Development of 4DEnVar at Météo-France

G. Desroziers, E. Arbogast, L. Berre

METEO FRANCE

Outline

1. Introduction
2. Operational configuration
3. 4DEnVar development
4. Localization with the time dimension
5. Conclusion

Outline

1. Introduction
2. Operational configuration
3. 4DEnVar development
4. Localization with the time dimension
5. Conclusion

Introduction

- 4D-Var

✓ Possible improved representation of **B** with an ensemble of 4D-Vars (Météo-France, ECMWF):
"flow-dependent" error variances and
correlations, with a wavelet **B** (Fisher, 2003; Varella et al, 2011).

✓ However, difficult development and maintenance of TL/AD.
✓ Low scalability of TL/AD at low resolution.

- 4D-Var based on a 4D ensemble: 4DEnVar

✓ Avoids TL/AD forecast models.
✓ Similar to 4D-Var (**H**, global analysis, additional terms, outer loops, …).
✓ Localization in model space (versus observation space for EnKF).
✓ Lower cost than 4D-Var and parallelization possibilities.

Outline

1. Introduction
2. Operational configuration
3. 4DEnVar development
4. Localization with the time dimension
5. Conclusion

METEO FRANCE

Current ARPEGE global assimilation configuration

- ”Deterministic” 4D-Var

✓ 6 hour time window.

✓ 2 outer loops

T1198 C2.2 (7.5 km min) L105 / T149 (~135 km), T399 (~ 50 km).

✓ Jc-DFI, VarBC.

✓ $\mathbf{B}^{1/2} = \mathbf{K}^b \Sigma^b \mathbf{C}^{1/2}$, wavelet **C**, $\mathbf{K}^b$ = spectral + non-linear balances.

- Ensemble assimilation

✓ 25 perturbed 4D-Vars.

✓ 1 outer loop T479 C1.0 (40 km) / T149 C1.0.

✓ Multiplicative inflation of 3h forecast perturbations.

✓ Gives

➢ filtered Σ^b from the last 25 perturbations, updated every 6 h,

➢ wavelet **C** from the last 6 x 25 perturbations (last 30 h), updated every 6 h.

Outline

1. Introduction
2. Operational configuration
3. 4DEnVar development
4. Localization with the time dimension
5. Conclusion

4DEnVar formulation

- Minimization of

$$J(\underline{\delta\mathbf{x}}) = \underline{\delta\mathbf{x}}^{T}\, \underline{\mathbf{B}}^{-1}\, \underline{d\mathbf{x}} + (\underline{\mathbf{d}}\text{-}\underline{\mathbf{H}}\ \underline{\delta\mathbf{x}})^{T}\, \underline{\mathbf{R}}^{-1}\, (\underline{\mathbf{d}}\text{-}\underline{\mathbf{H}}\ \underline{\delta\mathbf{x}}), \text{ with } \underline{\mathbf{B}} = \underline{\mathbf{X}}^{b'}\, \underline{\mathbf{X}}^{b'T},$$

$\underline{\mathbf{X}}^{b'} = (\underline{\mathbf{x}}^{b'}_{1}, \ldots, \underline{\mathbf{x}}^{b'}_{Ne})$, and $\underline{\mathbf{x}}^{b'}_{ne} = \underline{\mathbf{x}}^{b}_{ne} - <\underline{\mathbf{x}}^{b}> / (N^{e}-1)^{1/2}$, ne =1, N^{e}.

$\underline{\mathbf{x}}^{b'}$ with dimension K+1 (times) x M (3D variables) x N (3D dimension)

(Liu et al, 2008, 2009; Buehner et al, 2010; Lorenc, 2012).

- Formulation developed at Météo-France

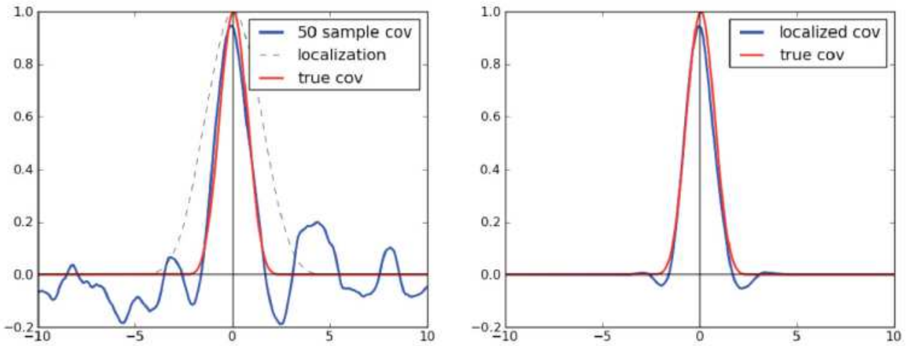
✓ $\underline{\delta\mathbf{x}}$ as a control variable (no α or χ variables) (Desroziers et al, 2014).

✓ DPCG minimizer (Derber and Rosati, 1989) with $\underline{\mathbf{h}} = \underline{\mathbf{B}}\ \underline{\mathbf{g}}$ operations.

✓ Also possible in observation space with $\underline{\mathbf{h}}^{y} = \underline{\mathbf{H}}\ \underline{\mathbf{B}}\ \underline{\mathbf{H}}^{T}\ \underline{\mathbf{g}}^{y}$ operations.

Localization of ensemble covariances

- Need for localization



(Whitaker, 2011)

- Simplication of the localization

Same **L** for all 3D variables and times: $\underline{\mathbf{B}} = \underline{\mathbf{X}}^{b'} \underline{\mathbf{X}}^{b'T} \circ \underline{\mathbf{L}}$, with

$$\underline{\mathbf{L}} = \begin{pmatrix} \mathbf{L} & \mathbf{L} & & \mathbf{L} \\ \mathbf{L} & & & \\ & & & \\ \mathbf{L} & & & \mathbf{L} \end{pmatrix} = \begin{pmatrix} \mathbf{I} \\ \mathbf{I} \\ \\ \mathbf{I} \end{pmatrix} \mathbf{L} \, (\mathbf{I} \;\; \mathbf{I} \qquad \mathbf{I}) = \underline{\mathbf{1}} \, \mathbf{L} \, \underline{\mathbf{1}}^T$$

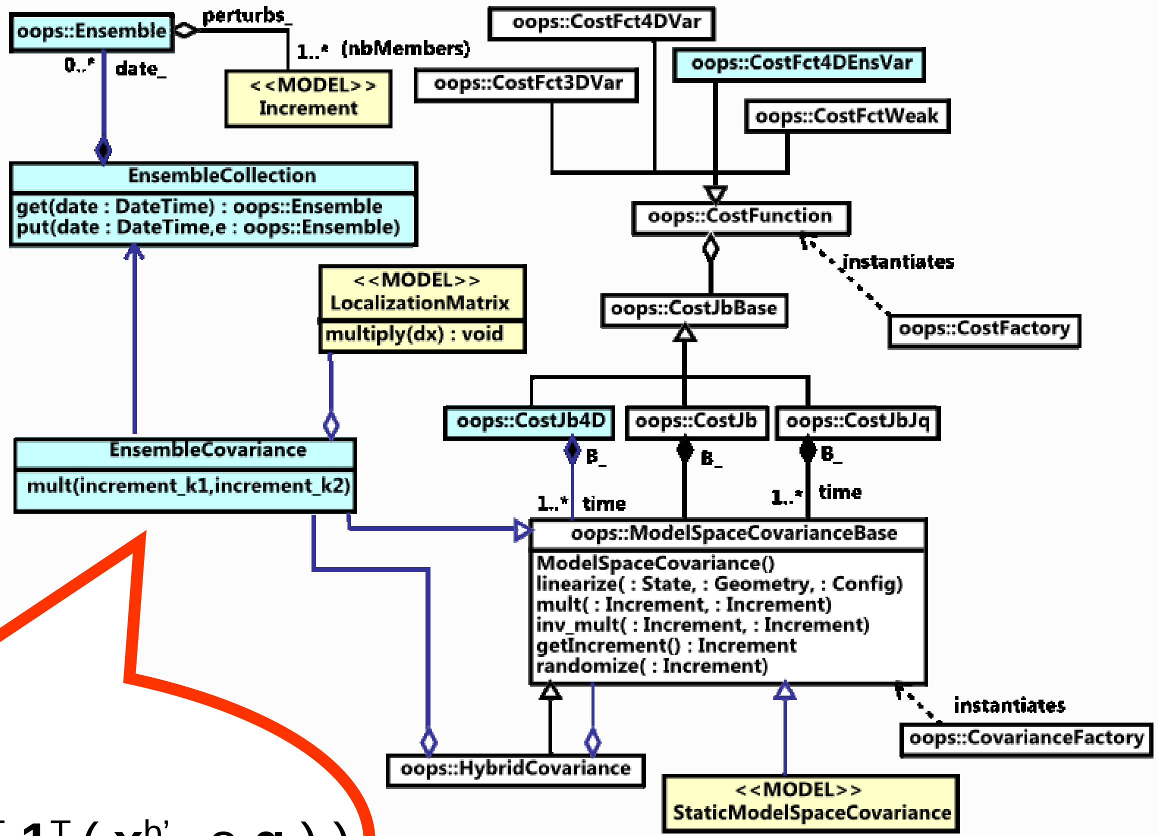
, and **I** is the NxN identity matrix.

Matrix $\underline{\mathbf{1}}$ contains K+1 (times) x M (var.) blocks **I** and **L** is a NxN correl. matrix.

- Applicat. of $\underline{\mathbf{B}}$: $\underline{\mathbf{h}} = \underline{\mathbf{B}} \, \underline{\mathbf{g}} = (\underline{\mathbf{X}}^{b'} \underline{\mathbf{X}}^{b'T} \circ \underline{\mathbf{L}}) \, \underline{\mathbf{g}} = \Sigma_{ne} \, \underline{\mathbf{x}}^{b'}_{ne} \circ (\underline{\mathbf{1}} \, \mathbf{S}^{-1} \, \mathbf{L} \, \mathbf{S}^{-T} \, \underline{\mathbf{1}}^T (\underline{\mathbf{x}}^{b'}_{ne} \circ \underline{\mathbf{g}})).$

4DEnVar under OOPS
(Object Oriented Prediction System, ECMWF/Météo-France)

Jb classes



$$\Sigma_{ne}\, \underline{\mathbf{x}}^{b'}_{ne} \circ \left(\underline{\mathbf{1}}\, \mathbf{S}^{-1}\, \mathbf{L}\, \mathbf{S}^{-T}\, \underline{\mathbf{1}}^{T} \left(\underline{\mathbf{x}}^{b'}_{ne} \circ \underline{\mathbf{g}} \right) \right)$$

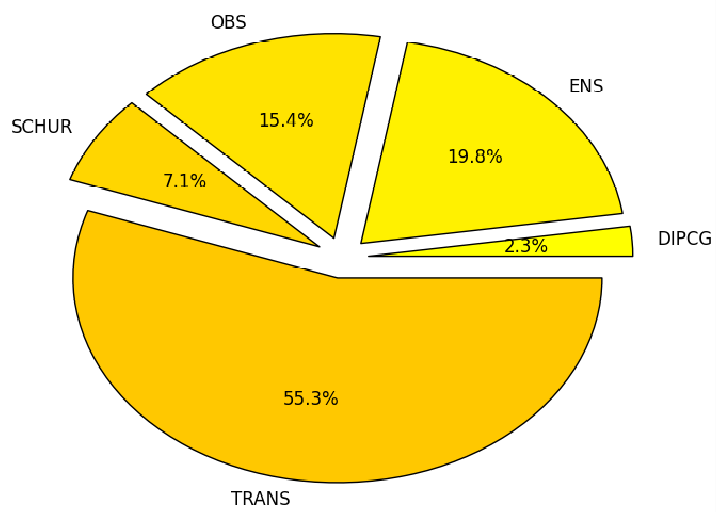
METEO FRANCE

Optimization of the computations

- Geographical parallelization with **MPI**.
- Gridpoint and spectral parallelization with **OpenMP**.
- $\underline{\mathbf{x}}^{b'}$ reading parallelization with **a pool of C++ threads**.
- Computation cost with $N^e = 200$, T399, 40 it,

  75 nodes, 150 MPI tasks, 12 OpenMP threads by task: **9'**

  with conventional observations only.



Optimization of the localization

- Transform ψ, χ and Ps to have better agreement between all variables

ψ —–> $\Delta^{1/2}\psi$

χ —–> $\Delta^{1/2}\chi$

T

q

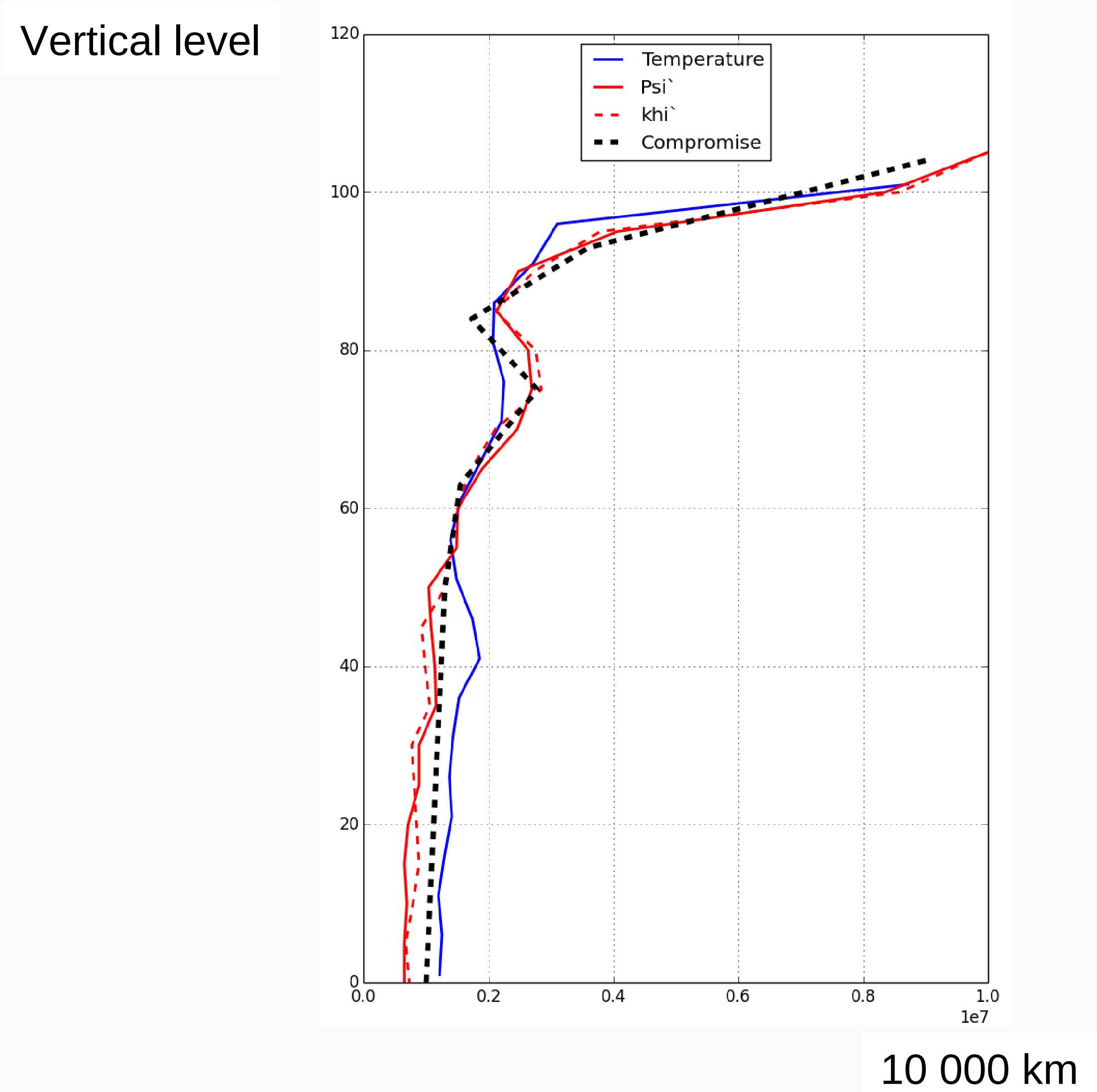
Ps —–> $\Delta^{1/2}$ Ps

- Localized matrix with transformed variables

$$\underline{\mathbf{B}} = \underline{\mathbf{U}}^{-1} \left(\left(\underline{\mathbf{U}}\, \underline{\mathbf{X}}^{b'} \right) \left(\underline{\mathbf{U}}\, \underline{\mathbf{X}}^{b'} \right)^{T} \circ \underline{\mathbf{L}} \right) \underline{\mathbf{U}}^{-T},$$

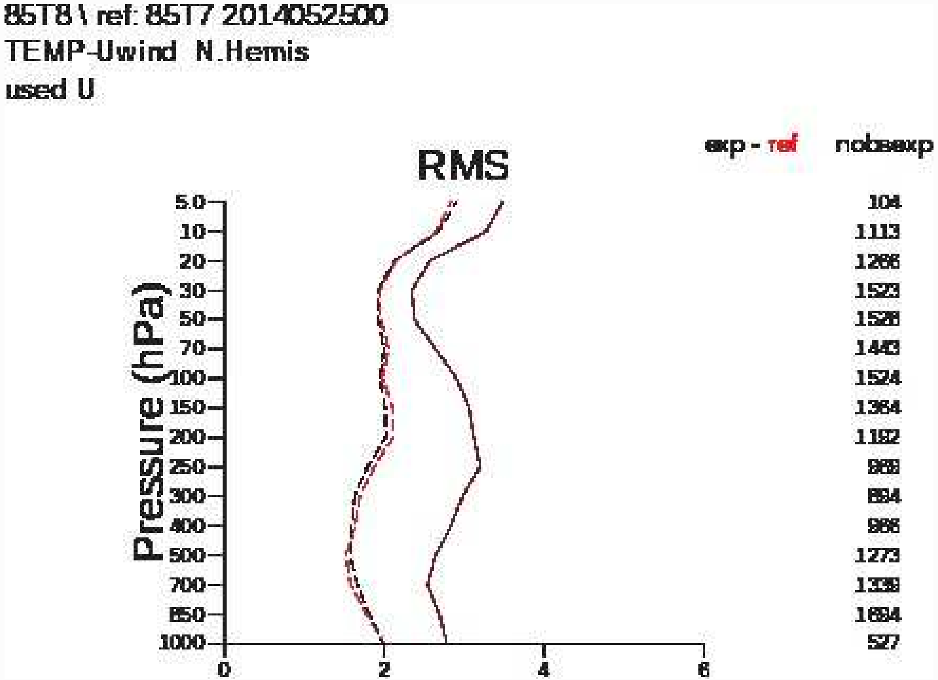
where $\underline{\mathbf{U}}$ is the change of variables.

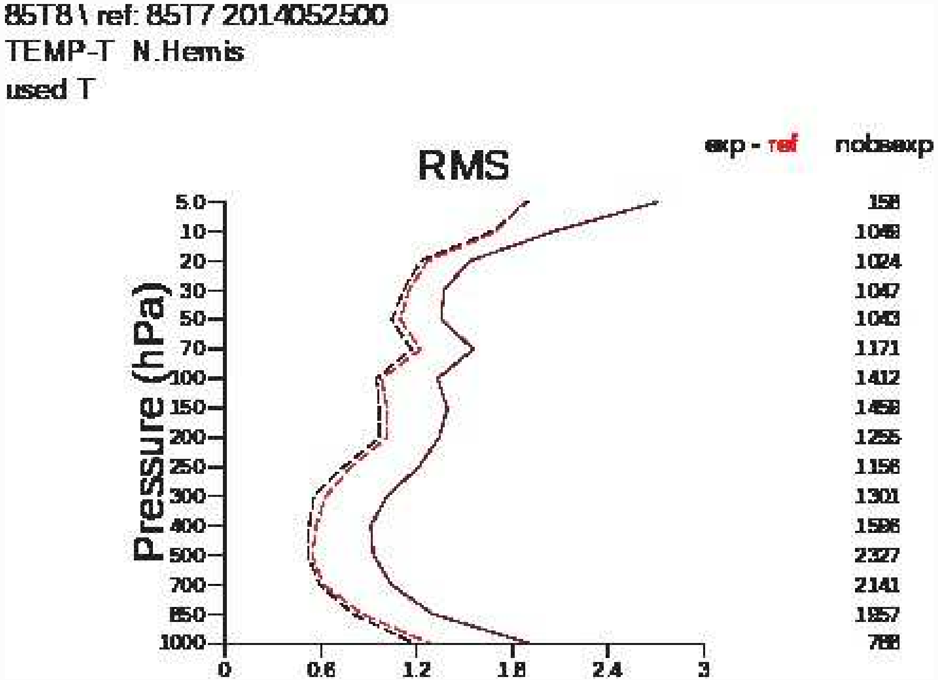
Optimization of horizontal localization length scales



$N^e = 200$

Results with conventional data
4D-Var / 4DEnVar ($N^e = 200$)





m/s °K

solid lines: RMS $y^o - H(x^b)$

dashed lines: RMS $y^o - H(x^a)$

Outline

METEO FRANCE

Advection of the localization

- Static $\underline{\mathbf{L}}$

$$\underline{\mathbf{L}} = \begin{pmatrix} \mathbf{L} & \mathbf{L} & & \mathbf{L} \\ \mathbf{L} & & & \\ & & & \\ \mathbf{L} & & & \mathbf{L} \end{pmatrix} = \begin{pmatrix} \mathbf{I} \\ \mathbf{I} \\ \\ \mathbf{I} \end{pmatrix} \mathbf{L} \begin{pmatrix} \mathbf{I} & \mathbf{I} & & \mathbf{I} \end{pmatrix} = \underline{\mathbf{1}}\, \mathbf{L}\, \underline{\mathbf{1}}^{T} .$$

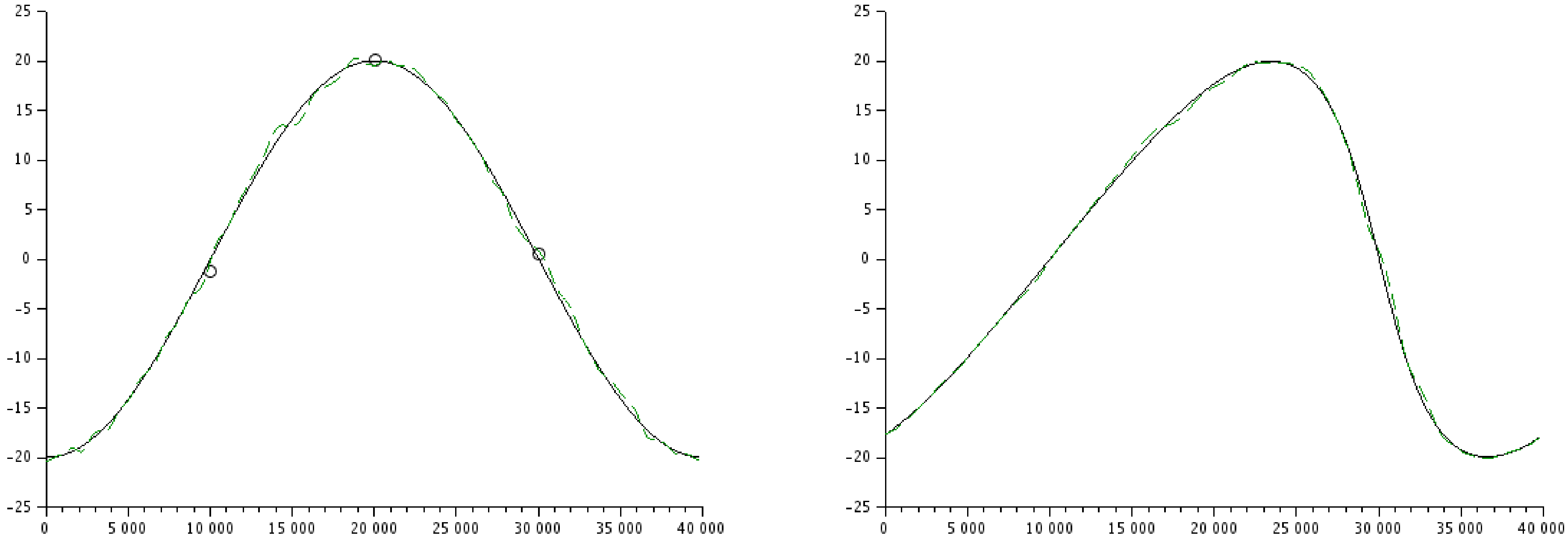
- Advected $\underline{\mathbf{L}}$

$$\underline{\mathbf{L}} = \begin{pmatrix} \mathbf{L} & \mathbf{L}\mathbf{A}_1^{T} & & \mathbf{L}\mathbf{A}_K^{T} \\ \mathbf{A}_1\mathbf{L} & & & \\ & & & \\ \mathbf{A}_K\mathbf{L} & & & \mathbf{A}_K\mathbf{L}\mathbf{A}_K^{T} \end{pmatrix} = \begin{pmatrix} \mathbf{I} \\ \mathbf{A}_1 \\ \\ \mathbf{A}_K \end{pmatrix} \mathbf{L} \begin{pmatrix} \mathbf{I} & \mathbf{A}_1^{T} & & \mathbf{A}_K^{T} \end{pmatrix} = \underline{\mathbf{A}}\, \mathbf{L}\, \underline{\mathbf{A}}^{T} .$$

Burgers'model

$$\partial u/\partial t + u\, \partial u/\partial x + \nu\, \partial^2 u/\partial x^2 = 0$$

m/s



t_0 km t_f=48h

____ u^t : true wind

----- u^b : simulated background

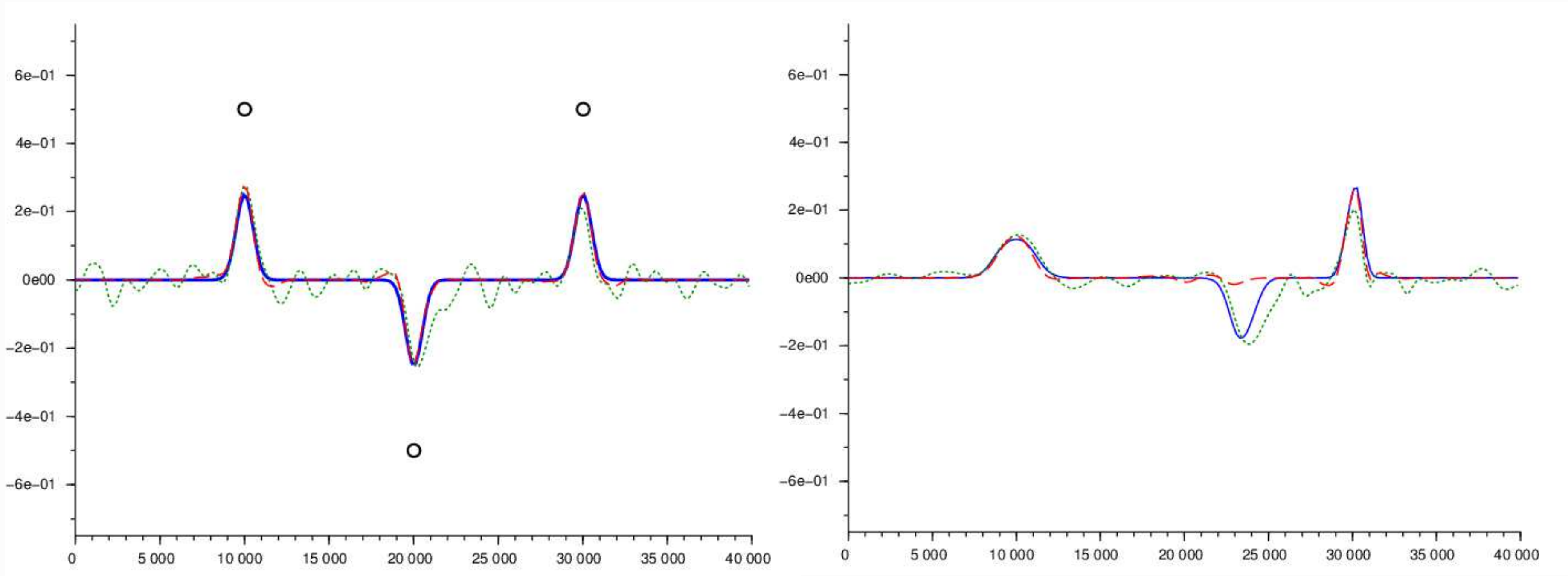
METEO FRANCE

Static L

N^e = 100. Lc =1500 km

$$\underline{\mathbf{C}} = \underline{\mathbf{X}}^{b'} \underline{\mathbf{X}}^{b'T} \circ \underline{\mathbf{1}}\, \mathbf{L}\, \underline{\mathbf{1}}^T$$

m/s



t_0 km t_f=48h

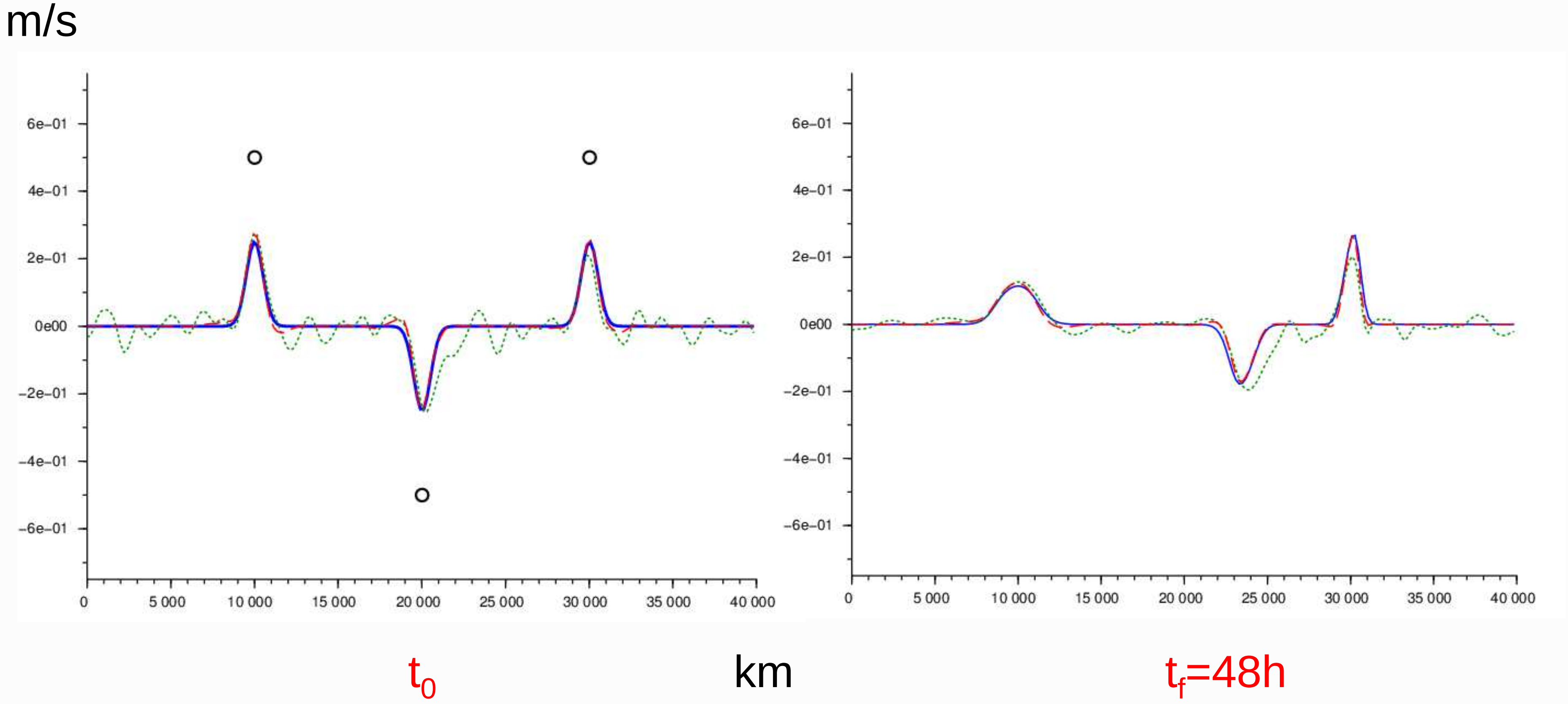
__ 4D-Var δu

…. 4DEnVar δu without localization

---- 4DEnVar δu with localization

Advected L
N^e = 100. Lc =1500 km

$$\underline{\mathbf{C}} = \underline{\mathbf{X}}^{b'} \underline{\mathbf{X}}^{b'T} \circ \underline{\mathbf{A}}\, \mathbf{L}\, \underline{\mathbf{A}}^T$$



__ 4D-Var δu

…. 4DEnVar δu without localization

---- 4DEnVar δu with adv localization

METEO FRANCE

Advection from t_0 to t_k
A Lagrangian point of view

- Lagrangian point of view

$$\alpha(s, t_k) = \alpha(s(t_0), t_0)$$

- Computation of deformed grid $s(t_0)$

$s(t_k\text{-}\delta t) = s$

for k' = k-1 : -1 : 0

$\quad s(t_{k'}\text{-}\delta t) = s(t_{k'}\text{-}\delta t) - u(s(t_{k'}\text{-}\delta t), t_{k'}\text{-}\delta t)) * \delta t$

end

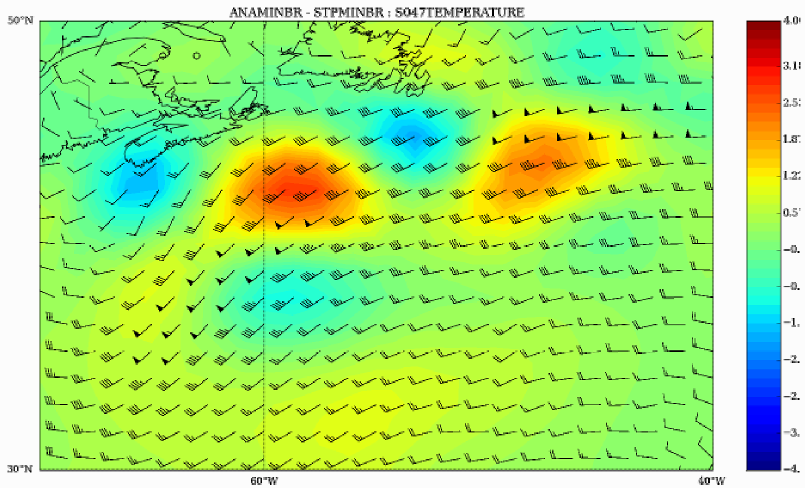
- Advection as a simple interpolation of $\alpha(t_0)$ at $s(t_0)$

$$\alpha(s, t_k) = \alpha(s(t_0), t_0) = I^{nt}\,(s(t_0))\ \alpha(s, t_0),$$

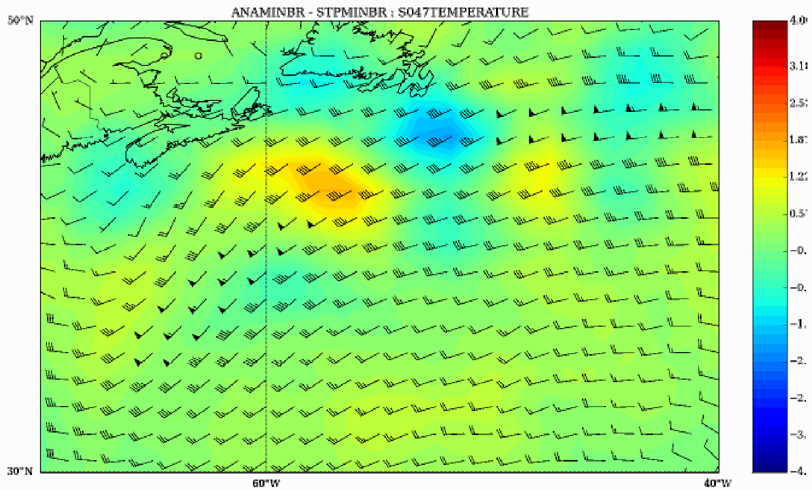
where $I^{nt}\,(s(t_0))$ are interpolation coefficients.

4D-Var / 4DEnVar

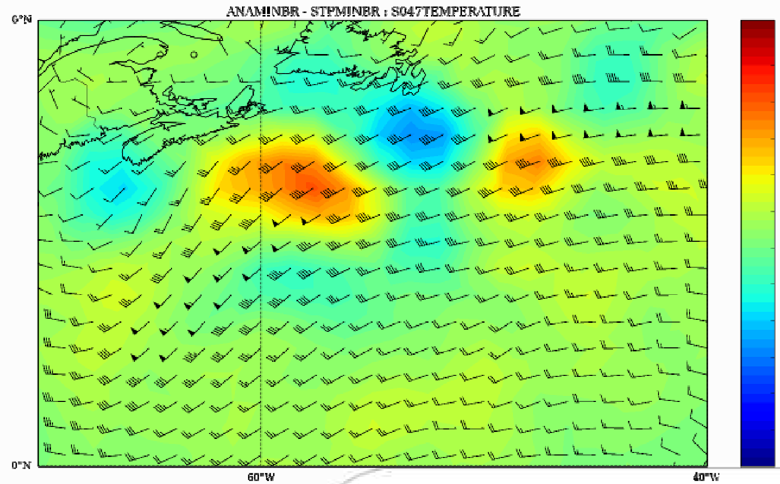
Temperature increment (°K) at t_0 with conv. obs. (@10 km)



4D-Var



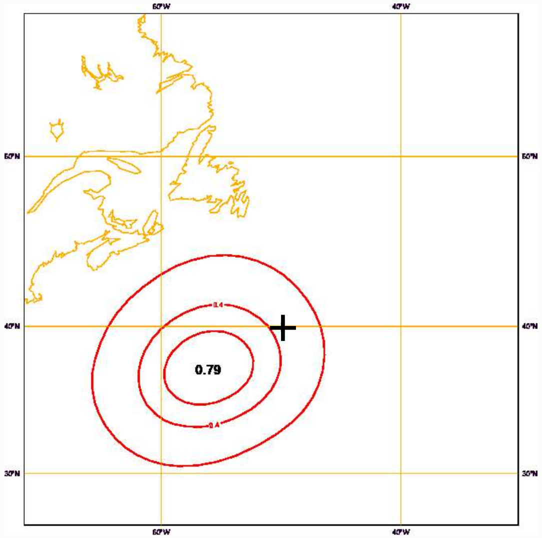
4DEnVar (N^e = 200)



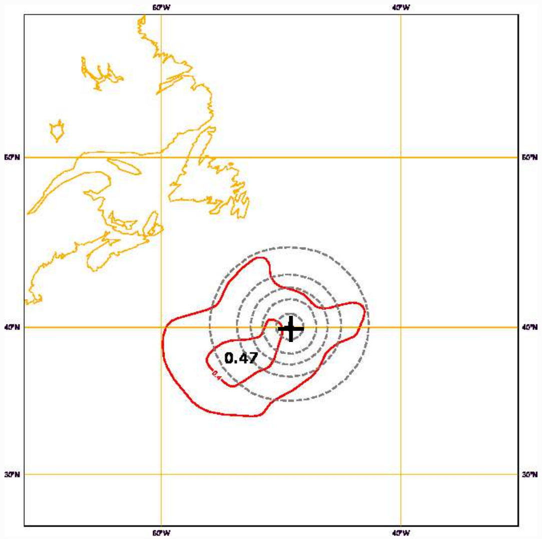
4DEnVar + localization advection

4D-Var / 4DEnVar

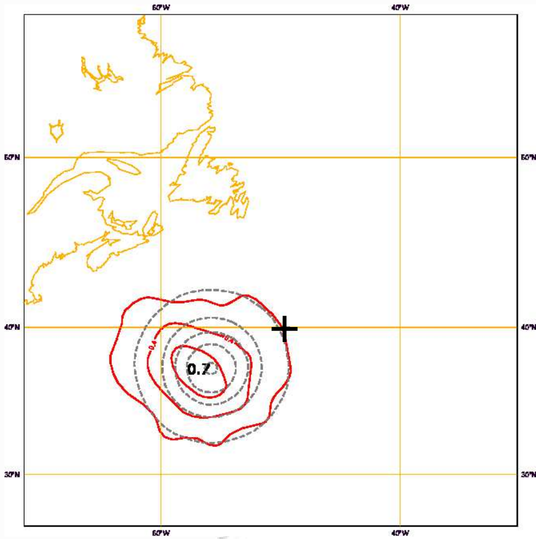
Temperature increment (°K) at t_0 with 1 obs at t_f (@10 km)



4D-Var



4DEnVar (N^e = 200)



4DEnVar + localization advection

Hybrid formulation

- Hybrid matrix: $\underline{\mathbf{B}}^h = \gamma^{c2}\, \underline{\mathbf{B}}^c + (1- \gamma^{c2})\, \underline{\mathbf{B}}^e.$

- Static $\underline{\mathbf{B}}^c$

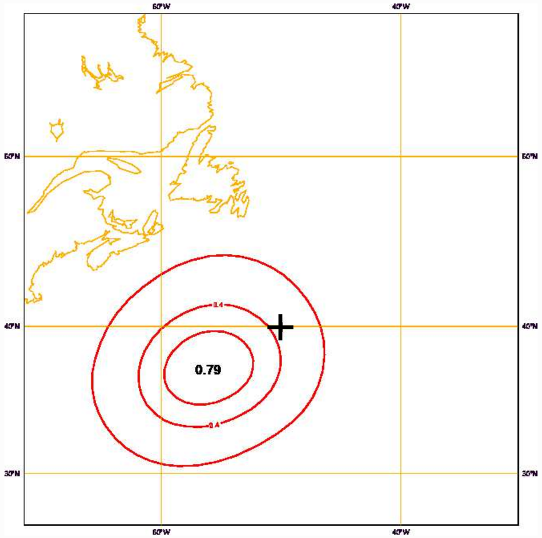
$$\underline{\mathbf{B}}^c = \begin{pmatrix} \mathbf{B}^c & \mathbf{B}^c & & \mathbf{B}^c \\ \mathbf{B}^c & & & \\ & & & \\ \mathbf{B}^c & & & \mathbf{B}^c \end{pmatrix} = \begin{pmatrix} \mathbf{I} \\ \mathbf{I} \\ \\ \mathbf{I} \end{pmatrix} \mathbf{B}^c \,(\,\mathbf{I} \;\; \mathbf{I} \qquad \mathbf{I}\,) = \underline{\mathbf{1}}\, \mathbf{B}^c\, \underline{\mathbf{1}}^T .$$

- Advected $\underline{\mathbf{B}}^c$

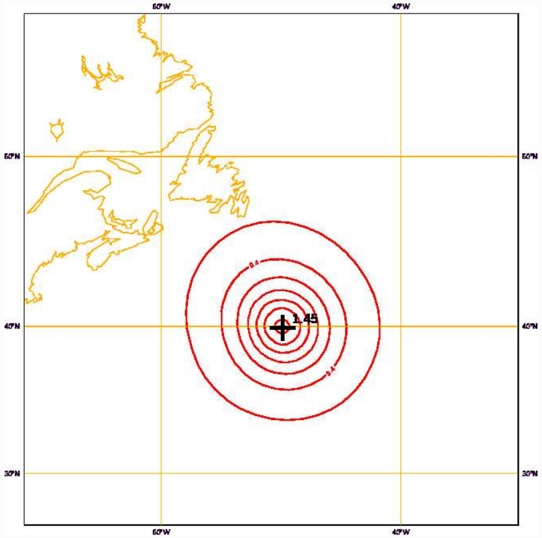
$$\underline{\mathbf{B}}^c = \begin{pmatrix} \mathbf{B}^c & \mathbf{B}^c\mathbf{A}_1^T & & \mathbf{B}^c\mathbf{A}_K^T \\ \mathbf{A}_1\mathbf{B}^c & & & \\ & & & \\ \mathbf{A}_K\mathbf{B}^c & & & \mathbf{A}_K\mathbf{B}^c\mathbf{A}_K^T \end{pmatrix} = \begin{pmatrix} \mathbf{I} \\ \mathbf{A}_1 \\ \\ \mathbf{A}_K \end{pmatrix} \mathbf{B}^c \,(\,\mathbf{I} \;\; \mathbf{A}_1^T \qquad \mathbf{A}_K^T\,) = \underline{\mathbf{A}}\, \mathbf{B}^c\, \underline{\mathbf{A}}^T .$$

4D-Var / 4DEnVar with B^c only (3D-Var FGAT)

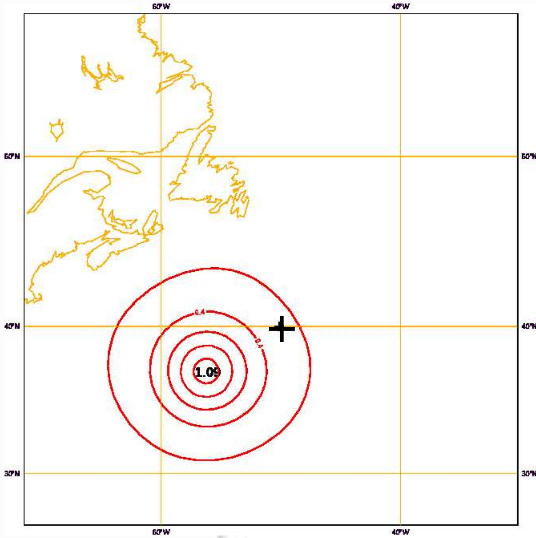
Temperature increment (°K) at t_0 with 1 obs at t_f (@10 km)



4D-Var



4DEnVar with static B^c



4DEnVar with advected B^c

Outline

1. Introduction
2. Operational configuration
3. 4DEnVar development
4. Localization with the time dimension
5. Conclusion

Conclusion and future work

- Conclusion

✓ Possible alternative to the use of α variables: 4D increment $\underline{\delta \mathbf{x}}$.

✓ Reading of perturbations is quick enough.

✓ First version of ARPEGE 4DEnVar at Météo-France.

✓ Localization is more difficult with the time dimension.

✓ The use of Lagrangian advection may help.

- Future work

✓ ARPEGE 4DEnVar: more observations, outer loops, Jc-DFI, …

✓ Improvement of spatial / spectral filtering, sensitivity to ensemble size.

✓ Ensemble of 4DEnVars.

✓ Development of 3D/4DEnVar for high resolution (1,3 km) LAM AROME.