Operational implementation of variational data assimilation

Pierre Gauthier
Department of Earth and Atmospheric Sciences
Université du Québec à Montréal (CANADA)

Presentation at the
2010 ESA Earth Observation Summer School on Earth system monitoring and modeling
Frascati, Italy, 2-13 August 2010

Historical perspective

- **Motivation for the development of 3D-Var**
  - Improve our capacity to use new types of observations particularly satellite radiances (Eyre, 1989; Thépaut and Moll, 1990)
  - New background-error statistics models without data selection
  - Extension to 4D-Var (Talagrand and Courtier, 1987)

- NCEP (1992), ECMWF (1996), Météo-France and CMC (1997), MetOffice (1999)
- HIRLAM and ALADIN (Météo-France) (for limited-area models)

Historical perspective (2)

- **Dual 3D-Var (Courtier, 1997)**
  - NASA's Global Modeling and Assimilation Office (GMAO) (Cohn et al., 1998)
  - Naval Research Laboratory (Daley and Barker, 2000)
- **4D-Var**
  - ECMWF (1997), Météo-France (2000), MetOffice (2004), JMA (2005), Meteorological Service of Canada (2005), NRL (2009)

Plan of presentation

- **3D-Var**
  - Introduction of the incremental formulation
  - First-Guess at Appropriate Time (FGAT)
- **4D-Var**
  - Extension from 3D to 4D-Var
  - Incremental formulation
  - Evaluation of the impact of the first implementation of 4D-Var at the Meteorological Service of Canada
- **Current issues**
  - Comparaison of 4D-Var with the Ensemble Kalman filter
  - Hybrid formulation
  - Taking into account model error: the weak-constraint 4D-Var

The variational problem

Bayes' Theorem: $$p(\mathbf{x}\,|\,\mathbf{y}) = \frac{p(\mathbf{y}\,|\,\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}$$

- **Example:**
  - Observation and background error have Gaussian distributions

$$p(\mathbf{y}\,|\,\mathbf{x}) = \frac{1}{C_3}\exp\left\{-\frac{1}{2}\left(\mathbf{y}-\mathbf{H}(\mathbf{x})\right)^T\mathbf{R}^{-1}\left(\mathbf{y}-\mathbf{H}(\mathbf{x})\right)\right\}$$

$$P(\mathbf{x}) = \frac{1}{C}\exp\left\{-\tfrac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)\right\}$$

  - $p(\mathbf{y}|\mathbf{x})$ is Gaussian only if **H** is linear
  - Maximum likelihood estimate (mode of the distribution):

$$J(\mathbf{x}) = -\ln p(\mathbf{x}\,|\,\mathbf{y})$$

$$= \tfrac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \tfrac{1}{2}(\mathbf{H}(\mathbf{x})-\mathbf{y})^T\mathbf{R}^{-1}(\mathbf{H}(\mathbf{x})-\mathbf{y})$$

- Reducing $J(\mathbf{x})$ implies an increase in the probability of **x** being the true value

Incremental approach

Successive linearizations with respect to the full model state is obtained

- Minimization of quadratic problems

$$J(\xi) = \tfrac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \tfrac{1}{2}\left(\mathbf{H}'\delta\mathbf{x} - \mathbf{y}'\right)^T\mathbf{R}^{-1}\left(\mathbf{H}'\delta\mathbf{x} - \mathbf{y}'\right)$$
$$\equiv \tfrac{1}{2}\xi^T\xi + \tfrac{1}{2}\left(\mathbf{H}'\mathbf{B}^{1/2}\xi - \mathbf{y}'\right)^T\mathbf{R}^{-1}\left(\mathbf{H}'\mathbf{B}^{1/2}\xi - \mathbf{y}'\right)$$

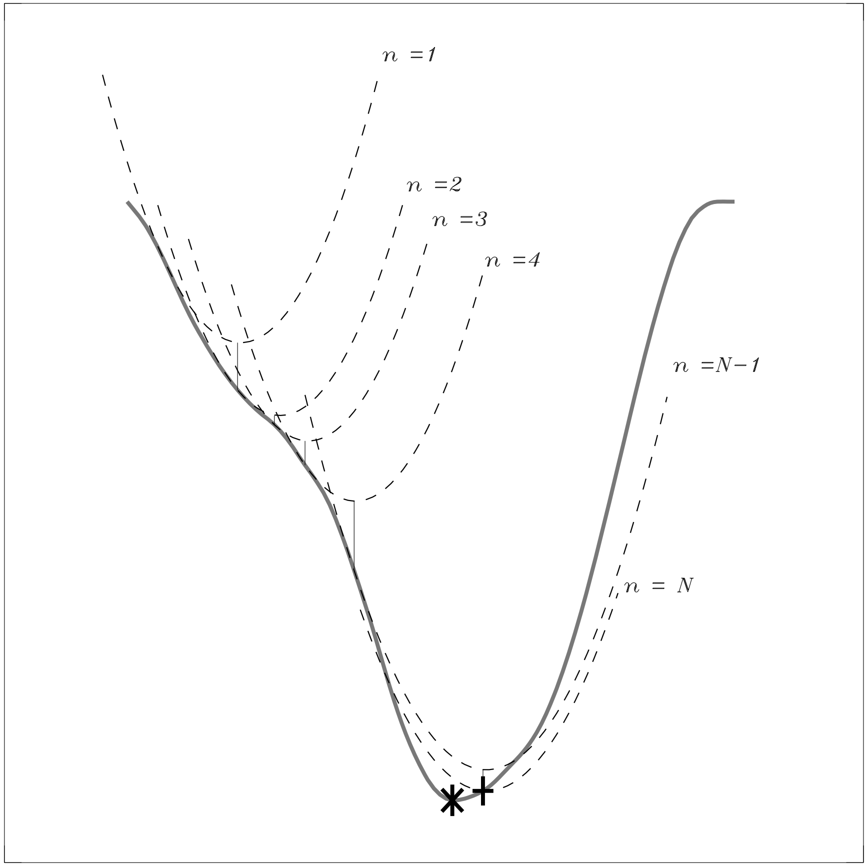
where

$\delta\mathbf{x} = \mathbf{B}^{1/2}\xi$

$\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$: increment

$\mathbf{H}' = \partial\mathbf{H}/\partial\mathbf{x}$: tangent-linear of the observation operator

$\mathbf{y}' = \mathbf{y} - \mathbf{H}(\mathbf{x}_b)$:
innovation vector (observation departure with respect to the high resolution background state)



From Laroche and Gauthier (1998)

3D-Var: variational formulation of the statistical estimation problem

Minimization of the cost function

$$J(\xi) = \frac{1}{2}\xi^T\xi + \frac{1}{2}\left(\mathbf{H}'\mathbf{B}^{1/2}\xi - \mathbf{y}'\right)^T \mathbf{R}^{-1}\left(\mathbf{H}'\mathbf{B}^{1/2}\xi - \mathbf{y}'\right)$$

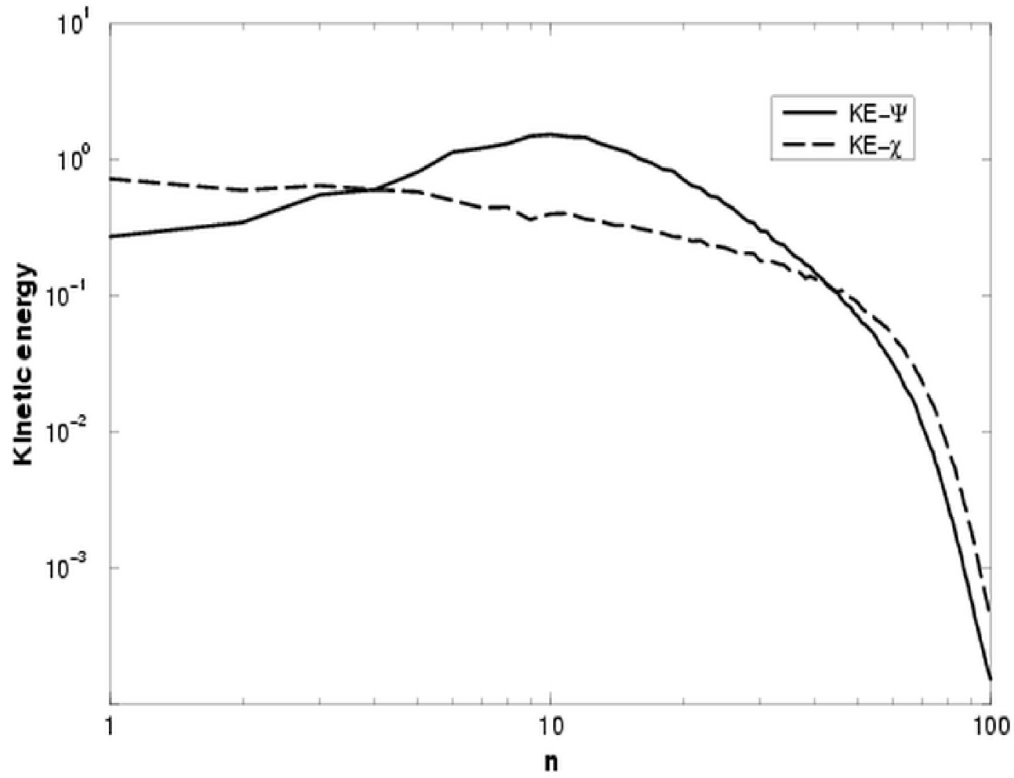
$$\delta\mathbf{x} = \mathbf{B}^{1/2}\xi$$

where

$\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$	:	increment
$\mathbf{H}' = \partial\mathbf{H}/\partial\mathbf{x}$	:	tangent-linear of the observation operator
$\mathbf{y}' = \mathbf{y} - \mathbf{H}(\mathbf{x}_b)$	:	innovation vector (observation departure) (computed with respect to the high resolution background state)

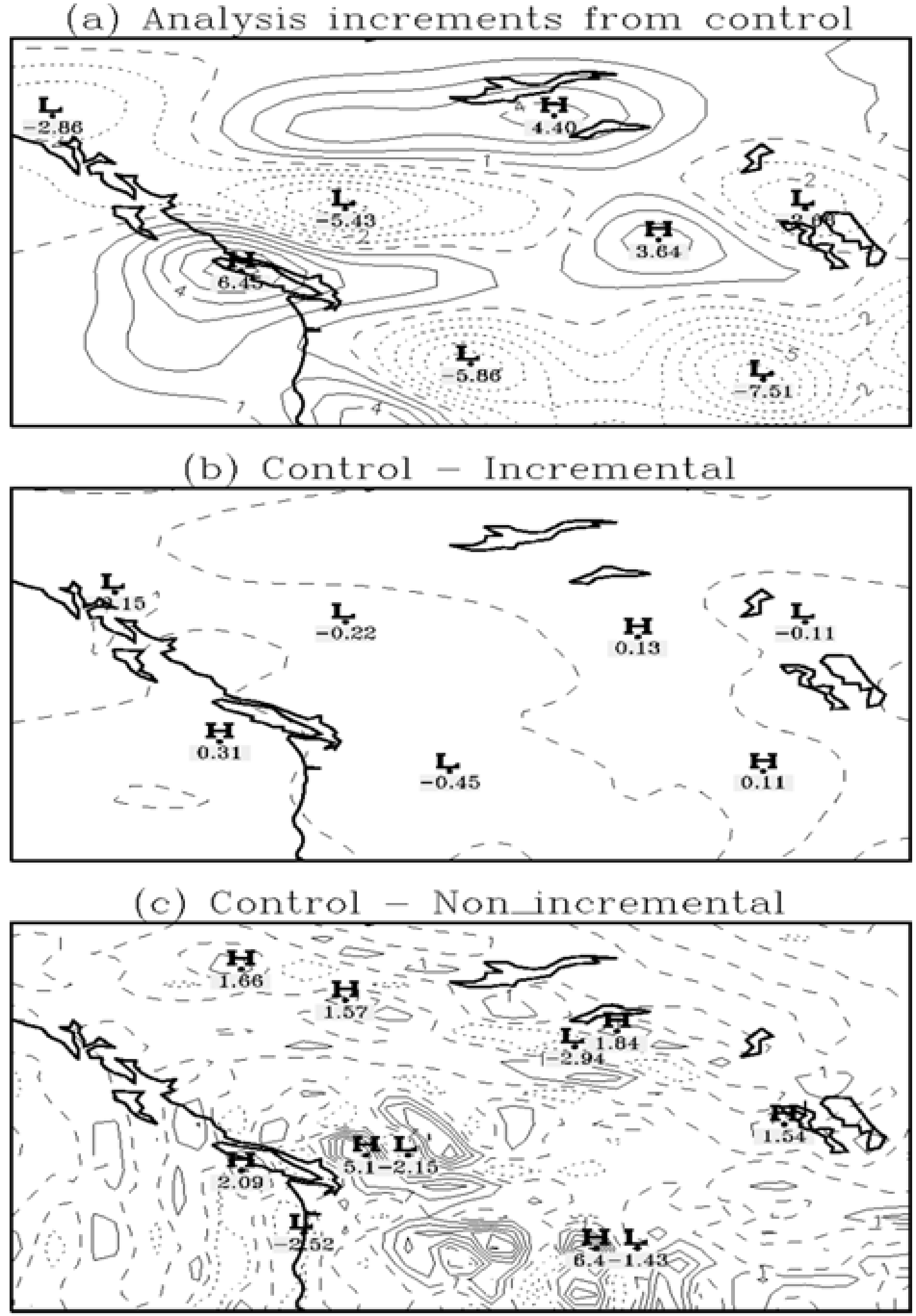
Autocorrelation spectra of rotational and divergent components of background-error

A triangular truncation n ≤ 108 (T108) resolution is sufficient to represent the whole autocorrelation spectra

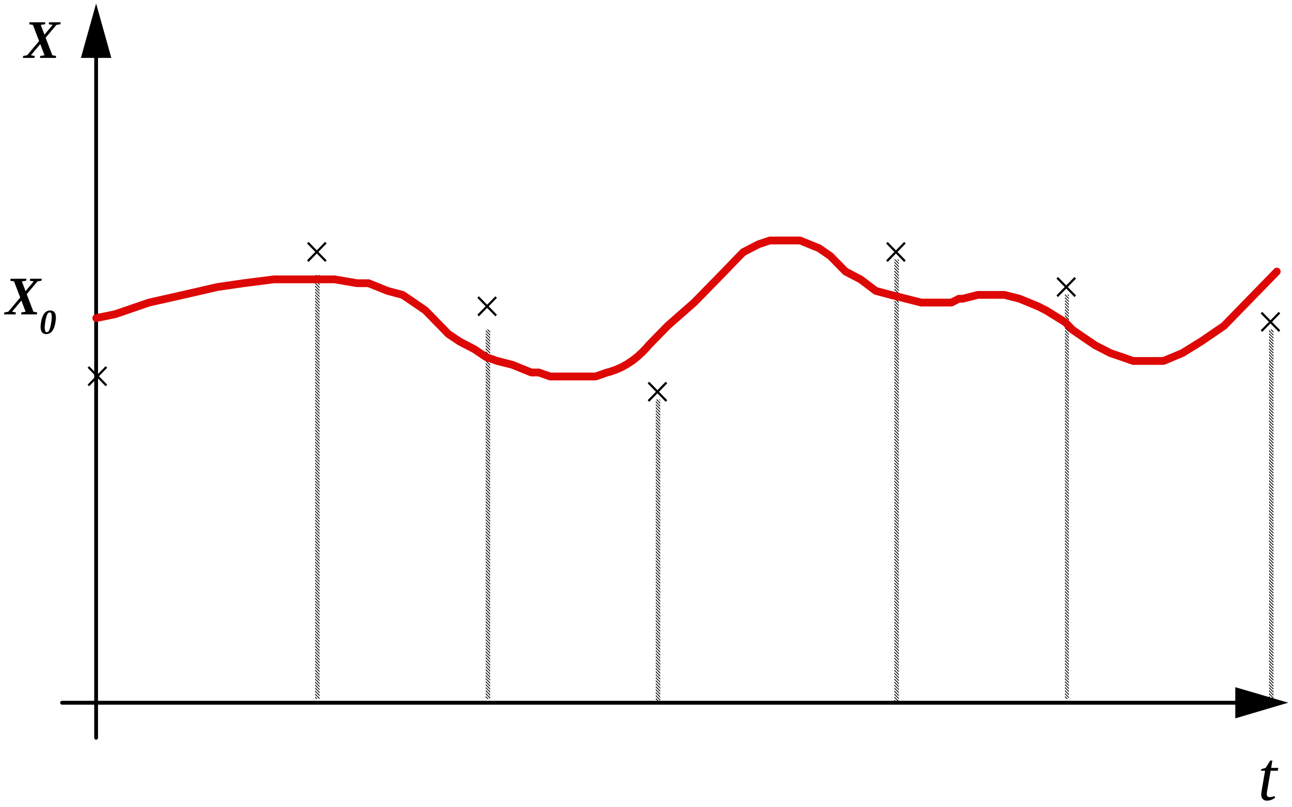


Regional analysis increment

- **Analysis increment produced at full resolution (~ 50 km) (control)**
- **Control - Incremental (increment has a resolution of ~200 km)**
- **Control - Non-incremental (innovations produced with respect to the low resolution background state)**



4D variational data assimilation (4D-Var)



- **Observation operator now involves a model integration that carry the initial conditions up to the time of the observations**

Extension of 3D-Var to 4D-Var

Cost function

$$J(\xi) = \frac{1}{2}\xi^T \xi + \frac{1}{2}\left(\mathbf{H}'\mathbf{L}(t_o, t)\mathbf{B}^{1/2}\xi - \mathbf{y}'\right)^T \mathbf{R}^{-1}\left(\mathbf{H}'\mathbf{L}(t_o, t)\mathbf{B}^{1/2}\xi - \mathbf{y}'\right)$$

- Representation of the covariances contained within the change of variables $\delta \boldsymbol{x}_0 = \mathbf{B}^{1/2}\xi$
-
- Each iteration of the minimization involves approximately 2-3 model integrations over the assimilation window ($0 < t < T$)

- Incremental formulation allows to reduce the cost of 4D-Var by using a simplified model, the tangent linear model linearized around the current model trajectory (Courtier *et al.*, 1994)

Tangent Linear model and Adjoint Model (LeDimet and Talagrand, 1986)

* **Direct Model :**

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X})$$

* **Tangent Linear Model :**

$$\frac{d\delta\mathbf{X}}{dt} = \left(\frac{\partial\mathbf{F}}{\partial\mathbf{X}}\left(\mathbf{X}_R(t)\right)\right)\delta\mathbf{X}$$

$$\equiv L\delta\mathbf{X}$$

* **Adjoint Model :**

$$\frac{d\delta^*\mathbf{X}}{dt} = -\left(\frac{\partial\mathbf{F}}{\partial\mathbf{X}}\left(\mathbf{X}_R(t)\right)\right)^* \delta^*\mathbf{X}$$

$$\equiv -L^*\delta^*\mathbf{X}$$

Example: the Lorenz (1963) model

- Direct Model

$$\frac{dX}{dt} = \sigma(-X + Y),$$
$$\frac{dY}{dt} = -XZ + rX - Y,$$
$$\frac{dZ}{dt} = XY - bZ,$$

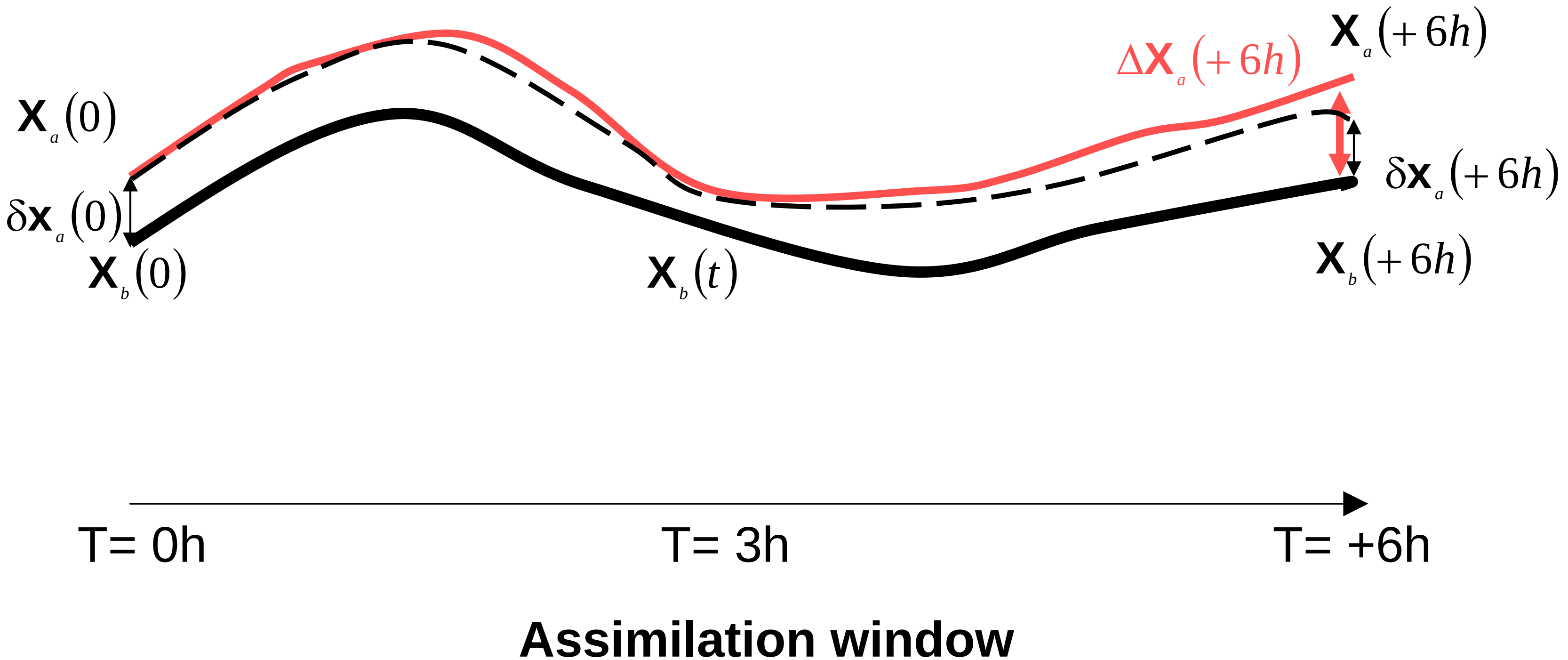
- Tangent Linear Model (TLM)

$$\frac{d}{dt}\begin{pmatrix}\delta X\\ \delta Y\\ \delta Z\end{pmatrix} = \begin{pmatrix}-\sigma & +\sigma & 0\\ -Z_R(t)+r & -1 & -X_R(t)\\ Y_R(t) & X_R(t) & -b\end{pmatrix}\begin{pmatrix}\delta X\\ \delta Y\\ \delta Z\end{pmatrix}$$

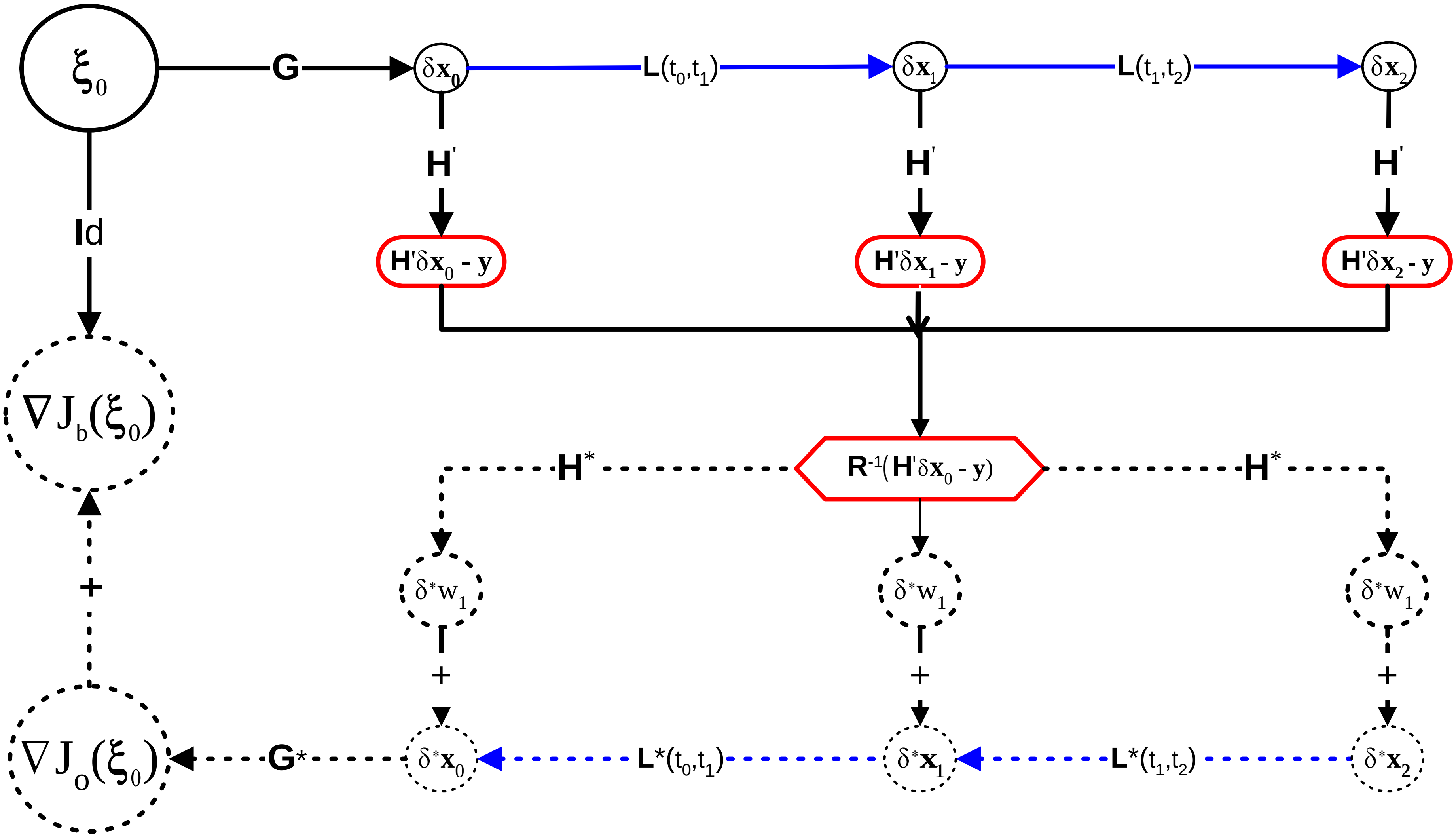
- Adjoint Model

$$\frac{d}{dt}\begin{pmatrix}\delta^* X\\ \delta^* Y\\ \delta^* Z\end{pmatrix} = -\begin{pmatrix}-\sigma & -Z_R(t)+r & Y_R(t)\\ +\sigma & -1 & X_R(t)\\ 0 & -X_R(t) & -b\end{pmatrix}\begin{pmatrix}\delta X\\ \delta Y\\ \delta Z\end{pmatrix}$$

Schematic of the incremental 4D-Var



Operations involved in a single iteration of 4D-Var



Outer and inner iterations of an incremental 4D-Var

- Integration of the operational model
  - Initial Conditions: $X_0^{(k)} = X_0^{(k-1)} + \Delta X_0^{(k-1)}$
  - Computation of observation departures $y' = y - HX^{(k)}(t)$
  - Definition of the trajectory ***X***(t) that defines the TLM and the adjoint model

Minimization of the incremental problem

- Use a simplified model (resolution and physical parameterizations)

$$\Delta X_0^{(k)} = h^{-I}\, \delta x_0^{(k)}$$

Dynamical considerations about 4D-Var

- **Predictability**
  - Limit to our ability to improve the fit to observations that are too distant in time
- **Justification of the incremental approach**
  - Can we obtain a good analysis by only approximately correcting the initial conditions?
- **Impact of model error**
  - Misfit to the observations is interpreted as an error in initial conditions
  - Error in the model may be the cause

Illustration with a barotropic model on the β-plane (Tanguay *et al.*, 1995)

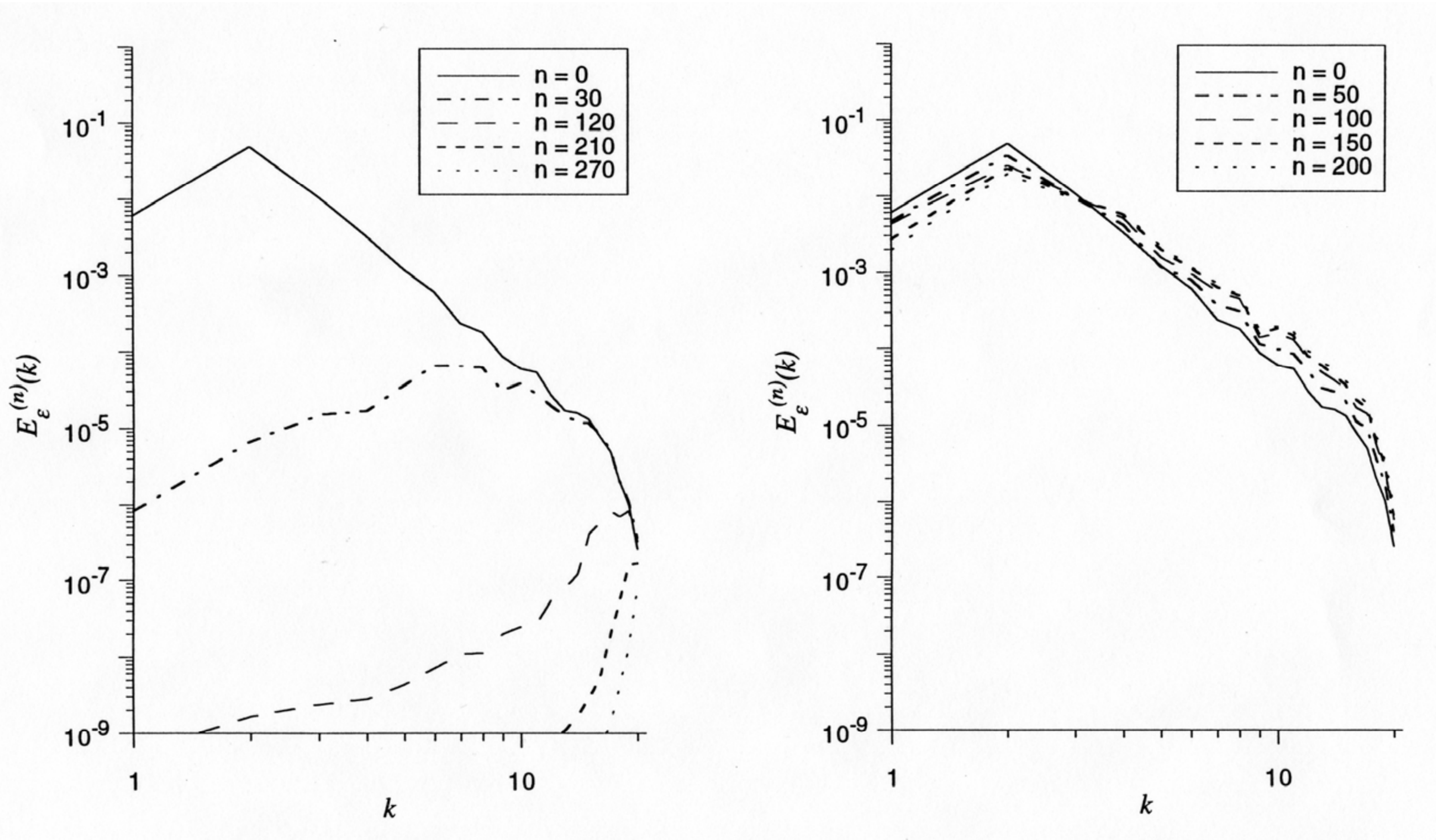
- **Model equations**

$$\frac{\partial \zeta}{\partial t} = -\frac{\partial(\Psi, \zeta + \beta y)}{\partial(x, y)} + F - E\zeta - \nu \nabla^{16} \zeta$$

	Model	Atmosphere
Length Scale	2π	7000 km
Forcing Scale	$2\pi/3$	2700 km
Rayleigh Friction	50	40 days
Advection time scale	1	0.3 days
Turnover Time	9	3 days

Error spectra as a function of iteration

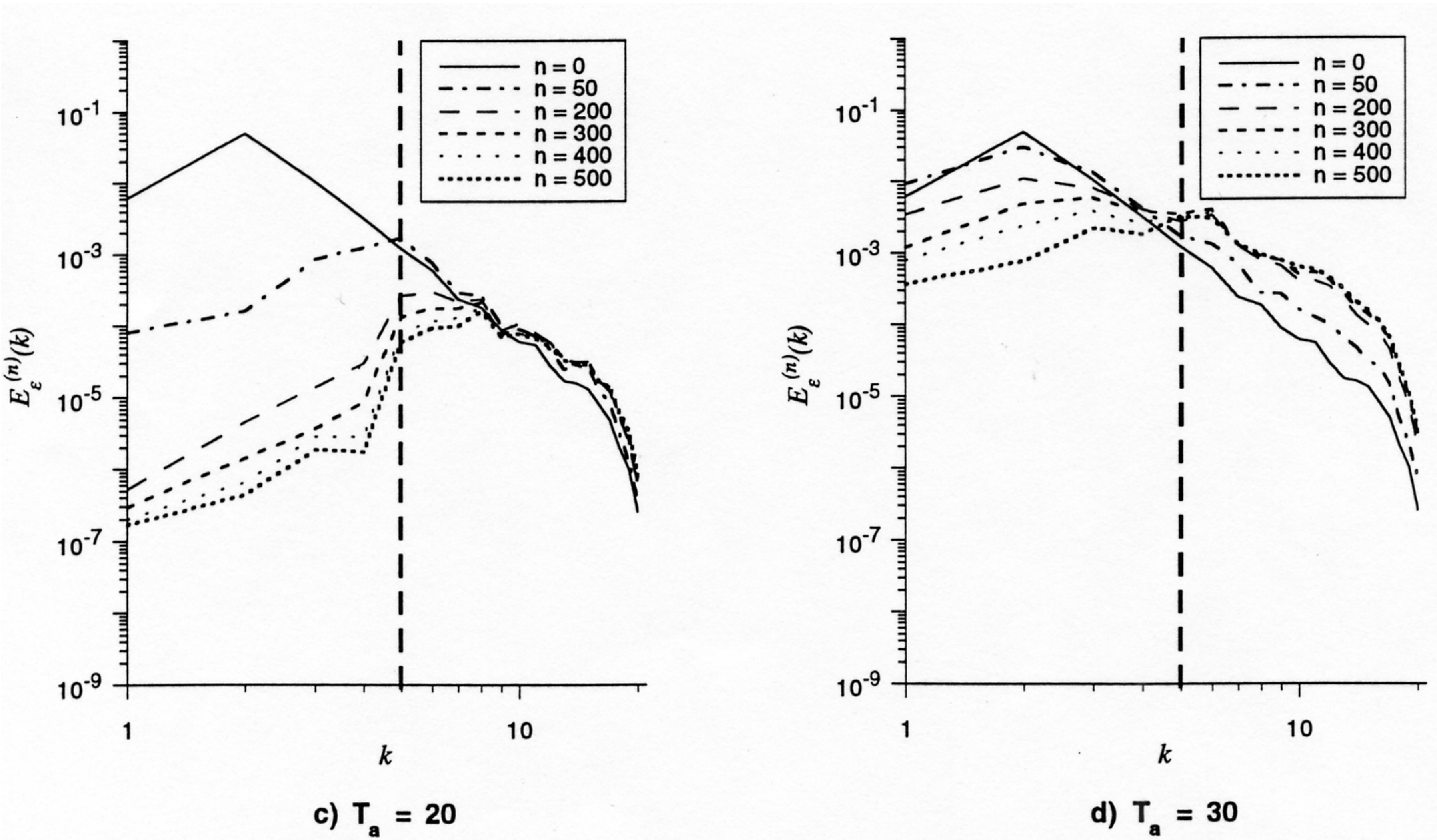
Perfect model integrations with perfect observations at all times and at every grid point (no background term)



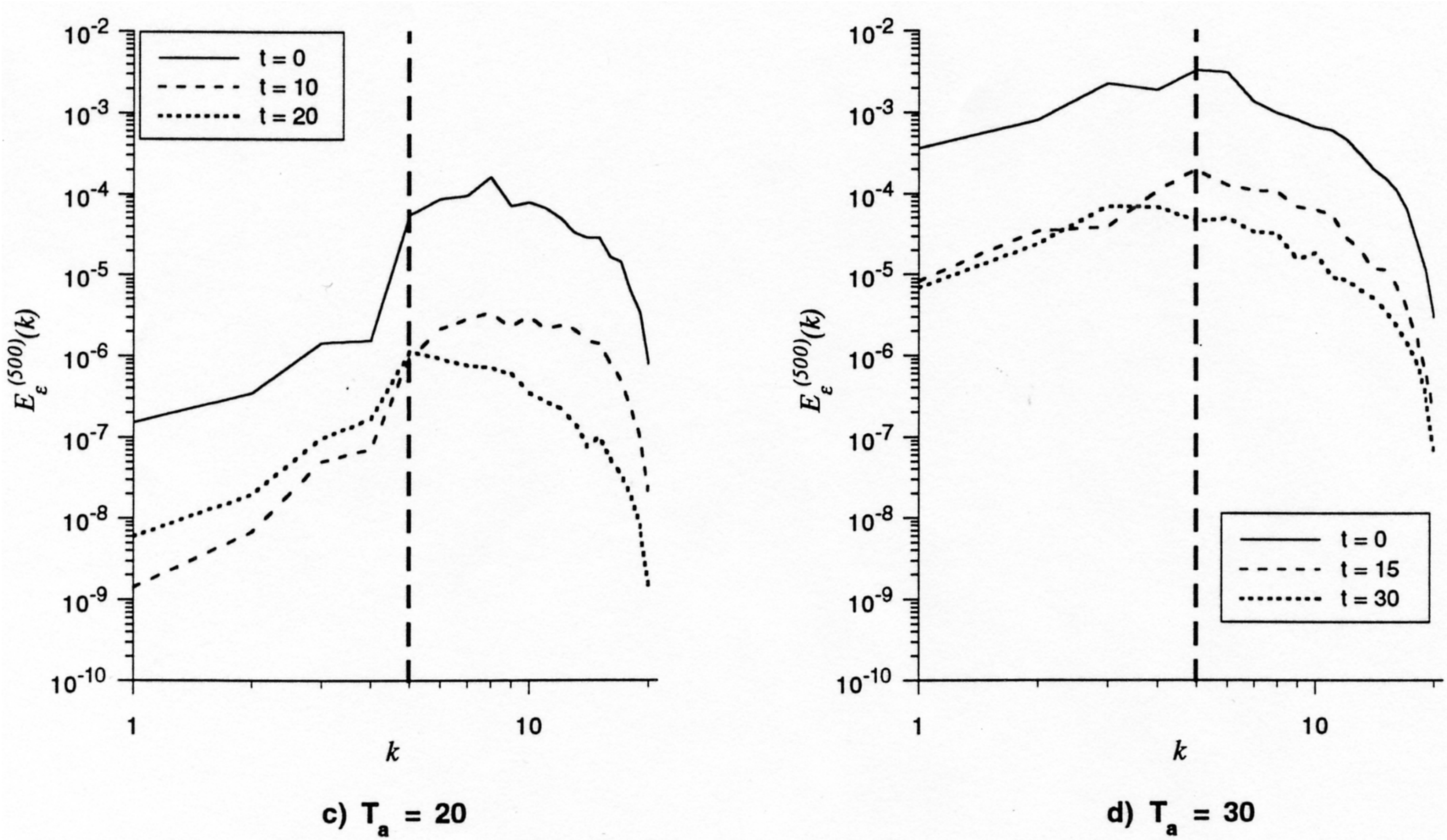
a) T_a = 10

b) T_a = 40

Error spectra as a function of iteration
Large scale perfect observations

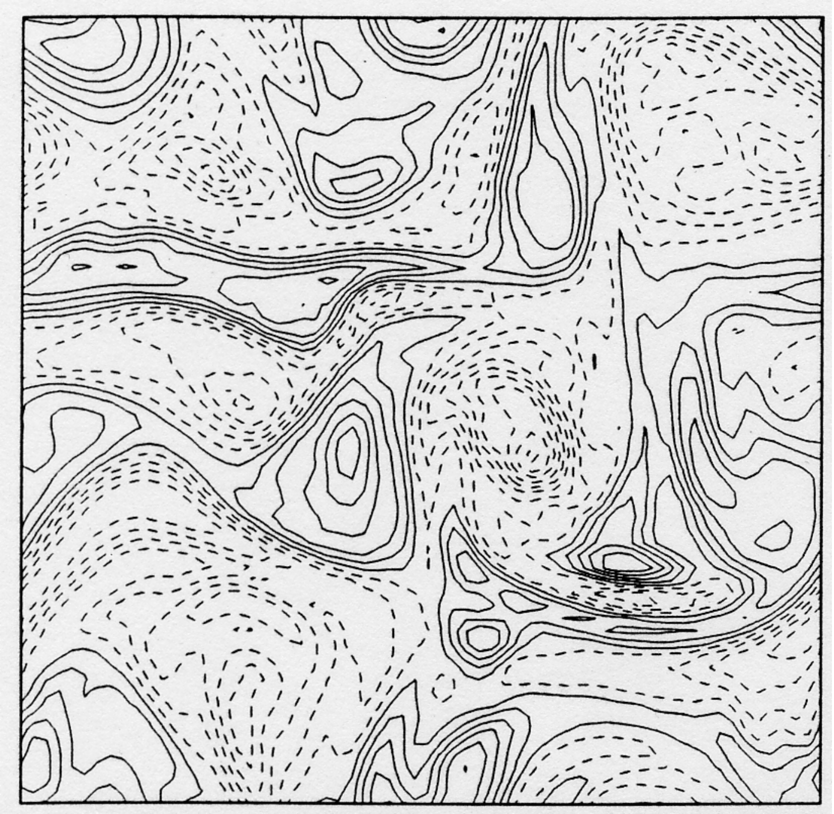


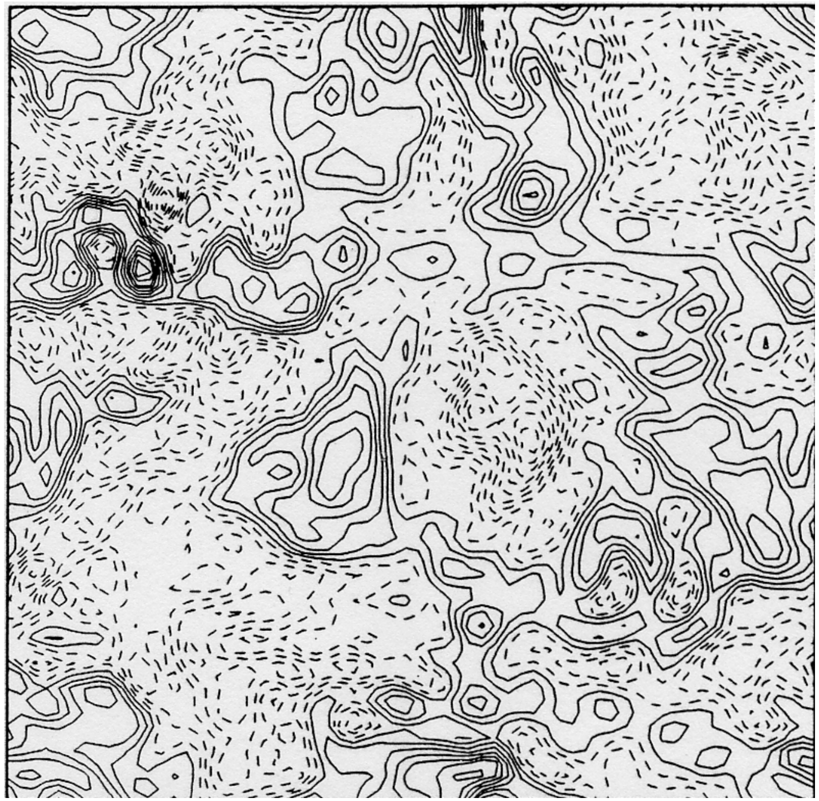
Time evolution of the error spectrum



Perfect Model 4D-Var assimilation (T_a = 20)

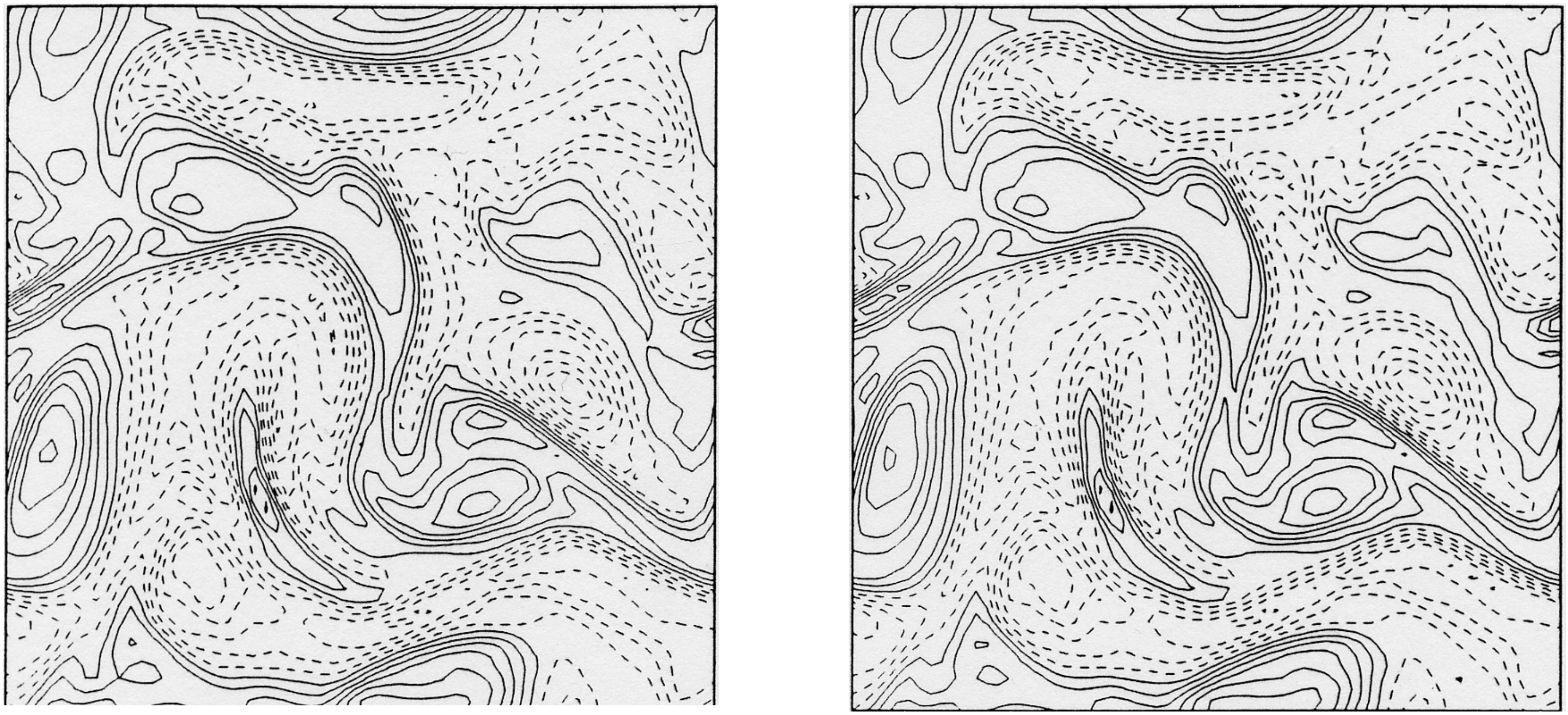
True State T = 0 4D-Var





Perfect Model 4D-Var assimilation (T_a = 20)

True State T = 20 4D-Var



Incremental approach

- **Use a low resolution version of the model**

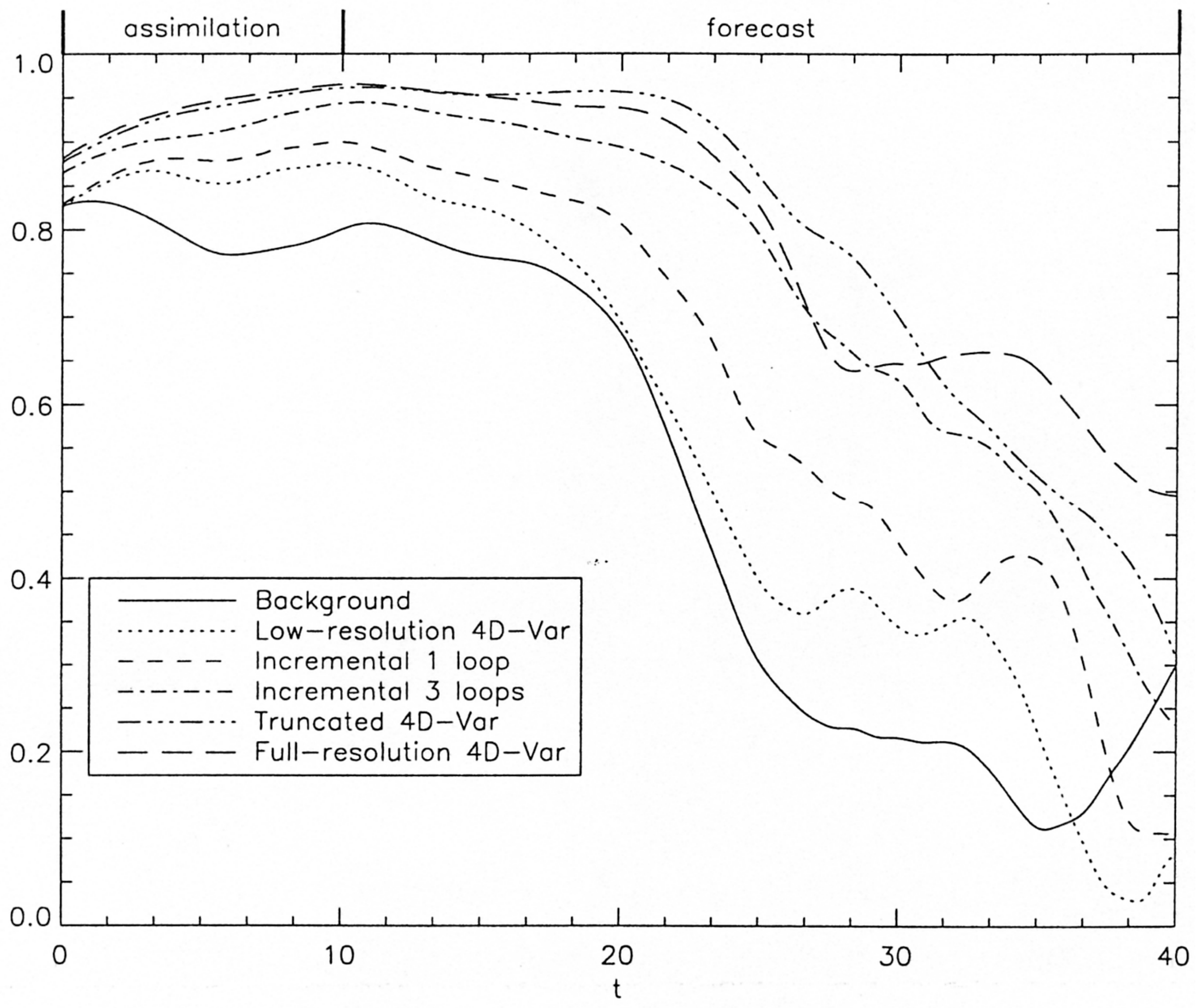
- **Experiments with a coarse observation network every time-step and with random error added**

  – *Truncated 4D-Var:* full 4D-Var except that the gradient is truncated to the resolution of the incremental model

  – *Low-Res 4D-Var:* full 4D-Var at low resolution

  – *Incremental method* with varying number of outer loops

Incremental approach: correlation between the true solution and the assimilation/forecast



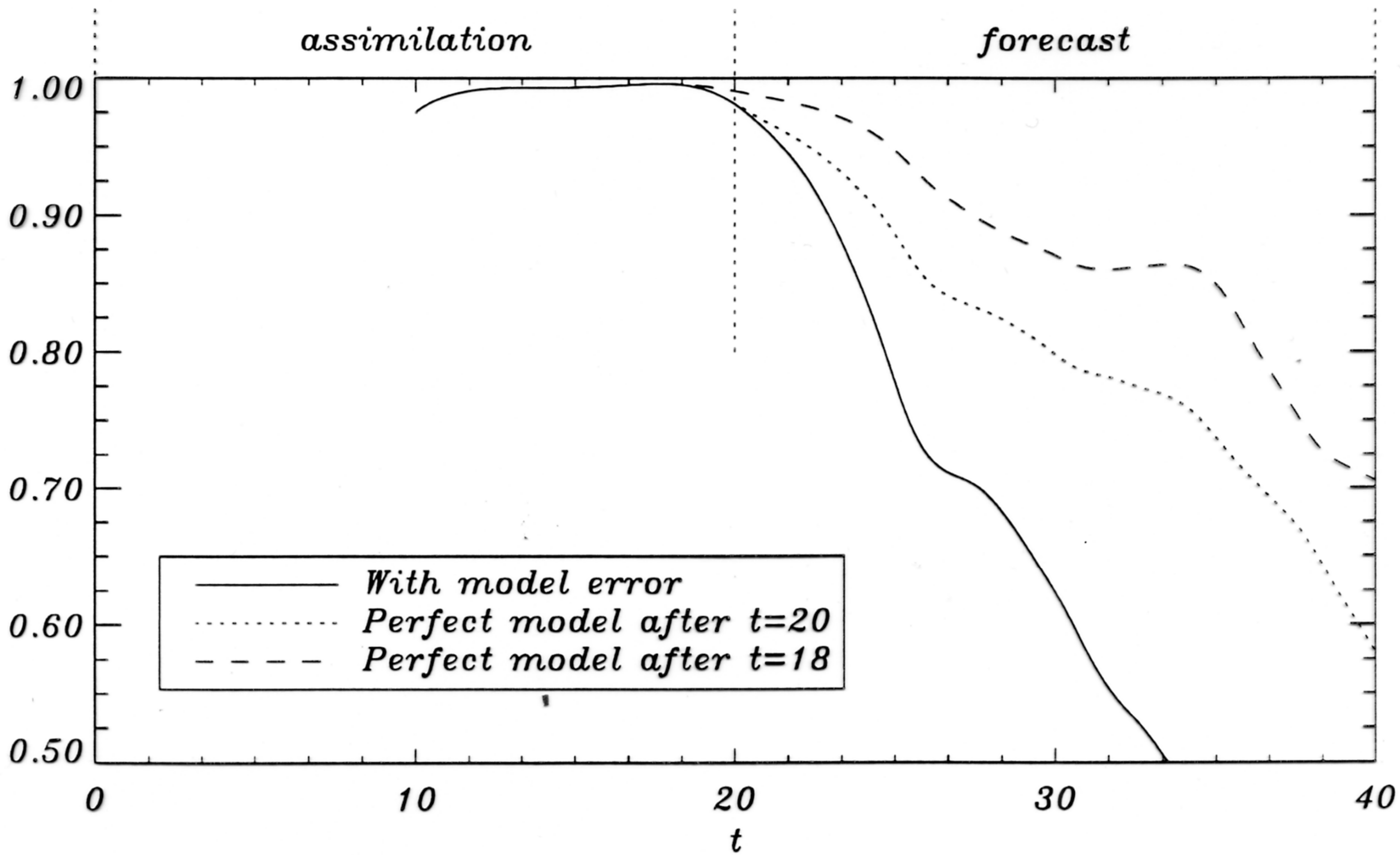
Model error: variation of the β parameter

- **Introduce error in the phase speed of propagation of Rossby waves**
  - Mimicks phase error that often occur in NWP models

- **Observations were generated with $\beta = 0.5$**

- **Assimilation performed with $\beta = 0.4$**

Experiment with phase error:
correlation between the true solution and the assimilation/forecast

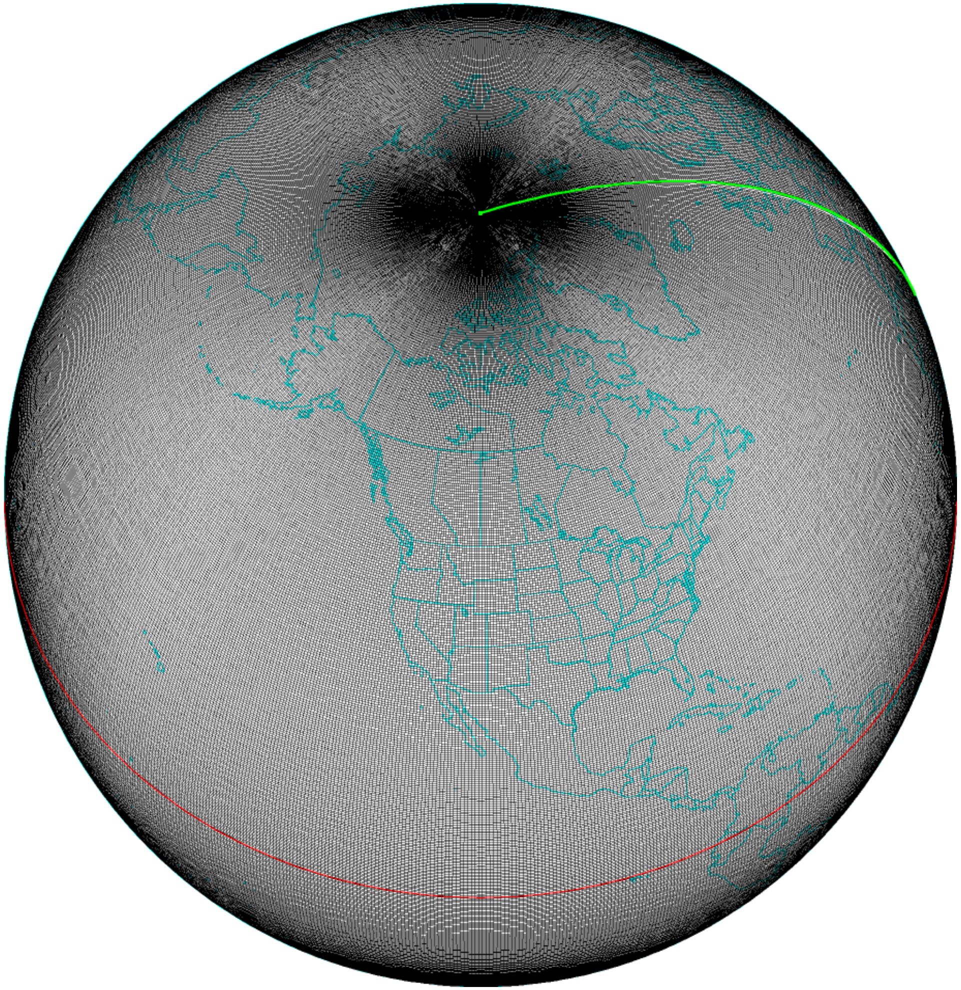


Impact of 4D-Var in the Canadian operational assimilation and forecasting system

Computational Grid for the operational global and regional forecast systems

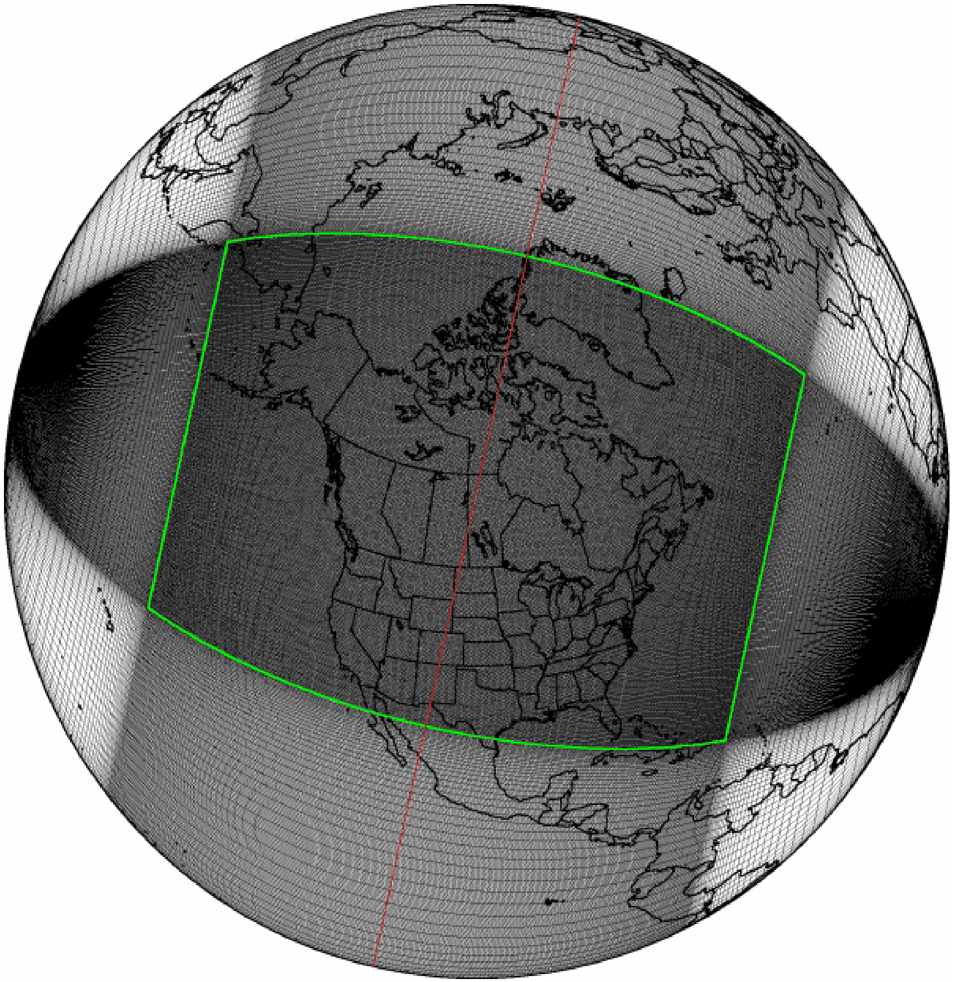
Global

Resolution over the globe: ~35 km

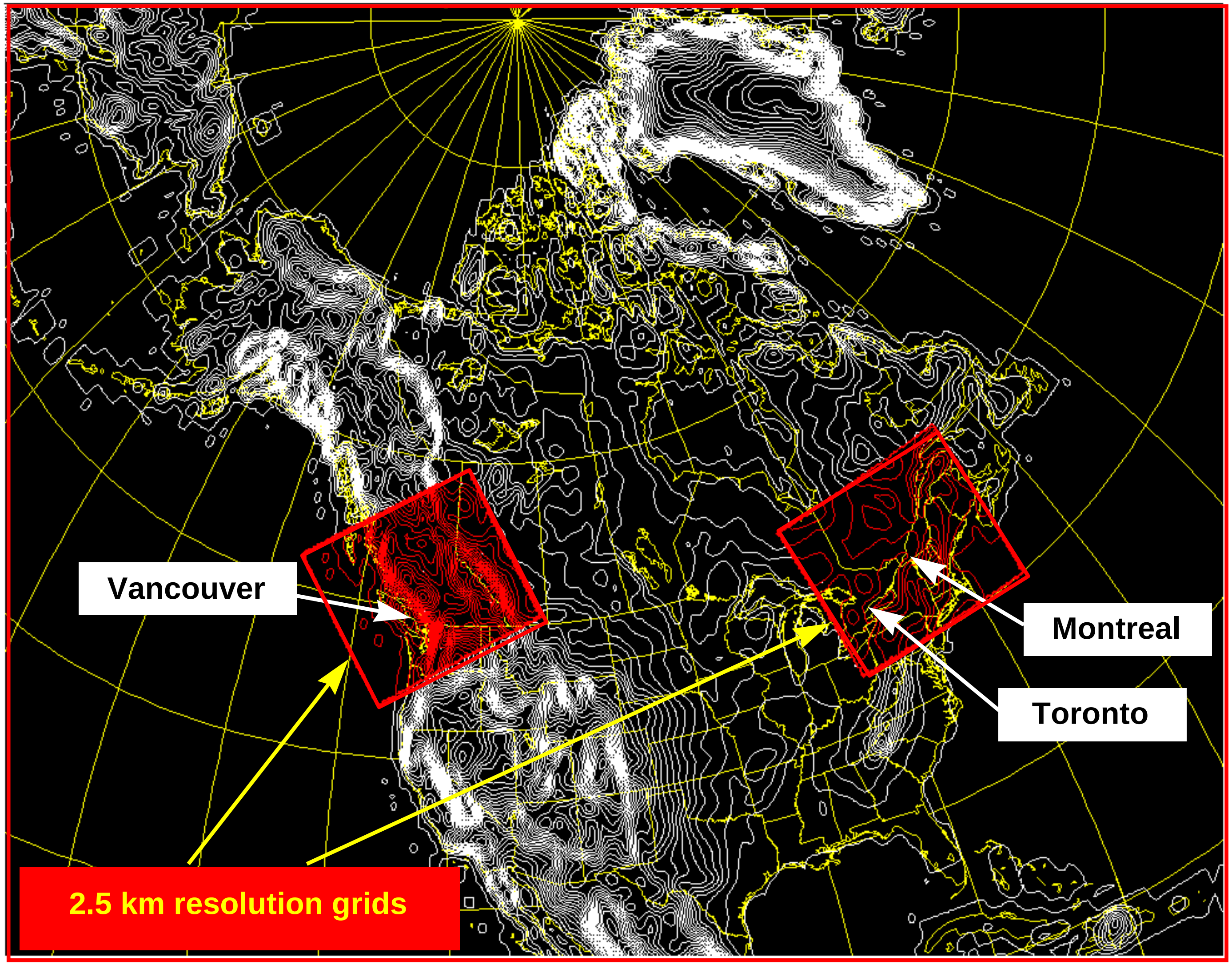


Regional

Resolution over North America: ~15 km



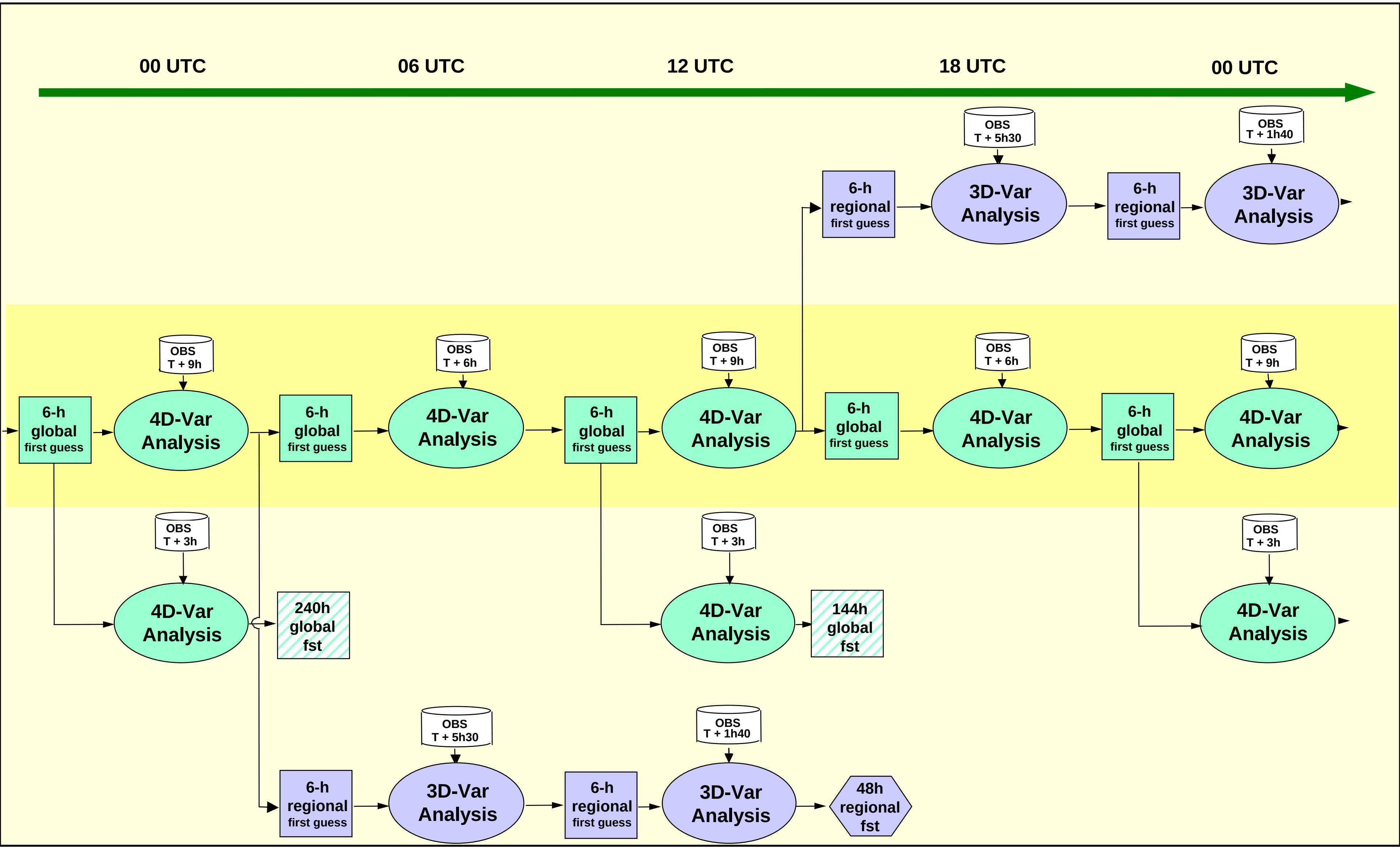
Computational Grid for the experimental mesoscale (local) forecast systems



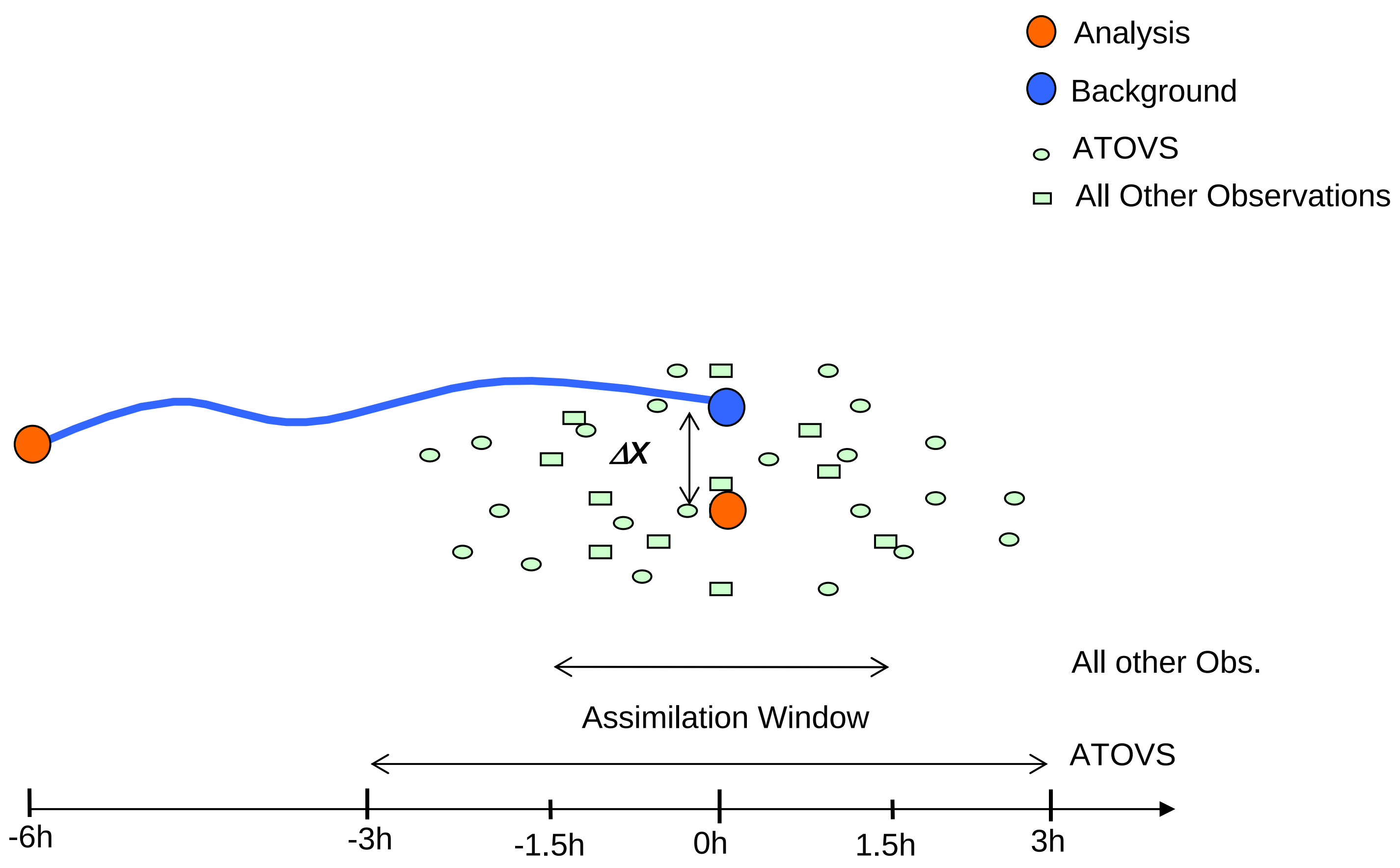
Main Features of Model versions

	Global	Regional	Mesoscale
Resolution	35 km L58	15 km L58	2.5 km L58
Primitive equations	Hydrostatic	Hydrostatic	Non-Hydrostatic
Time integration	Semi-implicit Semi-Lagrangian	Semi-implicit Semi-Lagrangian	Semi-implicit Semi-Lagrangian
Timestep	15.0 min	7.5 min	1.0 min
Land scheme	ISBA	ISBA	ISBA
PBL	TKE	Moist TKE	Moist TKE
Cloud and precipitation	-Kain-Fritsch deep conv. -Sundqvist cloud -Kuo transient shallow cloud	-Kain-Fritsch deep conv. -Sundqvist cloud -Kuo transient shallow conv	-Explicit One moment scheme (Kong-Yau)
Sub-grid orographic effects	-Gravity wave drag -Low-level blocking	-Gravity wave drag -Low-level blocking	none
Data assimilation	4D-Var	3D-Var	none
Initialization	DFI	DFI	none

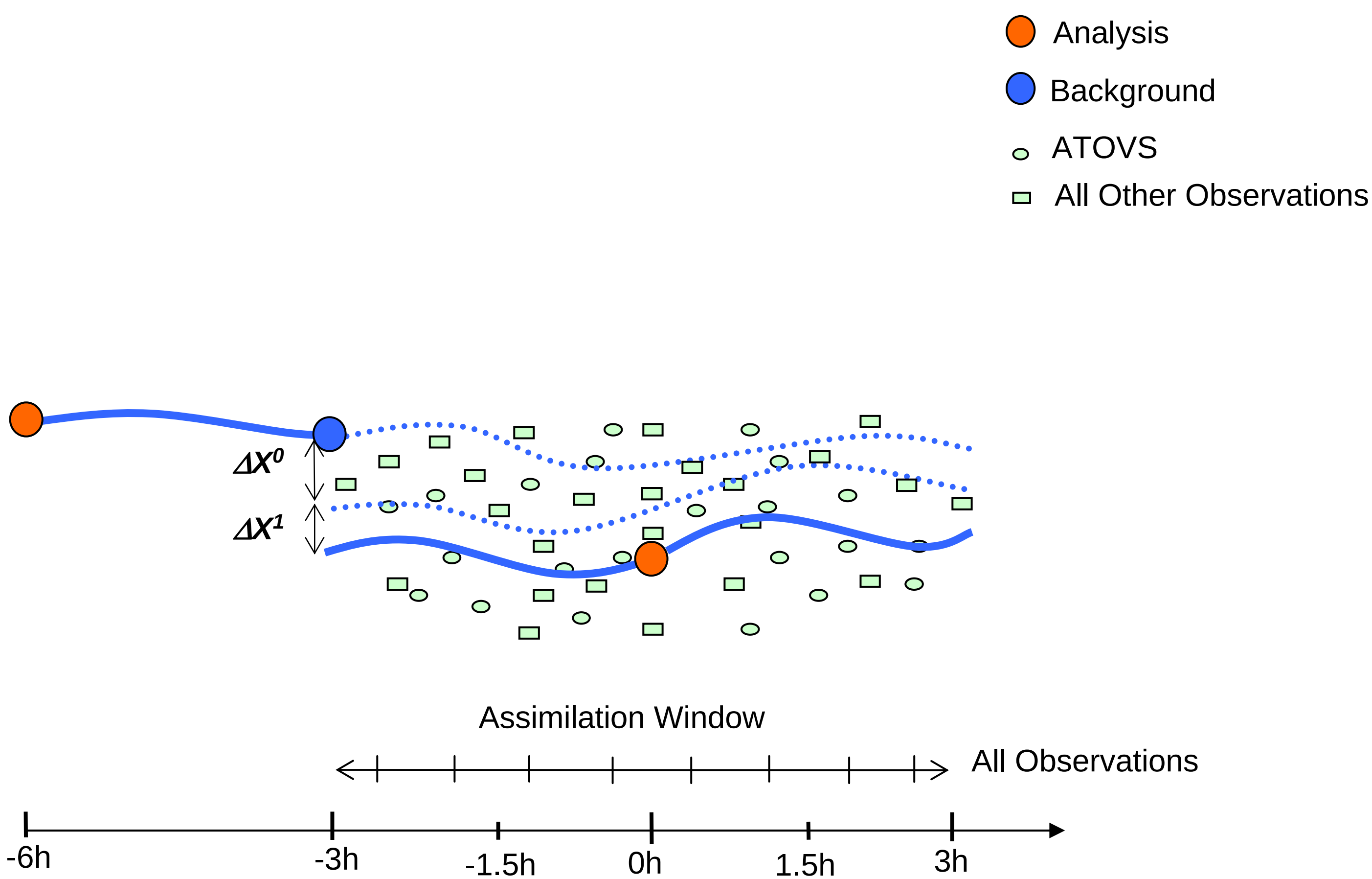
Data assimilation cycles at CMC



3D-Var


Analysis
Background
ATOVS
All Other Observations
ΔX
All other Obs.
Assimilation Window
ATOVS
-6h
-3h
-1.5h
0h
1.5h
3h

4D-Var


Analysis
Background
ATOVS
All Other Observations
ΔX⁰
ΔX¹
Assimilation Window
All Observations
-6h
-3h
-1.5h
0h
1.5h
3h

Configurations

Regional

	Outer loop	Number of inner loops	Simplified physics	Low-resolution Analysis increments	High-resolution trajectory
3D-Var	1	~ 90	-	1.5^{o} (T108) L58	~15 km L58

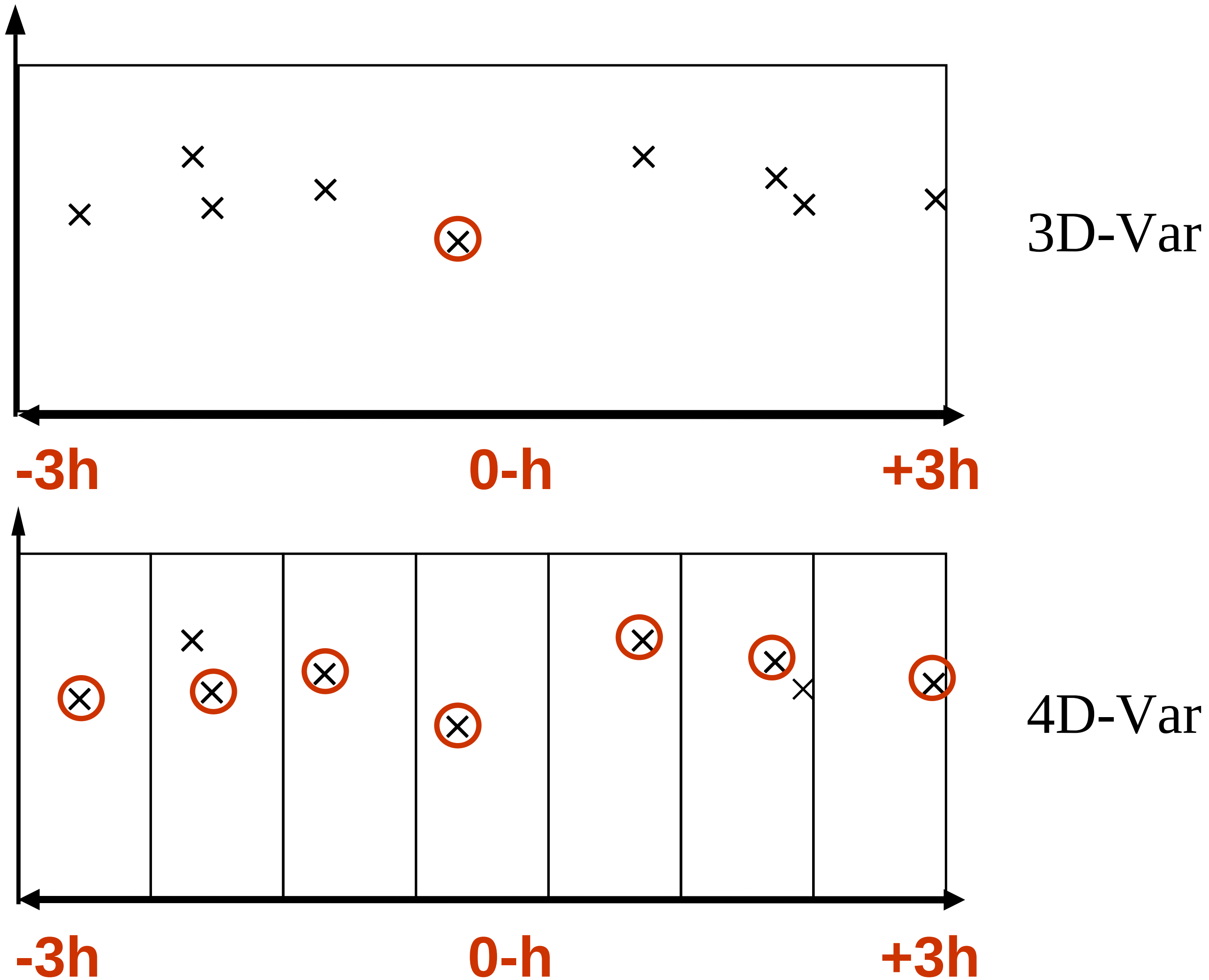
Global

	Outer loop	Number of inner loops	Simplified physics	Low-resolution Analysis increments	High-resolution trajectory
4D-Var	1	30	-PBL	1.5^{o} (T108) L58	(0.3^{o} x 0.45^{o}) L58
	2	25	-PBL -SGO -Stratiform precip.	1.5^{o} (T108) L58	(0.3^{o} x 0.45^{o}) L58

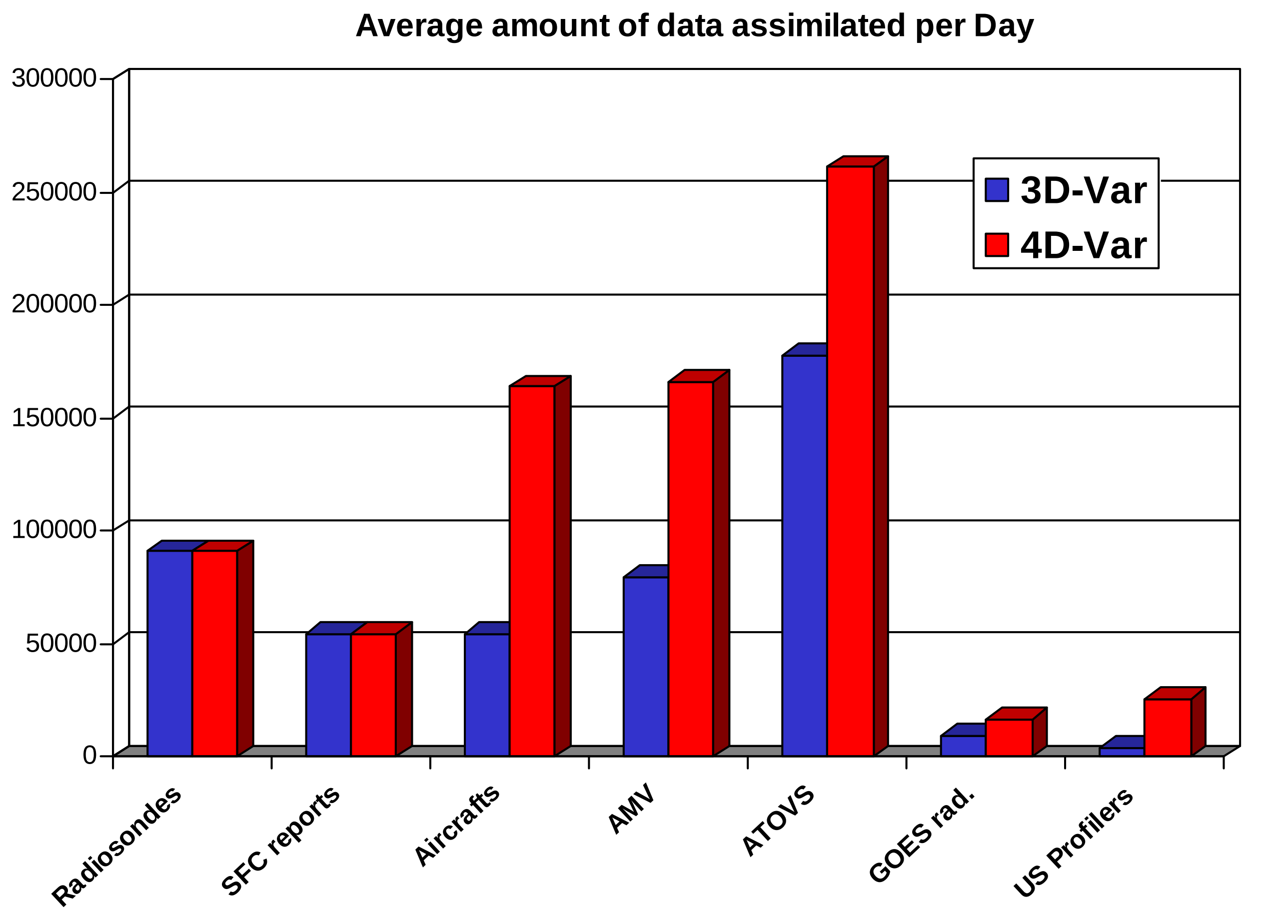
Observations assimilated at the CMC

<table>
<tr><th>Type</th><th colspan="3">Variables</th><th>Thinning</th></tr>
<tr><td>radiosonde/dropsonde</td><td colspan="3">U, V, T, (T-T_d), p_s</td><td>28 levels</td></tr>
<tr><td>Surface report</td><td colspan="3">T, (T-T_d), p_s, (U, V over water)</td><td>1 report/6h</td></tr>
<tr><td>Aircraft
(BUFR, AIREP, AMDAR, ADS)</td><td colspan="3">U, V, T</td><td>1^o x 1^o x 50 hPa</td></tr>
<tr><td rowspan="3">ATOVS
NOAA , AQUA</td><td></td><td>Ocean</td><td>Land</td><td rowspan="3">250 km x 250 km</td></tr>
<tr><td>AMSU-A</td><td>3-10</td><td>6-10</td></tr>
<tr><td>AMSU-B</td><td>2-5</td><td>3-4</td></tr>
<tr><td>Water vapor channel
GOES</td><td colspan="3">IM3
(6.7 μ)</td><td>2^o x 2^o</td></tr>
<tr><td>AMV
(Meteosat, GOES, MTSAT)</td><td colspan="3">U,V
(IR, WV, VI channels)</td><td>1.5^o x 1.5^o</td></tr>
<tr><td>MODIS
(Aqua, Terra)</td><td colspan="3">U,V</td><td>1.5^o x 1.5^o</td></tr>
<tr><td>Profiler
(NOAA Network)</td><td colspan="3">U,V</td><td>(750 m) Vertical</td></tr>
</table>

Temporal thinning



Average amount of data assimilated per Day



Impact of the various components of 4D-Var

Type	Outer loops	Simplified Physics	Temporal thinning
3D-Var	1	-	3D
3D-Var (FGAT)	1	-	3D
4D-Var (1 loop)	1	(simpler)	4D
4D-Var (simpler)	2	(simpler, simpler)	4D
4D-Var (3D-thin)	2	(simpler, better)	3D
4D-Var	2	(simpler, better)	4D

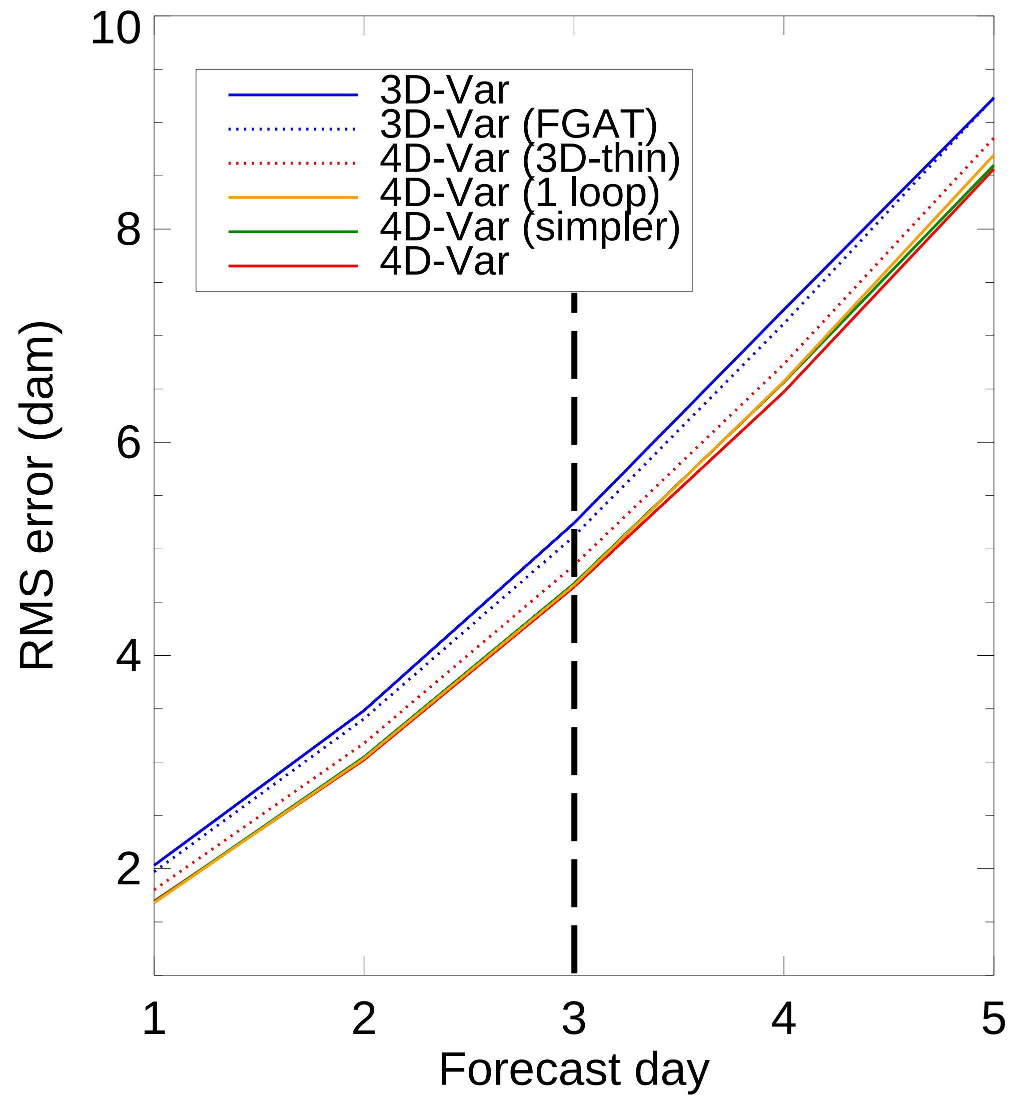
Impact of the various components of 4D-Var

August 2004

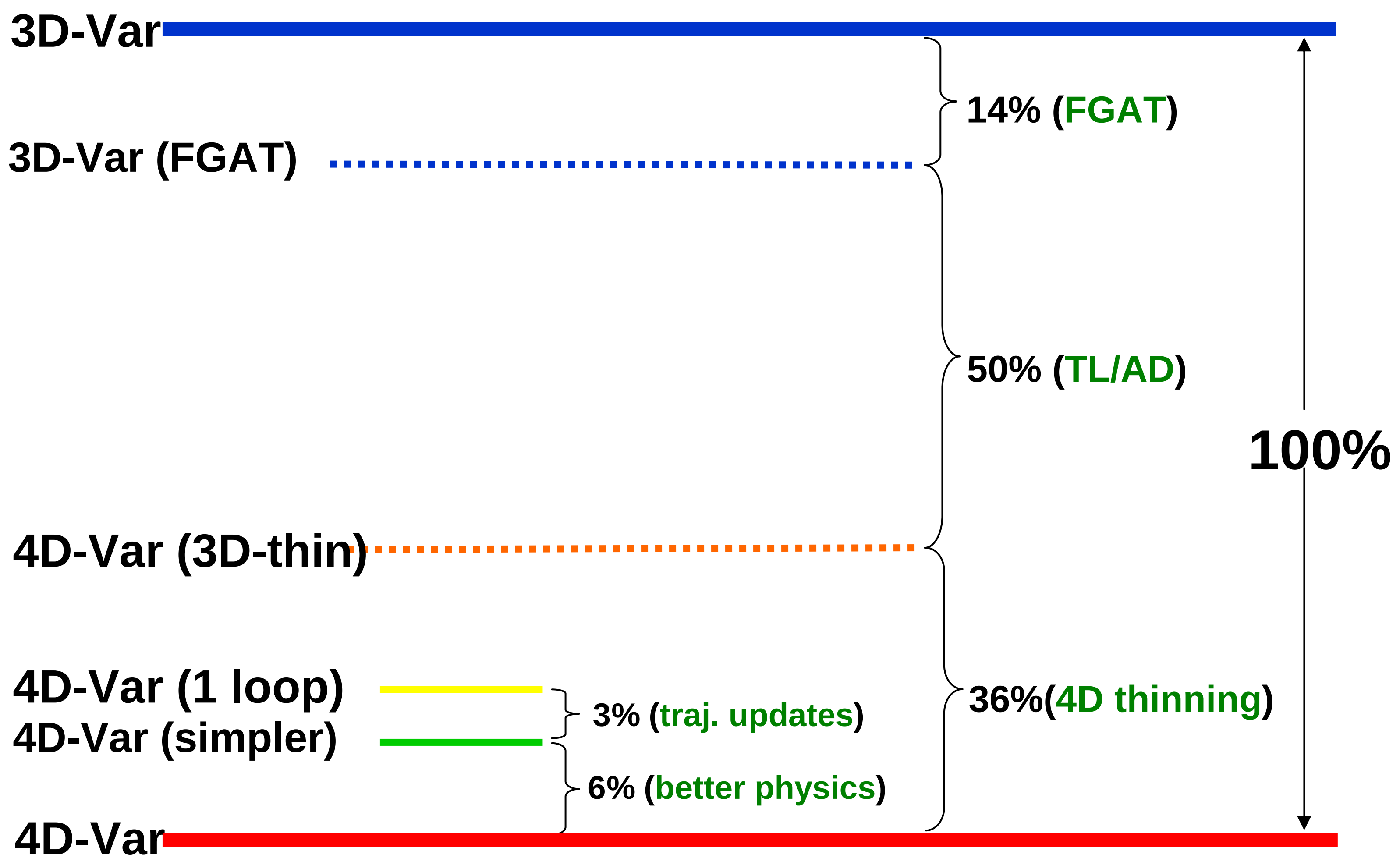
RMS error

GZ 500 hPa

Southern Hemisphere



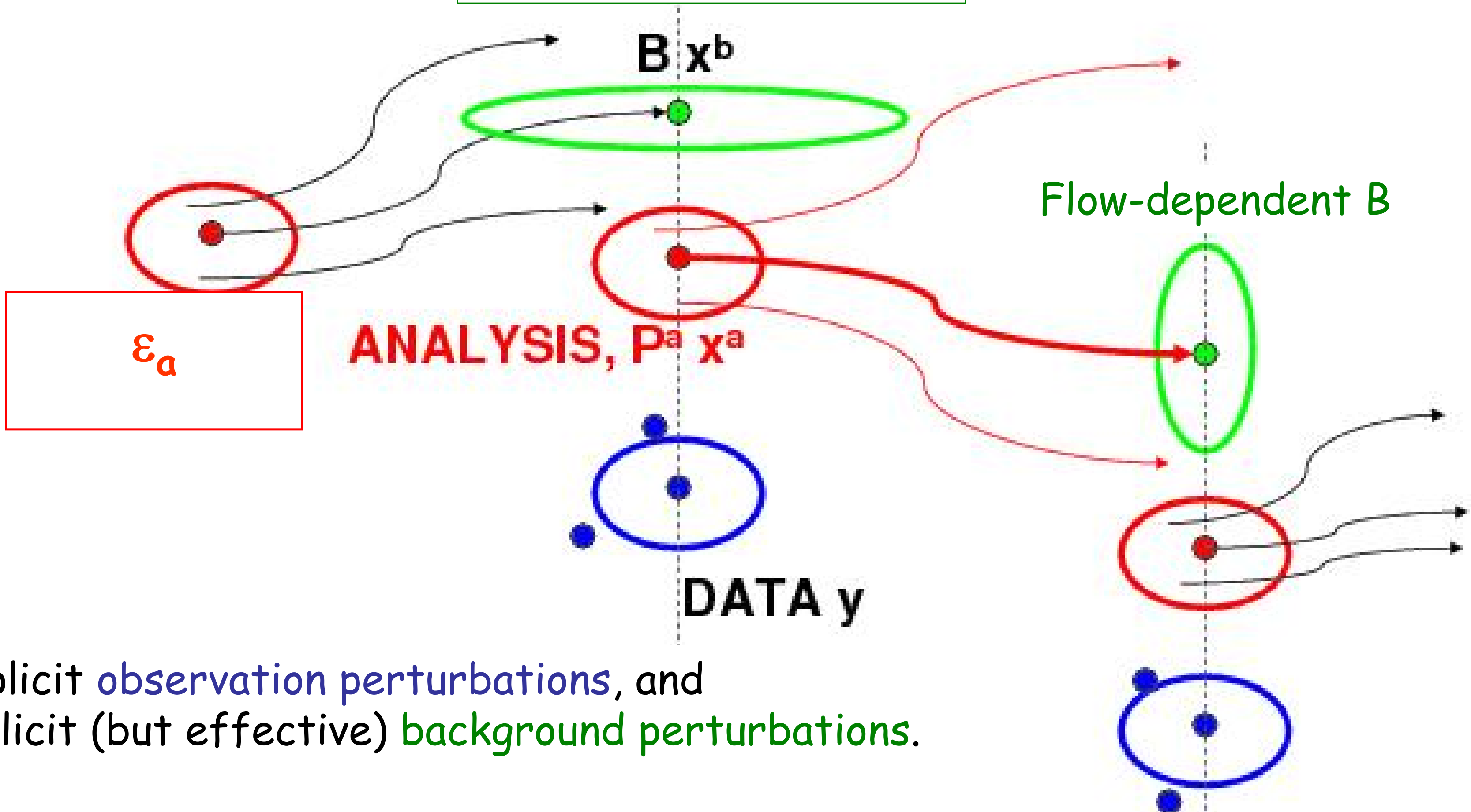
Impact of the various components of 4D-Var



4D-Var – EnKF intercomparison

Ensemble assimilation (EnDA = EnVar, EnKF, ...) : simulation of the error evolution

$$\varepsilon_b = M\ \varepsilon_a\ (+\ \varepsilon_m)$$



Explicit observation perturbations, and implicit (but effective) background perturbations.

(Houtekamer et al 1996; Fisher 2003 ; Ehrendorfer 2006 ; Berre et al 2006)

(from Berre et al., 2009)

Experimental Systems (Buehner et al., 2010)

Modifications to configurations operational during summer 2008

- **4D-Var**
  - incremental approach: ~35km/150km grid spacing, 58 levels, 10hPa top → Increased horizontal resolution of inner loop to 100km to match EnKF
- **EnKF**
  - 96 ensemble members: ~100km grid spacing, 28 levels, 10hPa top → Increased number of levels to 58 to match 4D-Var
- **Same observations assimilated in all experiments:**
  - radiosondes, aircraft observations, AMVs, US wind profilers, QuikSCAT, AMSU-A/B, surface observations
  - eliminated AIRS, SSM/I, GOES radiances from 4D-Var
  - quality control decisions and bias corrections extracted from an independent 4D-Var experiment

Experimental Configurations

- **Variational data assimilation system:**
  - 3D-FGAT and 4D-Var with **B** matrix nearly like operational system: NMC method
  - 3D-FGAT and 4D-Var with flow-dependent **B** matrix from EnKF at middle or beginning of assimilation window (same localization parameters as in EnKF)
  - Ensemble-4D-Var (En-4D-Var): use 4D ensemble covariances to produce 4D analysis increment without TL/AD models (most similar to EnKF approach)
- **EnKF:**
  - Deterministic forecasts initialized with EnKF ensemble mean analysis (requires interpolation from ~100km to ~35km grid)

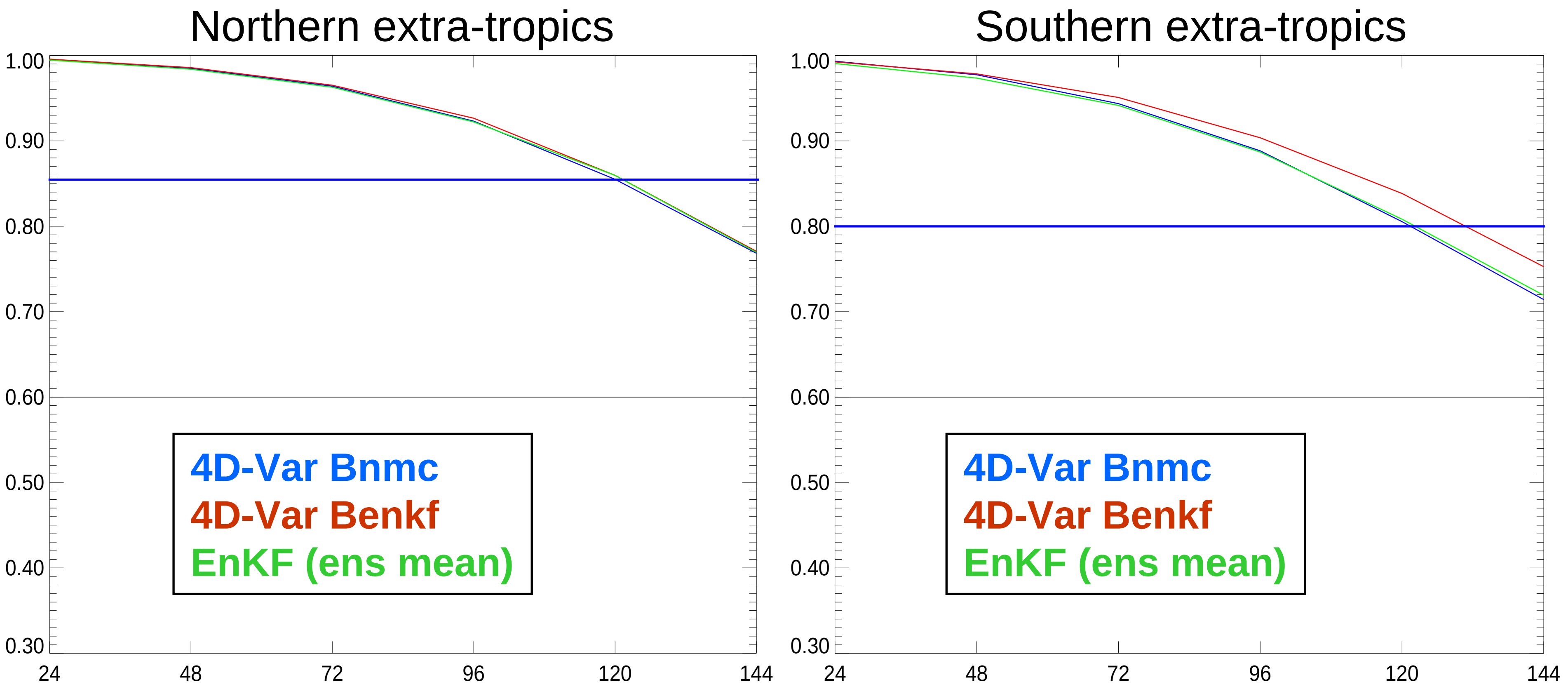
4D-Var – EnKF intercomparison

Main conclusions (presented at 5th WMO data assimilation symposium)

- **Both systems in operational suite, EnKF currently used only for initializing Ensemble Forecasts**
- Goal**: to compare the approaches (and combination approaches) in the context of high-resolution deterministic analyses/forecasts**
- **Deterministic forecasts initialized with 4D-Var with operational B and EnKF ensemble mean analyses have comparable quality: 4D-Var better in extra-tropics at short-range, EnKF better in the medium range and tropics**
- **Largest impact (~9h gain at day 5) in southern extra-tropics Feb 2007 for 4D-Var with flow-dependent EnKF B vs. 4D-Var with operational** B **and also better in tropics – smaller improvement during July 2008**
- **Continuing to test ways to make best use of EnKF ensemble covariances within the variational system**

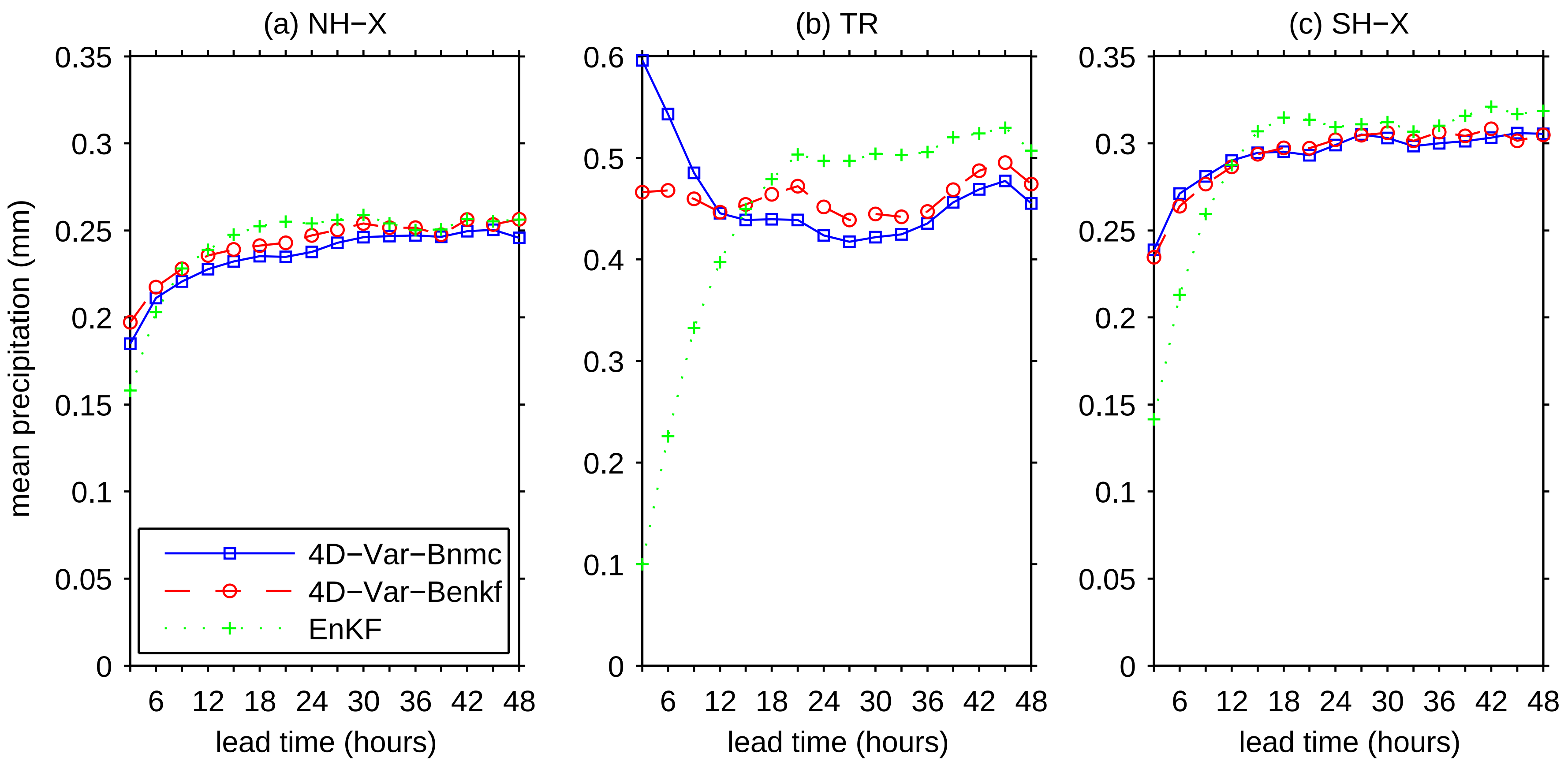
Results – 500hPa GZ anomaly correlation

Large improvement from using flow-dependent covariances in 4D-Var



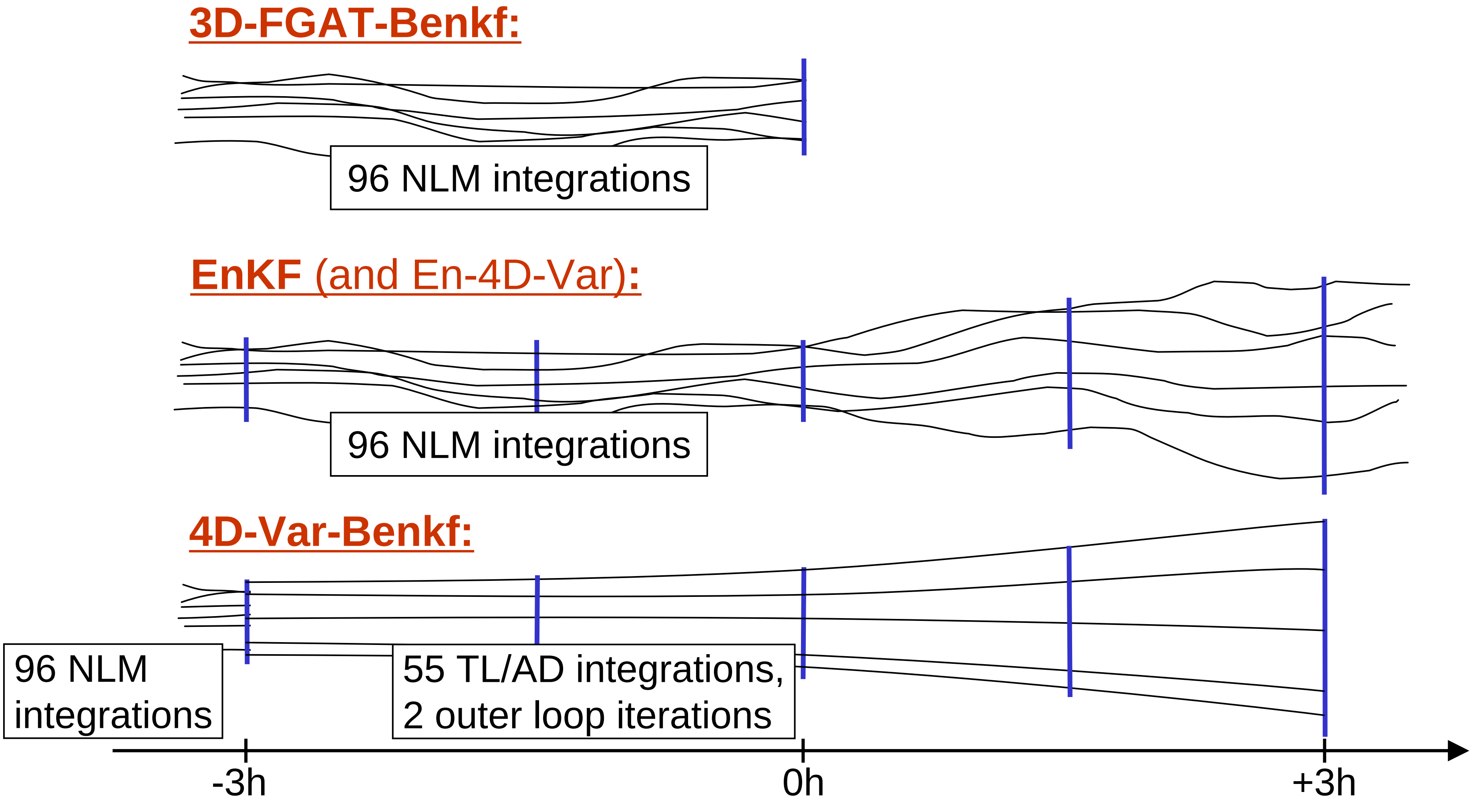
Forecast Results – Precipitation

Evolution of mean 3-hour accumulated precipitation



4D error covariances

Temporal covariance evolution (explicit vs. implicit evolution)

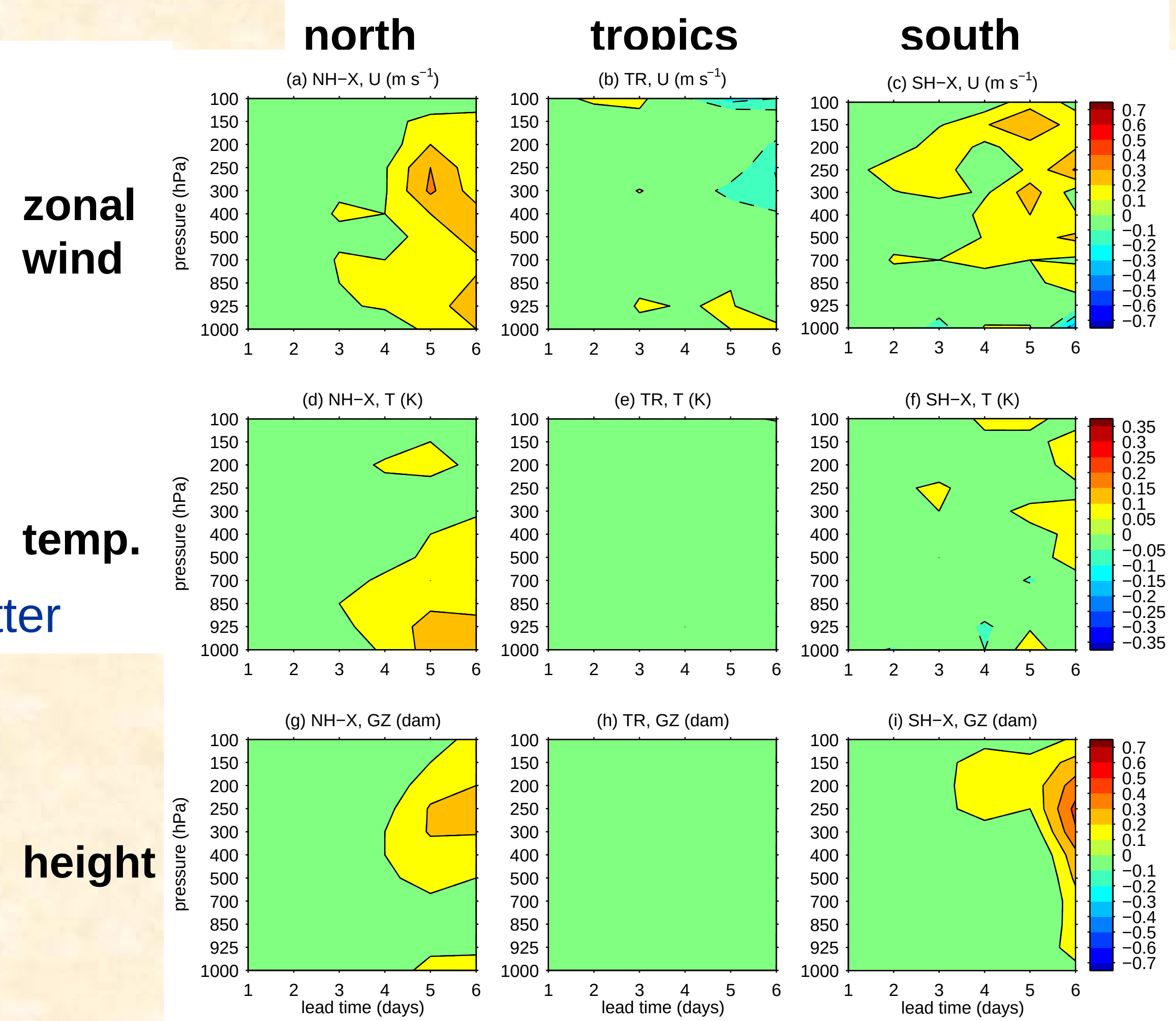


Forecast Results:
En-4D-Var vs. 3D-FGAT-Benkf

Difference in stddev relative to radiosondes:

Positive →
En-4D-Var better

Negative →
3D-FGAT-Benkf better

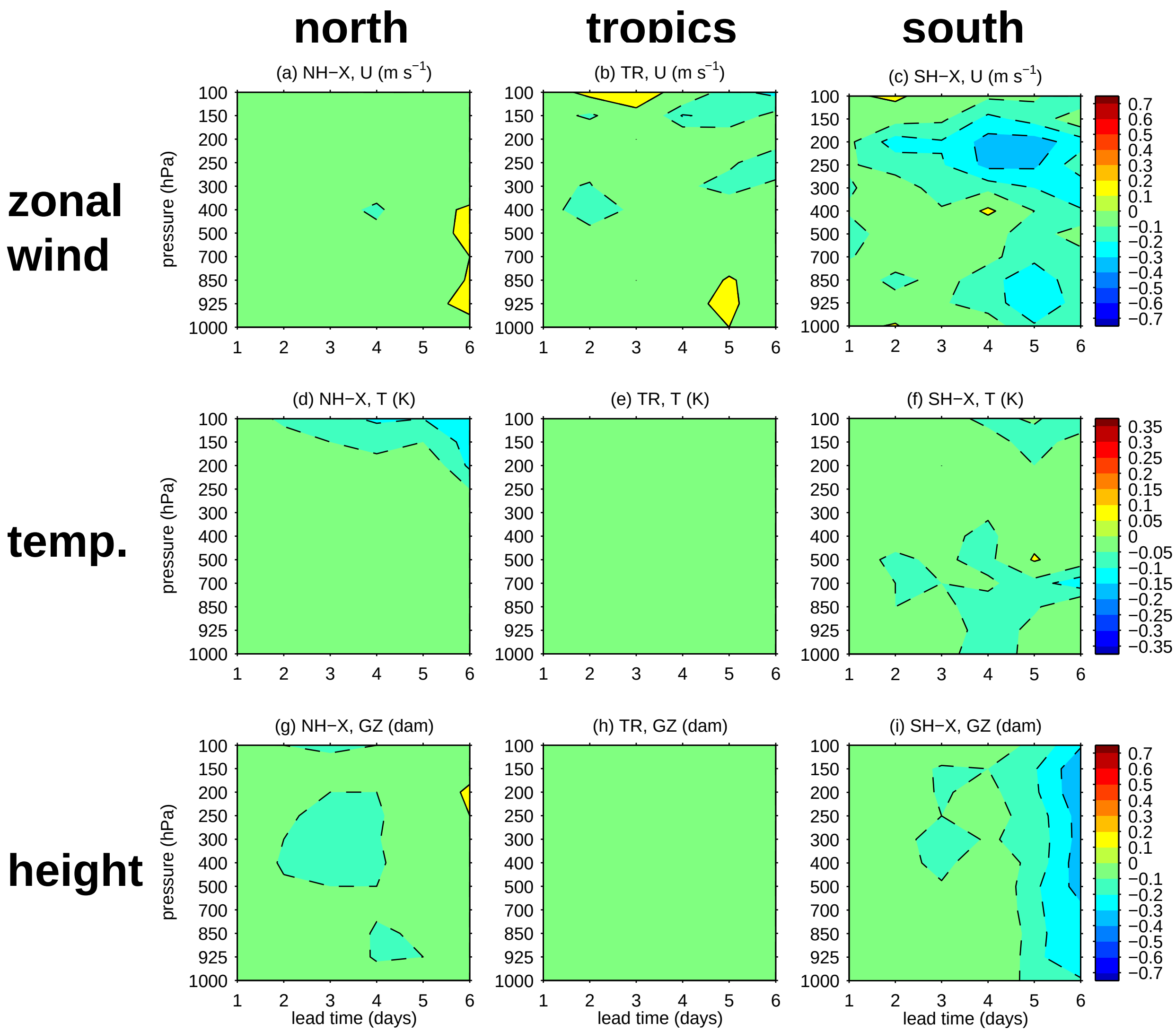


Forecast Results:
En-4D-Var vs. 4D-Var-Benkf

<u>Difference</u> in stddev relative to radiosondes:

Positive →
En-4D-Var better

Negative →
4D-Var-Benkf better



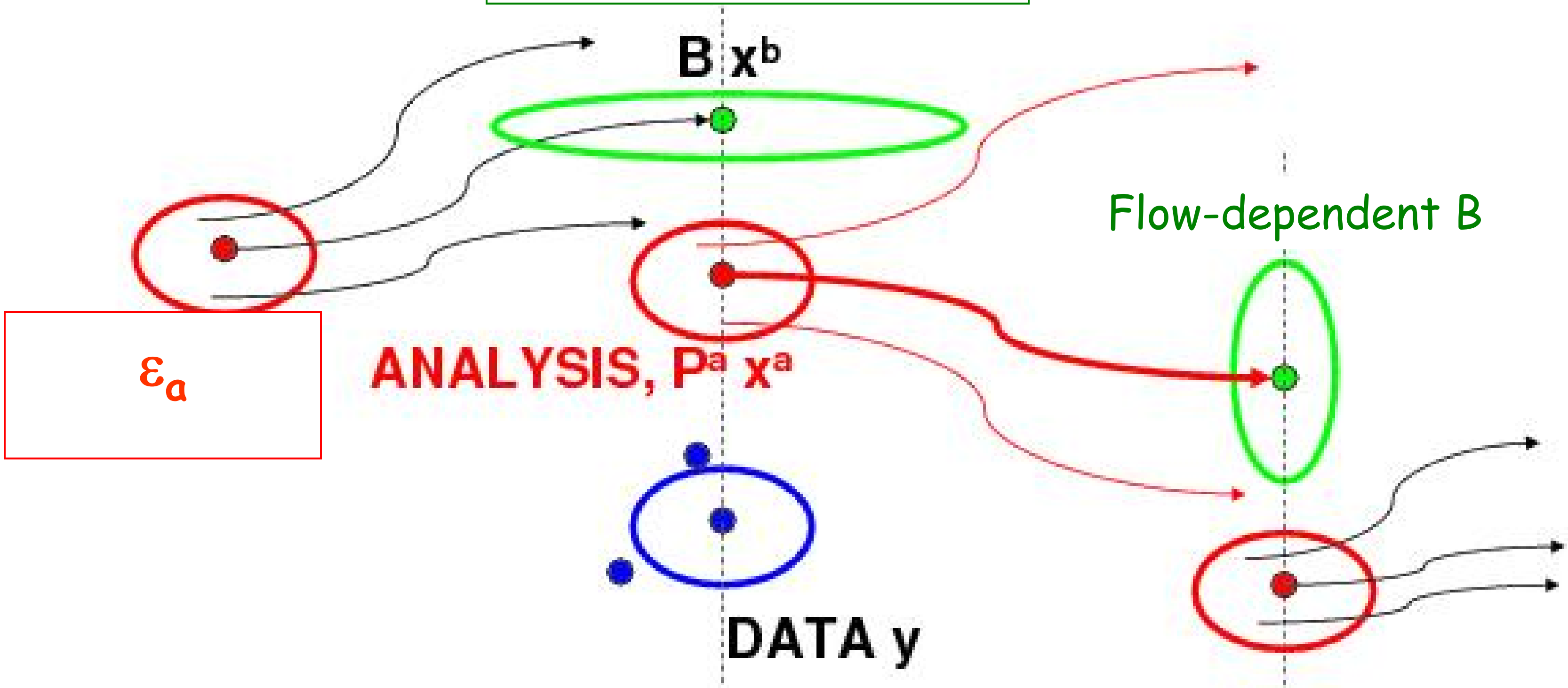
Calculation of ε_a :
«consistent ensemble 4D-Var »
or « hybrid EnKF/Var » ?

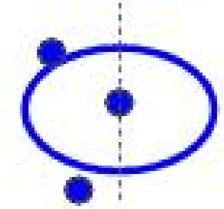
In the Météo-France context with 4D-Var,
a consistent variational approach is used
in both ensemble and deterministic components :

- simple to implement (perturbed Var ~ unperturbed Var).
- a full-rank hybrid B is consistently used in the 2 parts.
- non-linear aspects of 4D-Var can be represented in the analysis perturbation update.

Ensemble assimilation (EnDA = EnVar, EnKF, ...) : simulation of the error evolution

$$\varepsilon_b = M\, \varepsilon_a \; (+\; \varepsilon_m)$$





Explicit observation perturbations, and
implicit (but effective) background perturbations.

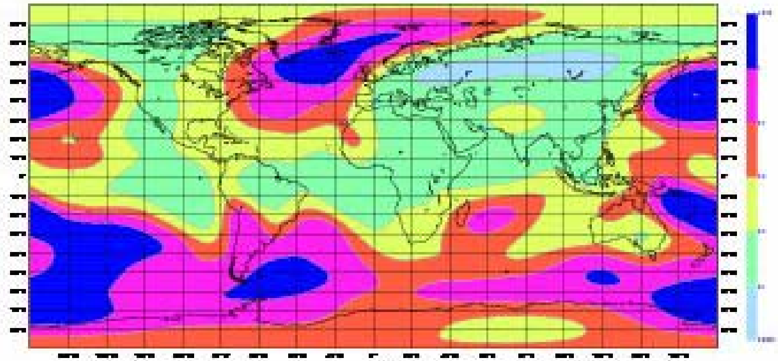
(Houtekamer et al 1996; Fisher 2003 ;
Ehrendorfer 2006 ; Berre et al 2006)

(from Berre et al., 2009)

"OPTIMIZED" SPATIAL FILTERING

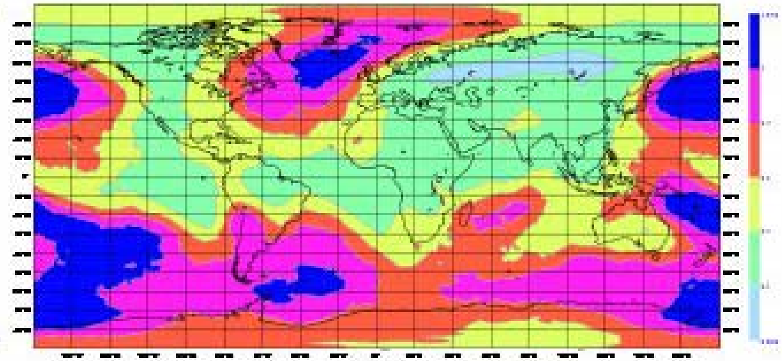
OF THE VARIANCE FIELD

« TRUE » VARIANCES

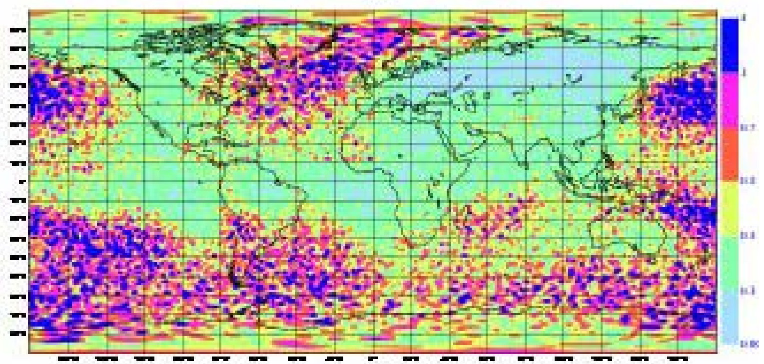


(a)

FILTERED VARIANCES (N = 6)



(b)



(c)

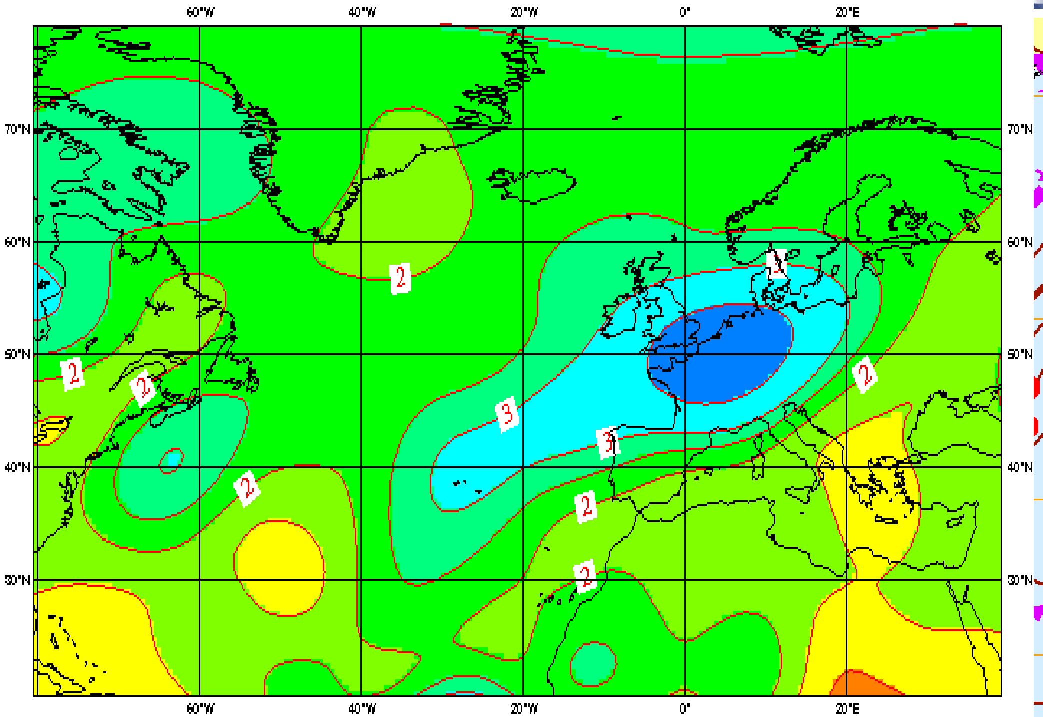
RAW VARIANCES (N = 6)

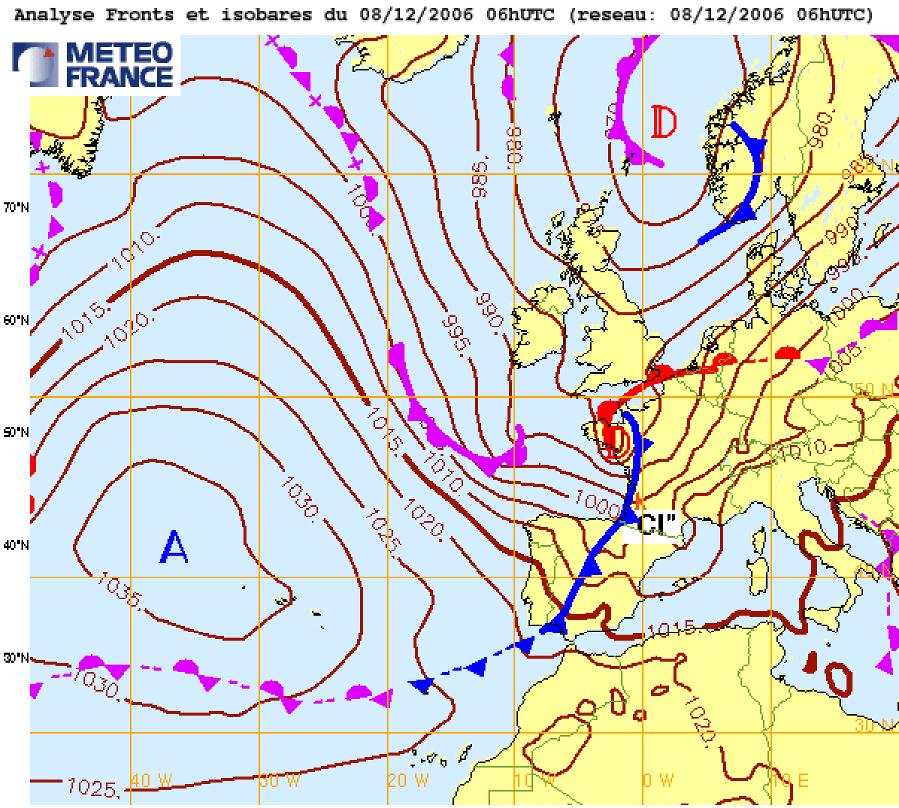
$$V_b^* \sim \rho \, V_b$$

where ρ = signal/(signal+noise)

(Berre et al 2007, Raynaud et al 2008,2009)

Connexion between large σ_b's and intense weather (08/12/2006 , 03-06UTC)





Ensemble spread:

large σ_b's over France

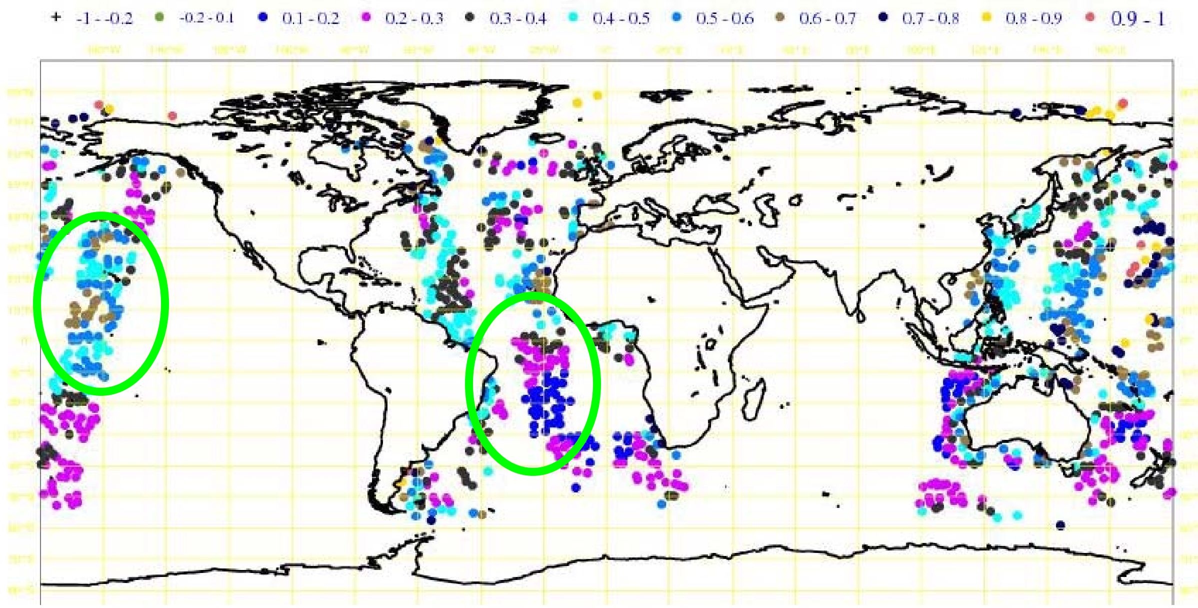
Mean sea level pressure :

storm over France

NB : changes in σ_b's are relatively localized.

Validation of ensemble σ_b's « of the day » in HIRS 7 space (28/08/2006 00h) (Berre et al 2007)

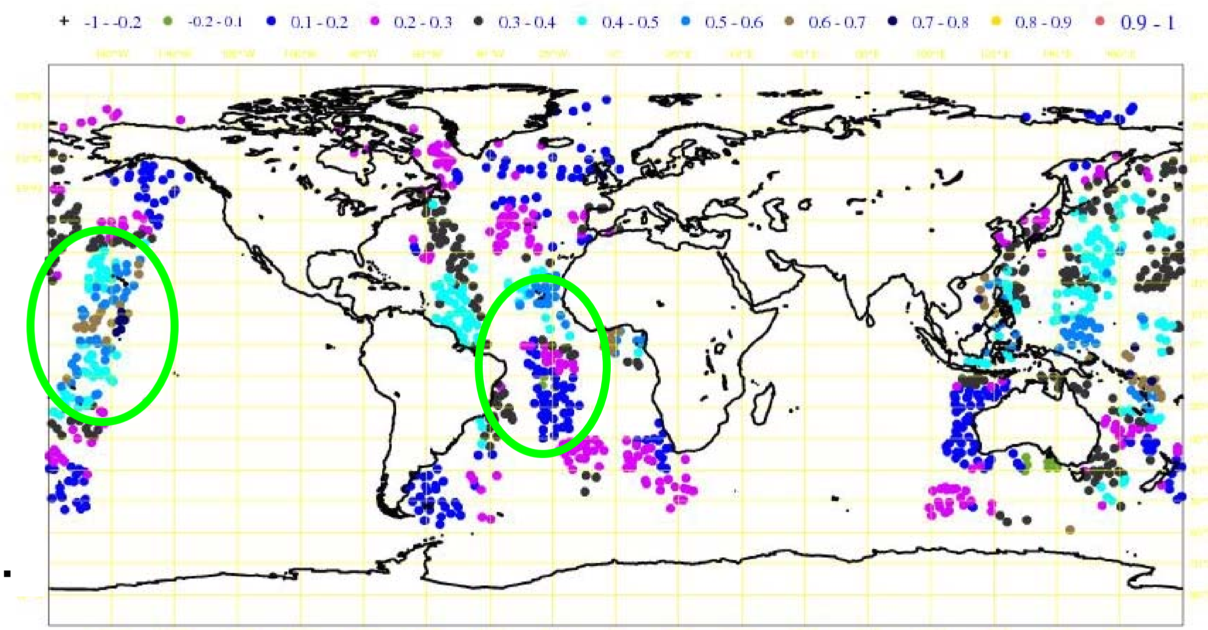
Ensemble sigmab's



« Observed » σ_b's

cov(H dx , dy) ~ H B H^T

(Desroziers et al 2005)



=> model error estimation.

Weak-constraint 4D-Var

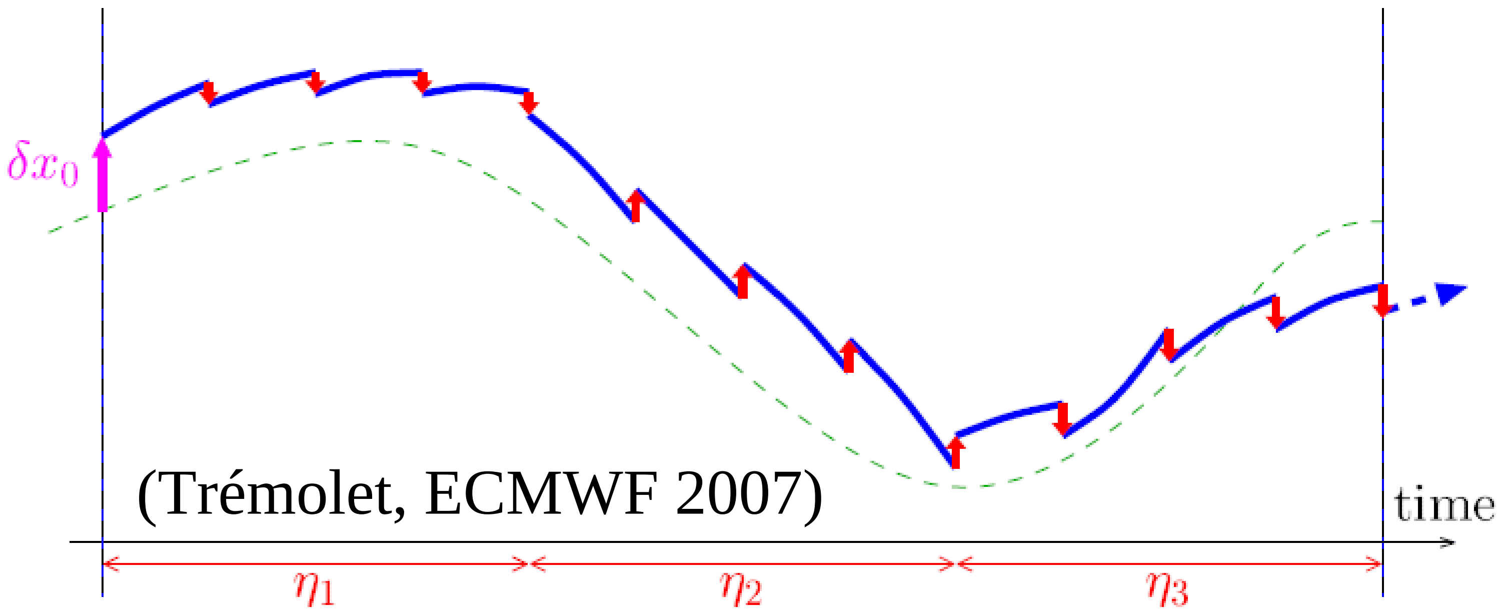
Cost function to minimize

$$J(\mathbf{x}_0,\eta)=\frac{1}{2}\sum_{i=0}^{n}\left(\mathbf{H}(\mathbf{x}_i)-\mathbf{y}_i\right)^T\mathbf{R}^{-1}\left(\mathbf{H}(\mathbf{x}_i)-\mathbf{y}_i\right)$$
$$+\frac{1}{2}(\mathbf{x}_0-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}_0-\mathbf{x}_b)+\eta^T\mathbf{Q}^{-1}\eta$$

Model error is added at regular intervals (e.g., timestep)

$$\mathbf{x}(t_i)\equiv\mathbf{x}_i=M\left(\mathbf{x}_{i-1}\right)+\eta_i$$

Allow to extend the assimilation window but the control variable is now a full model trajectory



(Trémolet, ECMWF 2007)

Conclusions

- **Hybrid methods permit to cycle a 4D-Var assimilation**
  - Measure of the accuracy of the background takes into account the information gained in the previous assimilations and flow-dependent error growth
  - Ensemble methods require some filtering to remove sampling noise
- **Results from two centres indicate that this has a positive impact on the forecasts**
  - Deterministic forecasts initialized with 4D-Var with operational **B** and EnKF ensemble mean analyses have comparable quality: 4D-Var better in extra-tropics at short-range, EnKF better in the medium range and tropics
  - Largest impact (~9h gain at day 5) in southern extra-tropics for 4D-Var with flow-dependent EnKF **B** vs. 4D-Var with operational **B** and also better in tropics – smaller improvement in July 2008
  - Use of *4D* ensemble **B** in variational system (i.e. En-4D-Var):
    - → improves on 3D-FGAT, but inferior to 4D-Var (both with 3D ensemble **B**), least sensitive to covariance evolution in tropics
    - → comparable with EnKF