vertical model wind profiles combining a tracking error and a vertical positioning error, by satellite, by channel and by AMV height calculation method.
  - Revised AMV quality control based on the use of producer IQs (reducing dependence on model information), on orographic-type criteria (instead of geographical criteria), or on the tropopause level.
  - The specified computed error is also used as an additional thinning criterion.*

Model (S. Antoine) :

The deposition of cloud droplets on vegetation is now taken into account, with a constant deposition speed (2cm/s) - Antoine et al. 2023 Expected impact: reduction of liquid water content in the first model levels.

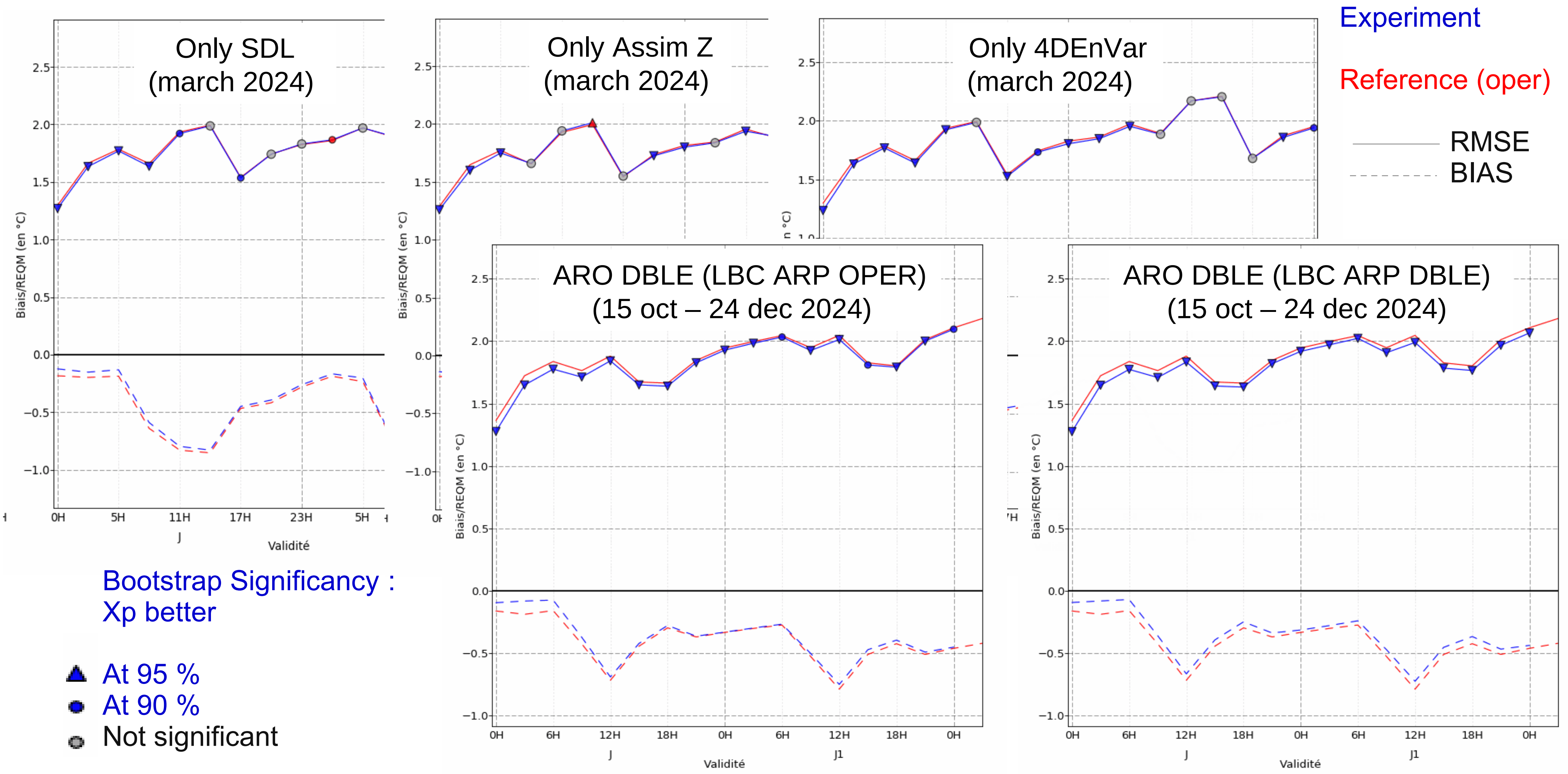
Outlines

- Descriptions of the main components :
  - Scale Dependant Localisation
  - Direct assimilation of radar reflectivities
  - 4DEnVar

- General performances

- Links between EDA, DA and EPS

Impact of the different contributions on T2m scores Versus 3DEnVar oper



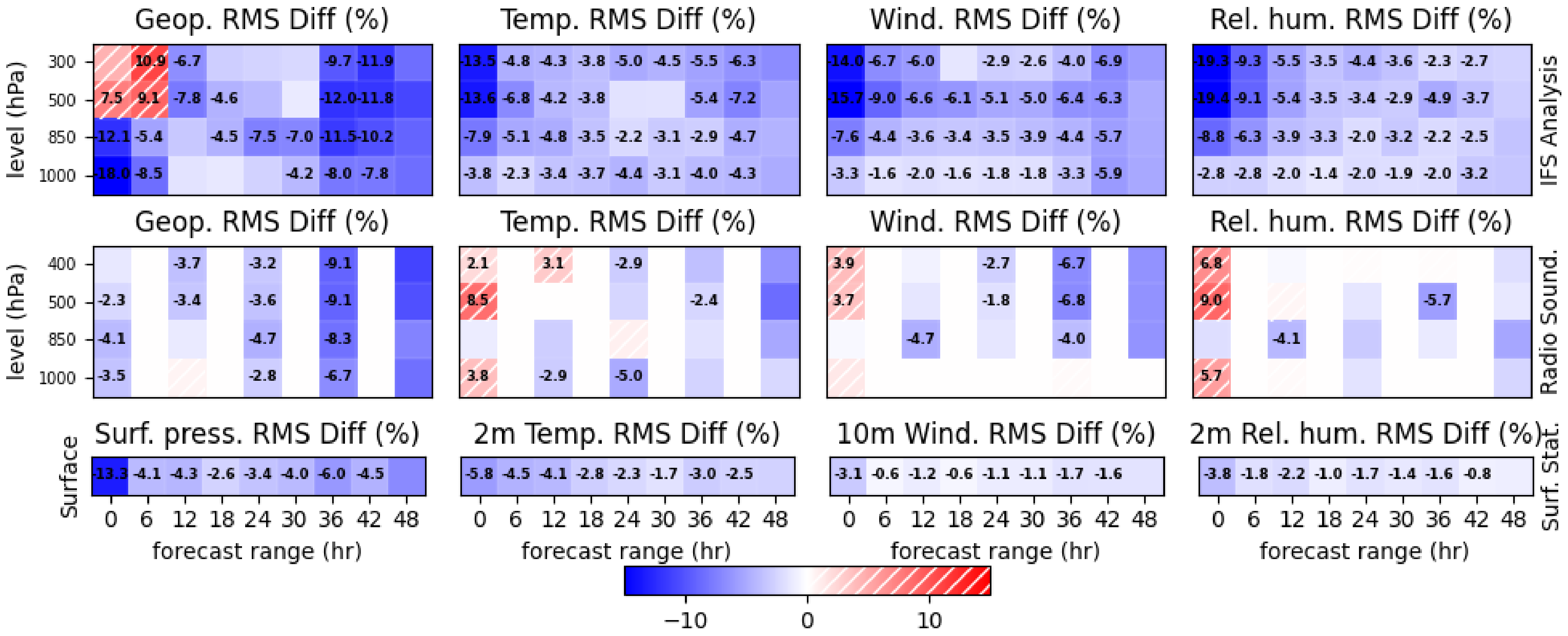
RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

METEO FRANCE

Scorecard 14 Oct. - 24 Dec 2024

Relative RMSE differences between oper and dble against different references (IFS analysis, Radiosounding and surface station measurements) for different parameters depending on the forecast range :

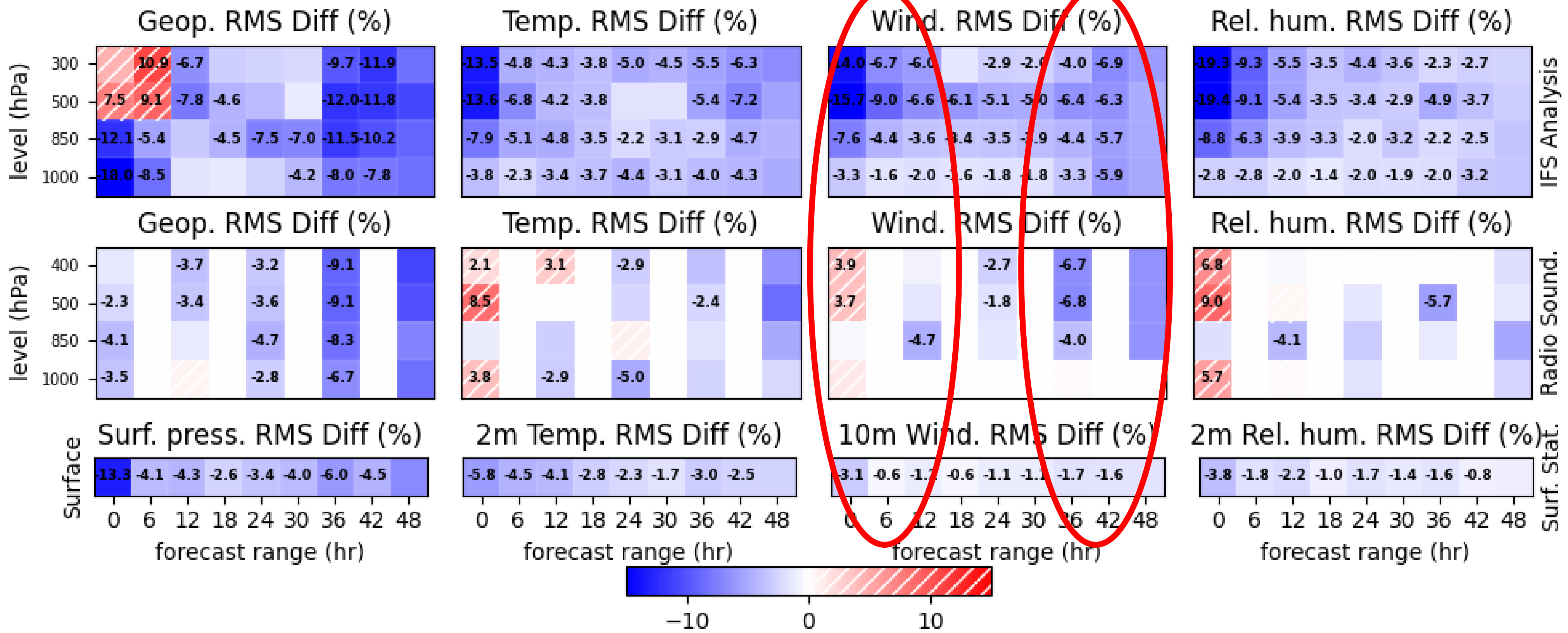
- blue boxes : dble better than oper
- numerical values in boxes : the difference is statistically significant at 95 % level of confidence using a bootstrap test



Scorecard 14 Oct. - 24 Dec 2024

Relative RMSE differences between oper and dble against different references (IFS analysis, Radiosounding and surface station measurements) for different parameters depending on the forecast range :

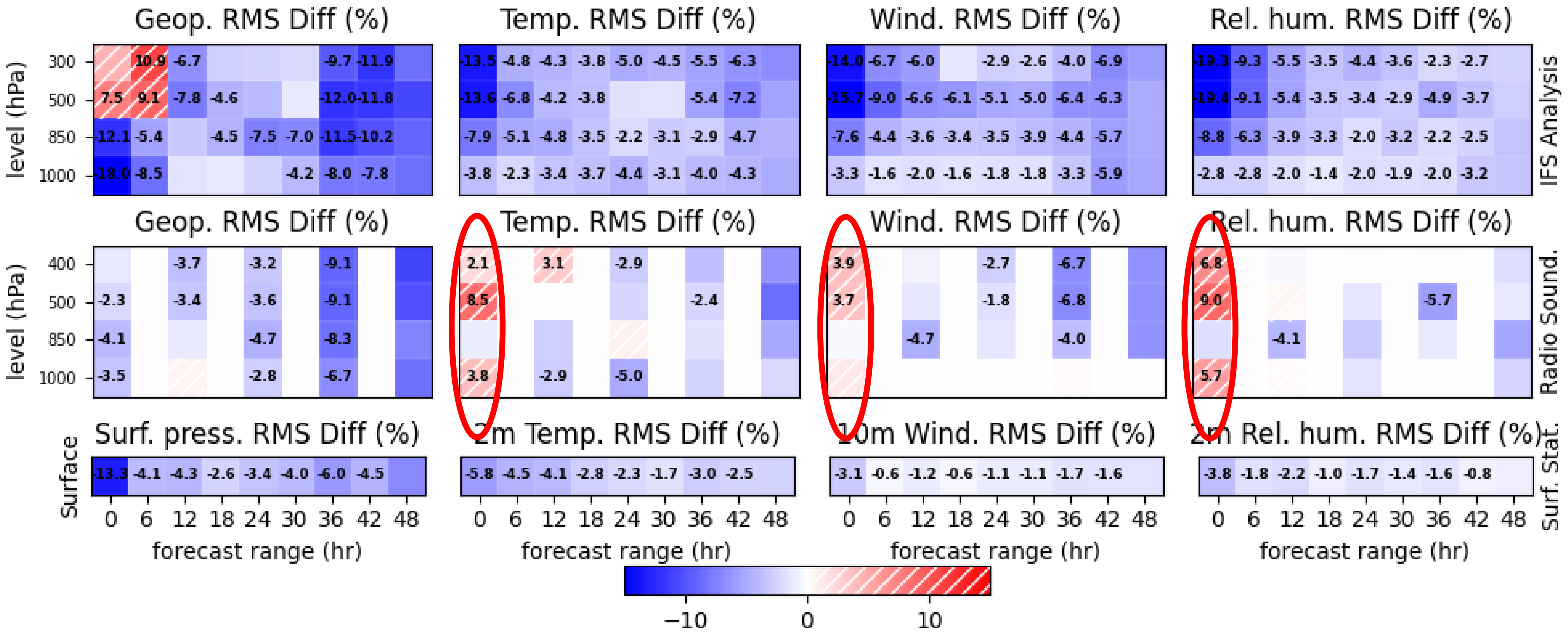
- blue boxes : dble better than oper
- numerical values in boxes : the difference is statistically significant at 95 % level of confidence using a bootstrap test



Scorecard 14 Oct. - 24 Dec 2024

Relative RMSE differences between oper and dble against different references (IFS analysis, Radiosounding and surface station measurements) for different parameters depending on the forecast range :

- blue boxes : dble better than oper
- numerical values in boxes : the difference is statistically significant at 95 % level of confidence using a bootstrap test



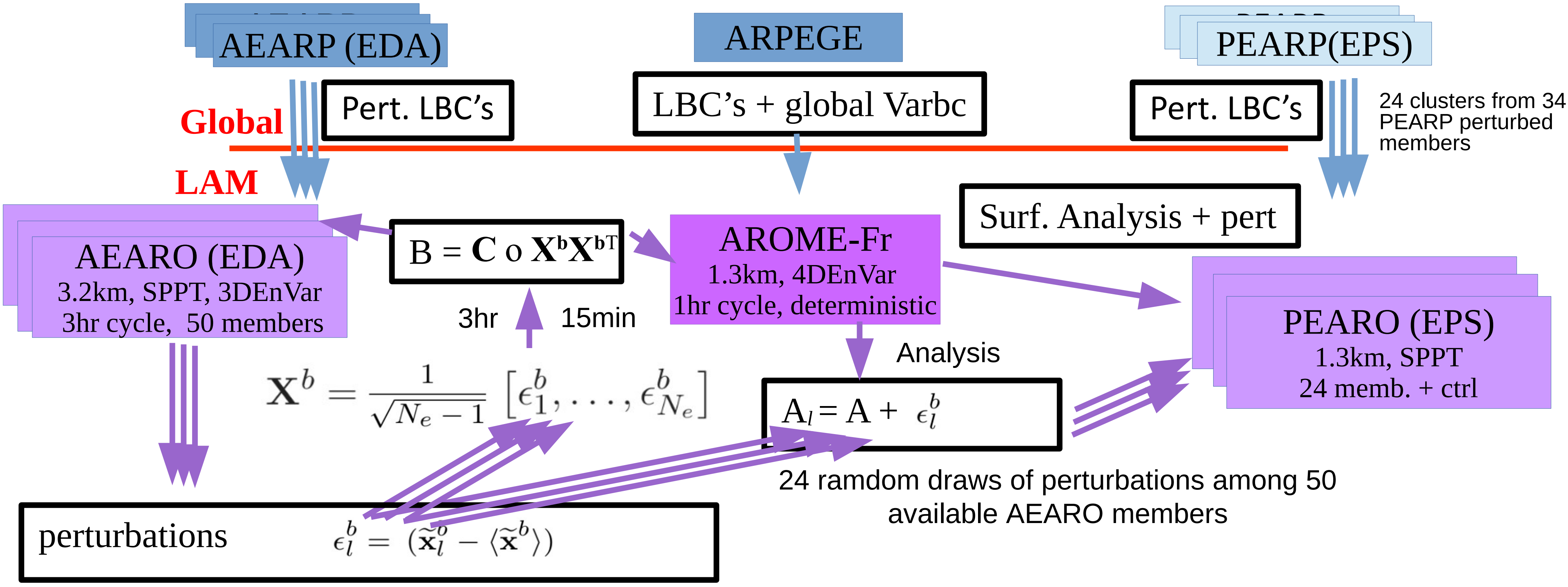
Outlines

- Descriptions of the main components :
  - Scale Dependant Localisation
  - Direct assimilation of radar reflectivities
  - 4DEnVar

- General performances

- Links between EDA, DA and EPS

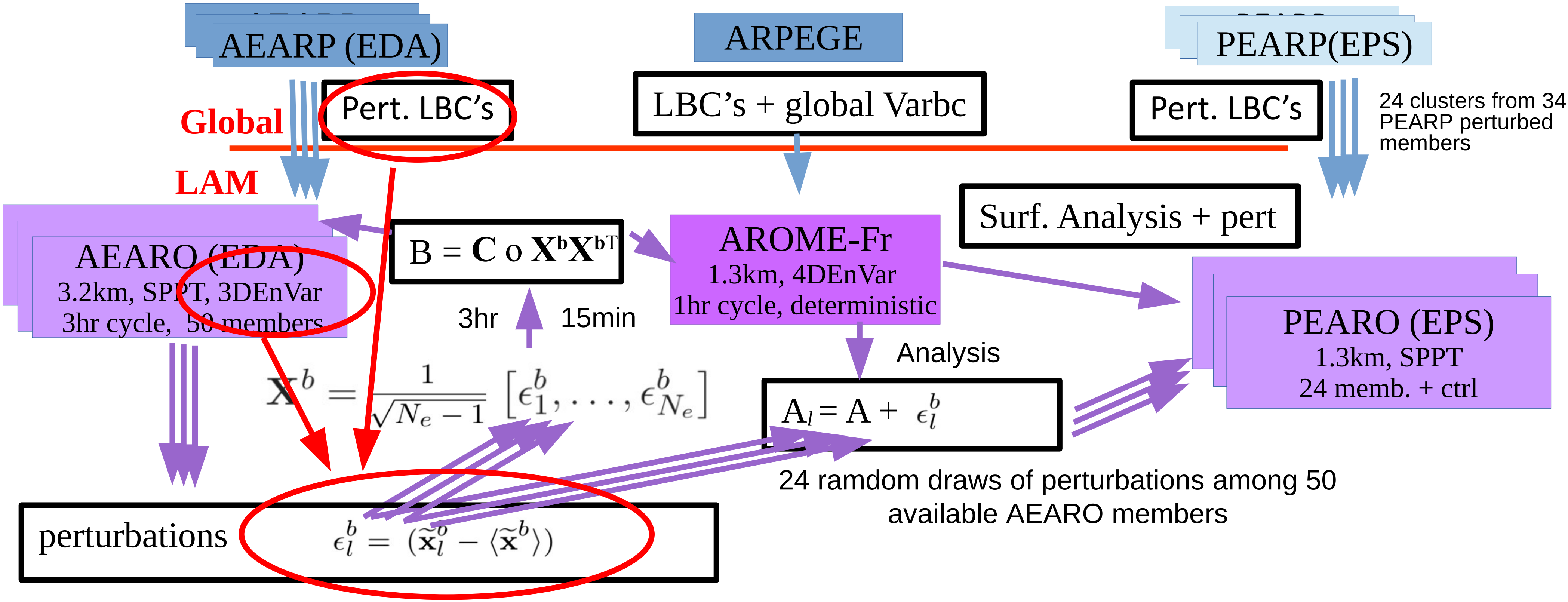
Links between EDA, DA and EPS



AROME EDA provides perturbations (deviation of an ensemble member with respect to the ensemble mean) :

- to estimate the B matrix of the EnVar scheme where C is the localization matrix, "o" the Schur product and X^b, the perturbation matrix
- to estimate perturbed initial conditions of the AROME EPS

Links between EDA, DA and EPS

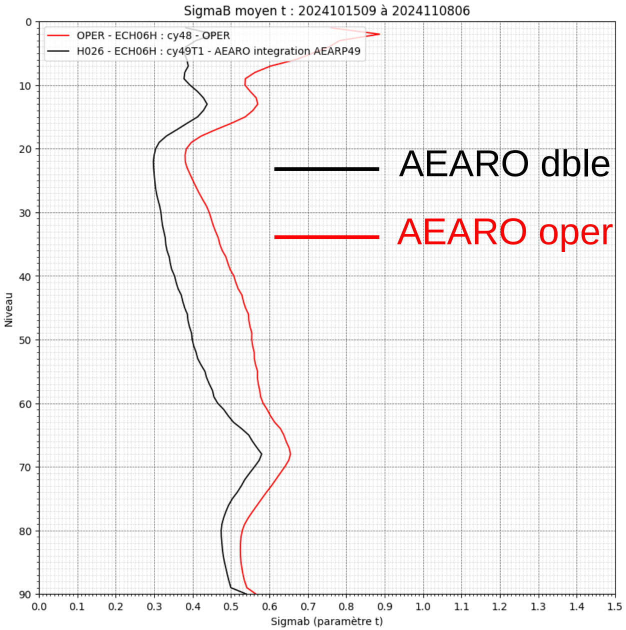


AROME EDA provides perturbations (deviation of an ensemble member with respect to the ensemble mean) :

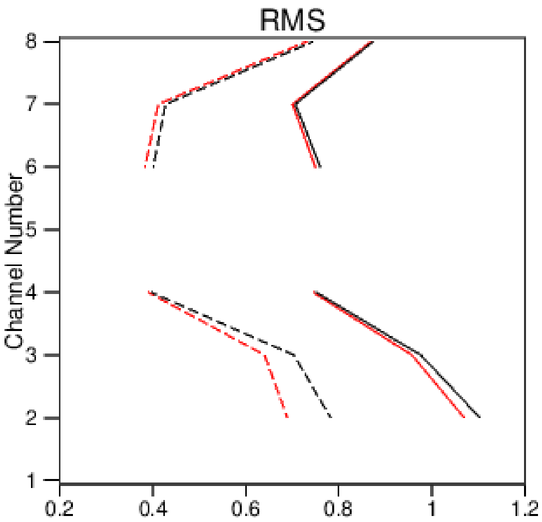
- to estimate the B matrix of the EnVar scheme where C is the localization matrix, "o" the Schur product and X^b, the perturbation matrix
- to estimate perturbed initial conditions of the AROME EPS

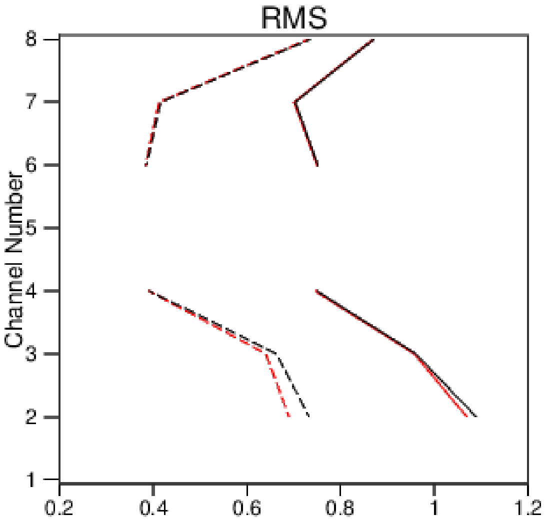
Links between EDA, DA and EPS : impact of EDA changes on DA and EPS

SEVIRI channels (3 weeks period)



Vertical profil of averaged sigma-b for temperature



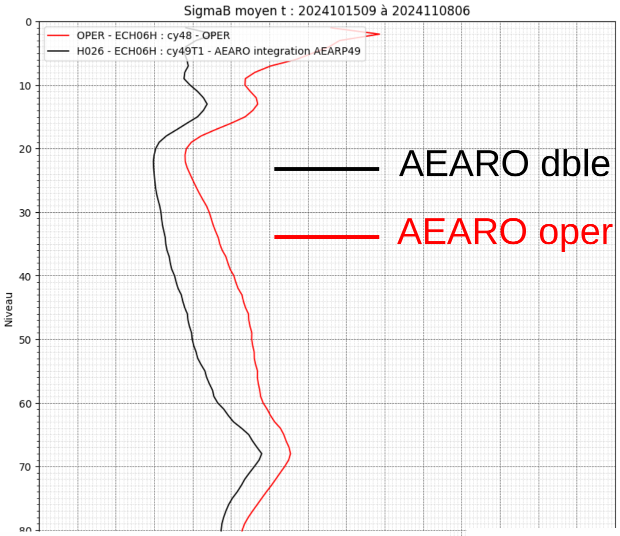


With $\epsilon_l^b = \quad 1.2 \quad \frac{1}{\sqrt{N_e - 1}} (\widetilde{\mathbf{x}}_l^b - \langle \widetilde{\mathbf{x}}^b \rangle)$

Deterministic 4DEnVar experiment

AEARO dble : lower spread due to 3DEnVar Scheme and RP in AEARP (through LBC's)
=> smaller sigma-b in AROME-France 4DEnVar : lesser analysis fit to the observations and degraded backgrounds
=> use of a 1.2 inflation factor to compensate this spread reduction

Links between EDA, DA and EPS : impact of EDA changes on DA and EPS

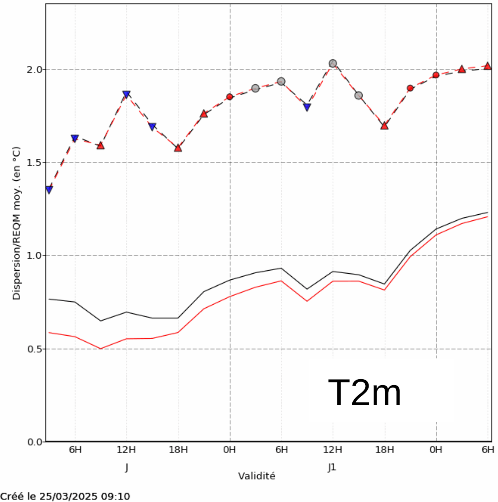


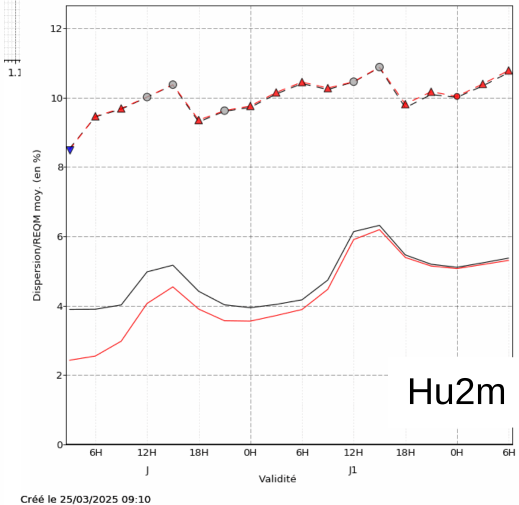
AEARO dble : lower spread due to 3DEnVar Scheme and RP in AEARP (through LBC's)
=> PEARO spread reduced, quite neutral on skill
=> need for an inflation factor to be defined (1.2 ?) in the E-suite accordingly to spread changes also in LBC's from ARPEGE EPS (PEARP)

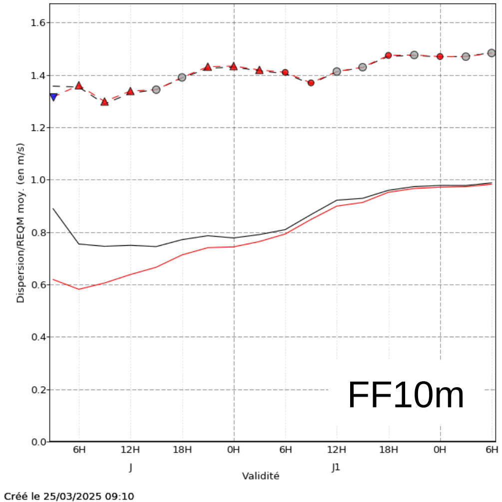
Bootstrap Significancy : Xp better

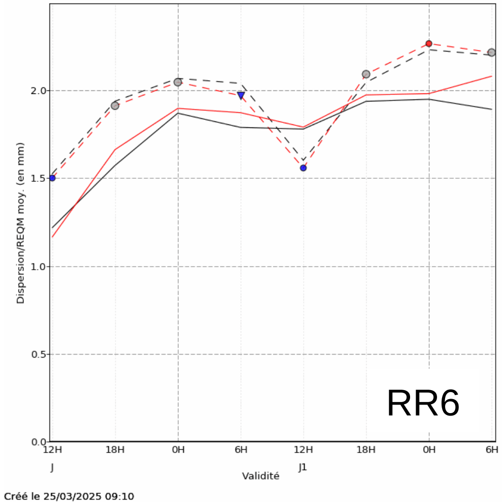
Spread/skill (20/10/2024-20/11/2024)
AEARO oper / AEARO dble
SPREAD
SKILL

At 95 %
At 90 %
Not significant









Conclusion

Thanks to the use of the OOPS framework and the move towards EnVar schemes, a new AROME-France E-Suite based on cy49 has been developed including :

- The use of Scale Dependant Localization in EnVar scheme : performances are improved and the need for localization retuning is reduced.

- The direct assimilation of radar reflectivities (and the use of hydrometeors in the control variable) instead of the 1D pseudo observation retrieval of relative humidity.

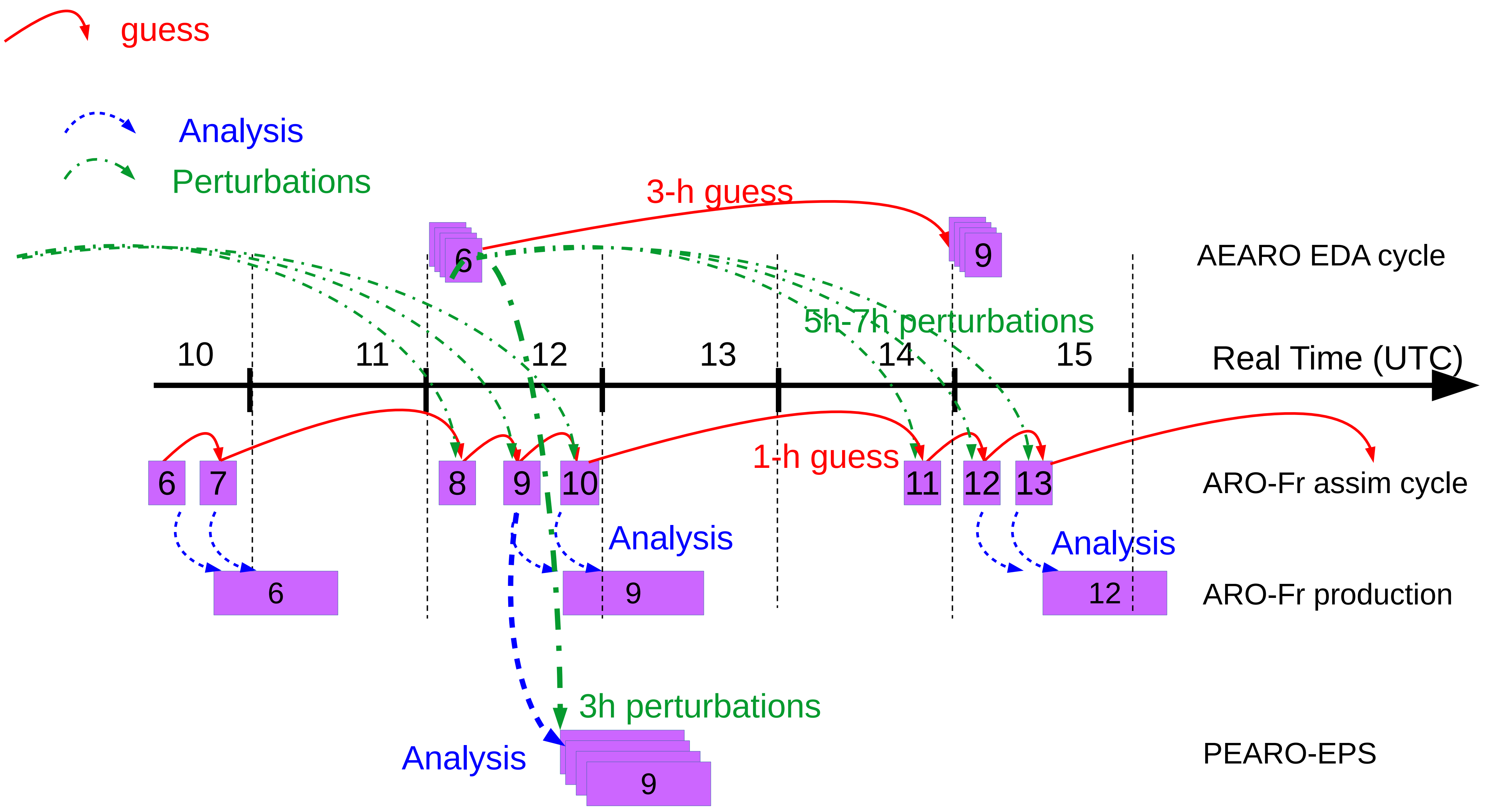
- The use of a 3DEnVar (instead of 3D-Var) in AEARO (EDA) and 4DEnVar (instead of 3DEnVar) in Arome-France : 4DEnVar allows the assimilation of 15 minutes observations (radar, Ground GNSS, SEVIRI, surface station)

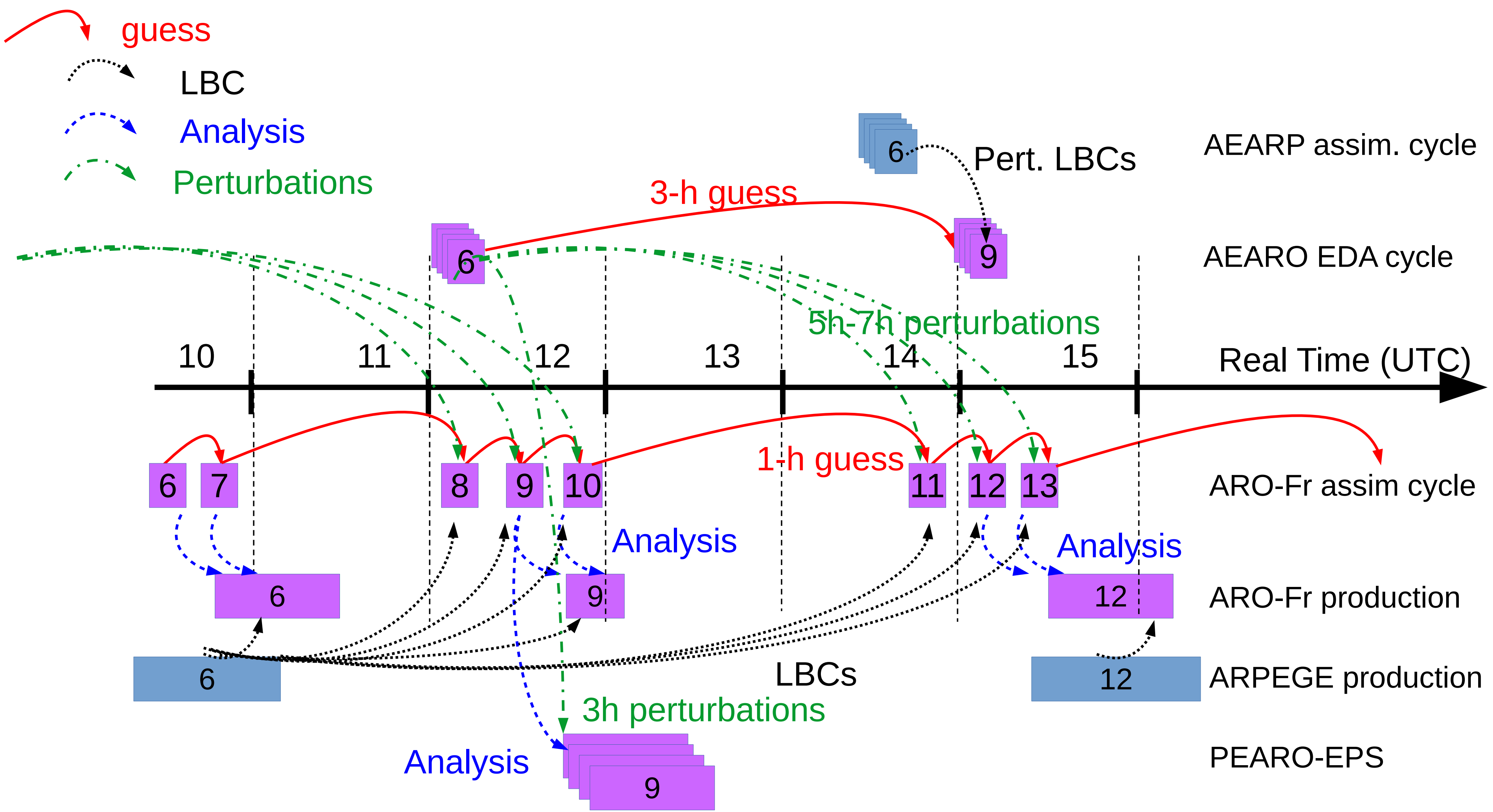
- Optimization of the assimilation of Mode-S and AMV observations

This new version presents improved general performances and better simulations of severe meteorological events (HPE, fog, strong winds …)

Finally, the links between EDA, DA and EPS have been depicted, and the adaptation of these interactions to a new behaviour of one component is also illustrated.

Thanks for your attention


guess
Analysis
Perturbations
3-h guess
6
9
AEARO EDA cycle
5h-7h perturbations
10
11
12
13
14
15
Real Time (UTC)
6
7
8
9
10
1-h guess
11
12
13
ARO-Fr assim cycle
Analysis
Analysis
6
9
12
ARO-Fr production
3h perturbations
Analysis
9
PEARO-EPS


guess
LBC
Analysis
Perturbations
AEARP assim. cycle
6
Pert. LBCs
3-h guess
AEARO EDA cycle
6
9
5h-7h perturbations
10
11
12
13
14
15
Real Time (UTC)
1-h guess
ARO-Fr assim cycle
6
7
8
9
10
11
12
13
Analysis
Analysis
ARO-Fr production
6
9
12
LBCs
ARPEGE production
6
12
3h perturbations
Analysis
PEARO-EPS
9