Meteorologisk institutt met.no

The ETKF rescaling scheme

in HIRLAM

Jelena Bojarova[1], Nils Gustafsson[2], Ole Vignes[1], Åke Johansson[2]

[1]*The Norwegian Meteorological Institute*

[2]*The Swedish Meteorological and Hydrological Institute*

Non-linear models

$x_t = f(x_{t-1}) + \eta_{t-1}$

$y_t = h(x_t) + \varepsilon_t$

$Y_t = \{y_1, y_2, \dots, y_t\}$

Deterministic data assimilation

$x^f_t = E(x_t | Y_{t-1});\ B^f_t = var(x_t | Y_{t-1});$

$x^a_t = E(x_t | Y_t);\ B^a_t = var(x_t | Y_t);$

Dynamic update

$B^f_t = var(f(x_{t-1}) | Y_{t-1}) + Q_{t-1}$

$x^a_t = x^f_t + K_t (y_t - h(x^f_t)) =$

$= x^f_t + cov(x_t, h(x_t))\, var(y_t - h(x^f_t))^{-1} (y_t - h(x^f_t))$

$B^a_t = B^f_t - K_t (F_t)^{-1} (K_t)^T$

Extended Kalman Filter

The first-order approximation

$x_t = x^f_t + \delta x^a_{t-1} \approx f(x^a_{t-1}) + \frac{\partial f(x^a_{t-1})}{\partial x_{t-1}} \delta x^a_{t-1}$

$h(x_t) = h(x^f_t) + \frac{\partial h(x^f_t)}{\partial x_t} \delta x^f_t$

$B^f_t = F B^a_{t-1} F^T + Q;$

$var(h(x_t)) = H B^f_t H^T$

$cov(x_t, h(x_t)) = B^f_t H^T$

Unscented Kalman filter

Ensemble Transform Kalman Filter

The second-order approximation

B^a_{t-1} → select $Z^a_{t-1,i} = X_{t-1,i} - x^a_{t-1}$ → apply $f(X_{t-1,i})$ → B^f_t → select $Z^f_{t,i} = X_{t,i} - x^f_t$ → apply $h(X_{t,i})$ → B^a_t

$Z^a_{t-1,i}$ from eig $(B^a_{t-1})_i$

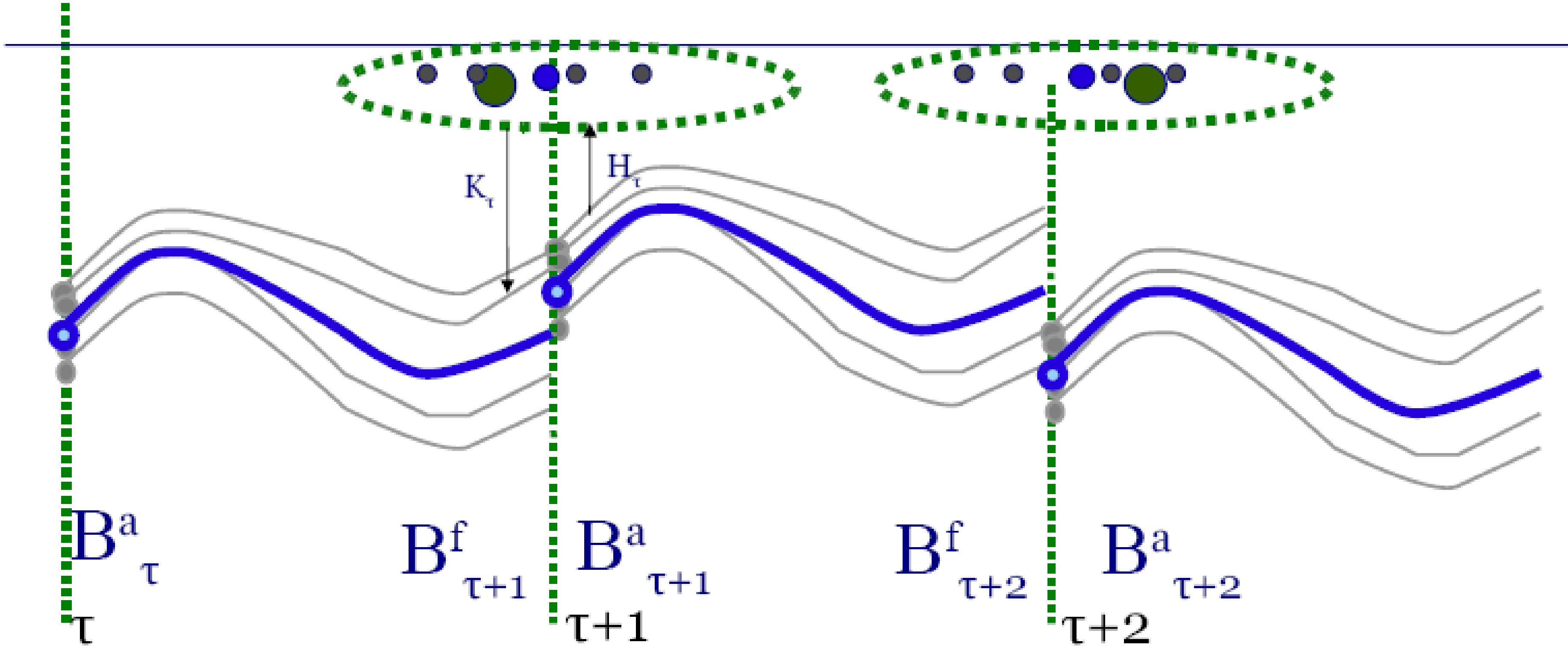
$Z^f_{t,i}$ from eig $(B^f_t)_i$

$B^a_{t-1} = Z^a_{t-1} (Z^a_{t-1})^T$

$B^f_t = Z^f_t (Z^f_t)^T$

$Z^a_t = Z^f_t C_t$

Initial perturbations : ETKF rescaling scheme



$$B^a_\tau = B^f_\tau (I-(K_\tau H_\tau)^T)$$

$$B^f_\tau = Z^f_\tau (Z^f_\tau)^T$$

$$B^a_\tau = Z^a_\tau (Z^a_\tau)^T$$

$$Z^a_\tau = Z^f_\tau T_\tau$$

Properties of the rescaling matrix T_τ

$$T_\tau = C_\tau (G_\tau + I)^{-1/2} C_\tau^T$$

Observation operator

$$C_\tau G_\tau C_\tau^T := B^f_{ens} := (Z_\tau^f)^T H_\tau^T R_\tau^{-1} H_\tau Z^f_\tau$$

Observation error variance

Dimensionality of B^f_{ens} is (N x N)

Trace(B^f_{ens}) ~ $O(p)$

Trace(B^a_{ens}) ≤ $N-1$

$$Z^a_\tau = Z^f_\tau \bar{T}_\tau = Z^f_\tau \Pi C_\tau (G_\tau + I)^{-1/2} A C_\tau^T$$

Inflation factor

Stabilisation of filter

ETKF rescaling was sequentially applied each 6 hours during the period

2007.08.12_00-2007.08.23_18

to simulate analysis perturbations

In the HIRLAM framework :

Coordinate of South pole : lat = -40.00, lon = 22.00

EPS71(south=-20.43,west=-46.48,north=31.37,east =14.53)

Resolution	0.20d, 40 vert. lev.
Ensemble size	N = 8
Total amount of obs	
00 UTC	p = 29721÷34931
06 UTC	p = 19454÷32100
12 UTC	p = 43141÷48821
18 UTC	p = 13049÷20721
Boundaries	EuroTEPS

1.Real observational network :

TEMP, PILOT, AIREP,SYNOP, SHIP, DRIBU

(satellite observations are not used in the construction of T_τ at present)

ETKF rescaling was sequentially applied each 6 hours during the period

2007.08.12_00-2007.08.23_18

to simulate analysis perturbations

In the HIRLAM framework :

Coordinate of South pole : lat = -40.00, lon = 22.00

EPS71(south=-20.43,west=-46.48,north=31.37,east =14.53)

Resolution		0.20d, 40 vert. lev.
Ensemble size		N = 8
Total amount of obs		
	00 UTC	p = 29721÷34931
	06 UTC	p = 19454÷32100
	12 UTC	p = 43141÷48821
	18 UTC	p = 13049÷20721
Boundaries		EuroTEPS

1.Real observational network :

TEMP, PILOT, AIREP,SYNOP, SHIP, DRIBU

(satellite observations are not used in the construction of T_τ at present)

Tuning for application in HIRLAM

1. Rescaling factor Π is adjusted to the estimated variance of innovations at 12 UTC

2. Time filter is applied on Π_τ to prohibit undesirable oscillations

$$\Pi_\tau^{tf} = 0.2\Pi_\tau + 0.6\Pi_{\tau-1}^{tf} + 0.2\Pi_{\tau-2}^{tf}$$

3. The HIRLAM ETKF perturbations are merged with the TEPS perturbations.

$$Z^{HIRLAM}_i = \alpha^{TEPS} Z^{TEPS}_i + (1-\alpha^{TEPS}) Z^{ETKF}_i$$

4. The HIRLAM ETKF perturbations are relaxed toward the TEPS perturbations on the boundaries and in the high stratosphere

Conclusions

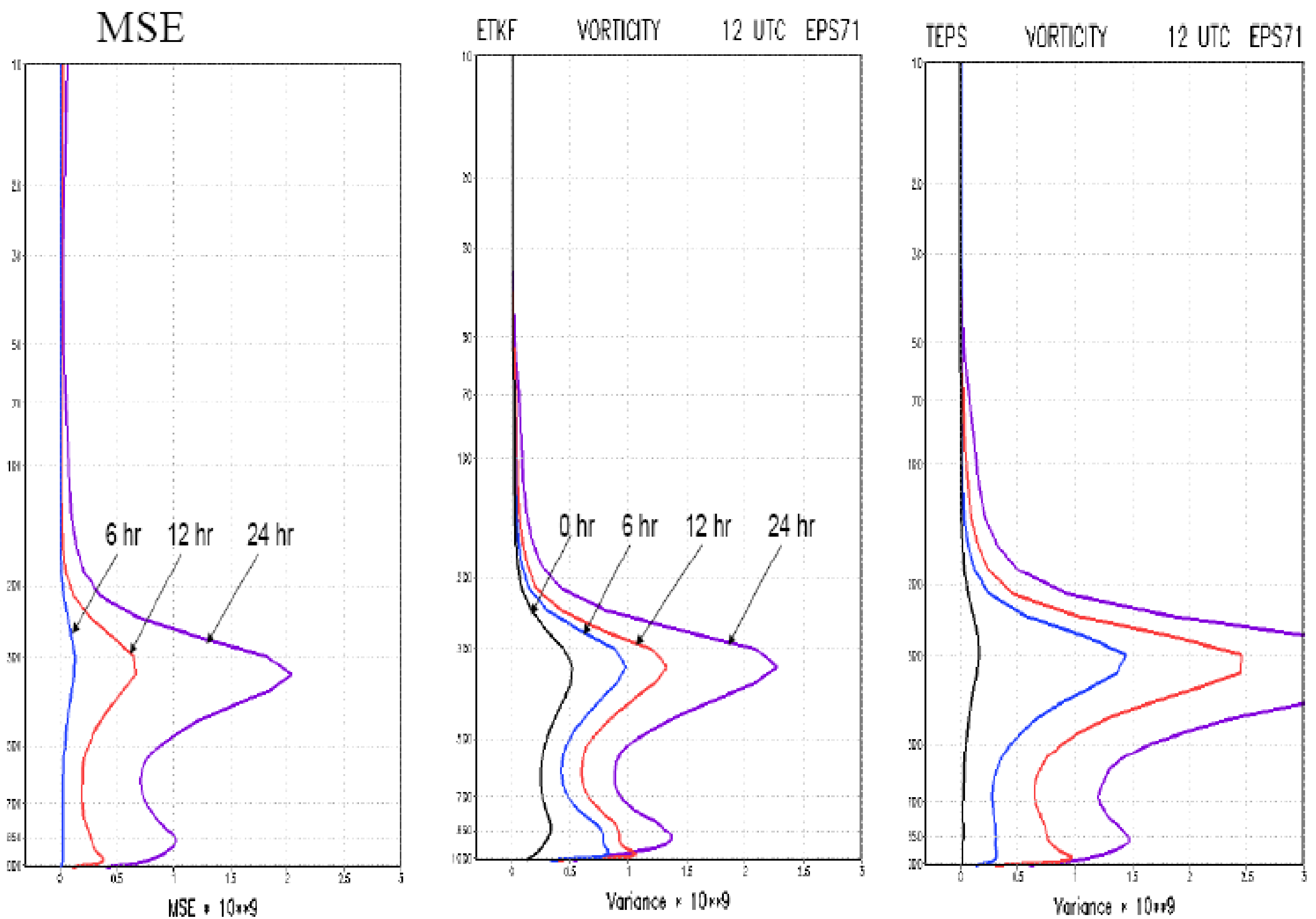
1. The spread of the ETKF perturbations reflects
   the amount of the assimilatedobservations;
   the spatial density of the observational network;

2. The ETKF rescaling scheme, being a linear one, preserves
   linear balances between the model state components and
   spectrum and spectral scales.

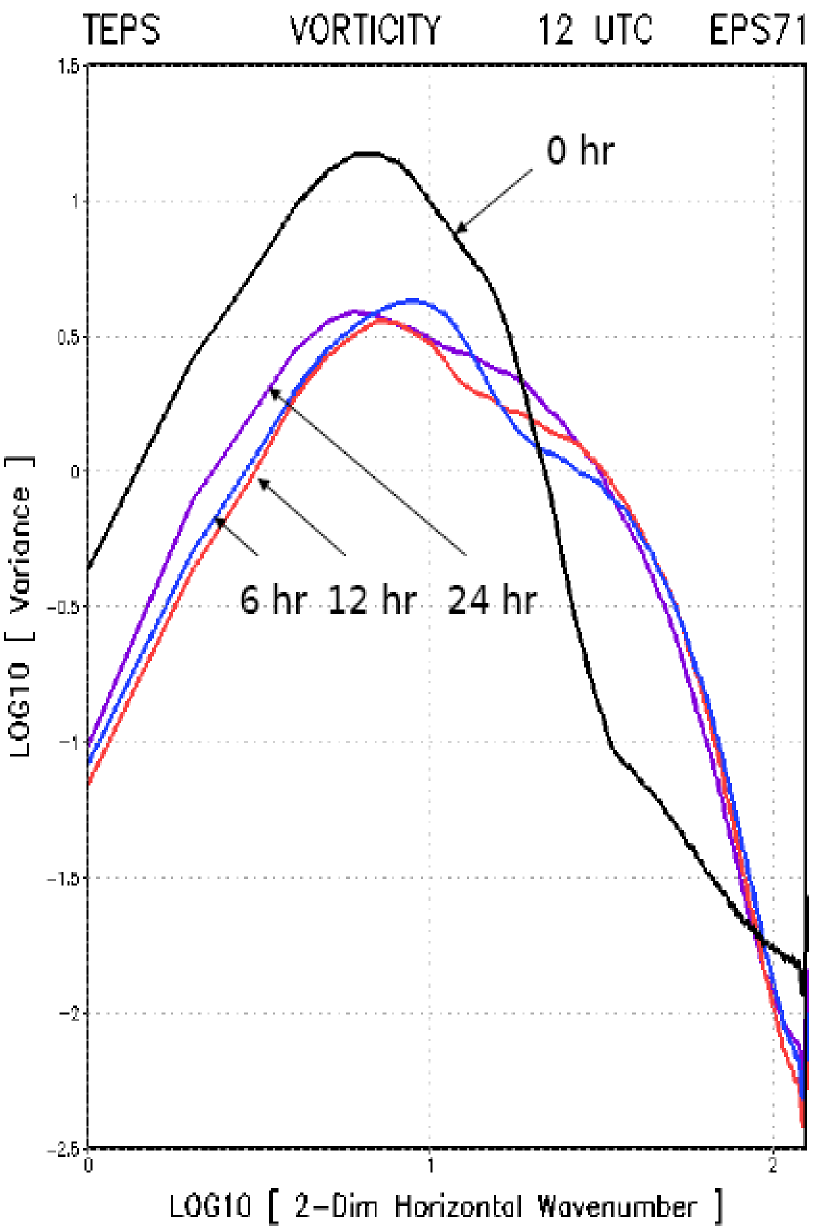
3. The spread of the ETKF perturbations reflects the growth of the uncertainty about the estimate ot the model state due to flow-dependent instabilities.

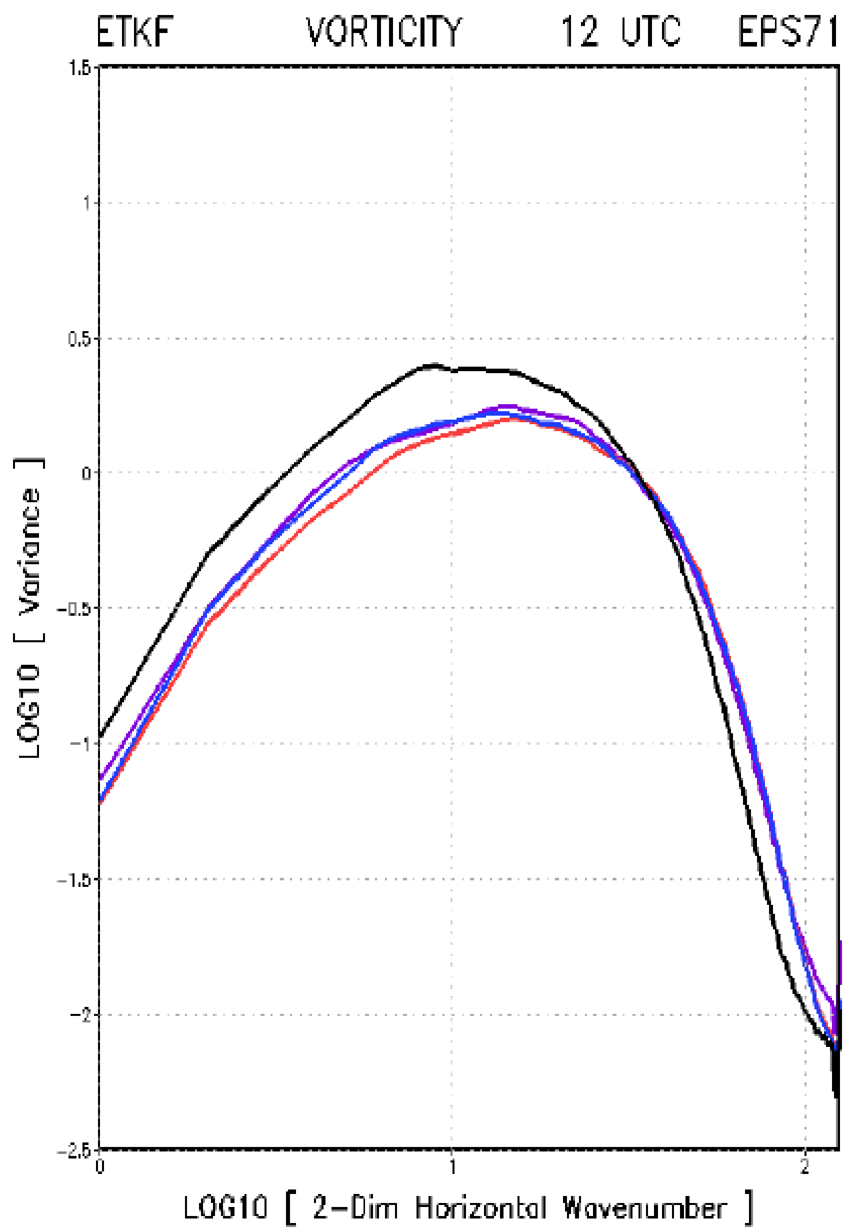
4. The dynamical structures of the ETKF perturbations are closer to those of the MSE of the model state for the short range forecasts than the dynamical structures of the TEPS (singular vectors).

Spatially and temporaly averaged MSE and the Ensemble estimate variance



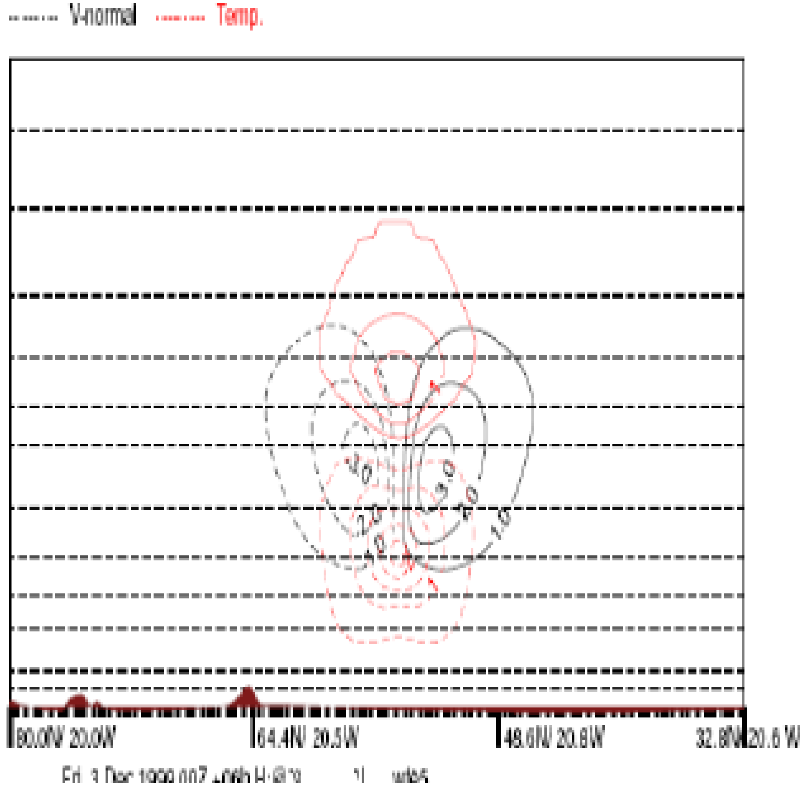
Temporally and spatially averaged spectral density



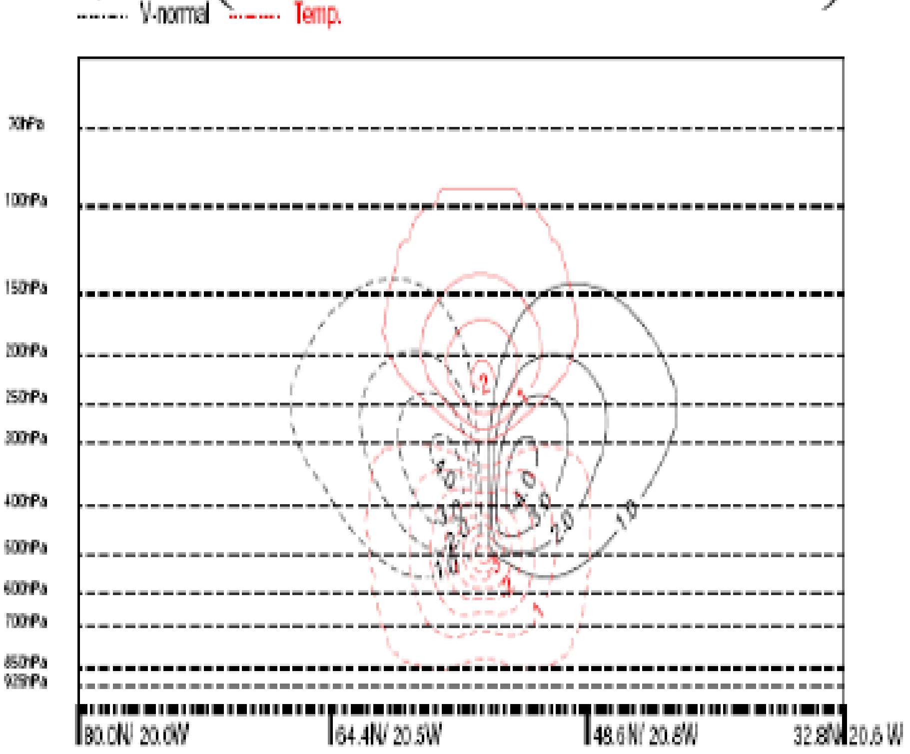


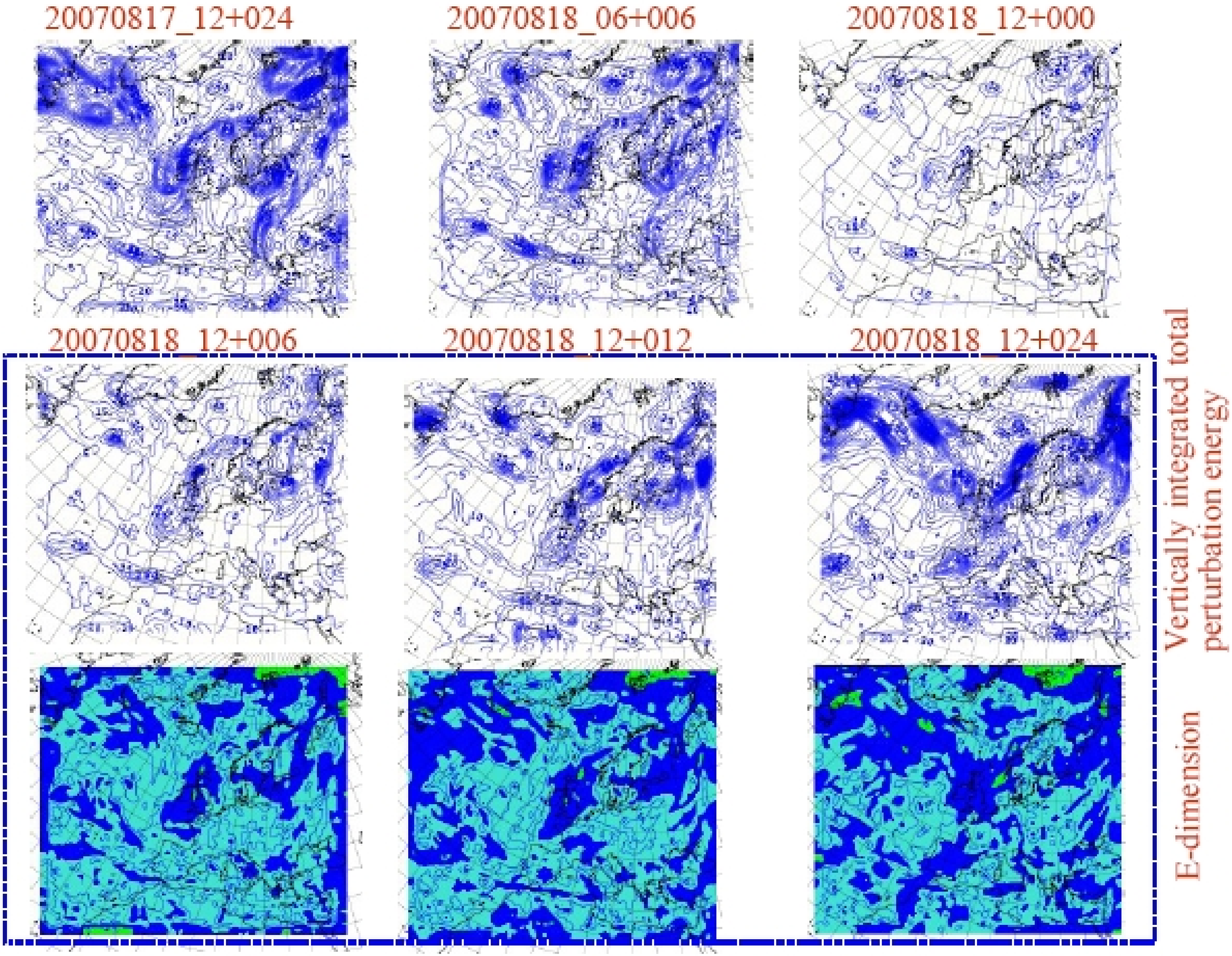
Single Observation Experiment

ETKF

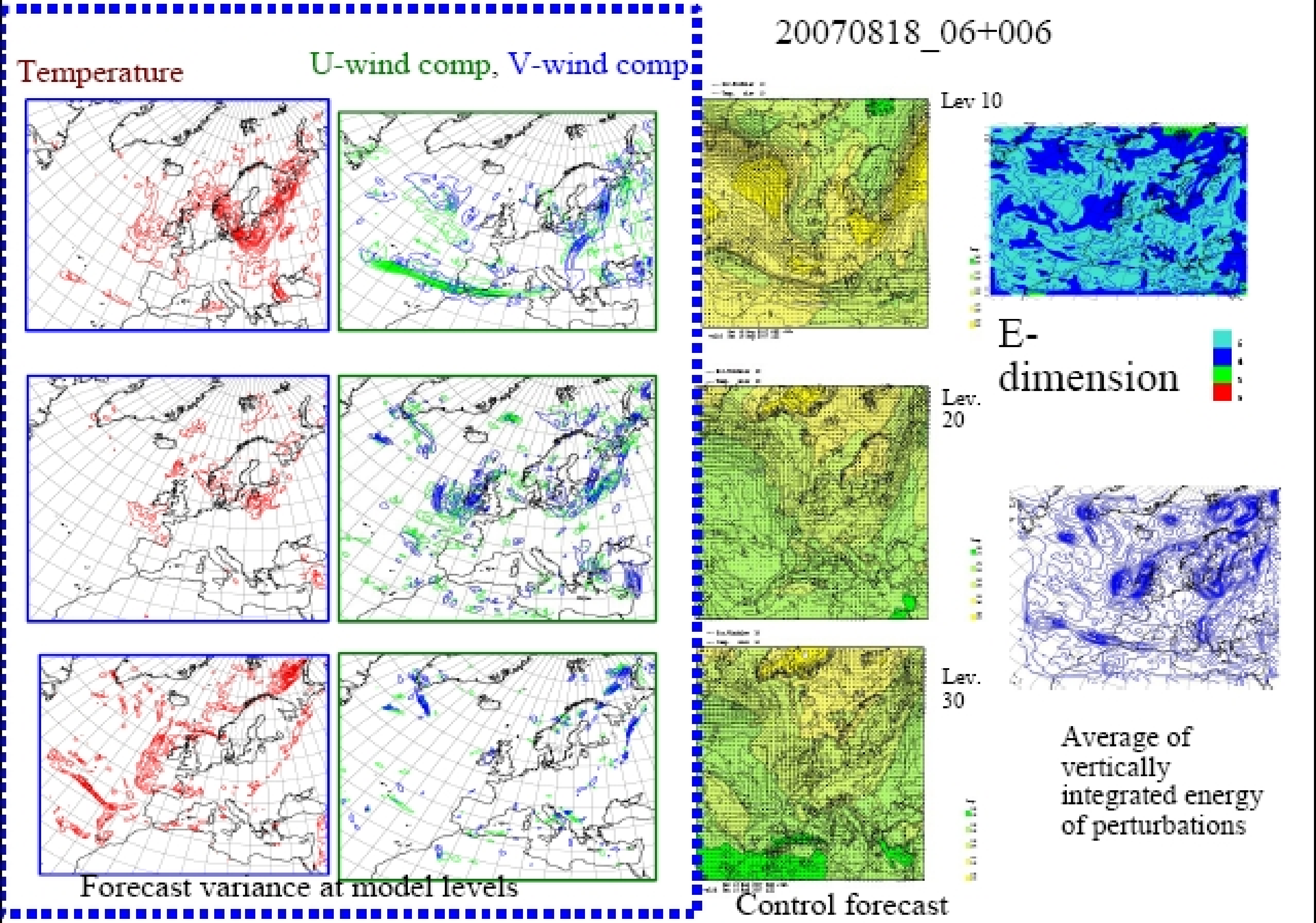


NMC(36-12hh forecast diff)

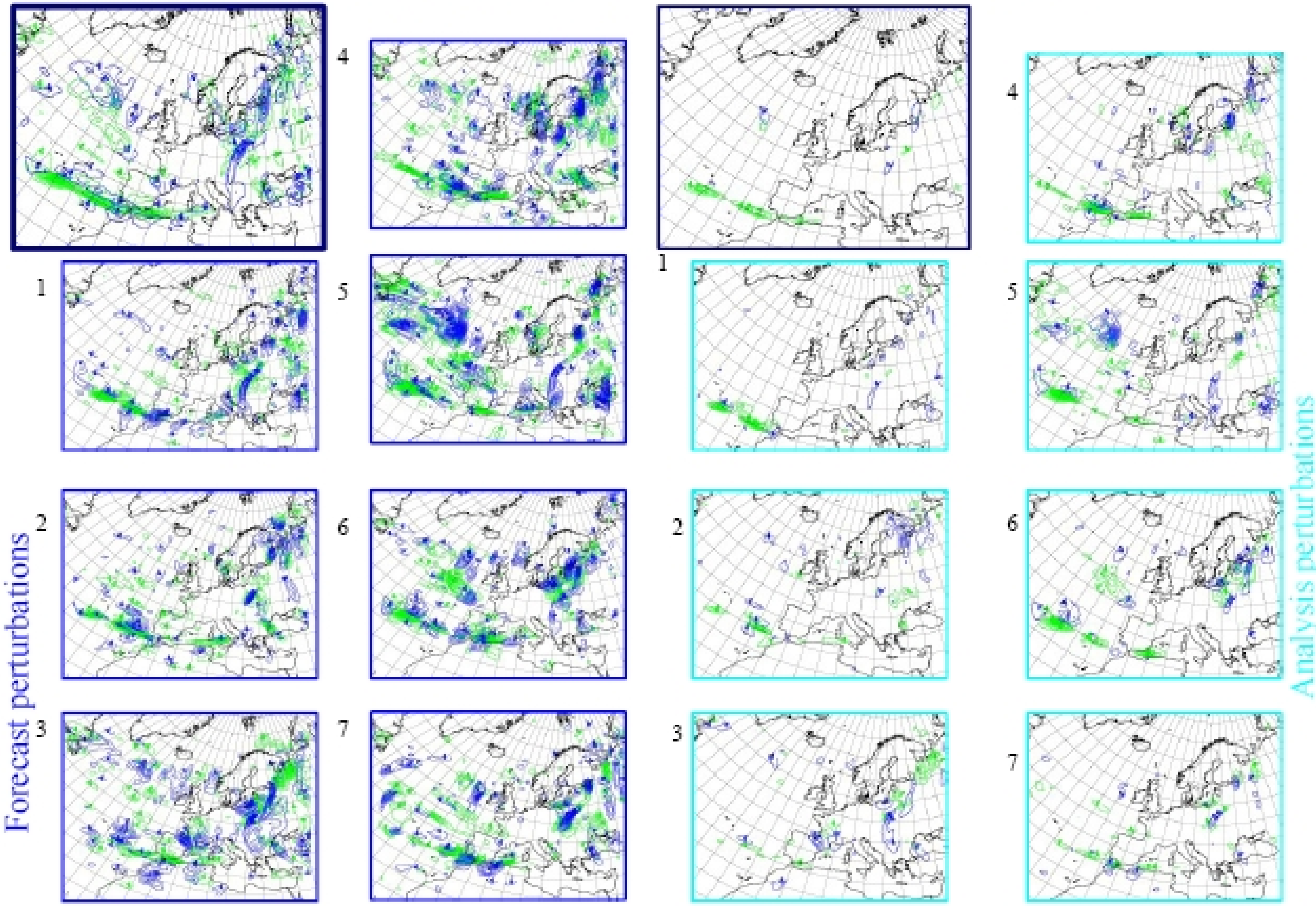




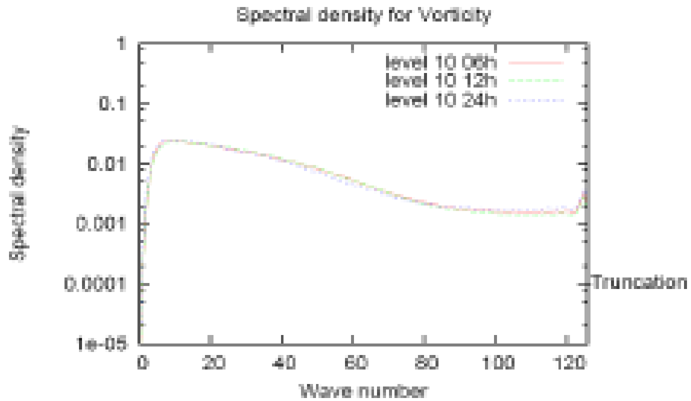
20070817_12+024
20070818_06+006
20070818_12+000
20070818_12+006
20070818_12+012
20070818_12+024
Vertically integrated total perturbation energy
E-dimension

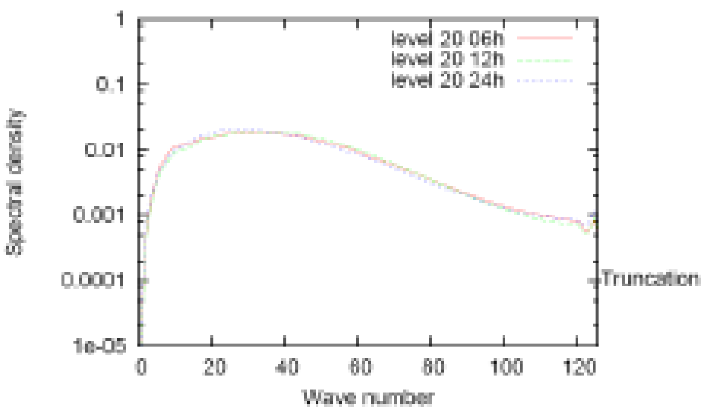

20070818_06+006
Temperature
U-wind comp, V-wind comp
Lev 10
Lev. 20
Lev. 30
E-dimension
Average of vertically integrated energy of perturbations
Forecast variance at model levels
Control forecast

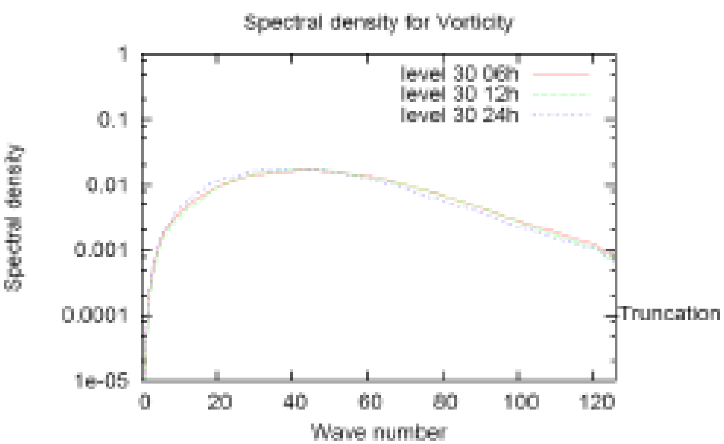
Squared projection on the eigen-vectors space



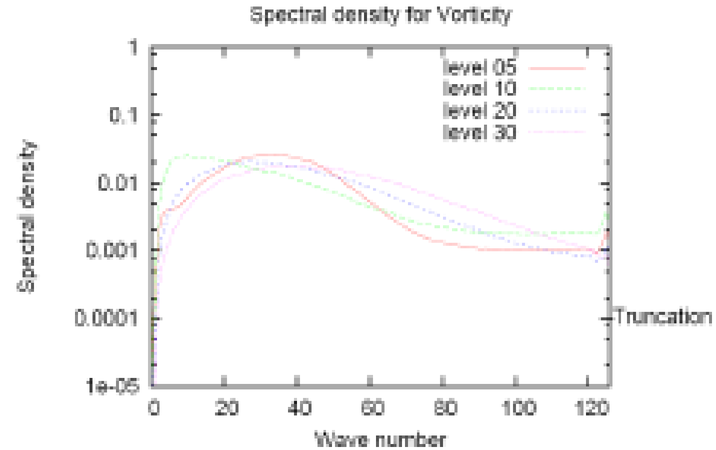
ETKF rescaling, 8 members

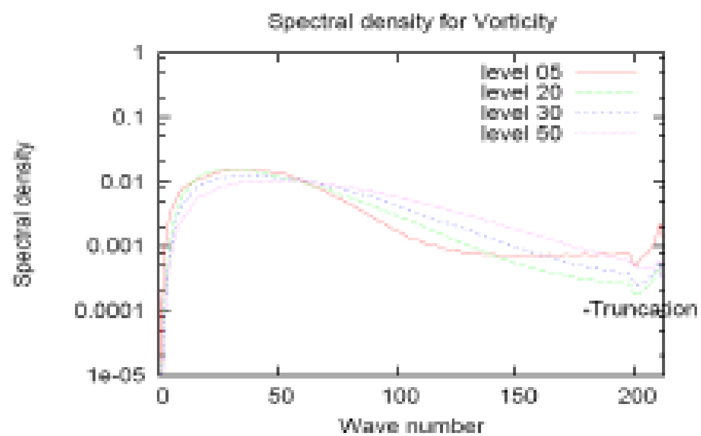






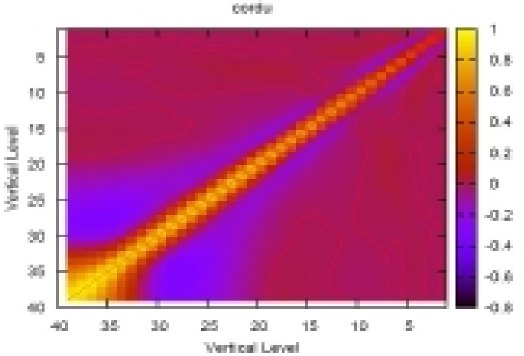
Time averaged spectral density for vorticity



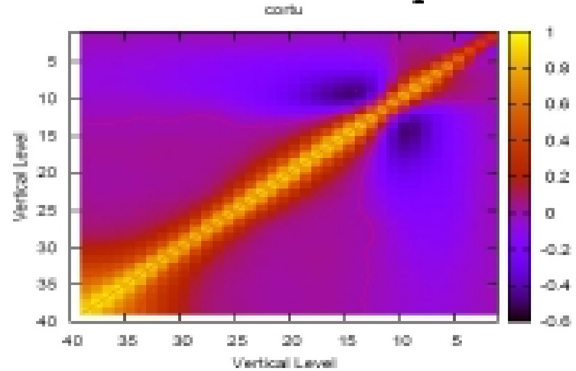


HIRLAM structures of B_{3DVar} diff. 36 – 12 hours forecasts

Cross-corr. comp u u

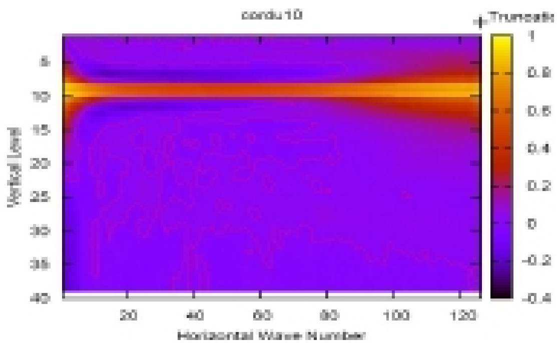


Cross-corr. comp t u

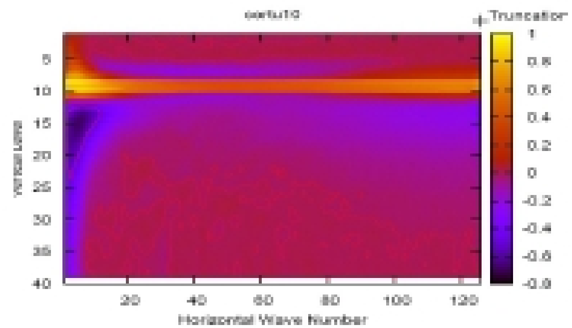


Time averaged cross-correlations, estimated from ETKF ensemble.
8 members, 40 model levels

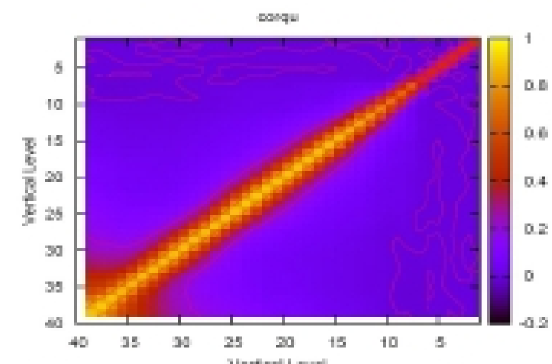
Cross-corr. comp u u 10



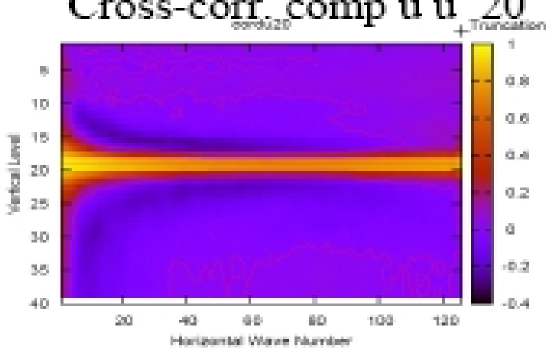
Cross-corr. comp t u 10



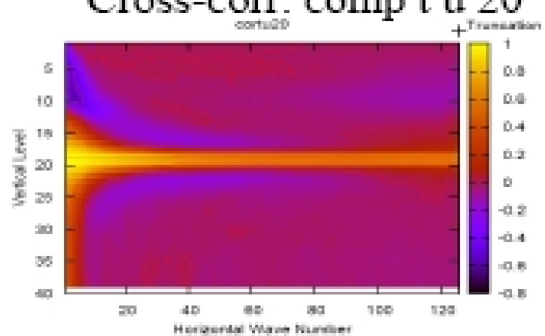
Cross-corr. comp q u



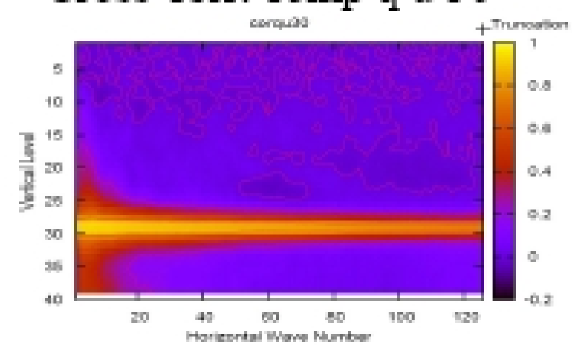
Cross-corr. comp u u 20



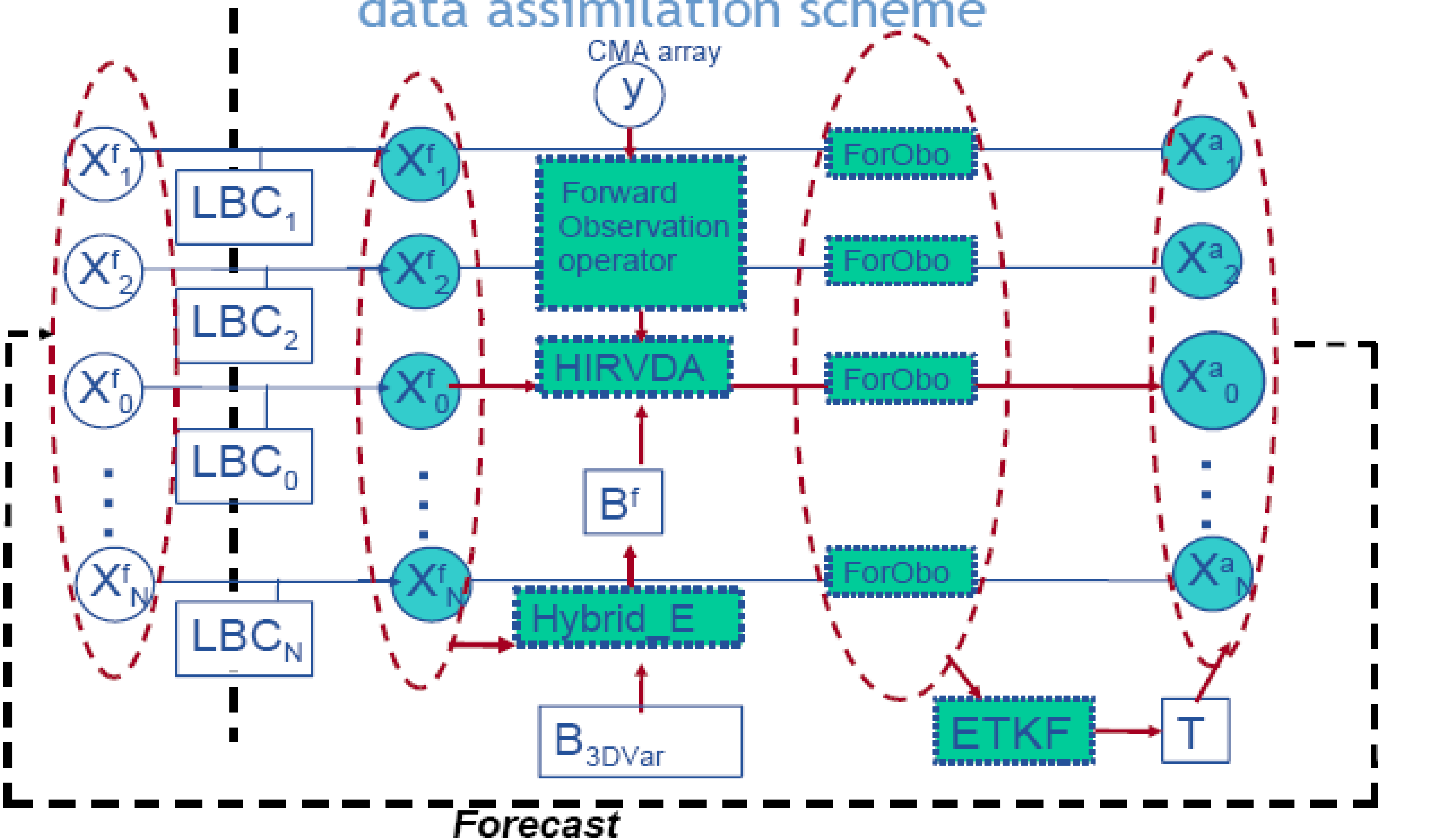
Cross-corr. comp t u 20



Cross-corr. comp q u 30



The Hybrid ETKF-Variational data assimilation scheme



The Hybrid Ensemble-Variational Data Assimilation

$$\Delta x^v_t : \min J(\Delta x^v_t) = 0.5 J_b(\Delta x^v_t) + 0.5 J_o(\Delta x^v_t) =$$
$$(\Delta x_t^v)^T (B)^{-1} \Delta x^v_t + (y_t - H(x^f_t + \Delta x^v_t))^T R^{-1} (y_t - H(x^f_t + \Delta x^v_t))$$

$$\Delta x^v_t \Rightarrow \Delta x_{t,ijl} = \Delta x^v_{t,ijl} + \Delta x^e_{t,ijl} = \Delta x^v_{t,ijl} + \Sigma_{1 \le n \le N} \Delta x^n_{t,ijl} \, a^n_{t,ij}$$

$$\Delta x^v_t : \min J(\Delta x_t) = 0.5 \beta_1 J_b(\Delta x^v_t) + 0.5 \beta_2 J_e(\Delta x^e_t) + 0.5 J_o(\Delta x_t) =$$
$$0.5 \beta_1 (\Delta x_t^v)^T (B)^{-1} \Delta x^v_t + 0.5 \beta_2 (\Delta x_t^e)^T (B_{et})^{-1} \Delta x^e_t +$$
$$0.5 (y_t - H(x^f_t + \Delta x^v_t + \Delta x^e_t))^T R^{-1} (y_t - H(x^f_t + \Delta x^v_t + \Delta x^e_t))$$

$$1/\beta_1 + 1/\beta_2 = 1$$

The Hybrid Ensemble-Variational data assimilation scheme

$$\min J_{HEV}(\Delta x_t) = 0.5\ \beta_1 \eta^{v\,T}_t \eta^v_t + 0.5\ \beta_2 \eta^{e\,T}_t \eta^e_t + 0.5\ J_o(x_b + \Delta x_{vt} + \Delta x_{et})$$

with

$$\Delta x_t = \Delta x_{vt} + \Delta x_{et} = B_v^{1/2}\ \eta^v_t + B_{et}^{1/2}\ \eta^e_t$$

<u>"produce" (???)</u> the same analysis increments as the Variational Data Assimilation scheme with the modified background covariance matrix B

$$\min J_{VAR}(\Delta x_t) = 0.5\ \Delta x_t^T B^{-1} \Delta x_t + 0.5\ J_o(x_b + \Delta x_t)$$

where

$$B = B_v/\beta_1 + B_{et}/\beta_2$$

Augmented set of control variables

$$J(\Delta x_t) = 0.5(\beta_1 (\eta^v_t)^T \eta^v_t + \beta_2 (\eta^e_t)^T \eta^e_t) +$$

$$0.5(y_t - H(x^f_t + U^v \eta^v_t + U^e_t \eta^e_t))^T R^{-1} (y_t - H(x^f_t + U^v \eta^v_t + U^e_t \eta^e_t))$$

where

$$U^v = B^{1/2},\ U^e_t = (B^e_t \bullet S)^{1/2}$$

$$(B^e_t \bullet S)_{ij} = \Sigma_{1\le n\le N} Z^e_{it} Z^e_{jt} S_{ij} = \Sigma_{1\le n\le N} Z^e_{it} Z^e_{jt} \Sigma_{1\le r\le R} s_{ir} s_{jr} =$$

$$\Sigma_{1\le k\le NR} W^n_{it} W^n_{jt}$$

$$\Delta x^e_t = U^e_t \eta^e_t = \Sigma_{1\le k\le NR} w^n_t \eta^e_{kt} = \Sigma_{1\le n\le N} Z^{en}_t \Sigma_{1\le r\le R} s_r \eta^e_{rnt} = \Sigma_{1\le n\le N} Z^{en}_t a^n_t$$

$$\Delta x^e_{ijl,t} = \Sigma_{1\le n\le N} Z^e_{ijl,t} a^n_{ij,t} \text{ where } a^n_{ij,t} = \Sigma_{1\le r\le R} s_{ijr} \eta^e_{rnt}$$

$$J(\Delta x_t) = 0.5(\beta_1 (\eta^v_t)^T \eta^v_t + \beta_2 \Sigma_{1\le n\le N} (a^n_t)^T S^{-1} a^n_t) +$$

$$(y_t - H(x^f_t + \Delta x^v_t + \Delta x^e_{tt}))^T R^{-1} (y_t - H(x^f_t + \Delta x^v_t + \Delta x^e_t))$$

A flow-dependent correction

$$\Delta x^{a}_{ijl,t} = (\Delta u^{v}_{ijl,t} + \Delta u^{e}_{ijl,t}, \Delta v^{v}_{ijl,t} + \Delta v^{e}_{ijl,t}, \Delta T^{v}_{ijl,t} + \Delta T^{e}_{ijl,t}, \Delta \ln ps^{v}_{ijl,t} + \Delta \ln ps^{e}_{ijl,t})$$

1. $\eta^{n}_{qp,t} \xrightarrow{S_p S_q} a^{n}_{ij,t}$

2. $Z^{n}_{ijl,t}\, a^{n}_{ij,t} = (\Delta \zeta^{n}_{ijl,t}\, a^{n}_{ij,t}, \Delta \delta^{n}_{ijl,t}\, a^{n}_{ij,t}, \Delta T^{n}_{ijl,t}\, a^{n}_{ij,t}, \Delta \ln ps^{n}_{ijl,t}\, a^{n}_{ij,t})$

3. $\Delta \zeta^{corr,n}_{ijl,t}, \Delta \delta^{corr,n}_{ijl,t} \rightarrow \Delta \zeta^{corr,n}_{pql,t}; \Delta \delta^{corr,n}_{pql,t}$

4. $\Delta \zeta^{corr,n}_{pql,t}, \Delta \delta^{corr,n}_{pql,t} \rightarrow \Delta u^{corr,n}_{pql,t}, \Delta v^{corr,n}_{pql,t}$

5. $\Delta u^{corr,n}_{pql,t}, \Delta v^{corr,n}_{pql,t} \rightarrow \Delta u^{corr,n}_{ijl,t}, \Delta v^{corr,n}_{ijl,t}$

Coding of the

Hybrid Ensemble Variational Assimilation Scheme

IS FINISHED !!!!!

But pudding still must be eaten ...

Nils Gustafsson, the 8th of May 2009