Environnement Canada Environment Canada

Canada

Ensemble-Variational Data Assimilation at Environment Canada

Jean-François Caron and Mark Buehner

Data Assimilation and Satellite Meteorology Research Section, Environment Canada,
Montréal, Quebec, Canada

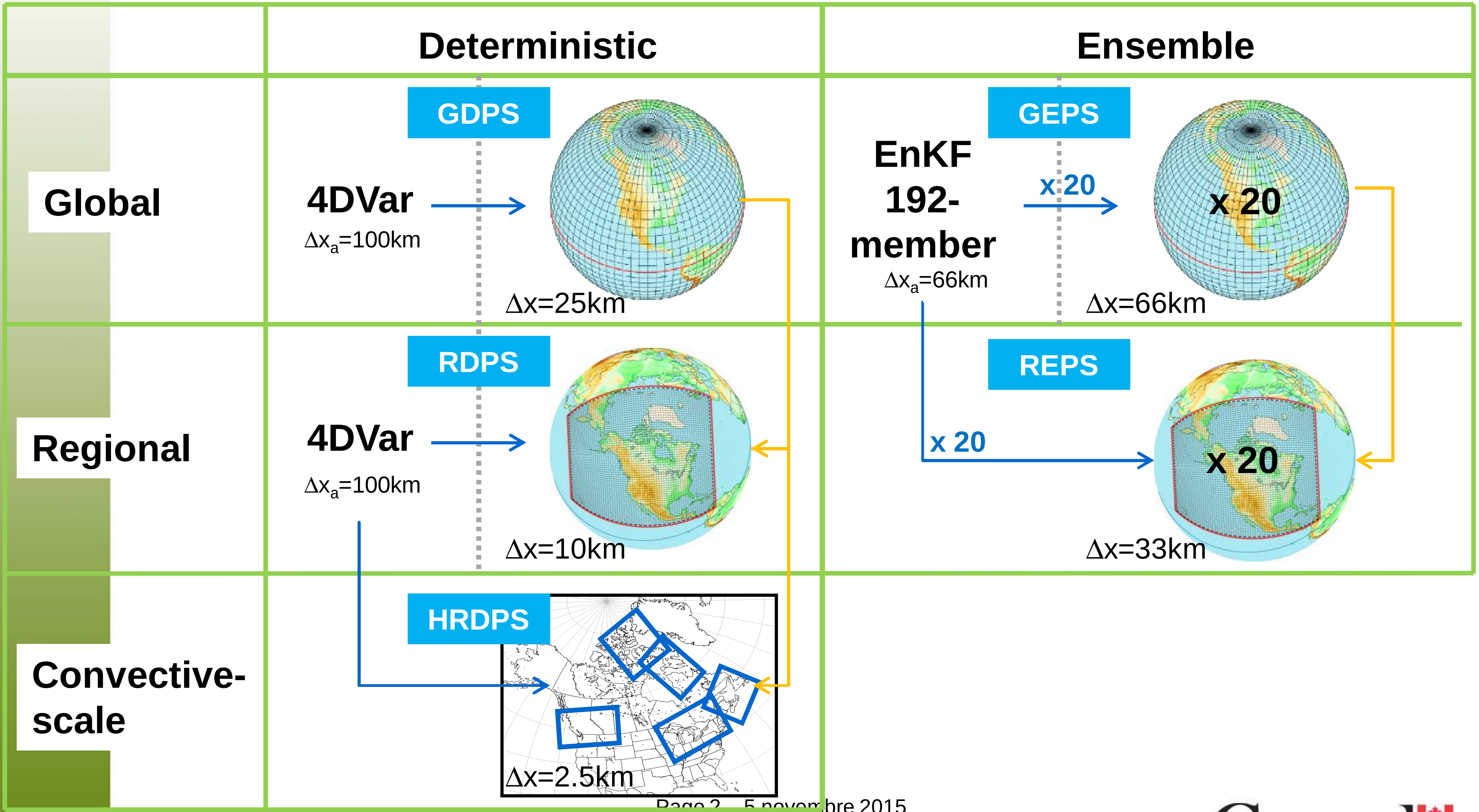
Outline

Part 1 From 4DVar to 4DEnVar at EC with a focus on the regional deterministic prediction system.

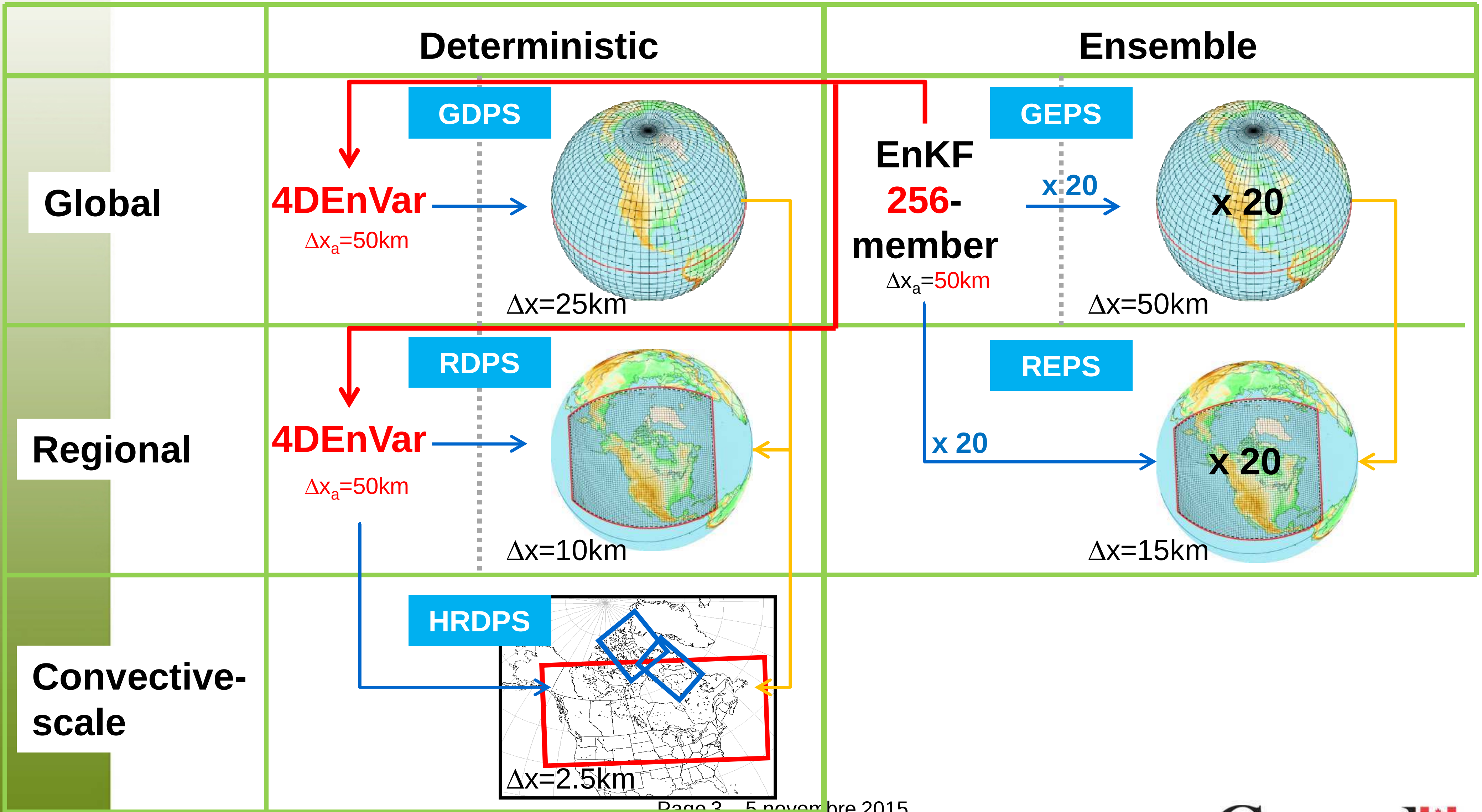
Part 2 Scale-dependent covariance localization in EnVar.

37th EWGLAM and 22nd SRNWP meetings – 5 October 2015 – Belgrade, Serbia.

EC's NWP systems before 11/2014



EC's NWP systems since 11/2014



4DEnVar

- 4DEnVar uses a **variational assimilation approach** in combination with the already available **4D ensemble covariances** from the EnKF
- By making use of the 4D ensembles, 4DEnVar performs a 4D analysis without the need of the tangent-linear and adjoint of forecast model
- Consequently, it is **more computationally efficient** and **easier to maintain/adapt** than 4DVar
- Future improvements to EnKF will benefit both ensemble and deterministic prediction systems
- Increased incentive to improve EnKF and **improve how ensemble members used within 4DEnVar**

4DEnVar Formulation

- **In 4D-Var…** the 3D analysis increment is evolved in time using the TL/AD forecast model (here included in $\mathbf{H}_{4D}$):

$$J(\Delta\mathbf{x}) = \frac{1}{2}(H_{4D}[\mathbf{x}_b] + \boxed{\mathbf{H}_{4D}\Delta\mathbf{x}} - \mathbf{y})^T \mathbf{R}^{-1}(H_{4D}[\mathbf{x}_b] + \boxed{\mathbf{H}_{4D}\Delta\mathbf{x}} - \mathbf{y}) + \frac{1}{2}\boxed{\Delta\mathbf{x}^T\mathbf{B}^{-1}\Delta\mathbf{x}}$$

- **In EnVar…** the background-error covariances and analysed state are explicitly 4-dimensional, resulting in cost function:

$$J(\Delta\mathbf{x}_{4D}) = \frac{1}{2}(H_{4D}[\mathbf{x}_b] + \boxed{\mathbf{H}\Delta\mathbf{x}_{4D}} - \mathbf{y})^T \mathbf{R}^{-1}(H_{4D}[\mathbf{x}_b] + \boxed{\mathbf{H}\Delta\mathbf{x}_{4D}} - \mathbf{y}) + \frac{1}{2}\boxed{\Delta\mathbf{x}_{4D}{}^T\mathbf{B}_{4D}^{-1}\Delta\mathbf{x}_{4D}}$$

EnVar is ~10x computationally cheaper than 4DVar

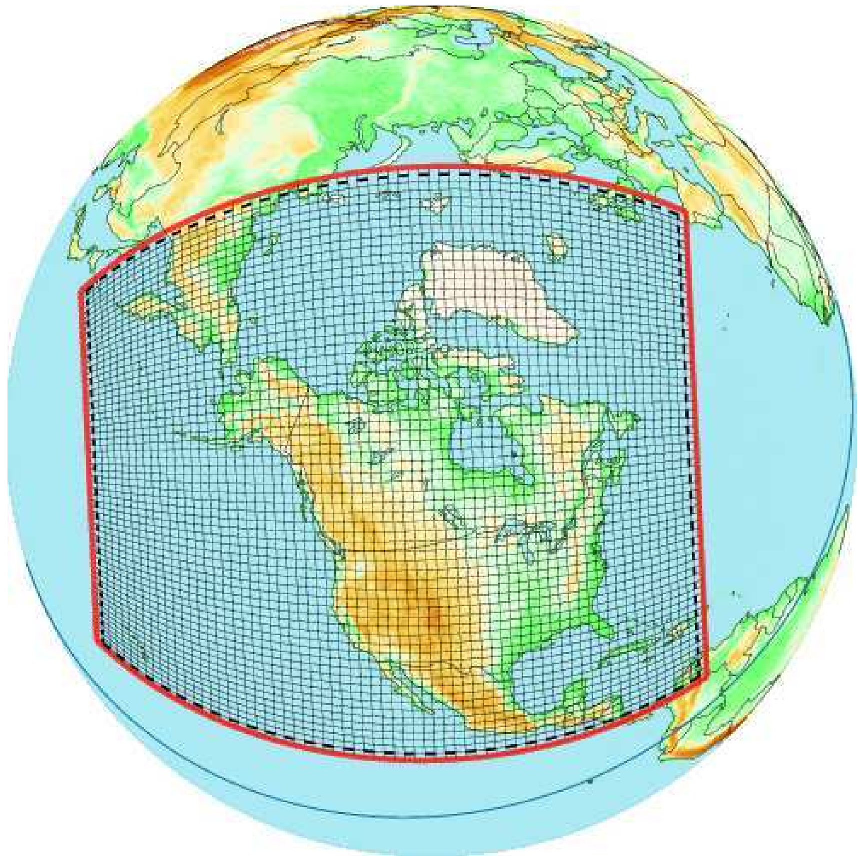
Hybrid **B** formulation

$$\mathbf{B}_{4D} = \beta_{\mathrm{nmc}}\mathbf{B}_{\mathrm{nmc}} + \beta_{\mathrm{ens}}\sum_{k=1}^{N_{\mathrm{ens}}}\left(\mathbf{e}_k\mathbf{e}_k^T\right)\circ\mathbf{L}$$

($\mathbf{e}_k$ is k^{th} ensemble perturbation divided by sqrt(N_{ens}-1))

The operational RDPS in a nutshell

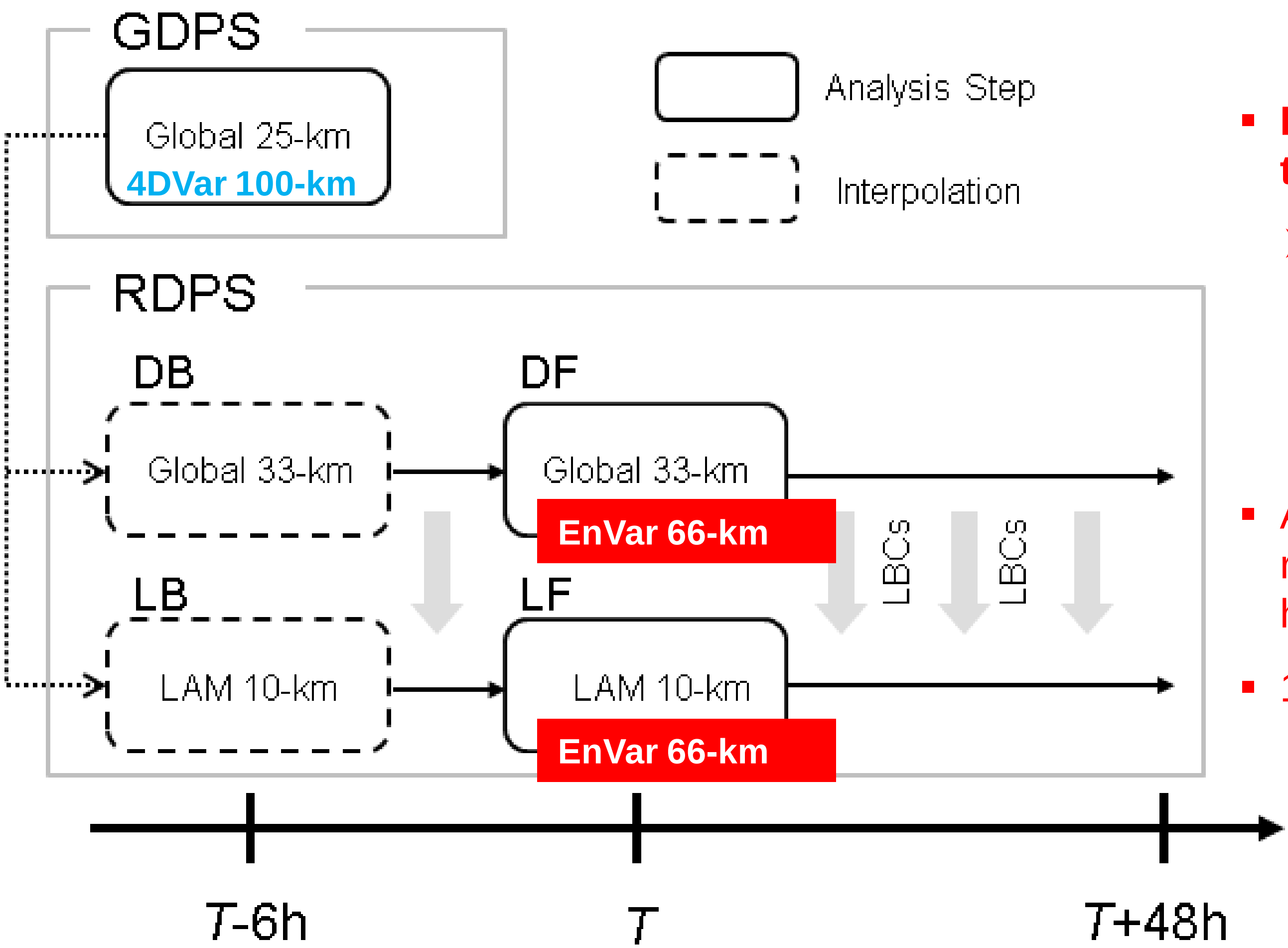
- Main NWP guidance at day 1 and 2 for MSC's forecasters
- Limited-area model
- 10-km, 80 levels
- Forecasts up t+48h
  (data available on http://dd.weather.gc.ca/)
- 4 runs per day (00,06,12,18 UTC)
- LBCs from a simplified GDPS
- Intermittent cycling



Replacing 4D/3D-Var by 4D-EnVar

proof of feasibility

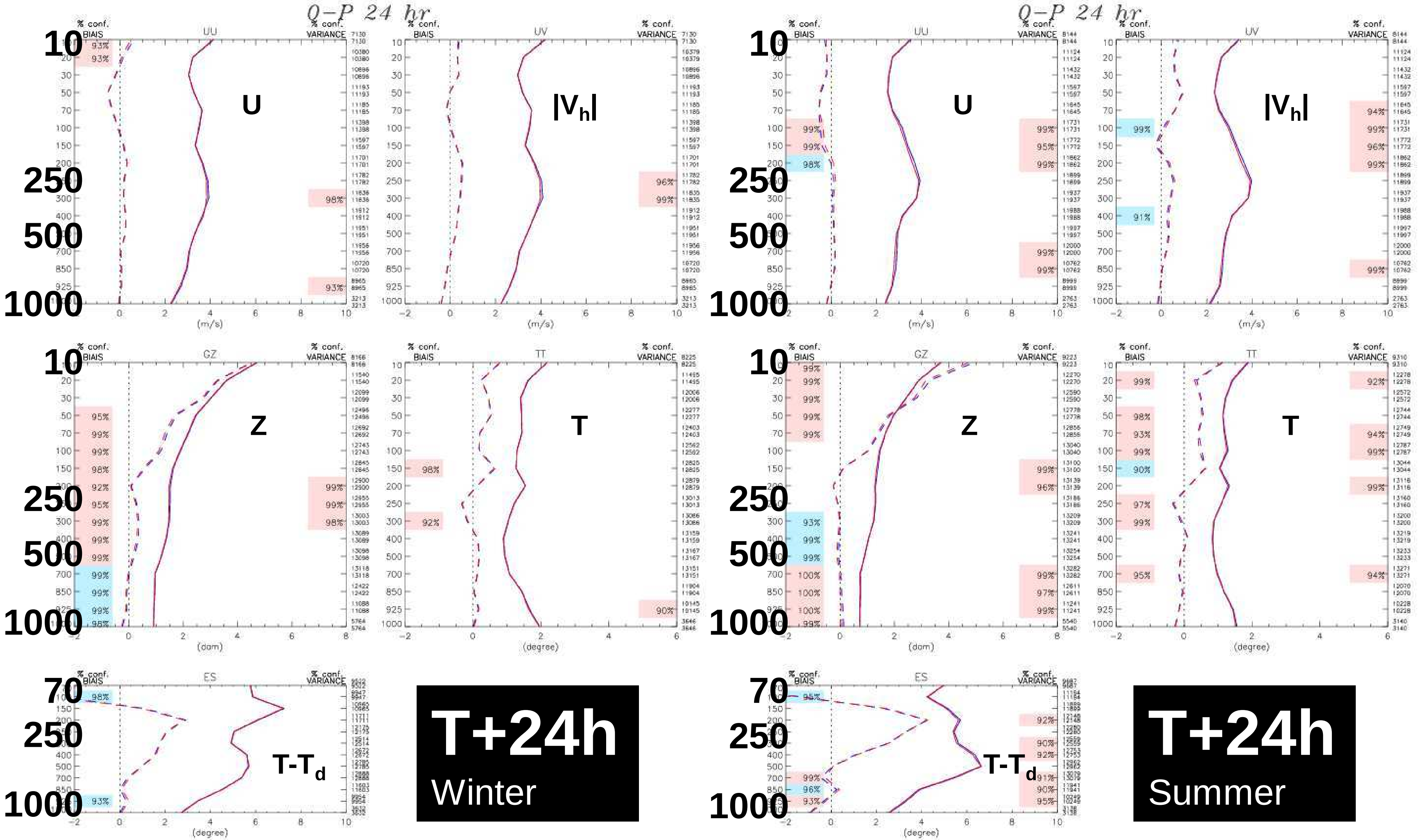
- EnVar based on global EnKF as in the GDPS



- **EnVar setup <u>identical</u> to the GDPS**
  - **B** is a mix (50 / 50) of static parameterized covariances and 4D global ensemble perturbations
- Analysis grid follow the resolution of the EnKF; here 66-km
- 192-member EnKF

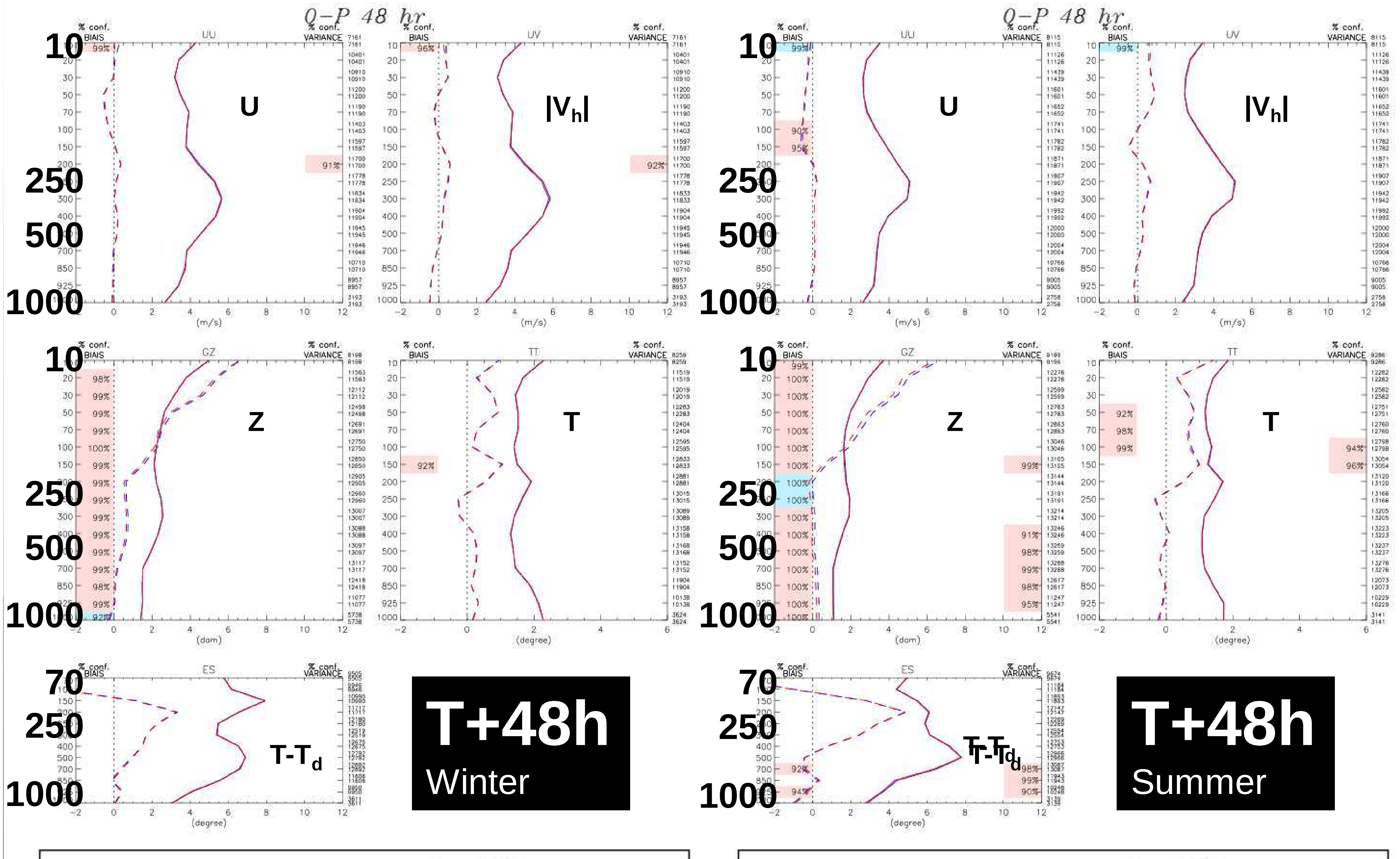
Replacing 4D/3D-Var by 4D-EnVar

proof of feasibility – verification against radiosondes



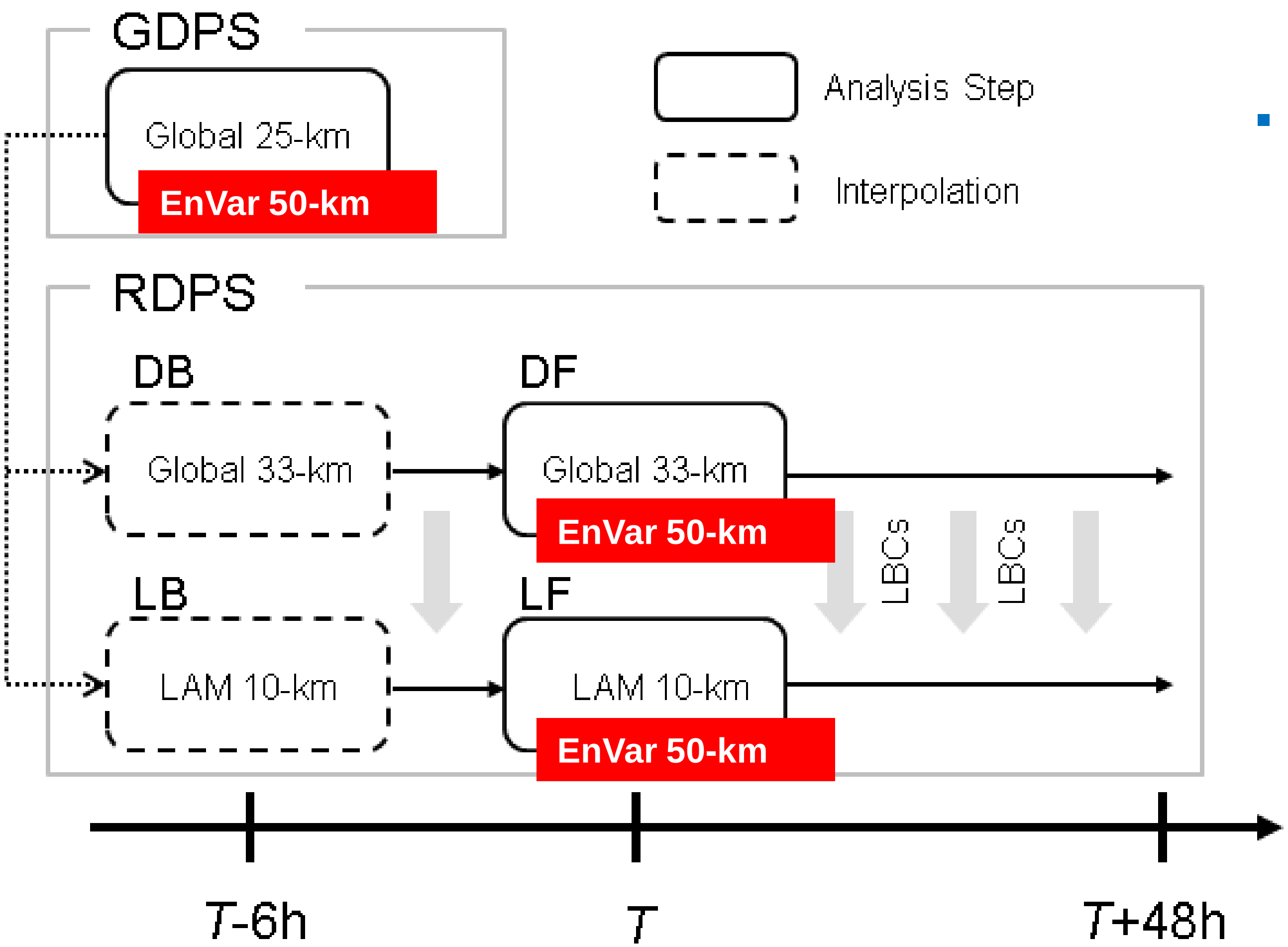
Replacing 4D/3D-Var by 4D-EnVar

proof of feasibility – verification against radiosondes



The new EnVar-based RDPS
operational since November 2014

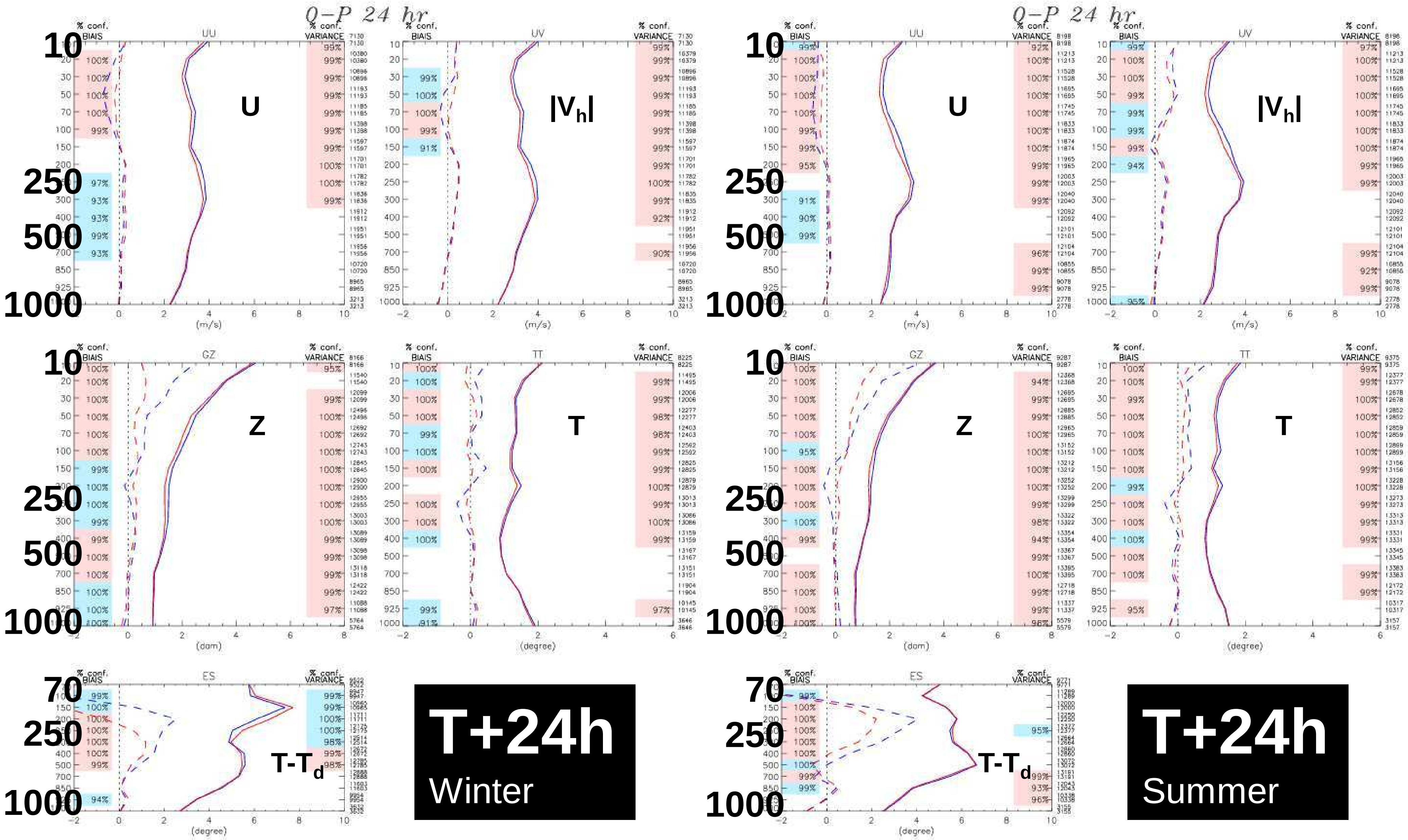
Additional changes



- **EnVar-based GDPS analysis at *T*-6h**
- **50-km and 256 members EnKF**
- **Upgrade to the observational data assimilated**
  - New radiances bias correction
  - Bias correction for aircraft T data
  - Radiosondes drift taken into account
  - Assimilation of ground-based GPS ZTD
  - Increased volume of data for AIRS and IASI

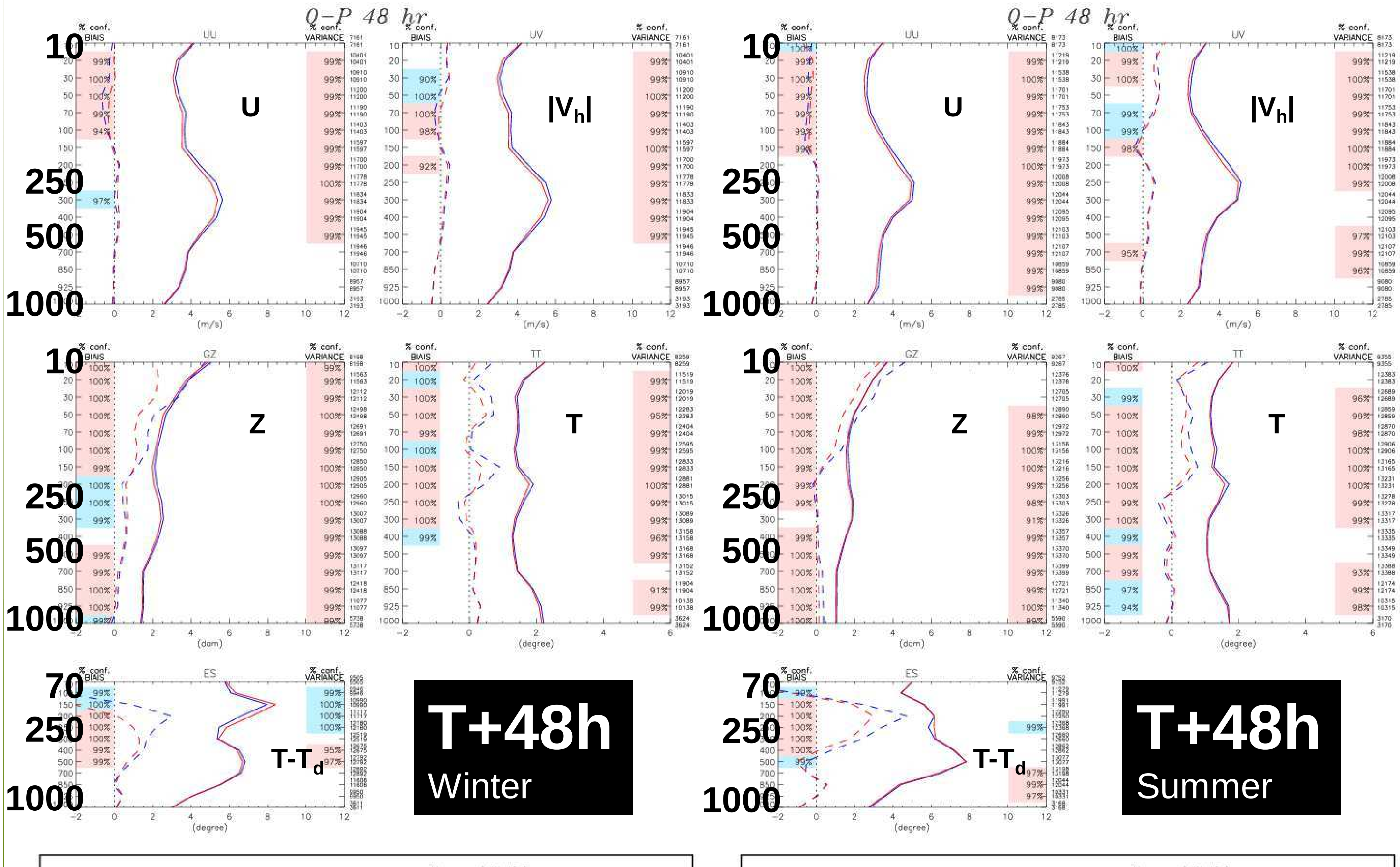
The new EnVar-based RDPS

verification against radiosondes



The new EnVar-based RDPS

verification against radiosondes



Timings and computational cost

On EC's IBM P7…

Analysis	DA Scheme	CPUs	Wall clock (min)	Cost (CPU min)
LAM	4D-Var *	2048	17	34816 *
	EnVar	**320**	**7**	**2240**
Driver	3D-Var	64	9	576
	EnVar	**320**	**9**	**2880**

➘ x15 !!!

➚ x5
but with the same wall clock time

* Not including the cost of the nonlinear forecast from *T*-3h to *T*

Summary

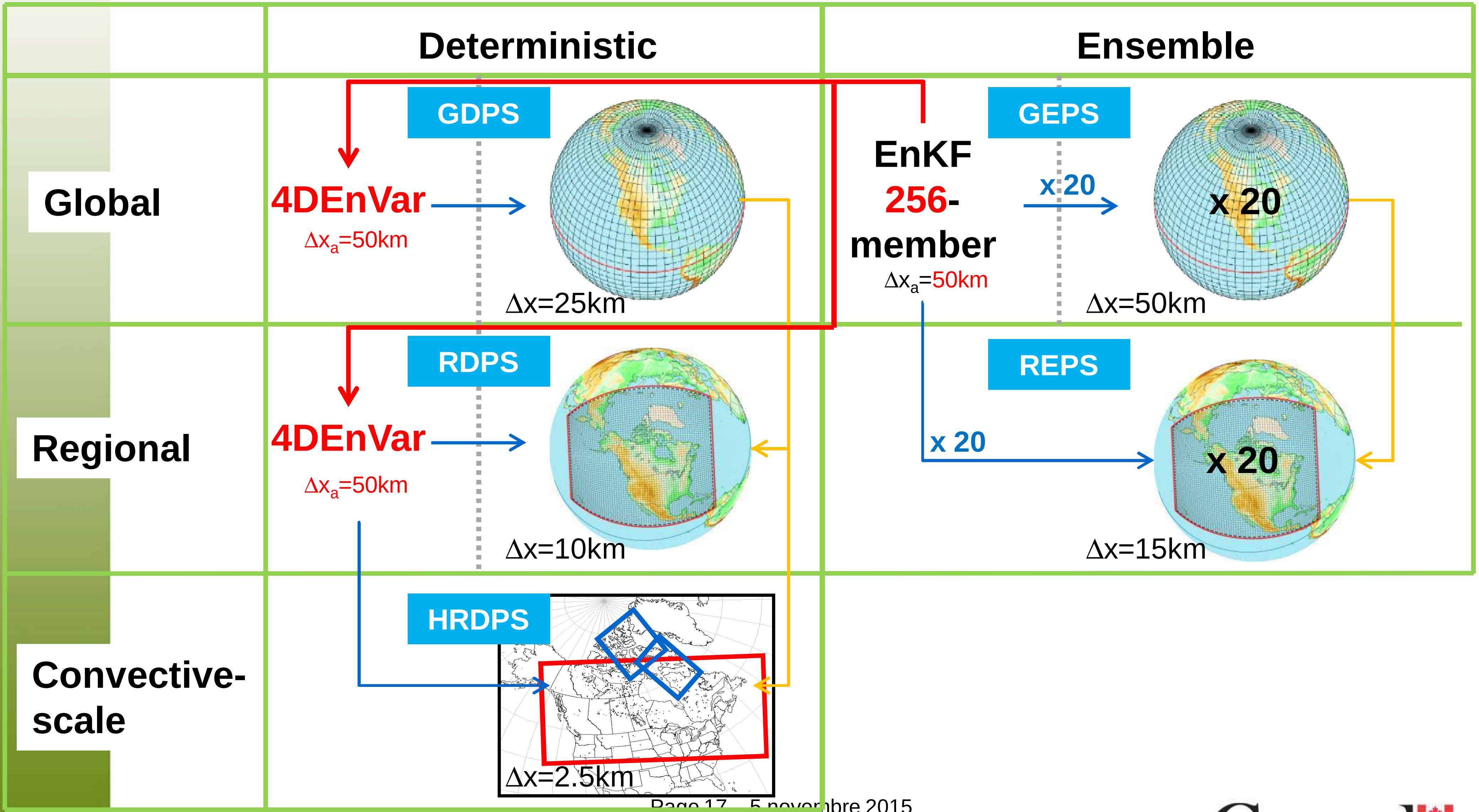
- We first showed that **the global-based EnVar approach** developed for the GDPS can **provide RDPS forecasts with either similar or slightly improved accuracy** compared to the operational configuration based on a limited-area 4D-Var for the LAM and a global 3D-Var for the driver

- A new operational RDPS configuration based on the EnVar approach was implemented in November 2014
  - The EnVar scheme is based on the ensemble of backgrounds from the upgraded global EnKF system (50-km, 256 members; as in the GDPS)
  - Several improvements to the processing and the volume of data assimilated were introduced (as in the GDPS)
  - The RDPS intermittent cycling strategy is now initialized with the improved EnVar-based GDPS analysis
  - The LAM analysis computational cost is reduced by an order of magnitude

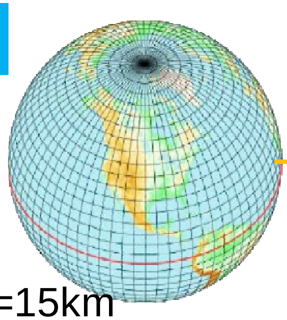
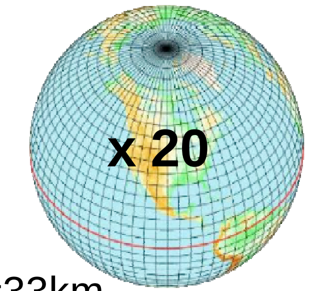
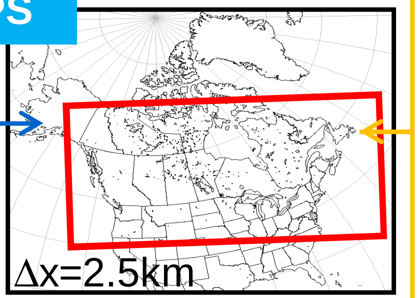
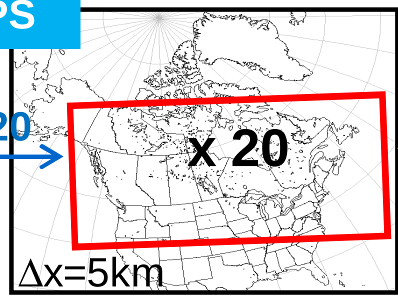
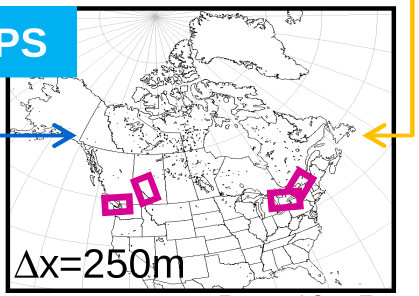
- To find out more see Buehner *et al.* and Caron *et al.* in the July 2015 issue of *Monthly Weather Review*.

! REPEAT OF A PREVIOUS SLIDE !

EC's NWP systems since 11/2014



EC's NWP systems in ~2020

	Deterministic	Ensemble
Global	GDPS 4DEnVar Δx_a=33km  Δx=15km	GEPS EnKF 512-member Δx_a=33km x 20  Δx=33km
Convective-scale	RDPS 4DEnVar Δx_a=5km  Δx=2.5km	REPS EnKF 256-member Δx_a=5km x 20  Δx=5km
Urban-scale	HRDPS  Δx=250m	

Disclaimer:

Tentative!

Part 2

Scale-Dependent Covariance Localization in EnVar

Scale-dependent covariance localization
Motivation

- Spatial covariance localization essential to obtain useful analyses with "small" ensembles (a 256-member ensemble is still "small"!).

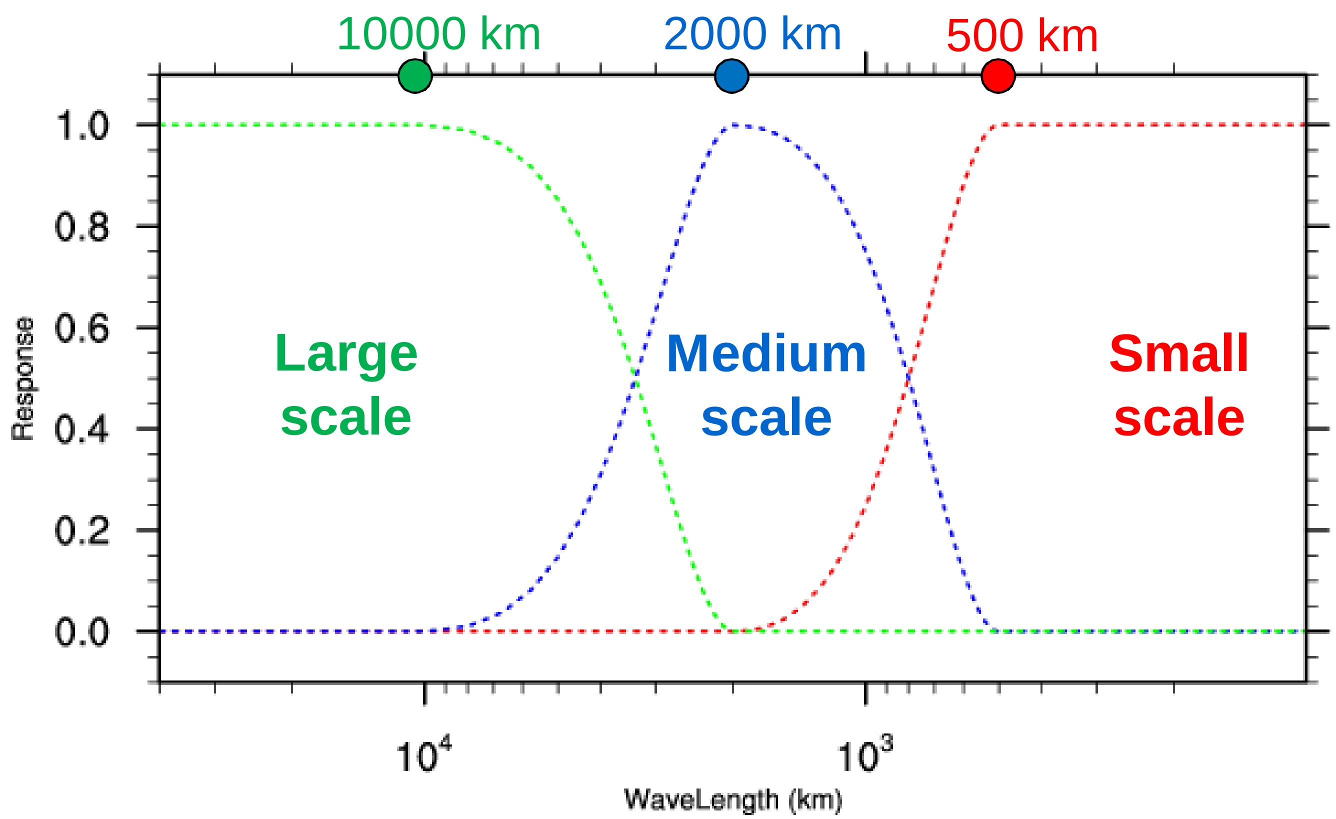
- Currently, EnVar uses simple localization of ensemble covariances, similar to EnKF: single length scale in both horizontal and vertical localizations based on Gaspari and Cohn (1999) 5th order piecewise rational function.

- Comparing various NWP studies, seems that the best amount of horizontal localization depends on application/resolution:
  - convective-scale assimilation: ~10km
  - mesoscale assimilation: ~100km
  - global-scale assimilation: ~1000km – 3000km

- Proposed approach: Simultaneously apply appropriate (i.e. different) localization to different range of scales: Scale-dependent localization.

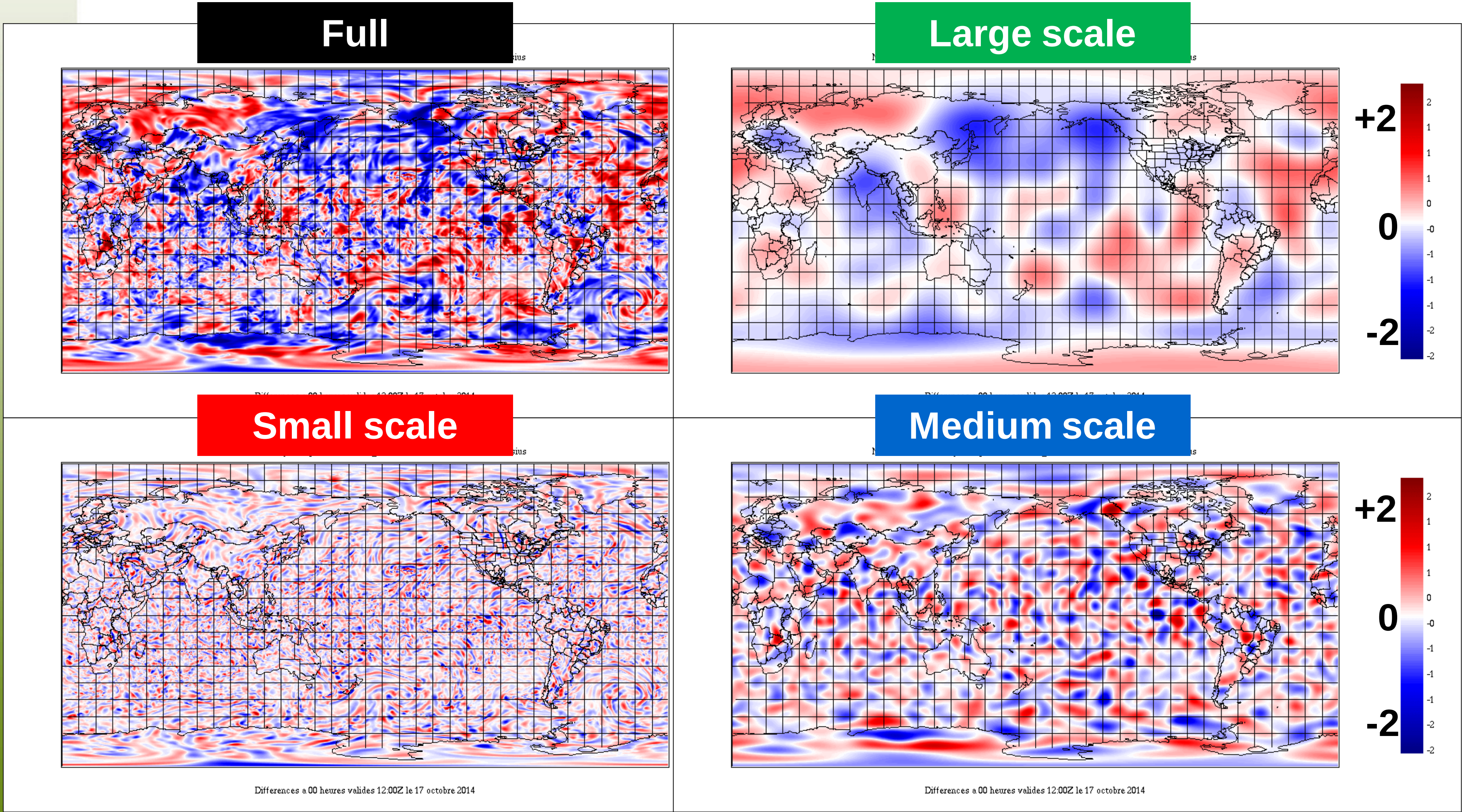
Horizontal Scale Decomposition

Filter response functions for decomposing with respect to 3 horizontal scale ranges



Horizontal Scale Decomposition

Perturbations for ensemble member #001 – Temperature at ~700hPa



Scale-dependent covariance localization

Implementation in EnVar

Current (standard) Approach

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

$$\Delta \mathbf{x} = \sum_k \mathbf{e}_k \circ \left(\mathbf{L}^{1/2} \boldsymbol{\xi}_k \right)$$

k: member index

Scale-dependent Approach (Buehner and Shlyaeva, 2015, submitted to *Tellus*)

- **Varying amounts of smoothing applied to same set of amplitudes for a given member**

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

k: member index
j: scale index

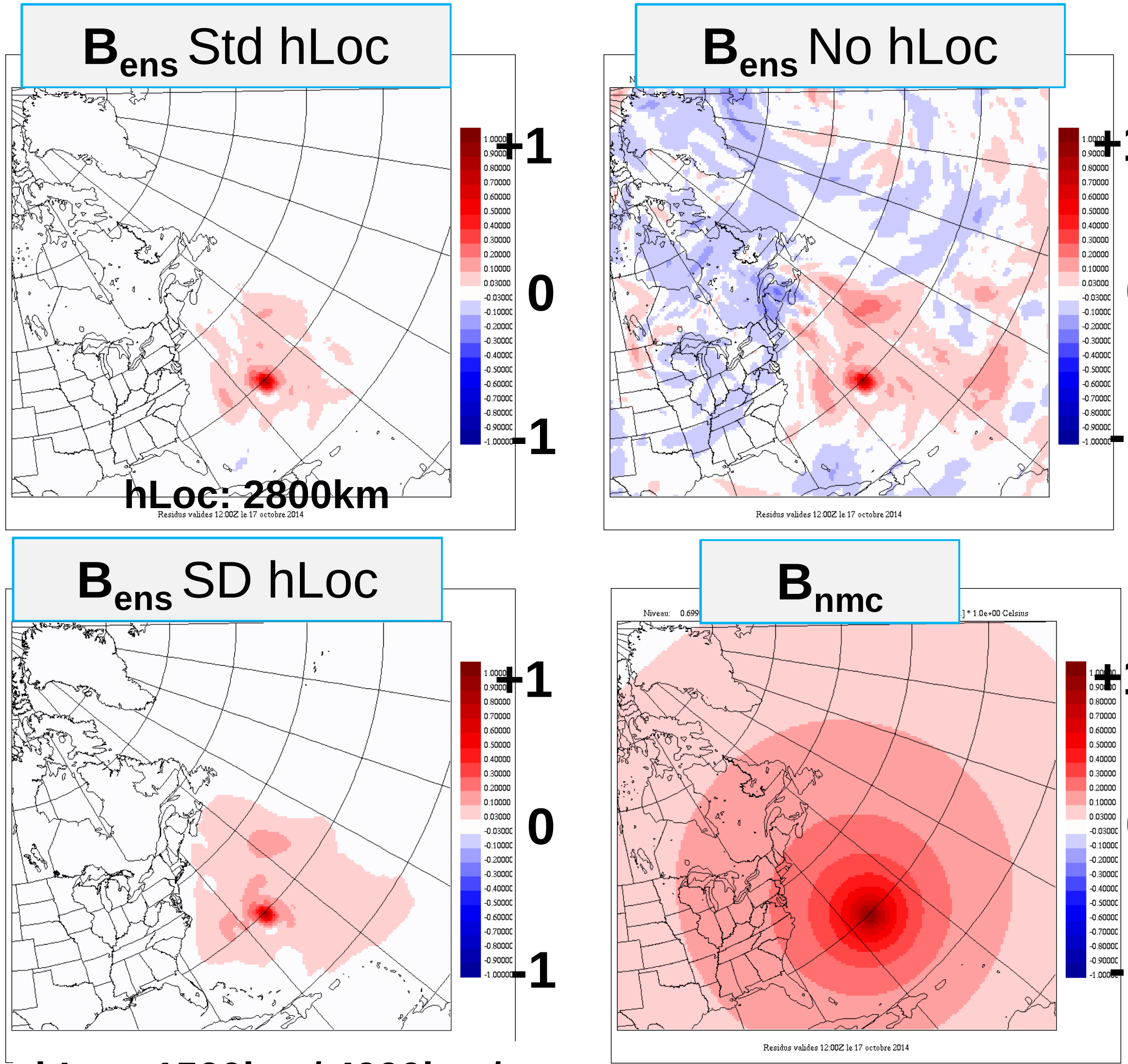
where $e_{k,j}$ is scale *j* of normalized member *k* perturbation

Scale-dependent covariance localization

Impact in single observation DA experiments

700 hPa T observation at the center of **Hurricane Gonzalo** (October 2014)

Normalized temperature increments (correlation-like) at 700 hPa resulting from various B matrices.



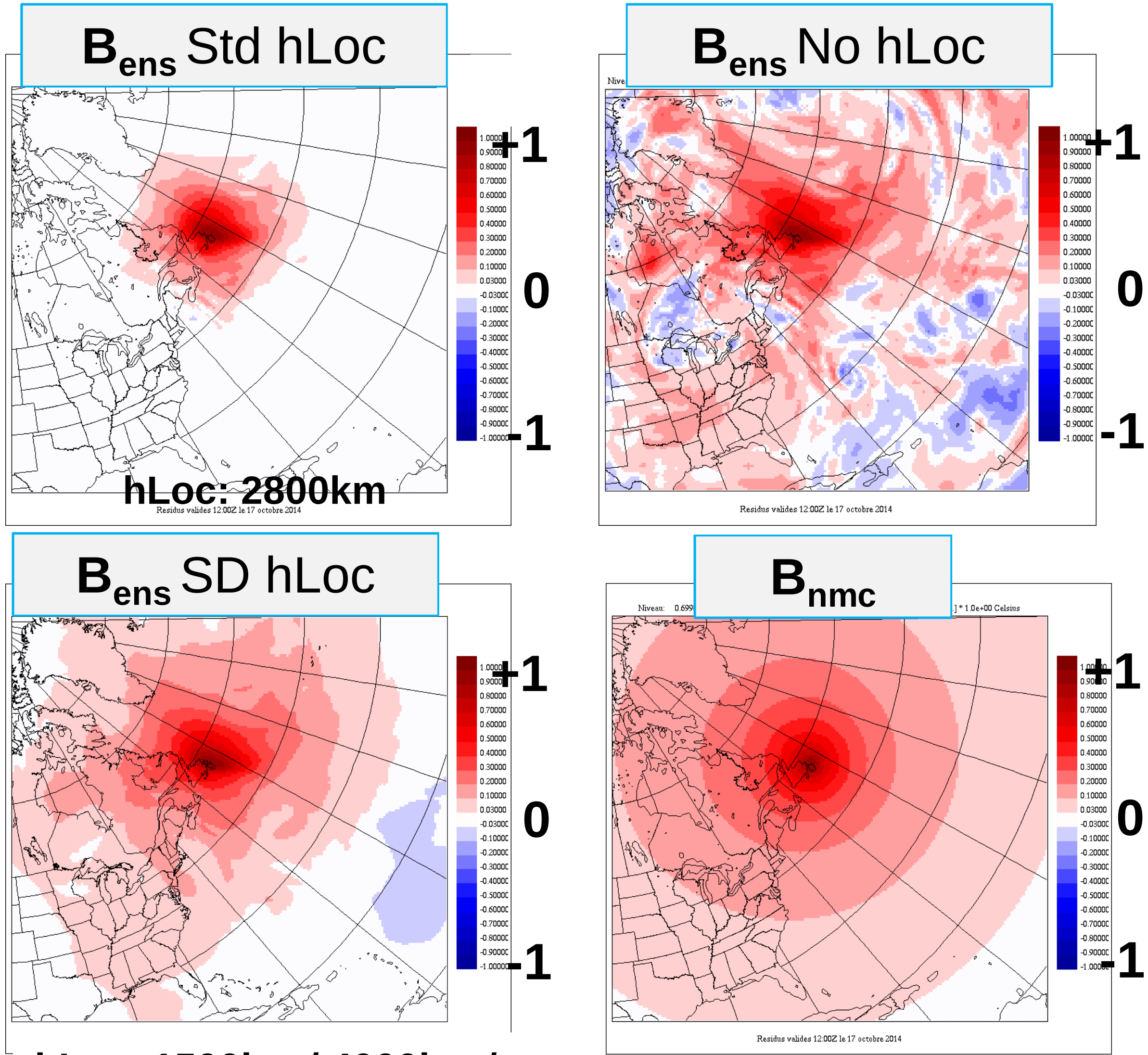
hLoc: 1500km / 4000km / 10000km

Scale-dependent covariance localization

Impact in single observation DA experiments

700 hPa T observation at the center of a **High Pressure**

Normalized temperature increments (correlation-like) at 700 hPa resulting from various B matrices.



hLoc: 1500km / 4000km / 10000km

Environnement Canada Environment Canada

Canada

Scale-dependent covariance localization

Forecast impact

- 1.5-month trialling (June-July 2014) in our global NWP system.
  - Why using the global system and not the regional system? Because the positive impact from the scale-dependent localization is likely to be greater in this system since…
    - An intermittent cycling strategy is used in the regional system
    - The global system has a wider range of horizontal scales

- 3DEnVar with 100% B_{ens} used in both experiments

1) **Control experiment** with hLoc = 2800 km, vLoc = 2 units of ln(p)

2) **Scale-Dependent experiment** with a 3 horizontal-scale decomposition

   I. Small scale uses hLoc = 1500 km
   II. Medium scale uses hLoc = 2400 km
   III. Large scale with uses = 3300 km

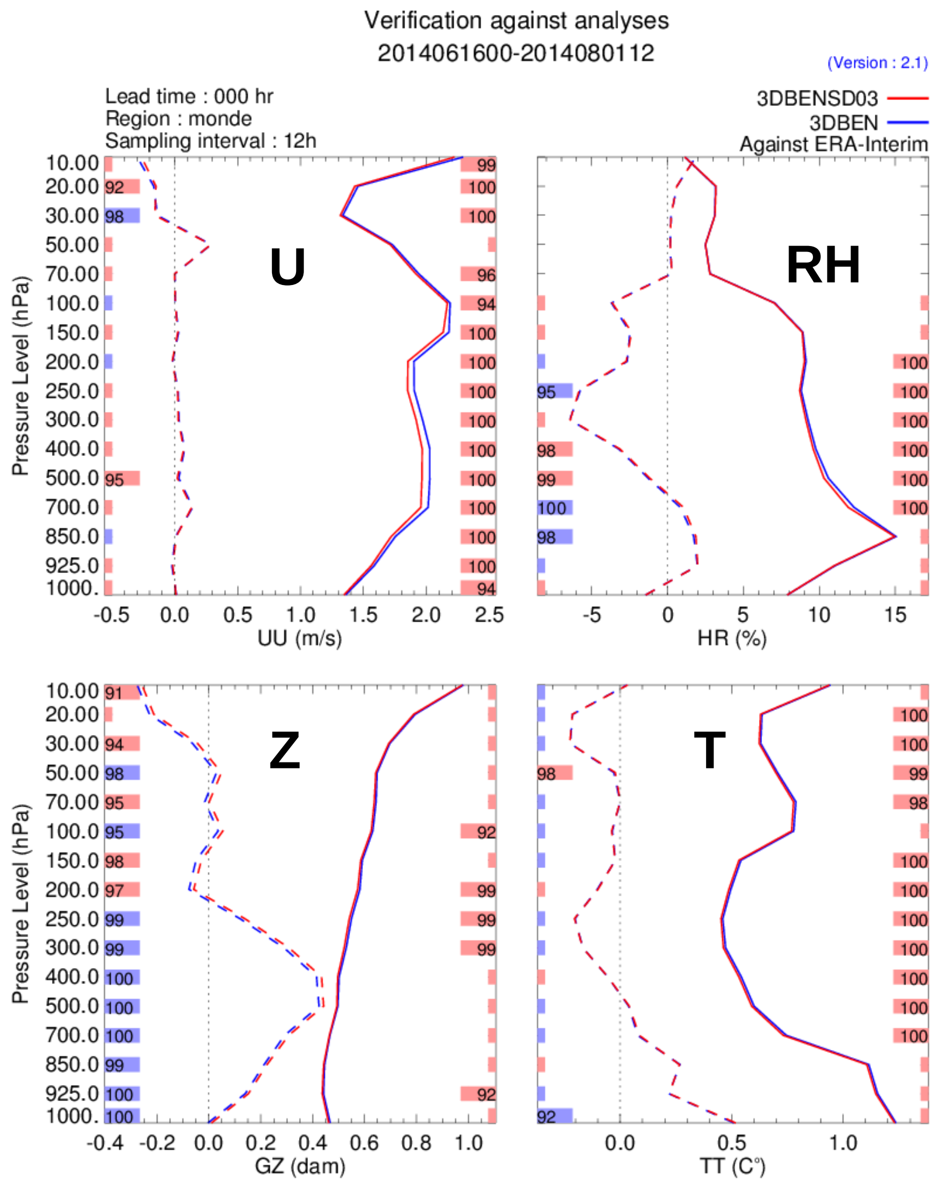
   Ad hoc values

Same vLoc (2 units of ln(p)) for every horizontal-scale

Scale-dependent covariance localization

Forecast impact – Comparison against ERA-Interim

- Control
- Scale-Dependent



T+0h

World

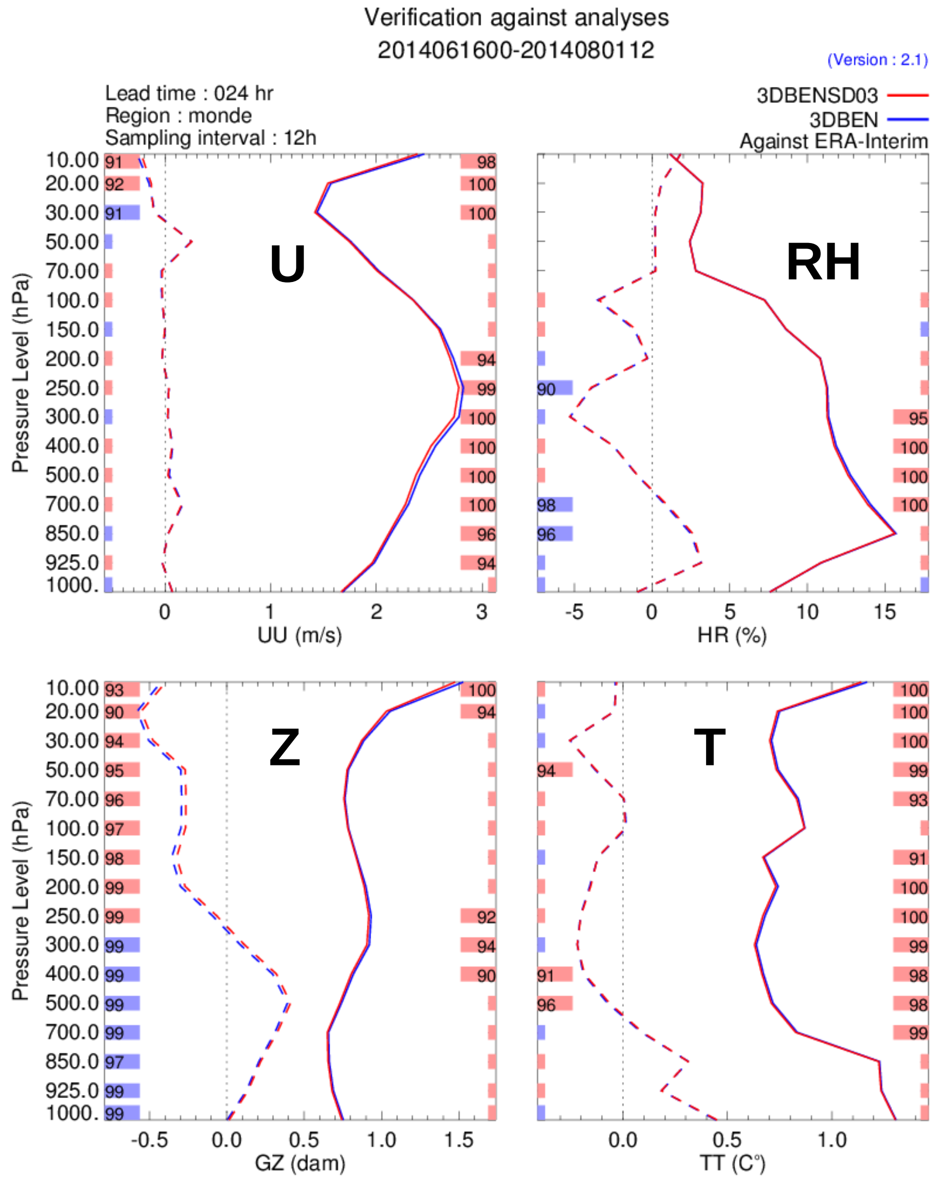
Scale-dependent covariance localization

Forecast impact – Comparison against ERA-Interim

- Control
- Scale-Dependent

T+24h

World

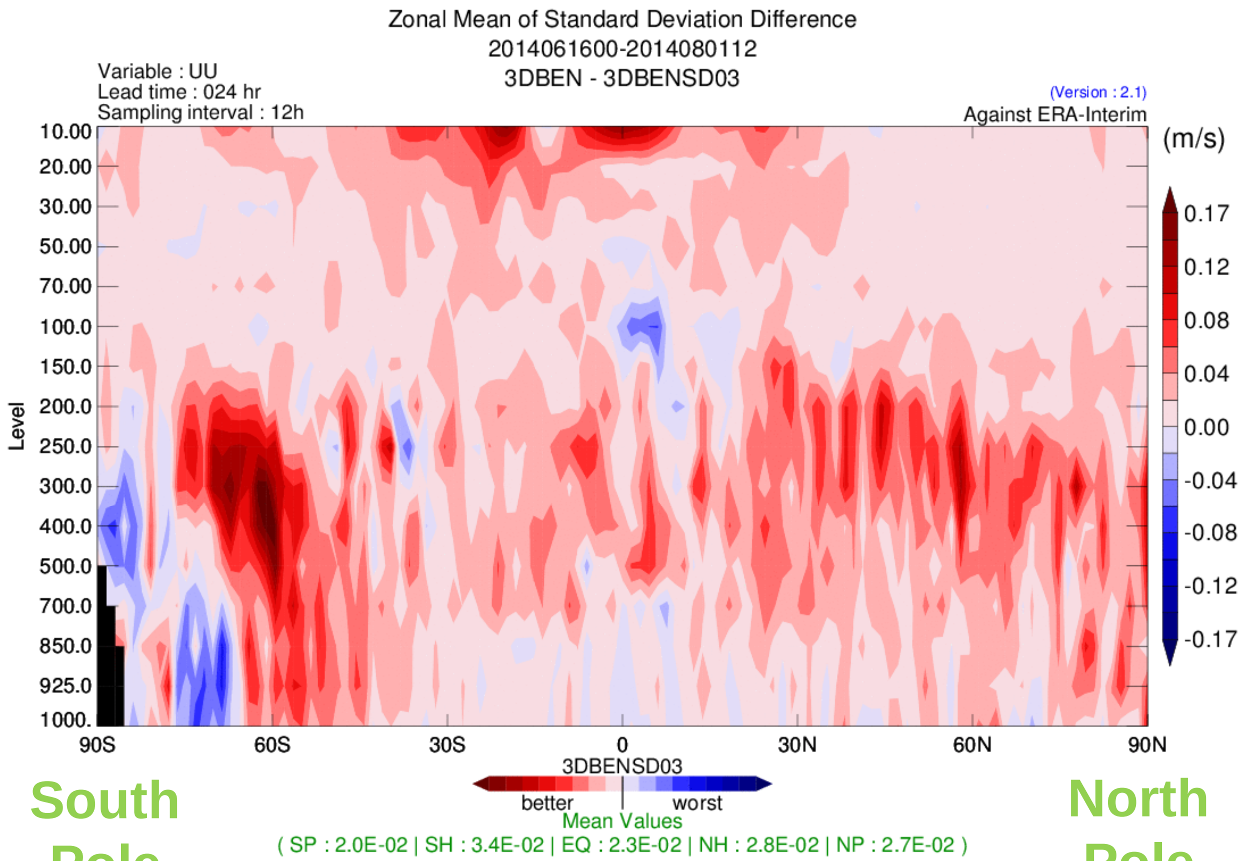


Scale-dependent covariance localization
Forecast impact – Comparison against ERA-Interim

- Control is better
- Scale-Dependent is better

T+24h
Zonal mean

Std Dev difference for U



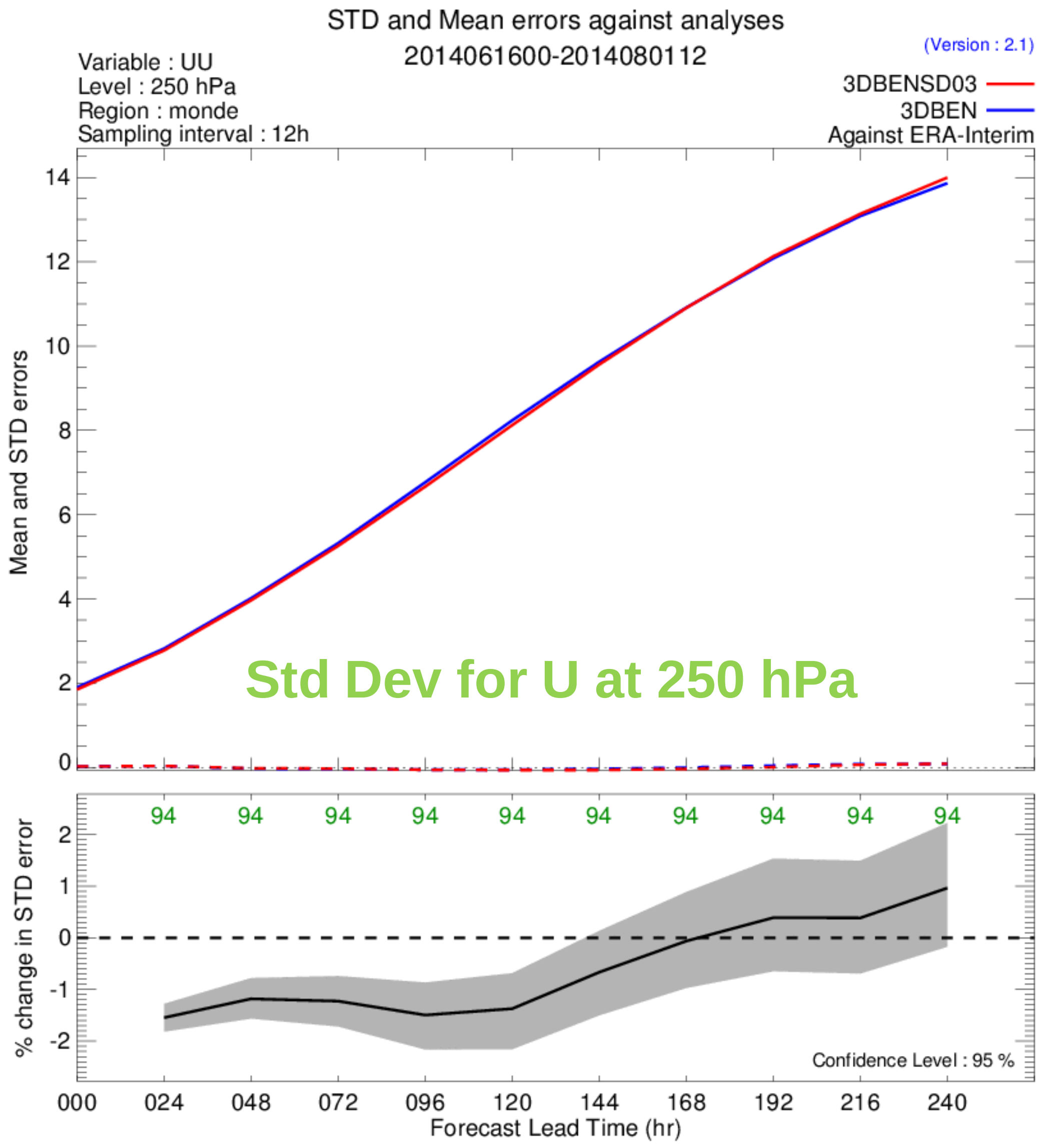
South Pole

North Pole

Scale-dependent covariance localization

Forecast impact – Comparison against ERA-Interim

- Control
- Scale-Dependent



Time Serie
World

Summary and Future Work

- Scale-dependent localization is feasible and straightforward to implement in EnVar, but more expensive than using single-scale localization.
- Preliminary results using a <u>horizontal</u>-scale-dependent <u>horizontal</u> localization indicate small forecast improvements in our global NWP system.
- Up next…
  - Optimize the horizontal localization length scales used for each horizontal-scale band. An objective evaluation approach is needed.
  - Examine the impact of <u>horizontal</u>-scale-dependent <u>vertical</u> localization
  - Examine the impact of <u>vertical</u>-scale-dependent <u>vertical</u> localization.

Questions?

EnVar Formulation - Preconditioning

- Preconditioned cost function formulation at EC:

$$J(\xi)=\frac{1}{2}\xi^T\xi+\frac{1}{2}(H_{4D}(\mathbf{x}_b)+\mathbf{H}\Delta\mathbf{x}(\xi)-\mathbf{y})^T\mathbf{R}^{-1}(H_{4D}(\mathbf{x}_b)+\mathbf{H}\Delta\mathbf{x}(\xi)-\mathbf{y})$$

- In EnVar with hybrid covariances, the control vector (ξ) is composed of 2 vectors:

$$[\xi]=\begin{bmatrix}\xi_{\mathrm{nmc}}\\ \xi_{\mathrm{ens}}\end{bmatrix} \rightarrow [\xi_{\mathrm{ens}}]=\begin{bmatrix}\xi_{\mathrm{ens}}^{1}\\ \vdots \\ \xi_{\mathrm{ens}}^{N_{\mathrm{ens}}}\end{bmatrix}$$

- The analysis increment is computed as (e_k is k'th ensemble perturbation divided by sqrt(N_{ens}-1)):

$$\Delta\mathbf{x}(\xi)=\beta_{\mathrm{nmc}}^{1/2}\mathbf{B}_{\mathrm{nmc}}^{1/2}\xi_{\mathrm{nmc}}+\beta_{\mathrm{ens}}^{1/2}\sum_{k=1}^{N_{\mathrm{ens}}}\mathbf{e}_k\circ\left(\mathbf{L}^{1/2}\xi_{\mathrm{ens}}^{k}\right) \rightarrow \mathbf{B}=\beta_{\mathrm{nmc}}\mathbf{B}_{\mathrm{nmc}}+\beta_{\mathrm{ens}}\sum_{k=1}^{N_{\mathrm{ens}}}\left(\mathbf{e}_k\mathbf{e}_k^T\right)\circ\mathbf{L}$$