A Comparison Between 4D-Var and 4D-EnVar in the Canadian Regional Deterministic Prediction System

Jean-François Caron[1], Mark Buehner[1], Luc Fillion[1], Thomas Milewski[2] and Judy St-James[2]

[1]Data Assimilation and Satellite Meteorology Research Section, Environment Canada.

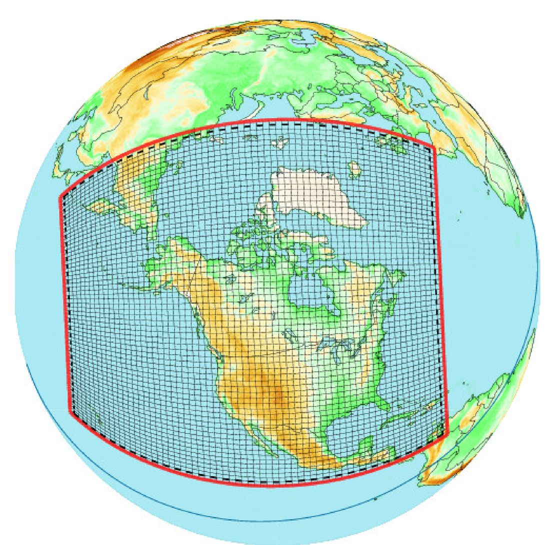
[2]Data Assimilation and Quality Control Development Section , Environment Canada.

Over the recent years, Environment Canada (EC) has devoted important resources to investigate the feasibility of replacing its 4D-Var data assimilation scheme in the global deterministic prediction system (GDPS) by a computationally cheaper variational scheme (4D-EnVar) where background error covariances are represented by a blend of climatological covariances and 4D flow-dependent covariances derived from an EnKF-based global ensemble prediction system (Buehner et al., 2013).

Following the positive results observed so far from 4D-EnVar in EC's GDPS, a similar effort was recently initiated in the regional deterministic prediction system (RDPS) which relies on a limited-area 4D-Var data assimilation scheme (Tanguay et al., 2011; operational since October 2012).

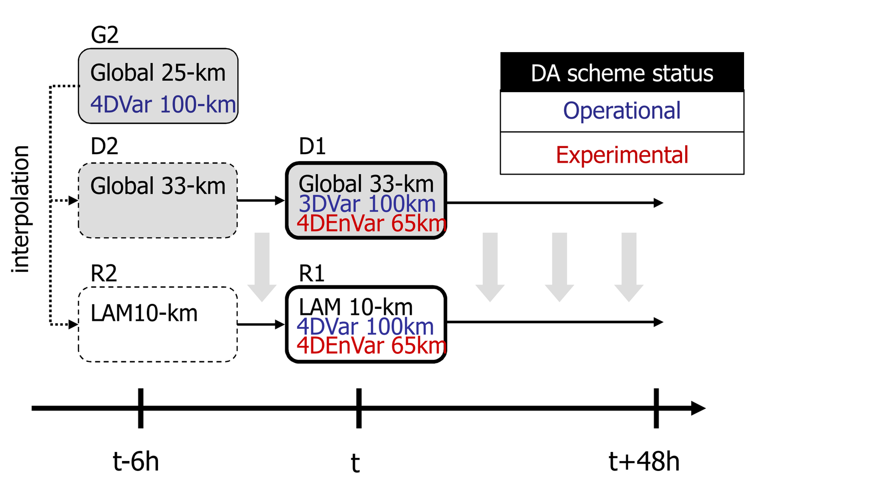
1. EC's operational RDPS in a nutshell

- Limited-area model covering North America
- 10-km grid spacing; 80 levels (lid at 0.1 hPa)
- Forecasts up t+48hrs
- 4 runs per day (00,06,12,18 UTC)
- LBCs from a simplified GDPS
- IC from 4D-Var
- Intermittent cycling



2. RDPS Data Assimilation Component

Intermittent cycling… example for an analysis at time t



Assimilated observations: All observations assimilated by the GDPS over the LAM 10-km domain

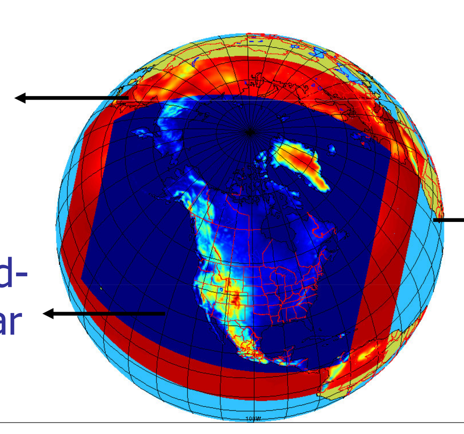
2.1. Operational Limited-Area 4D-Var

In 4D-Var the 3D analysis increment is evolved in time using GEM-LAM TL/AD forecast model (here included in H_{4D}):

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

100-km Limited-Area TL/Adjoint Model (dry)

10-km Limited-Area Nonlinear model



100-km Global Analysis Increments

B	Covariances representation (at t=0-h)	Forecast error estimates	Grid
	• Isotropic and homogeneous correlations • Control variables in terms of Helmholtz functions • Explicit Cross-covariances (no linear regression-based balance operators)	Season varying 120 x 48h–24h lagged Northern Hemisphere forecast differences	100 km / 80 levels (lid at 0.1 hPa)

2.2. Experimental Global-Based 4D-EnVar

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

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

4D-EnVar is ~10x computationally cheaper than 4D-Var

Hybrid **B** formulation

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

(e_k is k^{th} ensemble perturbation divided by sqrt(N_{ens}-1))

Since there is currently no operational equivalent to the global EnKF at the regional scale at EC, we simply used the 4D ensemble covariances derived from the global EnKF as in the GDPS experiments.

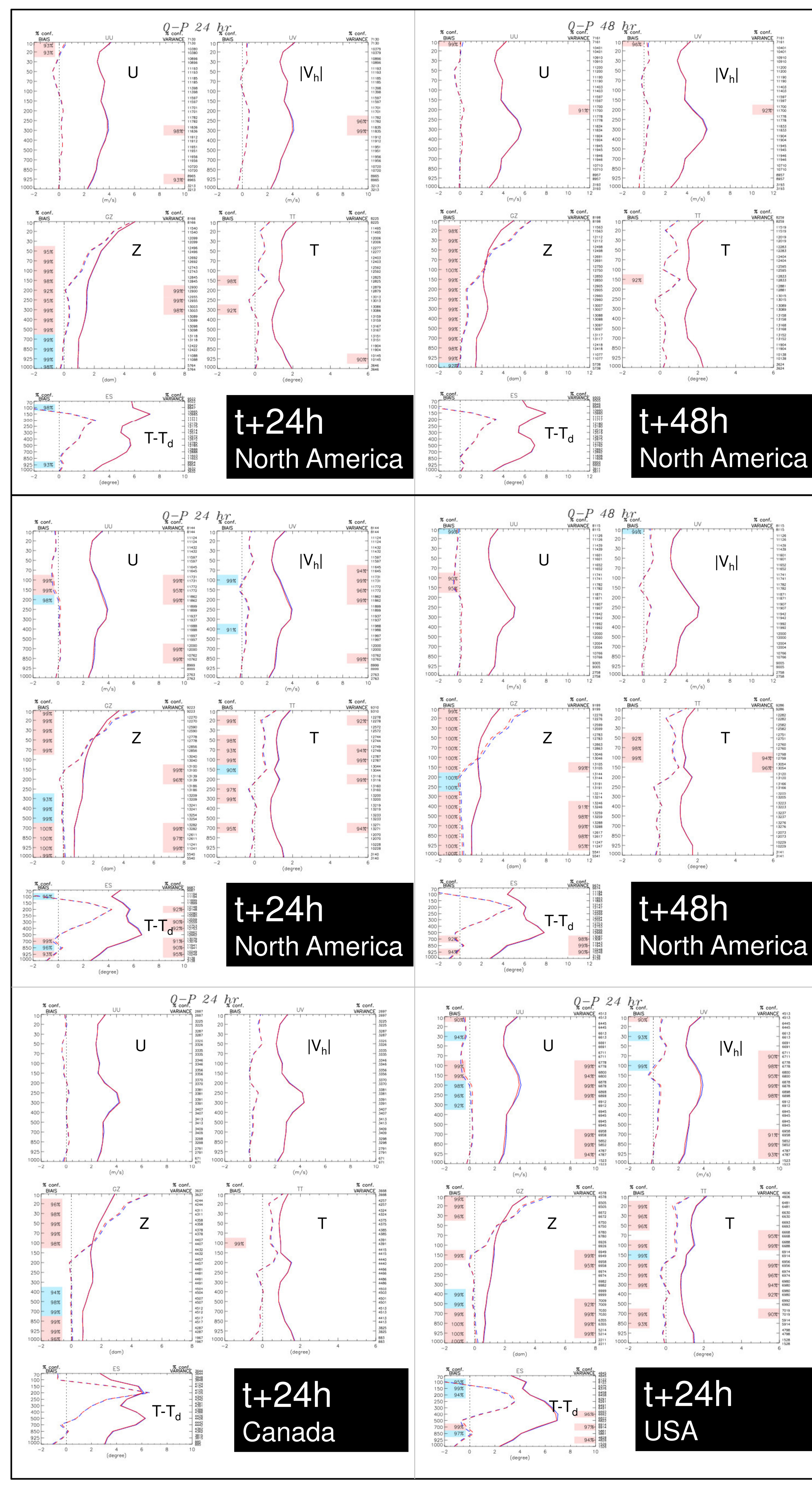
	β	Covariances representation	Forecast error estimates	Grid
B_{clim}	50%	• Isotropic and homogeneous correlations • Control variables in terms of Helmholtz functions • Cross-covariances through balance operators based on linear regression	Season varying 120 x 48h–24h lagged global forecast differences	65 km / 80 levels (lid at 0.1 hPa)
B_{ens}	50%	• Flow dependent with spatial covariance localization • Control variables in terms of model variables • No balance operators: explicit cross-covariances between each variables	192 x 3-h to 9-h global EnKF perturbations (every 1-h) valid at analysis time	65 km / 72 levels (lid at 2 hPa)

3. Verification

➢ Results showed that a global-based 4D-EnVar scheme can provide RDPS forecasts slightly improved compared to the operational limited-area 4D-Var scheme, particularly during the first 24-h of the forecasts and in summertime convective regime where the lack of moist physical processes representation in our TL/Ad model and the lower resolution of the analysis increments impede the performances of our 4D-Var scheme.

3.1. Against Radiosondes

Red (**Blue**) boxes indicate that forecasts from 4D-EnVar are statistically significantly better (worst) than forecasts from 4D-Var



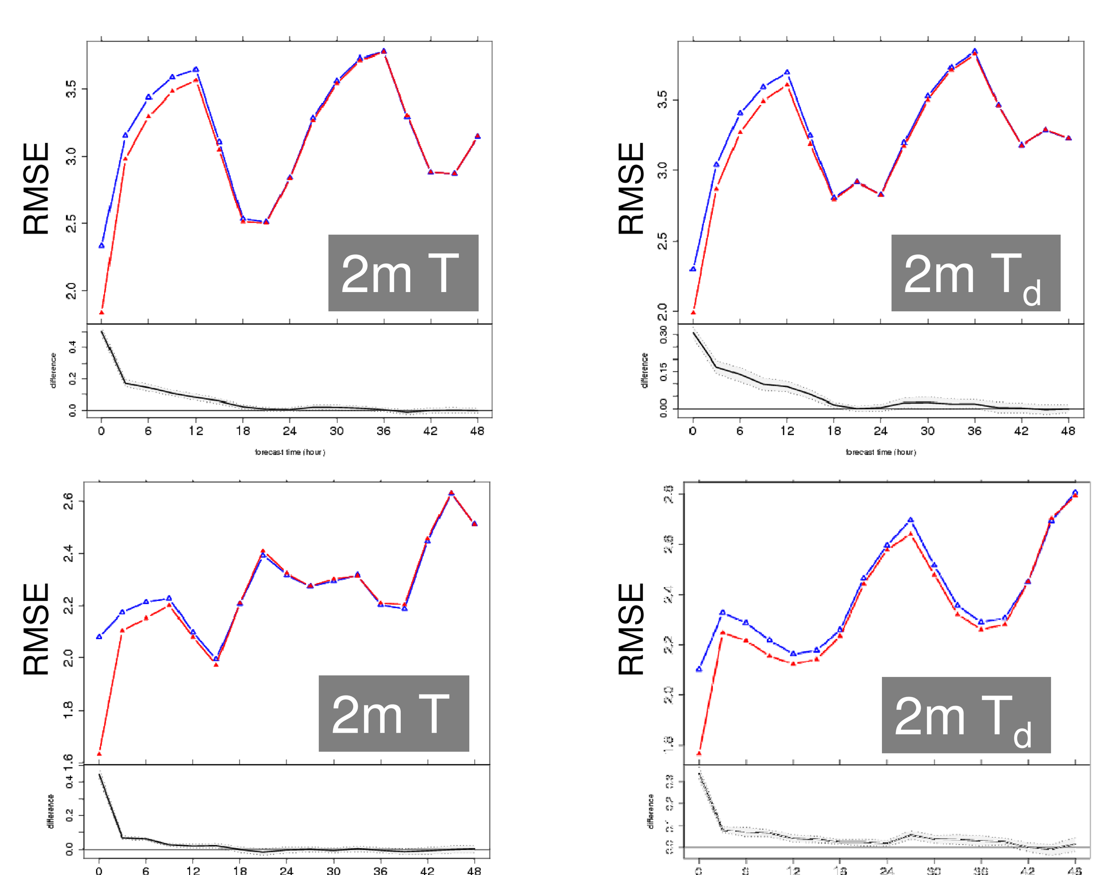
4D-EnVar vs. 4D-Var **Winter** 2011 (120 cases)

4D-EnVar vs. 4D-Var **Summer** 2011 (120 cases)

Solid Lines = Obs –Forecast **Standard Deviation**

Dashed Lines = Obs - Forecast **Mean**

3.2. Against Surface Reports



4D-EnVar vs. 4D-Var **Winter** 2011 (60 00Z cases)

4D-EnVar vs. 4D-Var **Summer** 2011 (60 00Z cases)

4. Looking Forward

Operational RDPS planned upgrade for Fall 2014

- 4D-EnVar using global 50-km EnKF with 256 members
- Improved satellite radiances bias correction method
- Balloon drift taken into account for all the assimilated radiosondes
- Additional observations
  - Ground-Based GPS
  - Additional IR channels from AIRS and IASI

References

Buehner, M., J. Morneau, and C. Charette, 2013: Four-Dimensional Ensemble-Variational Data Assimilation for Global Deterministic Weather Prediction, *Nonlin. Processes Geophys.*, **20**, 669-682.

Tanguay, M., L. Fillion, E. Lapalme, and M. Lajoie, 2011: Four-Dimensional Variational Data Assimilation for the Canadian Regional Deterministic Prediction System. *Mon. Wea. Rev.*, **140**, No. 5, 1517-1538.