Meteorologisk institutt met.no

High Resolution Limited Area Model Hirlam

The hybrid ETKF- Variational data assimilation scheme in HIRLAM

(current status, problems and further developments)

The Hungarian Meteorological Service, Budapest, 24.01.2011

Nils Gustafsson, Jelena Bojarova and Ole Vignes.

Thanks to Åke Johansson, Magnus Lindskog and Tomas Landelius

Numerical Weather Prediction (NWP)

To large extend processes in the atmosphere obey basic hydrodynamic and thermodynamic equations.

NWP models are based on the time integration of numerical approximation to these processes

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} - fv + \frac{\partial \phi}{\partial x} = 0$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + fu + \frac{\partial \phi}{\partial y} = 0$$

$$\frac{\partial \phi}{\partial t} + u\frac{\partial \phi}{\partial x} + v\frac{\partial \phi}{\partial y} + \phi(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}) = 0$$

$$X_t = T_t(X_{t-1}) + R_{t-1}\eta_{t-1}$$

Phenomena of interest

Simplifying assumptions

Numerical time integration scheme

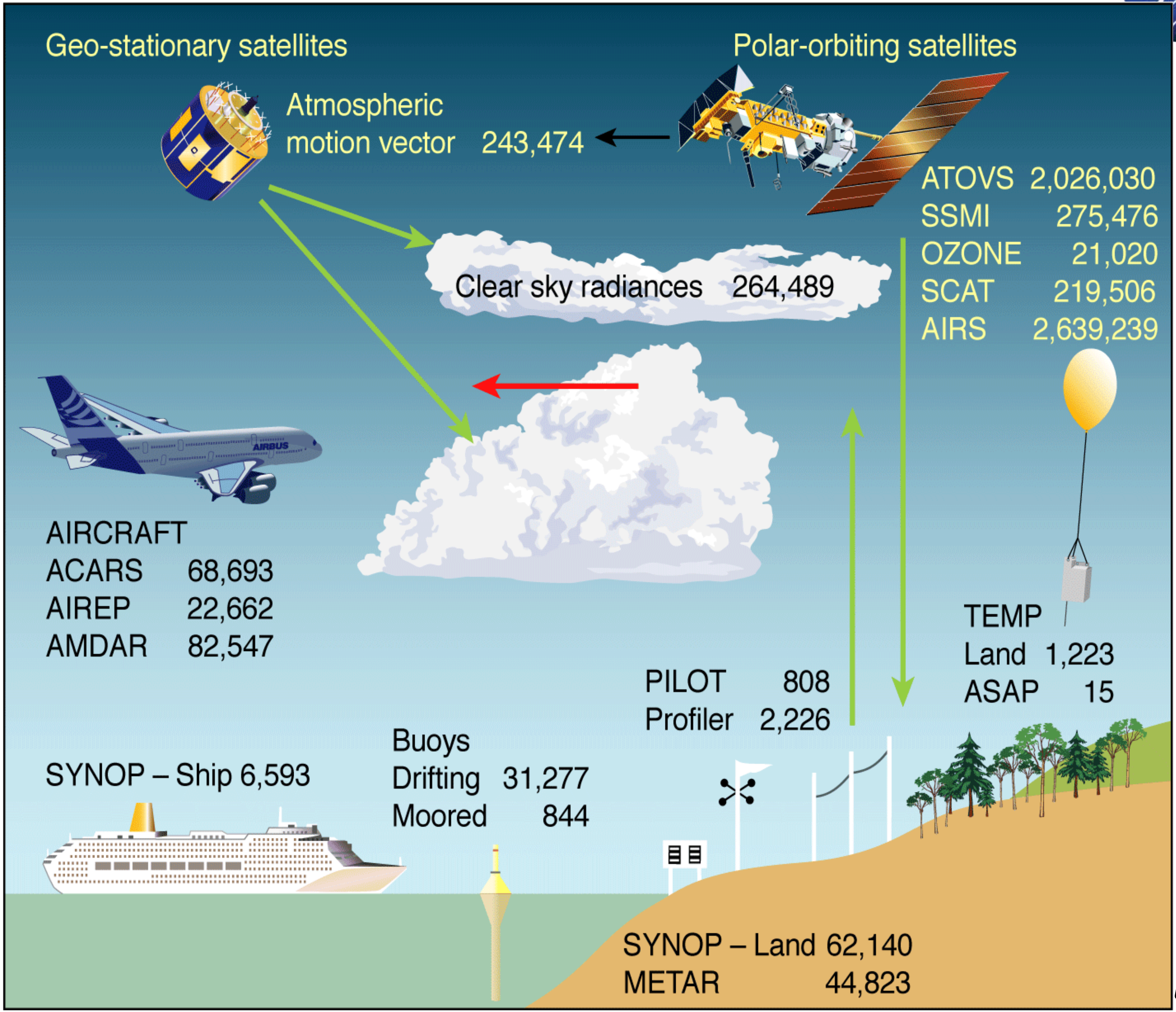
Scale of motion

Spatial discretization

Sub-grid variability

- Initial model state uncertainty;
- Limited resolution of numerical approximation;
- Lack of knowledge about complicated atmospheric processes;
- Unpredictability

Earth observing system



Data assimilation

A "true" model state is a reflection of the state of atmosphere projected on the discrete space of solution of differential equations which describe phenomena of interest.

Data assimilation prepares an initial state for NWP models.

Data assimilation provides a point estimate of the "true" model state conditional on the observed quantities.

Data assimilation feeds the NWP model with observations of nature in order to provide the best possible forecast of the phenomena of interest.

Filtering away of observation error

Interpolation of the observed information to other model state components

Balancing of model state components (explicit use of cross-dependencies)

Data assimilation tools

Prediction(linear)

The Kalman filter and the Kalman smoother tool

$$x_i^{ble} = E(X_i \mid \mathcal{Y}_{i-1}) + cov(X_i, v_i \mid \mathcal{Y}_{i-1})(var(v_i \mid \mathcal{Y}_{i-1}))^{-1} v_i$$
$$B_i^{ble} = Var(X_i \mid \mathcal{Y}_{i-1}) - cov(X_i, v_i \mid \mathcal{Y}_{i-1})(var(v_i \mid \mathcal{Y}_{i-1}))^{-1} cov(X_i, v_i \mid \mathcal{Y}_{i-1})^{-1}$$

Mode of posterior distribution (parametric)

Minimization of cost function

$$\tilde{a} = argmin L(X_0, \ldots, X_n)$$
$$= argmin\{-\log p(X_0) - \sum_{i=1}^{n} (\log p(y_i \mid X_i) + \log p(X_i \mid X_{i-1}))\}$$
$$\tilde{B} = \left[\left(\frac{\partial^2 L(X_0, \ldots, X_n)}{\partial X_i \partial X_j} \right)_{0 \le i,j \le n} \right]^{-1}$$

Simulations (non-parametric)

Importance sampling

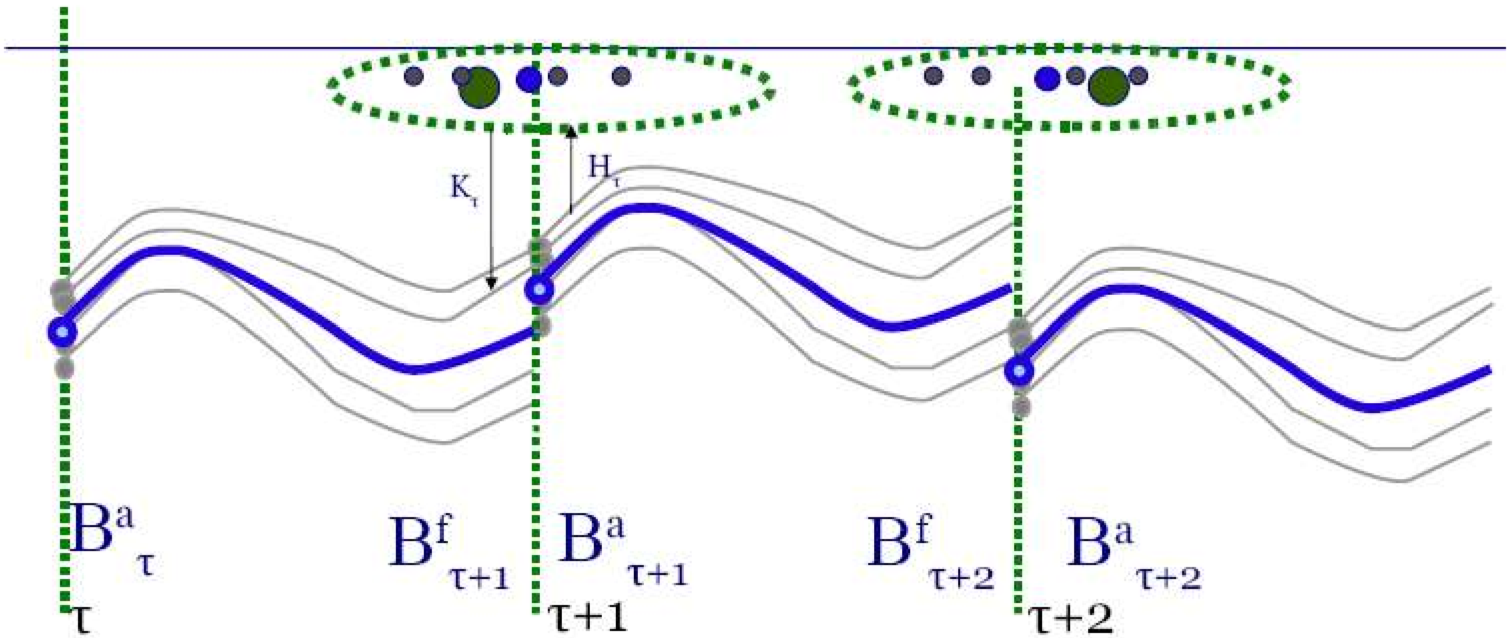
$$a = E(f(X_\tau) \mid \mathcal{Y}_\tau) \approx \sum_{i=1}^{N} f(x_{i,\tau}^a) w_{i,\tau}.$$

Sample from a convenient distribution;

Correct for properties of the original one

High resolution numerical weather prediction for synoptic scale phenomena

Sequential estimation of initial uncertainty



Approximate solution to the sequential update problem

non-linear state space model

$$y_{t_i} = \mathcal{H}_{t_i}(X_{t_i}) + \epsilon_{t_i}$$
$$X_{t_i} = \mathcal{M}(t_i, t_{i-1})(X_{t_{i-1}}) + T_i \xi_{t_i}$$

tangent-linear approximation

$$X_{t_i} = X^f_{t_i} + \delta x_{t_i}, \; X^f_{t_i} = \mathcal{M}(X_{t_{i-1}})$$
$$y_{t_i} = \mathcal{H}(X^f_{t_i}) + H \delta x_{t_i} + \epsilon_{t_i}$$
$$\delta x_{t_i} = M_{(t_{i-i}, t_i)} \delta x_{t_{i-1}} + T_i \xi_i$$

extended Kalman filter solution

$$\delta x_\tau = K_\tau (y_\tau - \mathcal{H}(X^f_\tau) - H_\tau \delta x_\tau)$$
$$K_\tau = B^f_\tau H^T_\tau (R_\tau + H_\tau B^f_\tau H^T_\tau)^{-1}$$
$$B_\tau = (I - K_\tau H_\tau) B^f_\tau$$

3D-Variational solution

$$\delta x_\tau = argmin L(\delta x_\tau \mid X^f_\tau, y_\tau)$$
$$L(\delta x_\tau \mid X^f_\tau, y_\tau) = 0.5 (\delta x_\tau)^T B^{-1} \delta x_\tau$$
$$+0.5 (y_\tau - \mathcal{H}(X^f_\tau) - H_\tau \delta x_\tau)^T R^{-1}_\tau (y_\tau - \mathcal{H}(X^f_\tau) - H_\tau \delta x_\tau)$$

Ensemble Kalman filter

$$B^f_\tau = Z^f_\tau (Z^f_\tau)^T$$
$$B_\tau = Z^a_\tau (Z^a_\tau)^T$$
$$Z^a_\tau = Z^f_\tau T_\tau$$

Hybrid Variational Ensemble Kalman filter

Different approaches for using ensembles in variational data assimilation

– **Covariance modelling with parameters of the covariance model determined from an ensemble.** Use for example a wavelet-based covariance model (Alex Deckmyn; Loik Berre et al. Meteo-France)
– **Use the ensemble-based covariances in a hybrid variational ensemble data assimilation** (Barker et al. WRF, UK Met.Office, HIRLAM)
– **Ensembles can also be used to determine static background error statistics**

Toward data assimilation using flow-dependent structures in HIRLAM

Rapid update cycle

(a short range forecasting of events of shorter temporal and spatial scales of variability)

Meso-scale data assimilation

(extraction from observations information about amospheric state related to processes with shorter temporal and spatial scales of variability)

Require

- Structures of background error covariance dependent of the observation network
- Structures of the background error covariance for meso-scale processes, which are dependent on the large scale forcing and are difficult to derive analytically

The ETKF rescaling perturbations in HIRLAM

ETKF (Ensemble Transform Kalman Filter) perturbations resemble structures of the analysis error covariance. The method can be viewed as a Generalisation of the Breeding Method.

Perturbations are grown through the non-linear model and downscaled afterwards.

A matrix of the ensemble size is constructed to perform downscaling.

Observational network is used to construct the matrix.

A scalar multiplicative inflation, based on the fit of the ensemble spread to the total variance of innovations, is applied in order to make amplitude of perturbations physically meaningfully

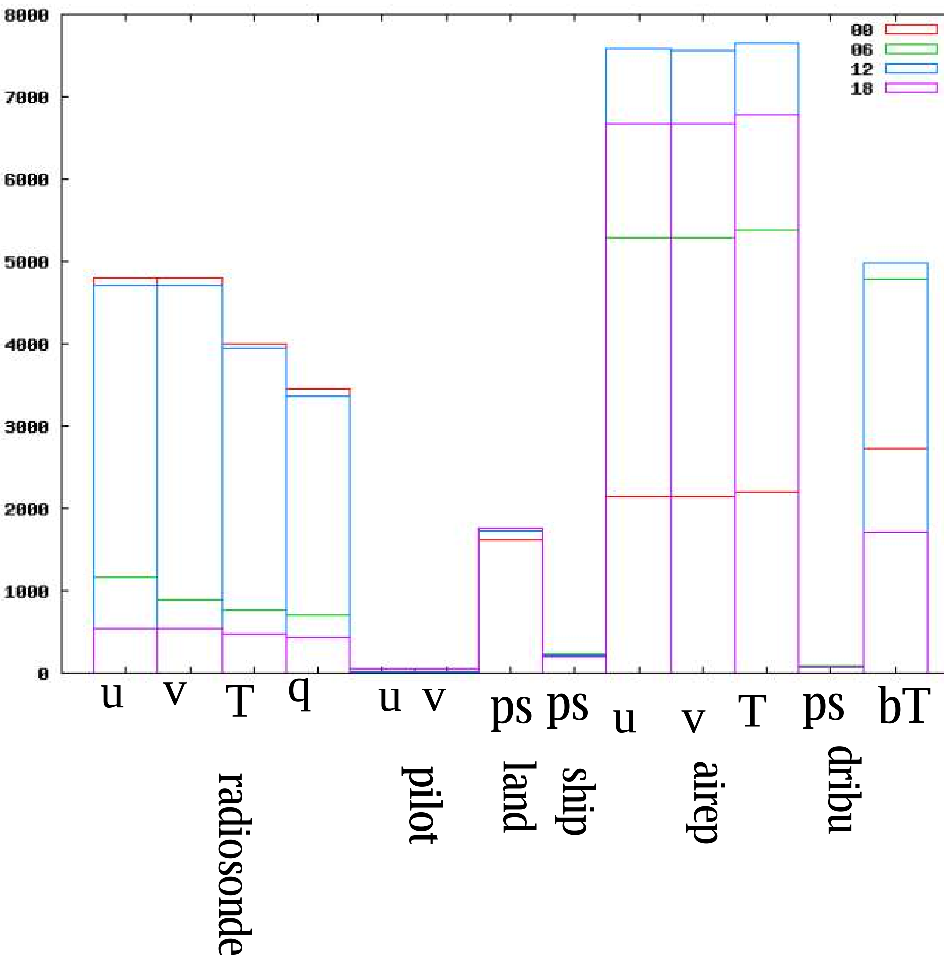
An ad-hoc downscaling of the non-leading eigenvectors of the ensemble analysis error covariance is performed

An additive inflation of the analysis error covariance is applied by mixing with the random perturbations with structures of $B_{3D\text{-}Var}$

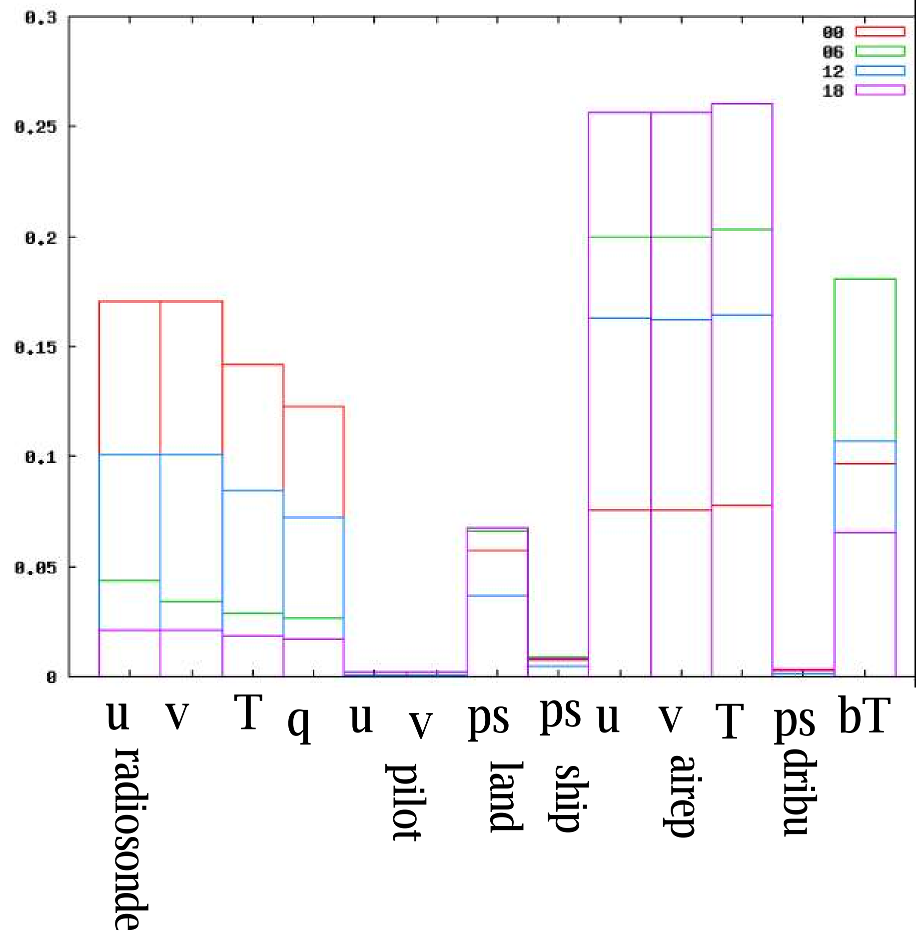
HIRLAM ETKF perturbations are mixed with the TEPS perturbations on the lateral boundaries and in the stratosphere.

Earth observing system

absolute amount of observations



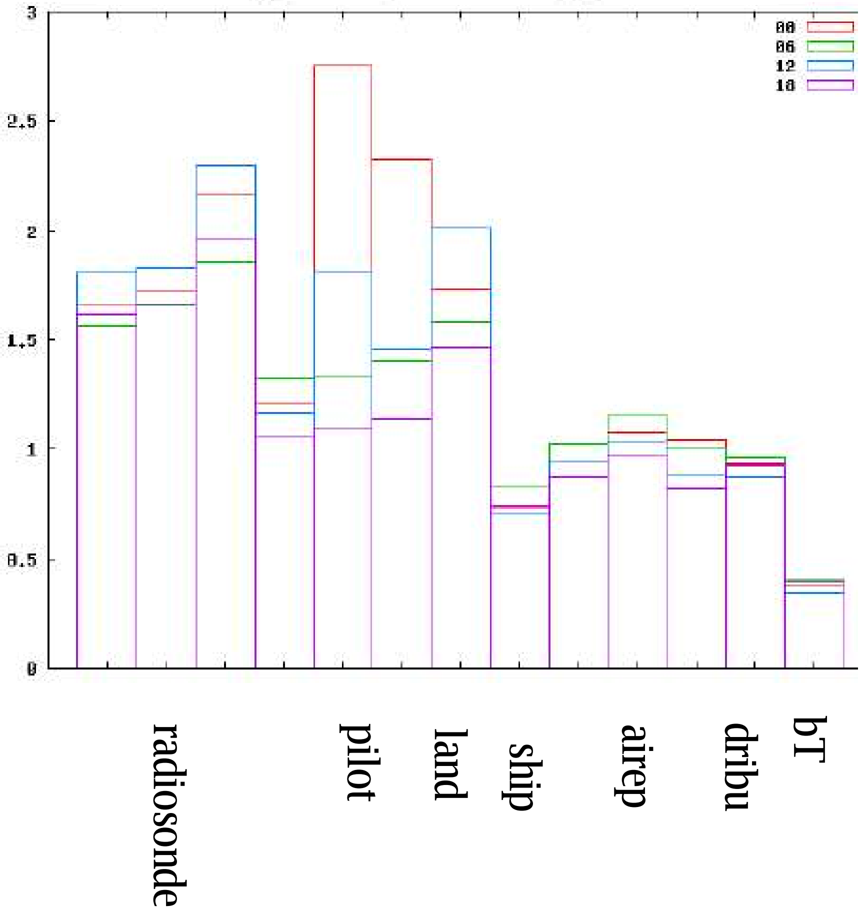
relative amount of observations

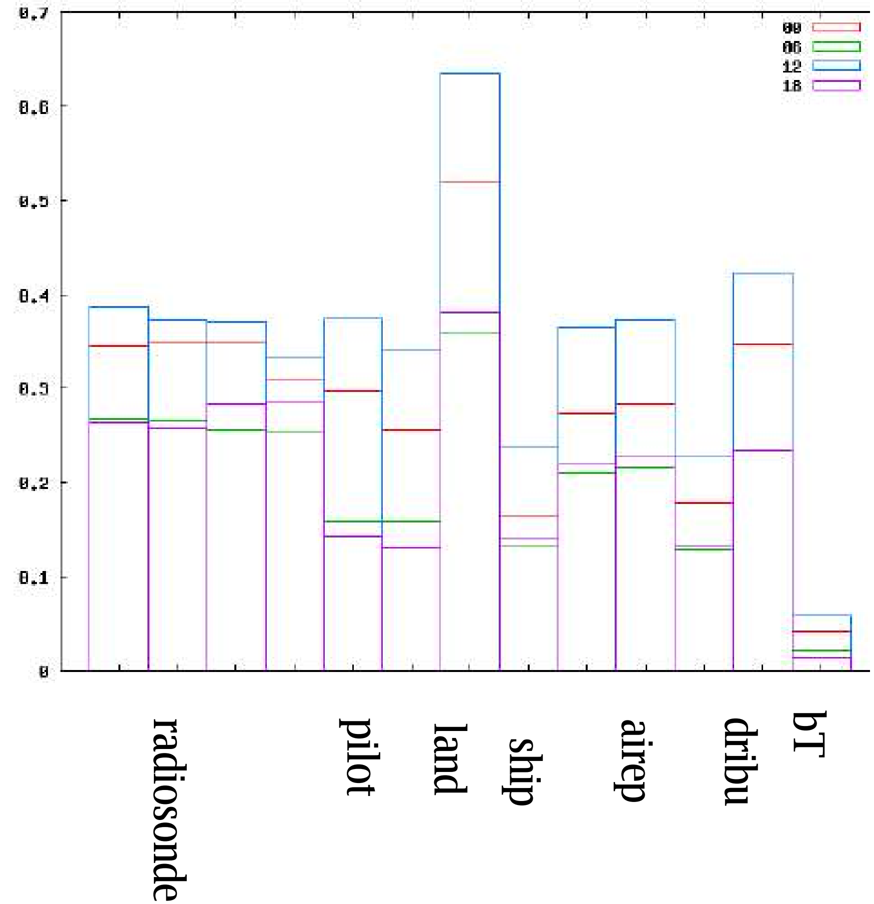


Relative squared innovation variance versus relative spread

$$d_i^2 := \sum_{j=1}^{p} \left(\frac{y_j - \mathcal{H}_{ji} x_i^b}{\sigma_{o,ji}}\right)^2 \approx \alpha_i \sum_{n=1}^{N-1} \lambda_{ni}^{N-1} + p$$

$$trace(B^f_{ens}) = p \frac{1}{p} \sum_{j=1}^{p} \frac{Var(\mathcal{H}_j(\delta X^f))}{\sigma^2_{o,j}}$$





HIRLAM approach to use ensembles in 3D-Var (and 4D-Var)

- ***Impose*** observation-network-dependent structures for analysis perturbations applying the ETKF rescaling scheme on a 6h forecast ensemble.
- ***Grow*** flow-dependent structures by integrating analysis ensemble forward in time to obtain the 6h forecast perturbations.
- ***Perform*** the variational data assimilation blending the structures of the full-rank statically and analytically deduced $B_{3D\text{-}Var}$ and the flow- and observation-network dependent structures of the rank-deficient B^f_{ens}.
- ***Repeat*** Steps 1-3

Lorenc (2003) augmentation of the control vector space:

$$J(\delta x_{3D-Var}, \alpha) = \beta_{3D-Var} J_{3D-Var}(\delta x_{3D-Var}) + \beta_{ens} J_{ens}(\alpha) + J_o$$

Spatial mean of $\alpha_k = 0$;
Spatial variance of $\alpha_k = 1/K$ is constant and controls amplitude;
Horizontal auto-correlation controls smoothness of α_k fields

$$\frac{1}{\beta_{3D-Var}} + \frac{1}{\beta_{ens}} = 1 \qquad J_{ens} = \frac{1}{2}\alpha^T A^{-1} \alpha$$

The same α_k fields for vertical levels and all types of model state components

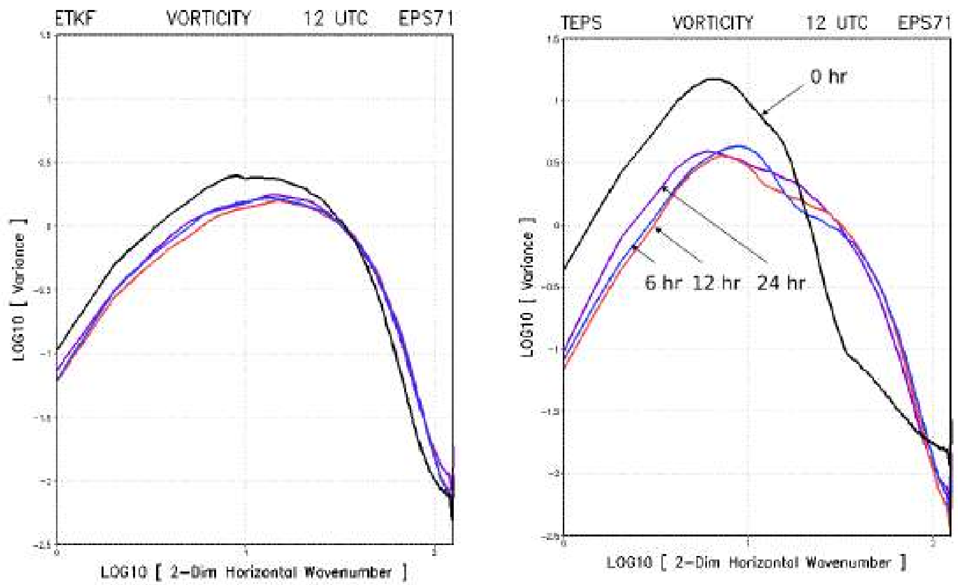
$$\delta x = \delta x_{3D-Var} + \sum_{k=1}^{K} (\alpha_k \circ \delta x_k^{ens})$$

Empirical matrix A contains spectral density of the horizontal auto-correlation of α_k fields

Spatial averaging is applied on vorticity, divergence, temperature, specific humidity and log of surface pressure in order to preserve a geostophic balance.

Diagnosis of ETKF perturbations-horizontal spectra

ETKF based | **Singular vector based**

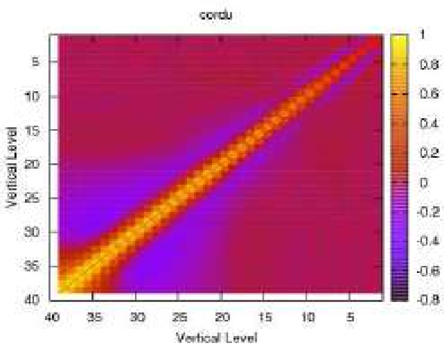


(a) (b)

Figure 10: The horizontal spectral density of the variance for the forecast error of vorticity at 00h (black), 06h (blue), 12h (red) and 24h (magenta), estimated from the ETKF perturbations (a) and from the TEPS perturbations (b)

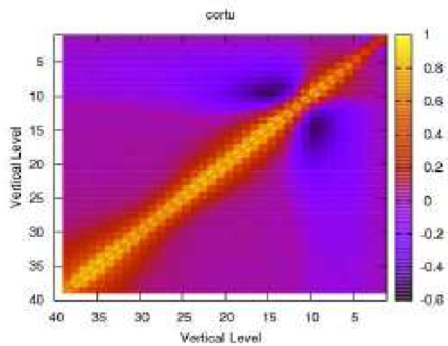
Diagnosis of ETKF perturbations - +12 h vertical correlations and balance relations

Auto-correlation Unb. divergence



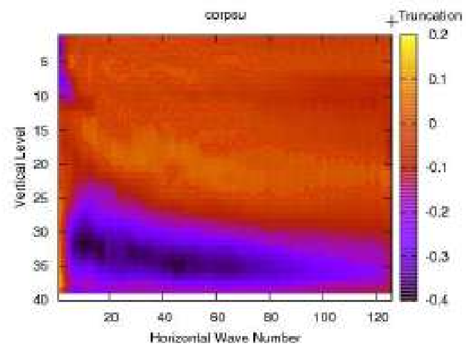
(a)

Auto-correlation Unb. temperatute



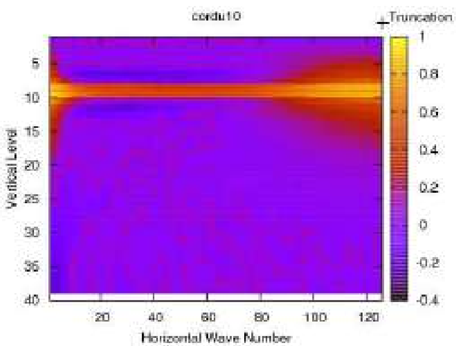
(b)

Cross-correlation Unb. Temperature Unb. Surface pressure

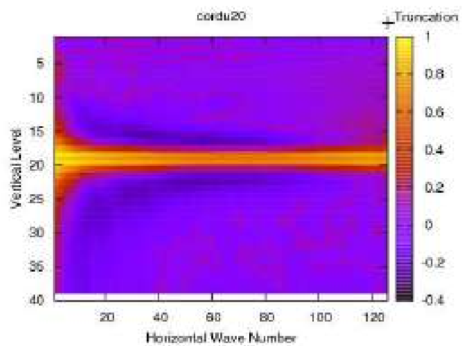


(c)

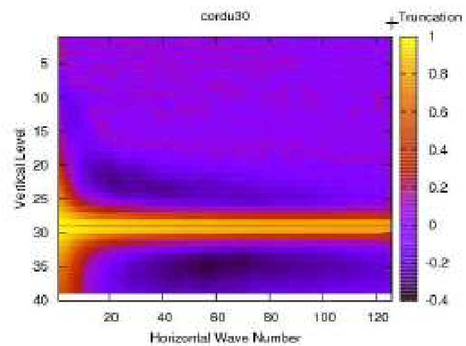
Auto-correlation of vorticity as a function of wave-number



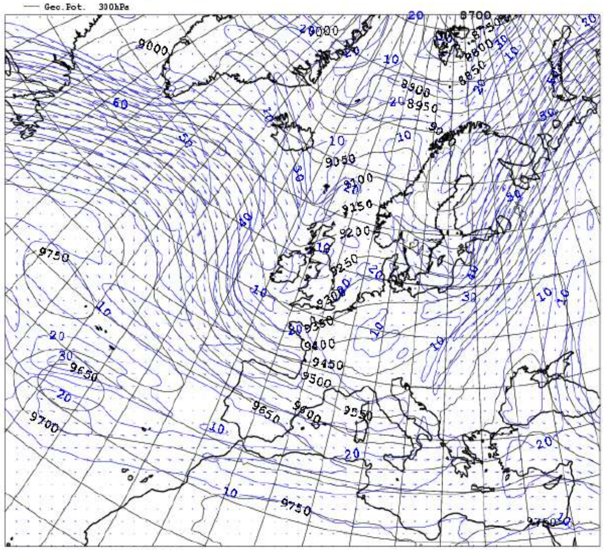
model level 10 (d)



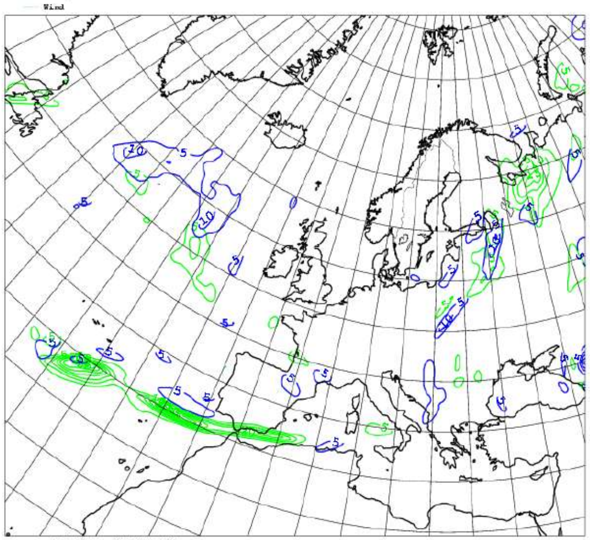
model level 20 (e)



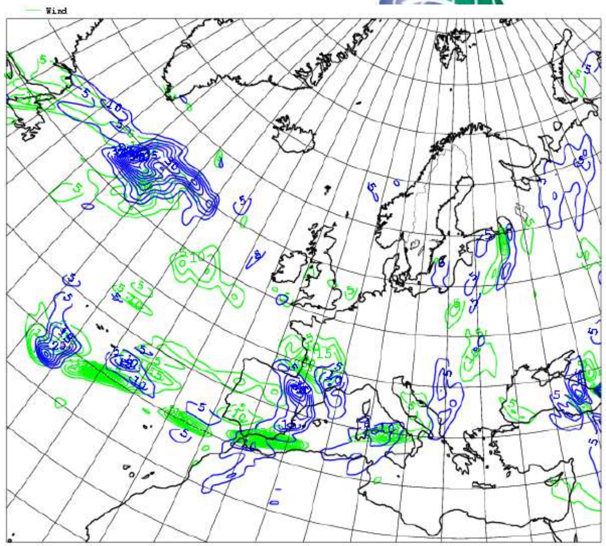
model level 30 (f)



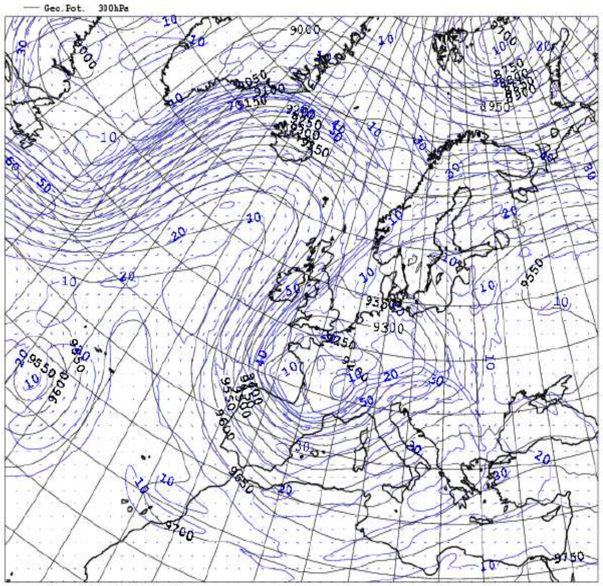
Sat 18 Aug 2007 06Z +06h
valid Sat 18 Aug 2007 12Z



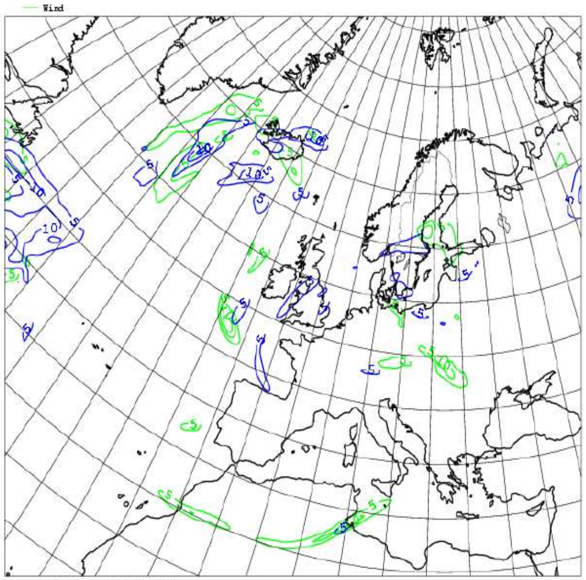
Sat 18 Aug 2007 06Z +00h
valid Sat 18 Aug 2007 06Z



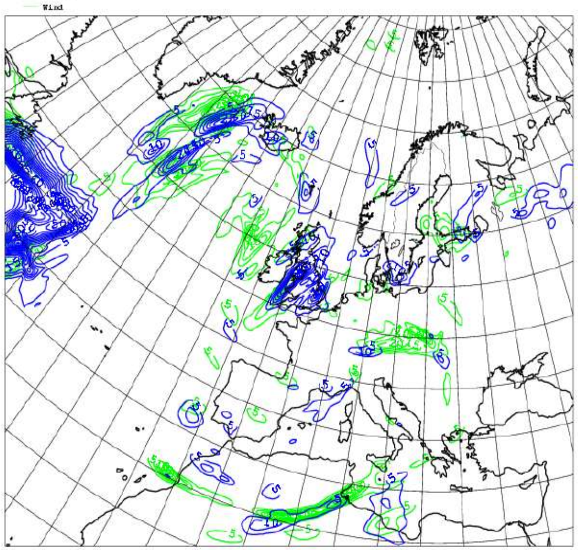
Sat 18 Aug 2007 06Z +00h
valid Sat 18 Aug 2007 06Z



Tue 21 Aug 2007 06Z +06h
valid Tue 21 Aug 2007 12Z



Tue 21 Aug 2007 06Z +00h
valid Tue 21 Aug 2007 06Z

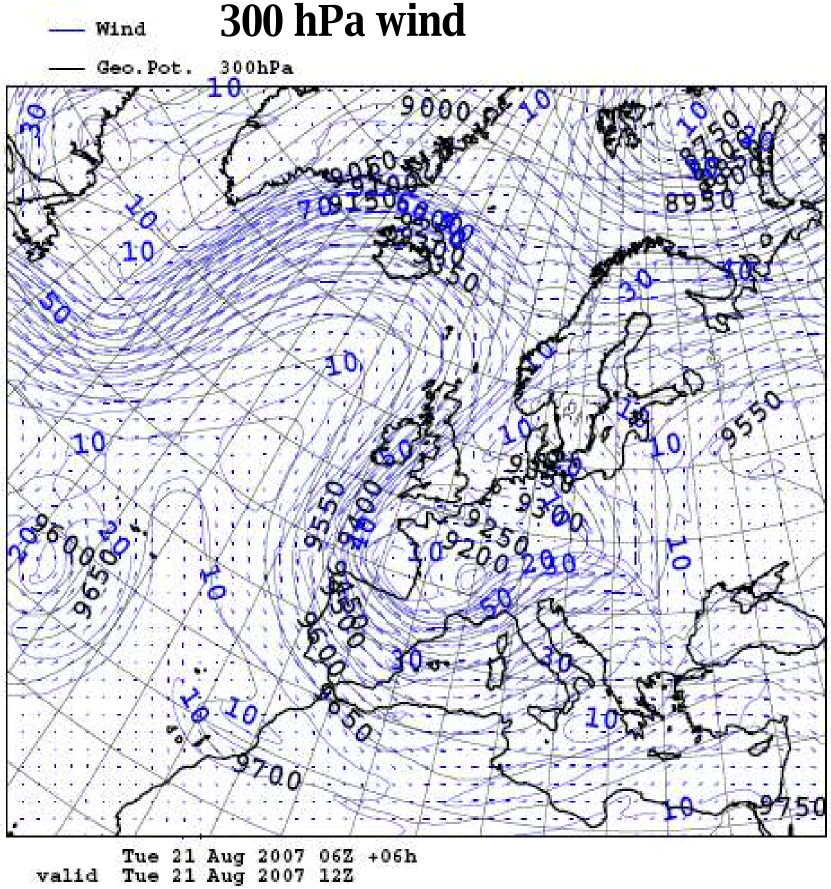


Tue 21 Aug 2007 06Z +00h
valid Tue 21 Aug 2007 06Z

Flow situation – 21 August 2007

Data assimilation experiment (12.08.2007-24.08.2007)

300 hPa wind



850 hPa temperature

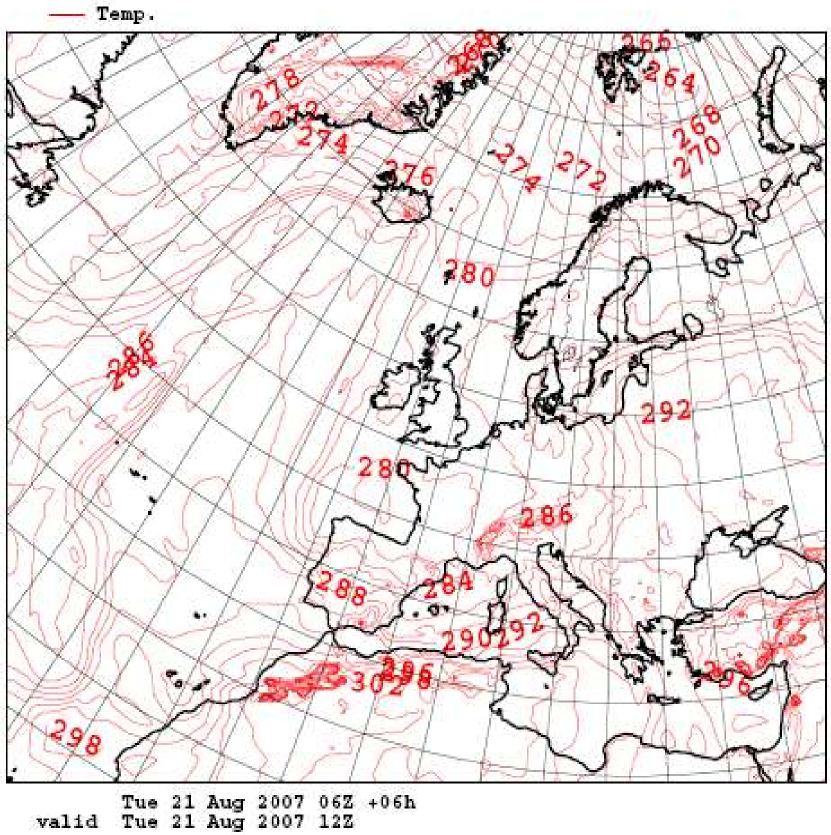


Figure 4. 300 hPa wind and geopotential (left) and 850 hPa temperature (right) taken from the background model state at 21 August 2007 06UTC + 6h; experiment based on equal static and ensemble contributions to the background error variance

Example of ensemble variance (spread) fields (12 members)

Wind components model level 15

Wind components model level 20

Temperature model level 30

Specific humidity model level 30

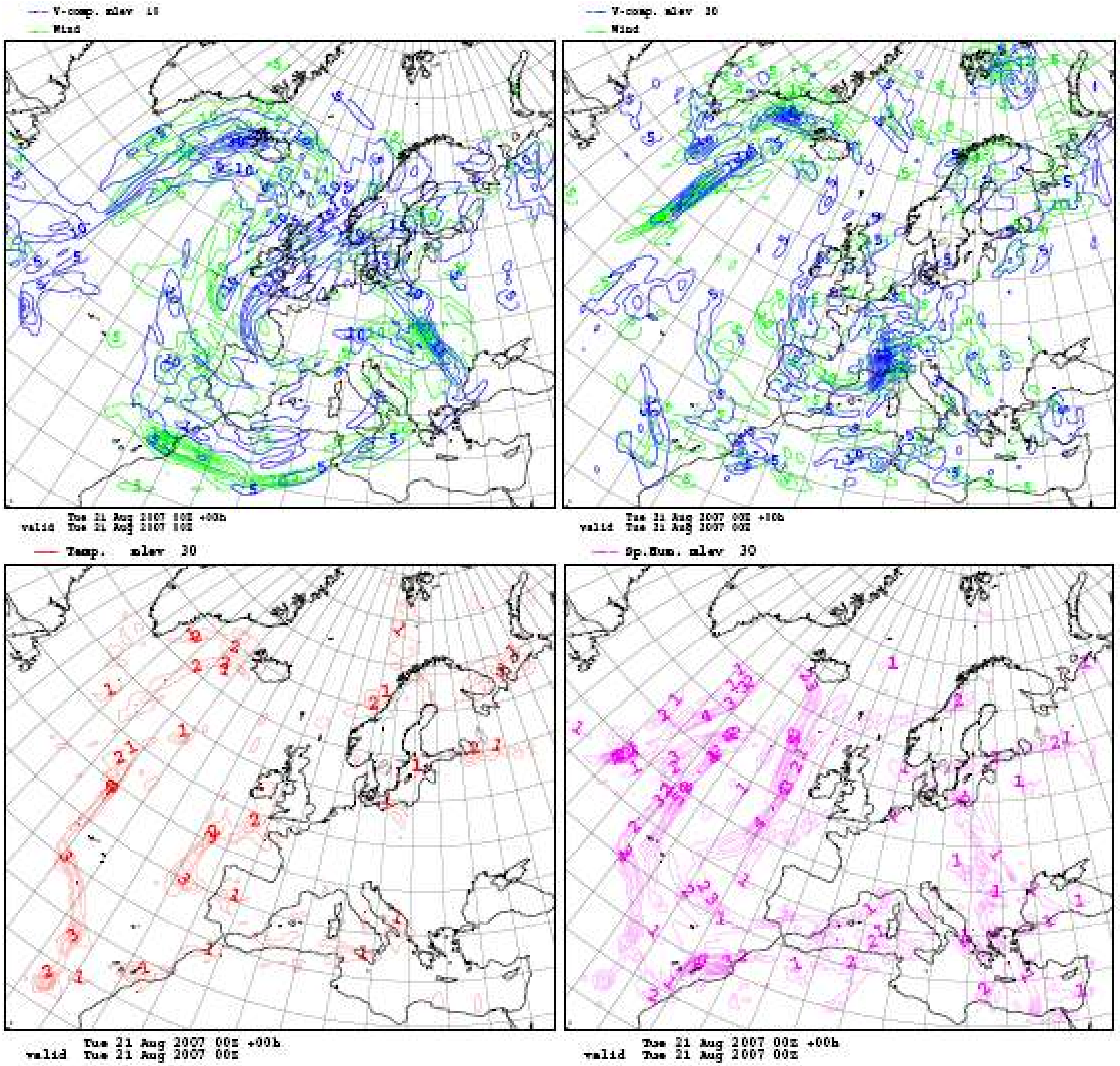
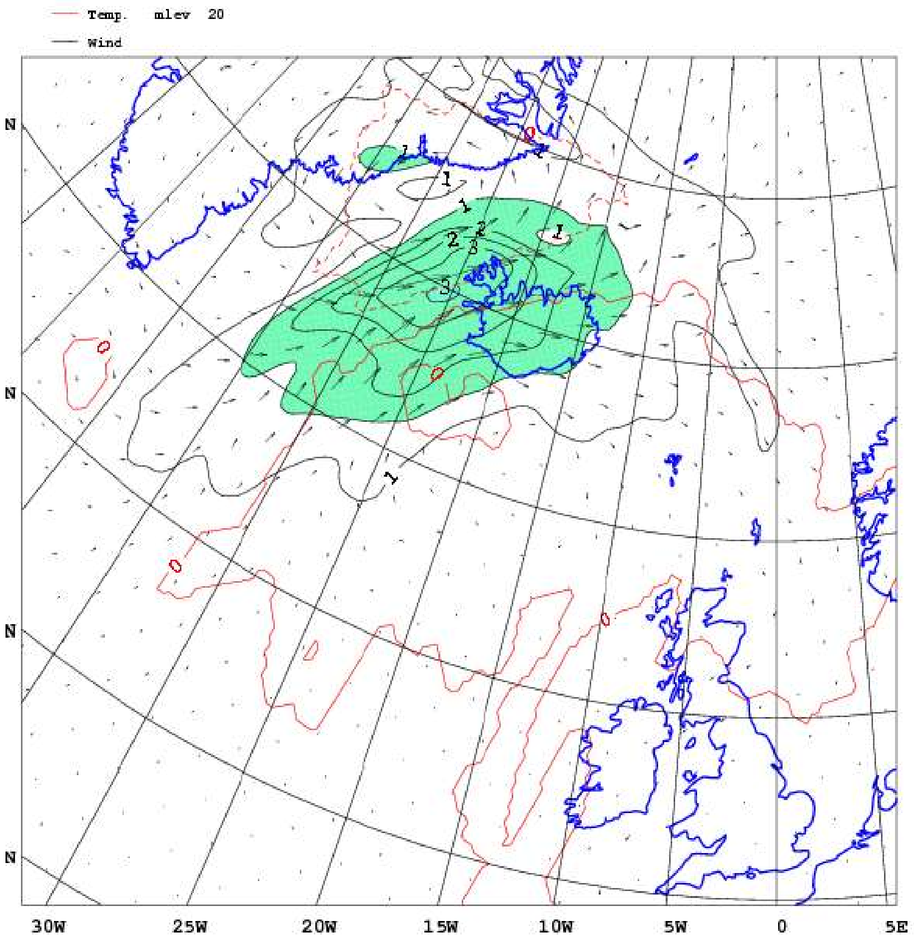


Figure 5. Examples of estimated background error variances based on the ensemble of +6h forecast valid at 21 August 2007 12UTC from the hybrid data assimilation experiment hybrid6_diag. Wind components at model level 10 (upper left), wind components at model level 20 (upper right), temperature at model level 30 (lower left) and specific humidity at model level 30 (lower right).

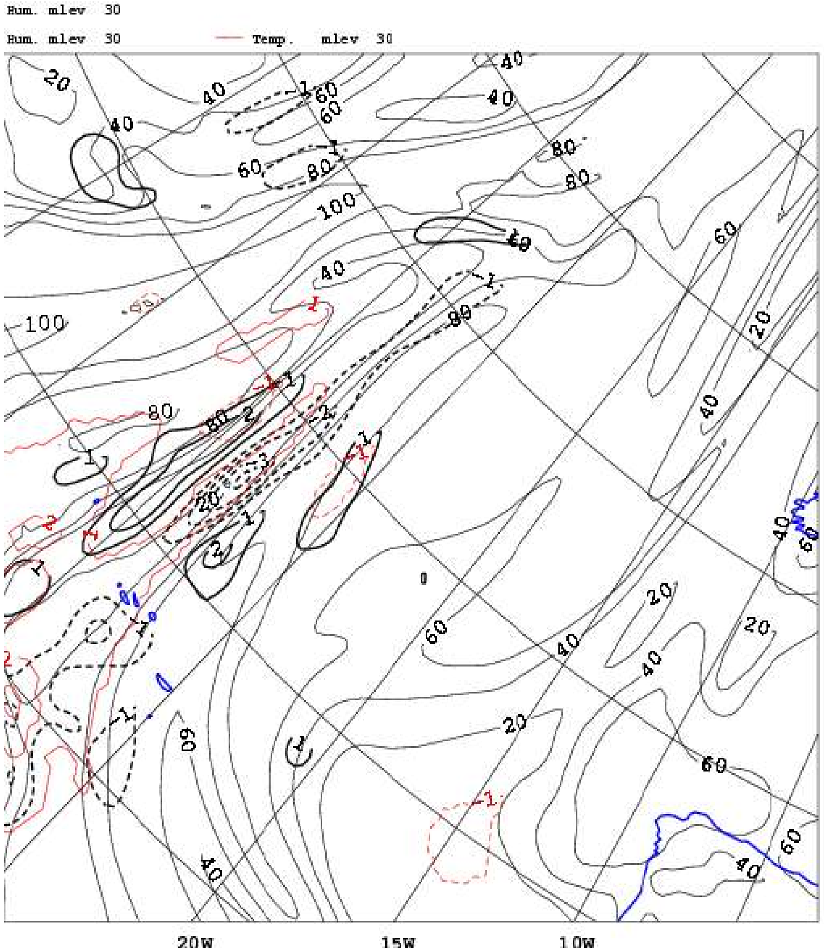
Single observation experiments

$$1/\beta_{3D\text{-}VAR} = 1/11$$

Wind increment (65N,25W) 300 hPa

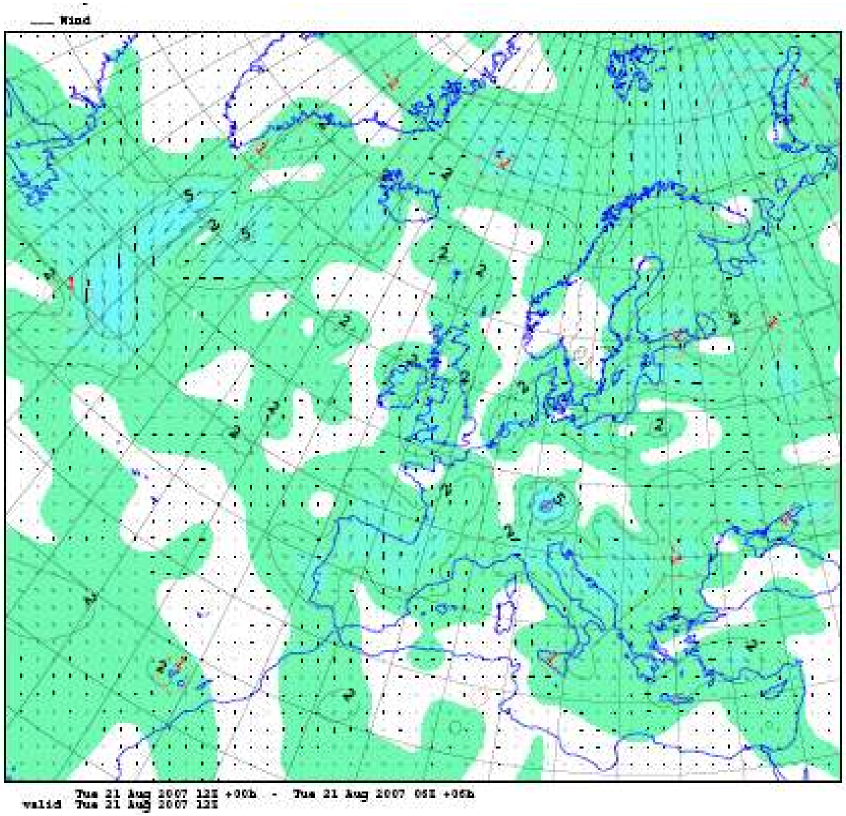


Tempture era increment (40N,30W) 850 hPa

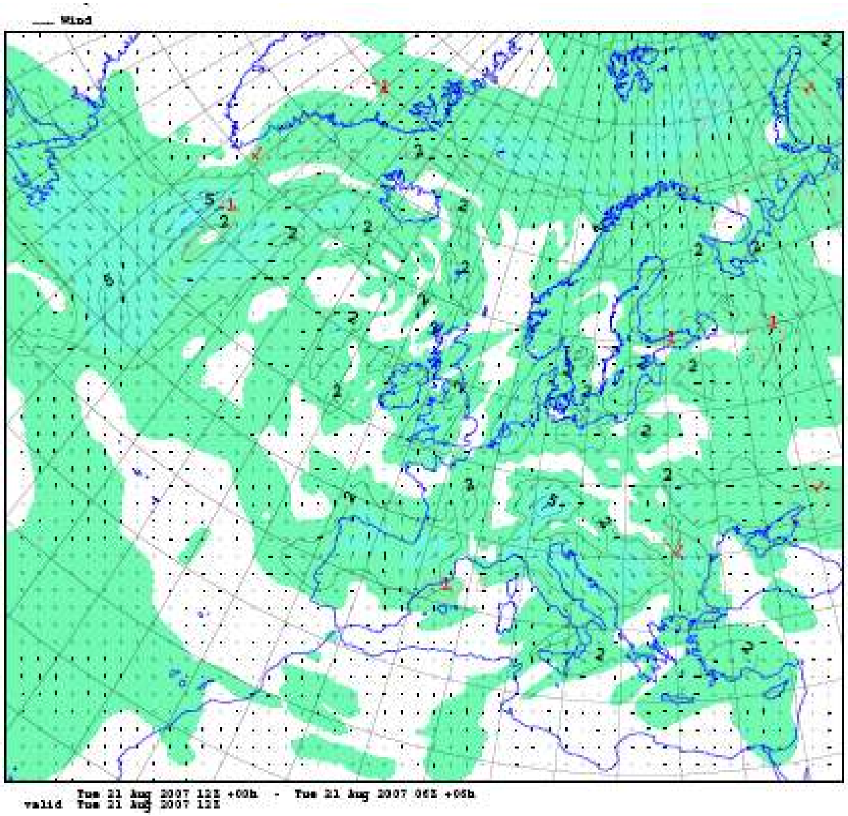


Example of assimilation increment – model level 15 wind

$1/\beta_{3D\text{-}VAR} = 1$

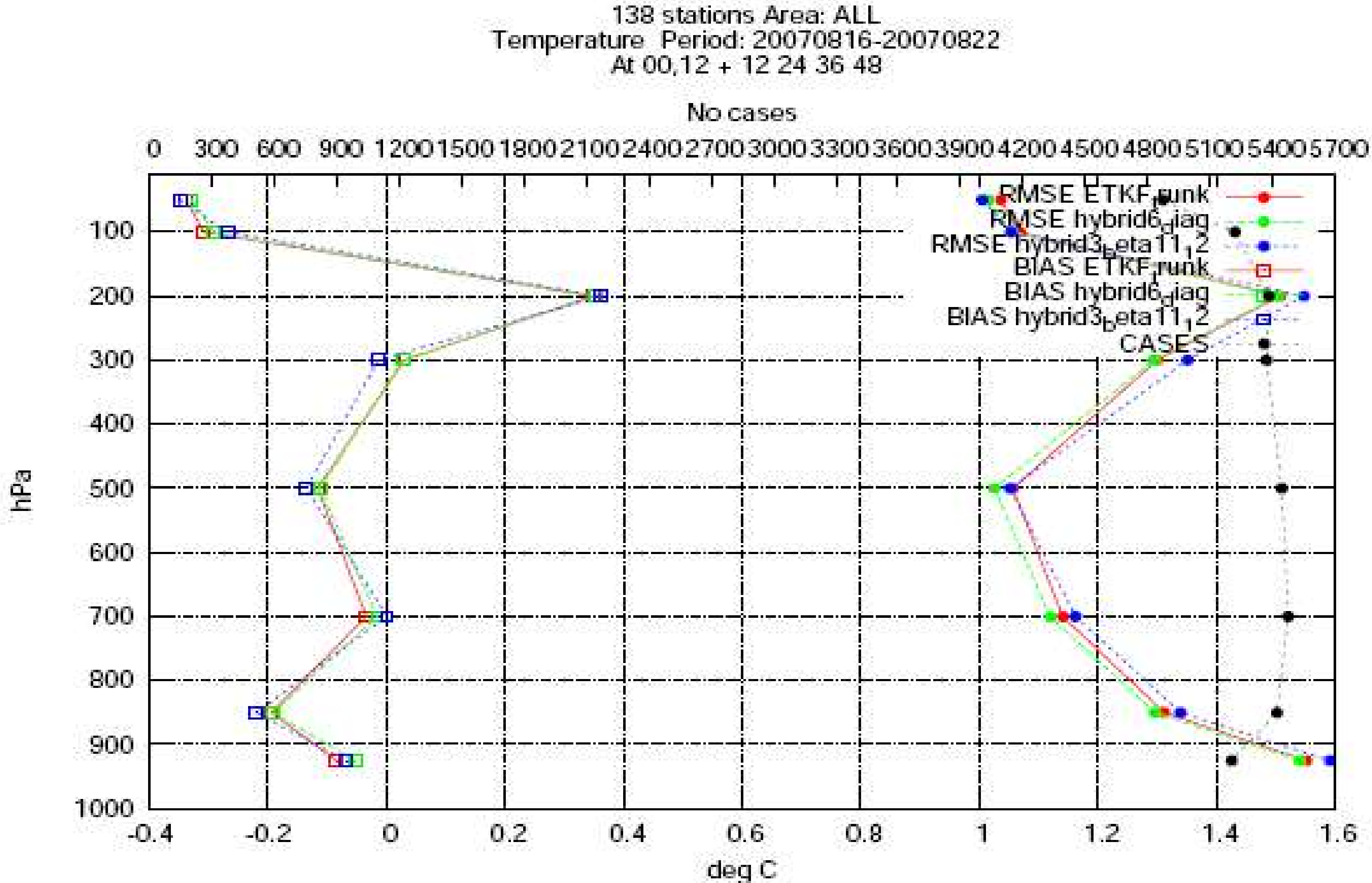


$1/\beta_{3D\text{-}VAR} = 1/2$



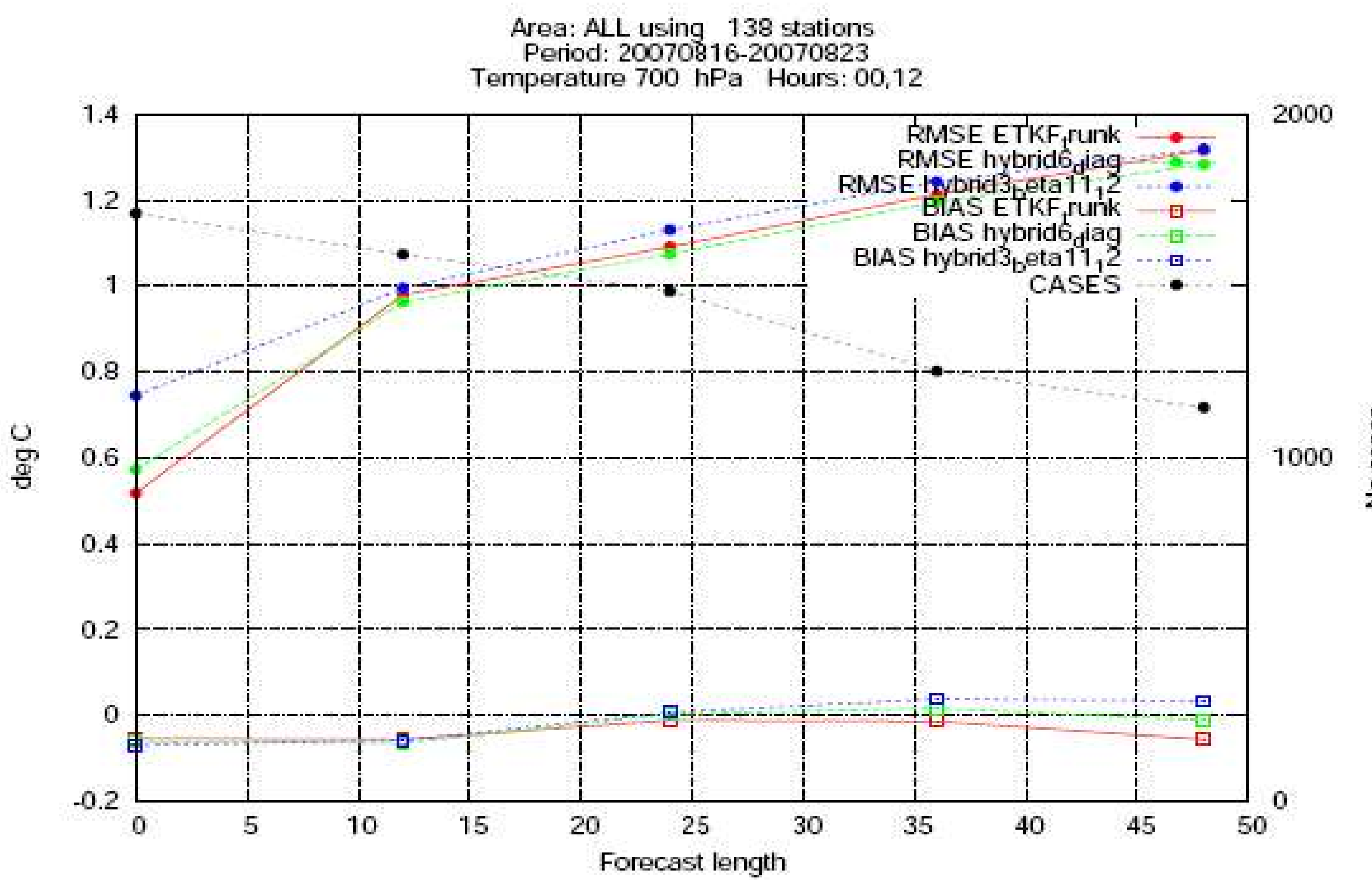
Verification of temperature profiles;

$1/\beta_{3D\text{-}VAR} = 1$ $1/\beta_{3D\text{-}VAR} = 1/2$ $1/\beta_{3D\text{-}VAR} = 1/11$



Verification of 700 hPa temperature

$1/\beta_{3D\text{-}VAR} = 1$ $1/\beta_{3D\text{-}VAR} = 1/2$ $1/\beta_{3D\text{-}VAR} = 1/11$



Thank You

for

Attention !..