Meteorologisk institutt
met.no

High Resolution Hirlam Limited Area Model

The hybrid ensemble variational data assimilation system in HIRLAM

(sensitivity to the ensemble generation approach)

HIRLAM All Staff Meeting, Norrköping, Sweden
5-8 April 2011

Nils Gustafsson and Jelena Bojarova

Thanks to Kai Sattler, Åke Johansson and Ole Vignes

Approximate solution to the sequential update problem

non-linear state space model

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

tangent-linear approximation

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

extended Kalman filter solution

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

3D-Variational solution

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

Ensemble Kalman filter

$$
\begin{aligned}
B^f_\tau &= Z^f_\tau (Z^f_\tau)^T \\
B_\tau &= Z^a_\tau (Z^a_\tau)^T \\
Z^a_\tau &= Z^f_\tau T_\tau
\end{aligned}
$$

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

Different approaches for the generation of ensembles of analyses perturbations

dynamics

The EuroTEPS perturbations (based on global singular vector perturbations targeted to Europe with 24h optimization time; both the initial and the evolved singular vectors are used to sample EuroTEPS)

dynamics+observation network

The ETKF rescaling perturbations (LAM perturbations with EuroTEPS boundaries based on the Generalized Breeding technique; both the dynamical instabilities and the density and the quality of the observation network are used to construct transformation; multiplicative and additive inflation are employed)

observation network

The EnsDA perturbations (LAM perturbations with EuroTEPS boundaries; the ensemble of data assimilation runs with stochastically perturbed observations; are not tuned well enough yet)

HIRLAM approach to use ensembles in 3D-Var

- ***Sample (or Construct)*** perturbations which reflect stuctures of the analysis error (*EuroTEPS*, *ETKF* or *EnsDA*)
- ***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:

spectral 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

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

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

grid-point space

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

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.

Extraction of the flow-dependent information
from the ensemble of forecast perturbations

HIRLAM approach

raw ensemble forecast-error variances +
Schur filtering of raw ensemble forecast error correlations in grid-point space +
mix with the climatological forecast error covariances

Meteo-France approach

filtering of ensemble forecast error variances and possibly covariances by spatial averaging (in spectral and/or wavelets space)

with the precision of Loik Berre!

What makes the hybrid ensemble variational approach attractive

- Ensemble of forecast error perturbations contains flow-dependent structures;
- Ensemble of the forecast error perturbations contains an-isotropic structures (both large- and small-scales are represented);
- Ensemble of the forecast error perturbations reflects relationship between large-scale forcing and meso-scale developments.

Flow situation – 21 August 2007

Data assimilation experiment (12.08.2007-24.08.2007)

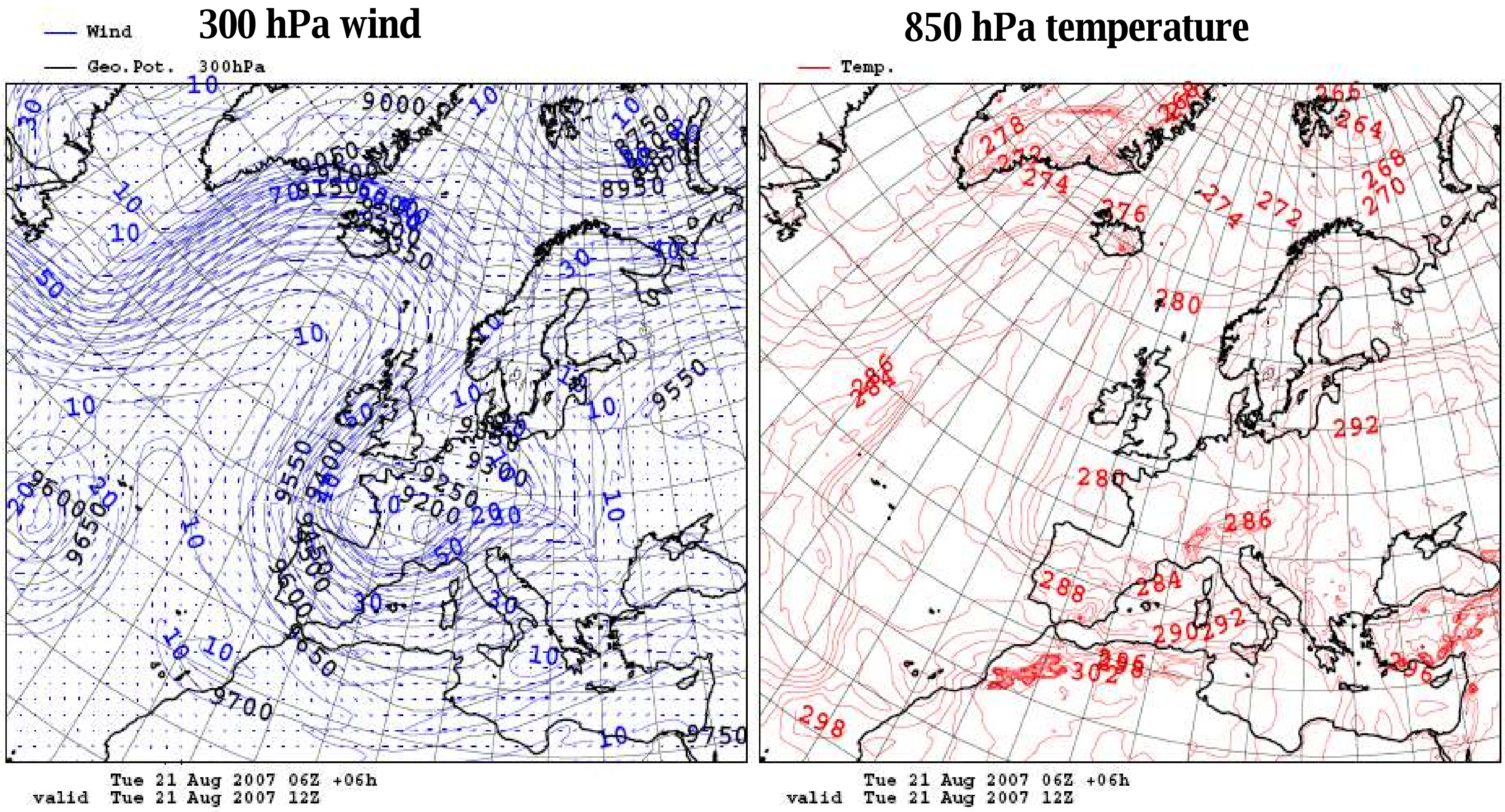
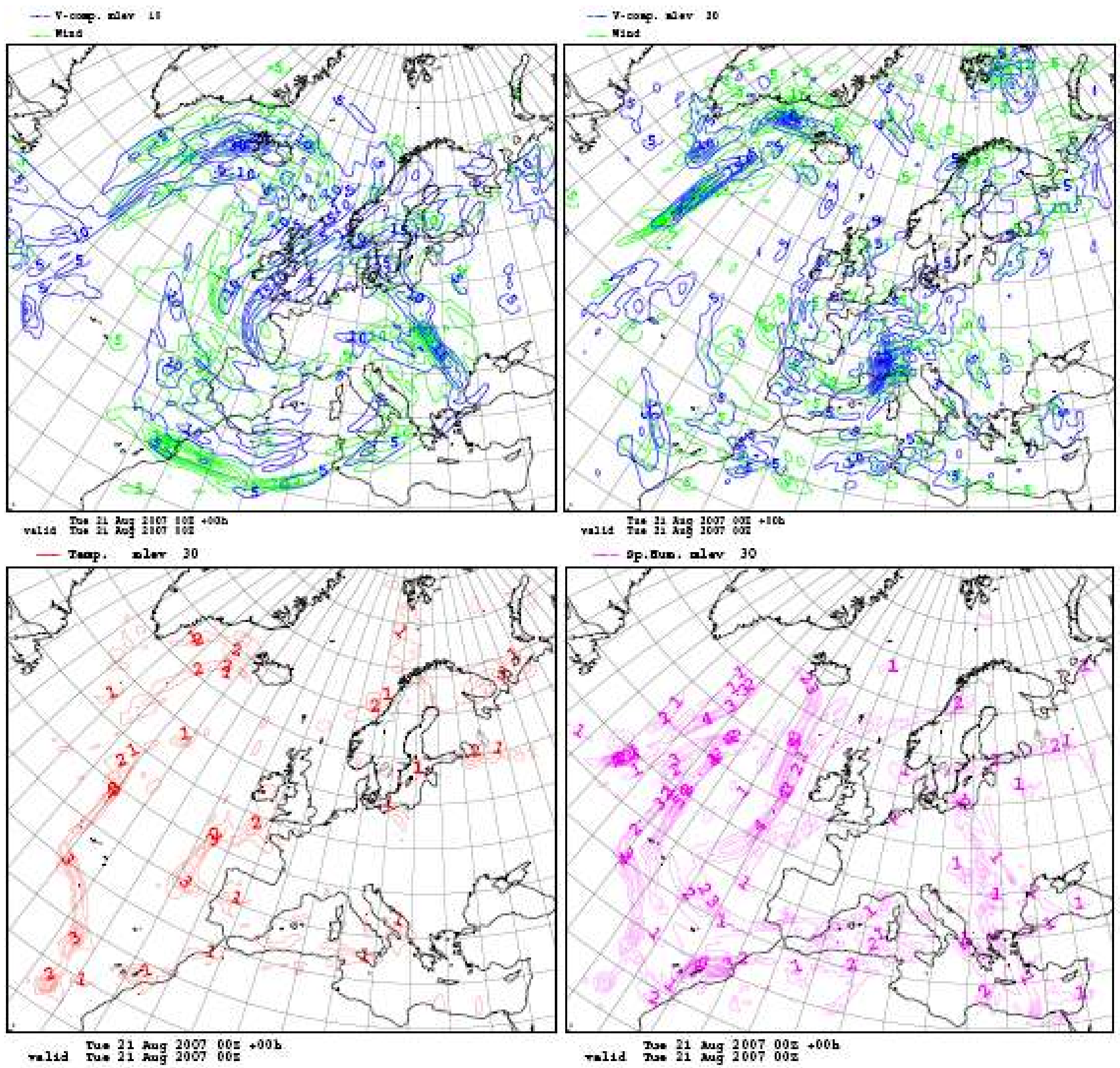


Figure 4. 300 hPa wind and geopotential (left) and 850 hPa temperature (right) taken from the background model state at 21 August 2007 06UTC + 6h; experiment based on equal static and ensemble contributions to the background error variance

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

Wind components model level 15

Wind components model level 20



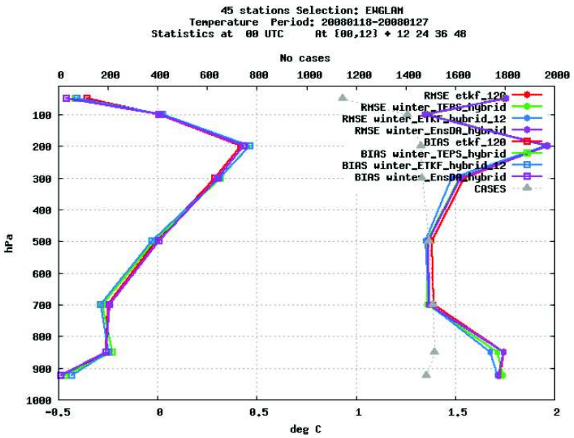
Temperature model level 30

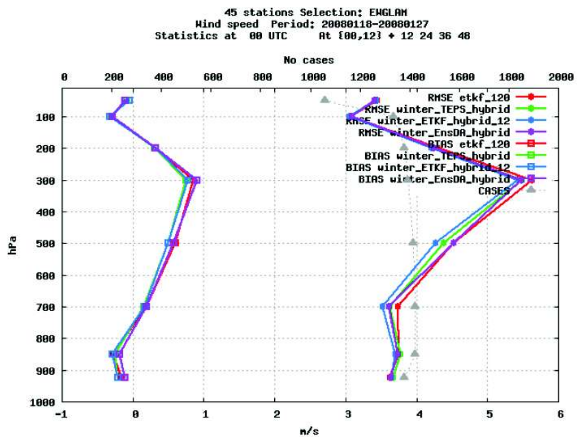
Specific humidity model level 30

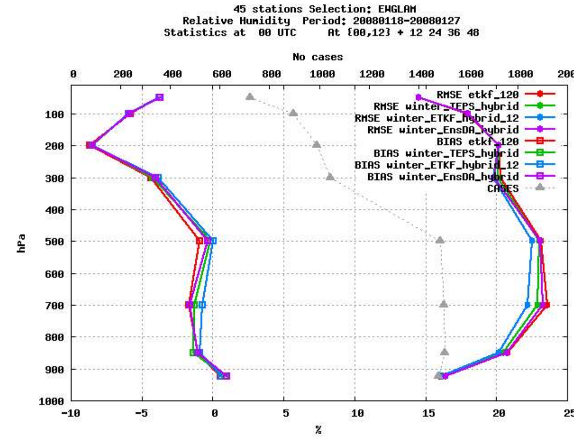
Figure 5. Examples of estimated background error variances based on the ensemble of +6h forecast valid at 21 August 2007 12UTC from the hybrid data assimilation experiment hybrid6_diag. Wind components at model level 10 (upper left), wind components at model level 20 (upper right), temperature at model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

Verification scores : “winter case” 17Jan2008-27Jan2008

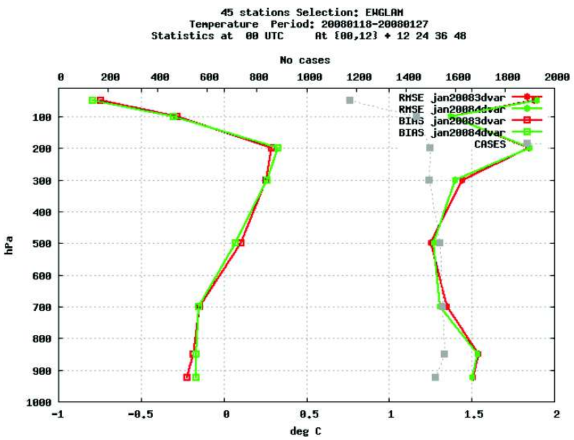
3D-Var versus hybrid approach (*ETKF*, *EnsDA*, *TEPS*)

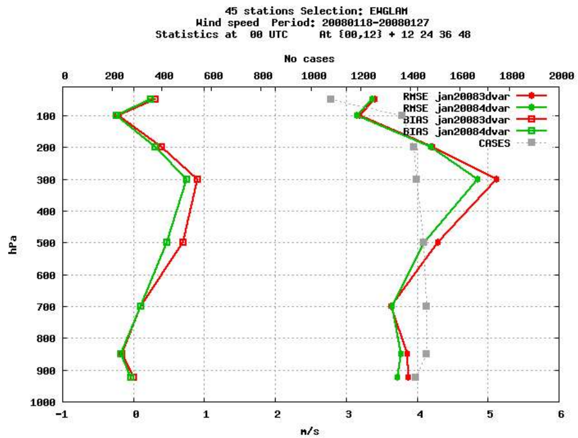


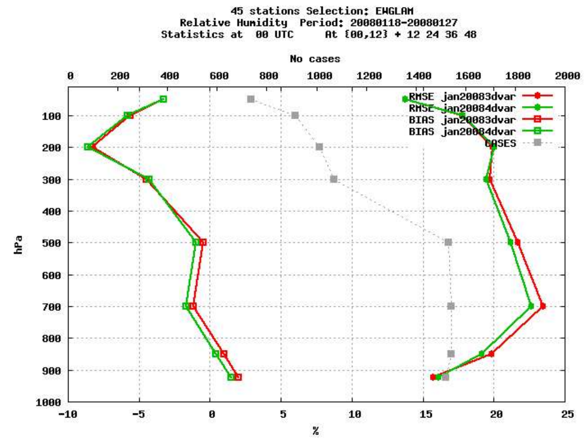




3D-Var versus *4DVAR* approach



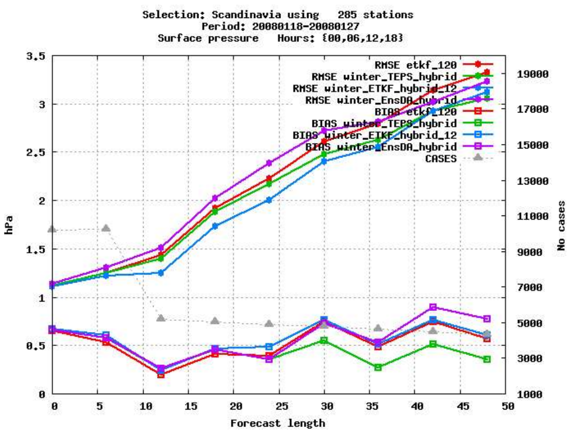




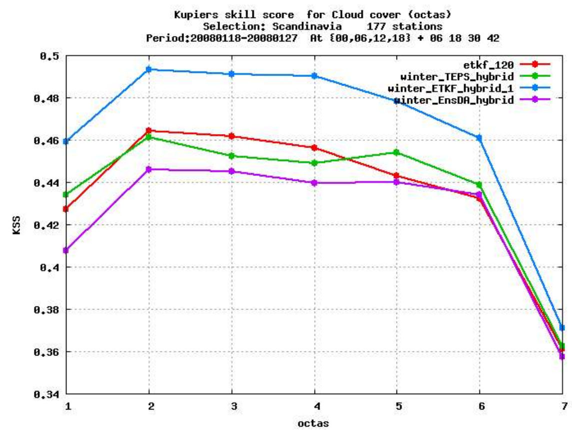
Temp. Wind speed Rel. humid.

Verification scores : “winter case” 17Jan2008-27Jan2008

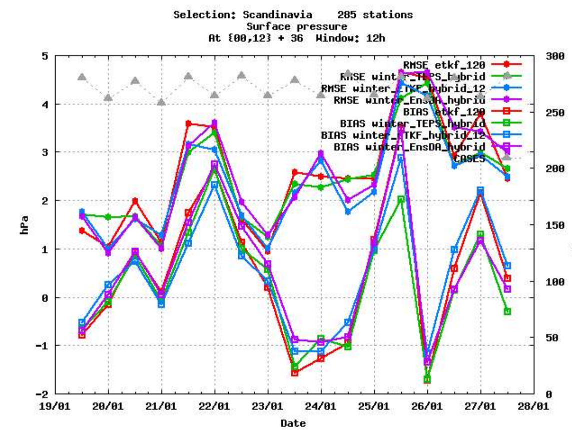
3D-Var versus hybrid approach (*ETKF*, *EnsDA*, *TEPS*)



Psml



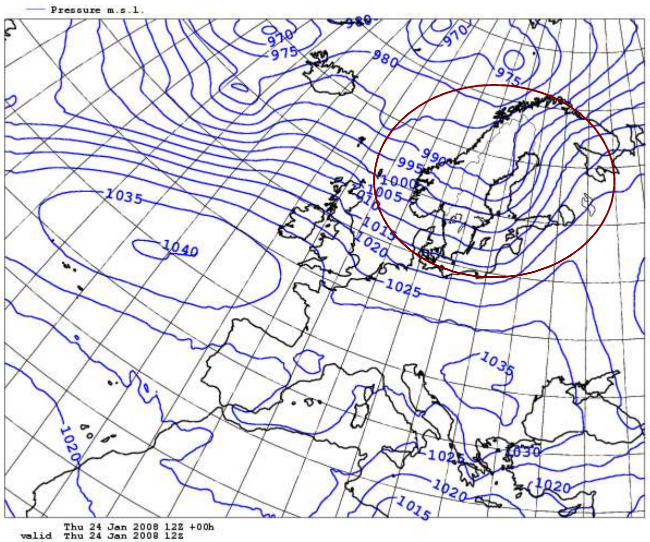
Cloud. Kuiper score

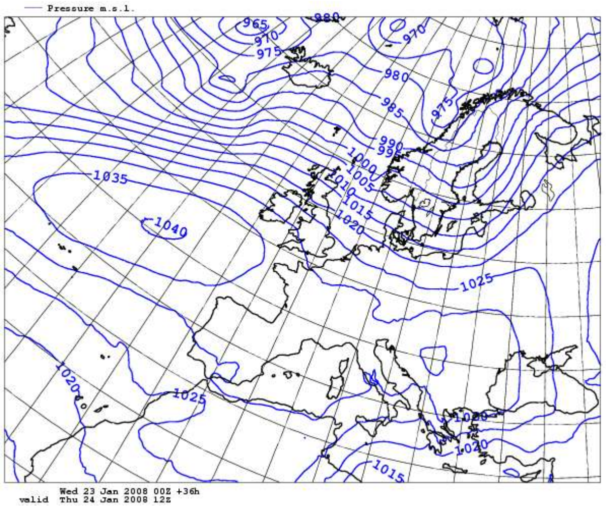


Surface pressure TS

3D-Variational data assimilation runs in non-optimal setup in these experiments (severe cold bias of SST). Experiments should be redone with optimal settings!!!

24 Jan 12 UTC

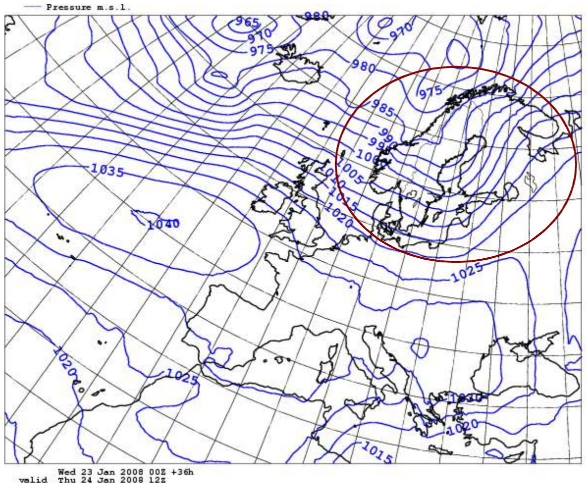




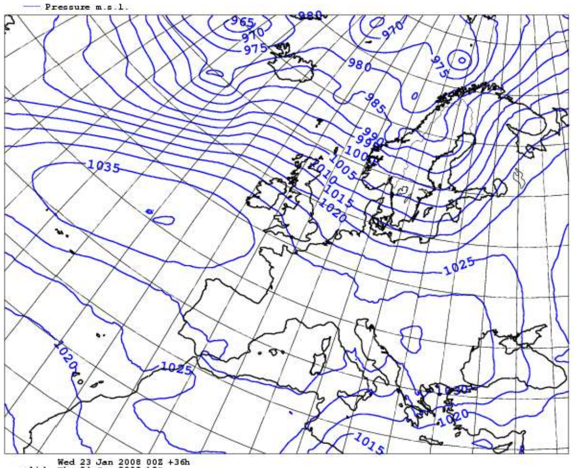
3DVAR

3d-Var analysis

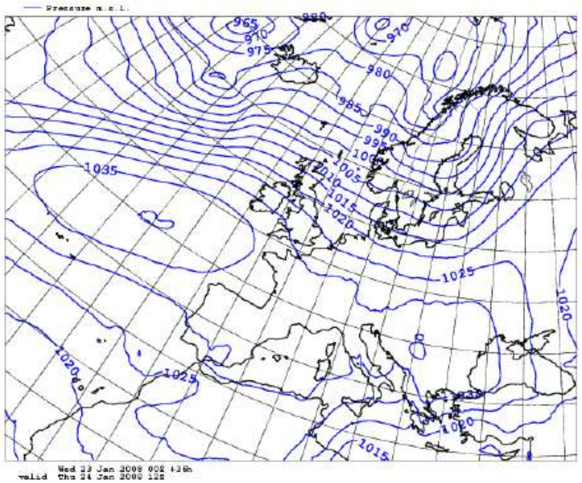
versus +36 h forecasts



3DVAR+ETKF



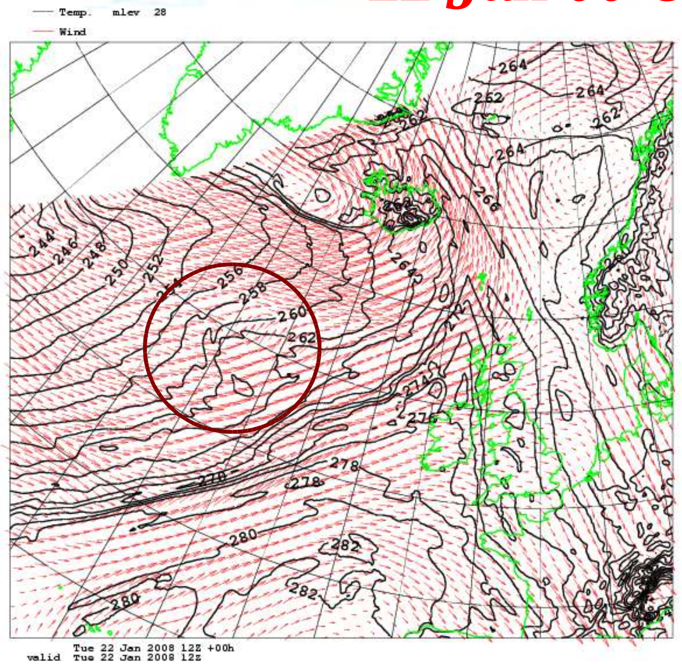
3DVAR+EnsDA



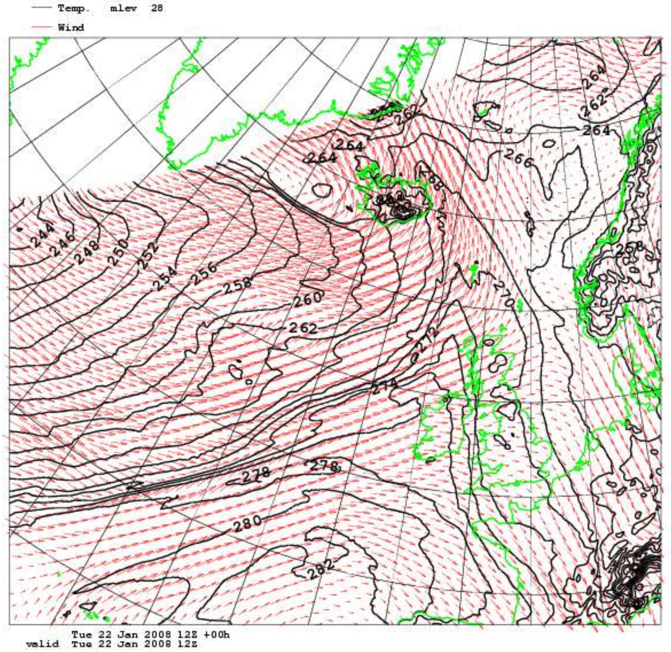
3DVAR+TEPS

22 Jan 00 UTC (front), mlev 28

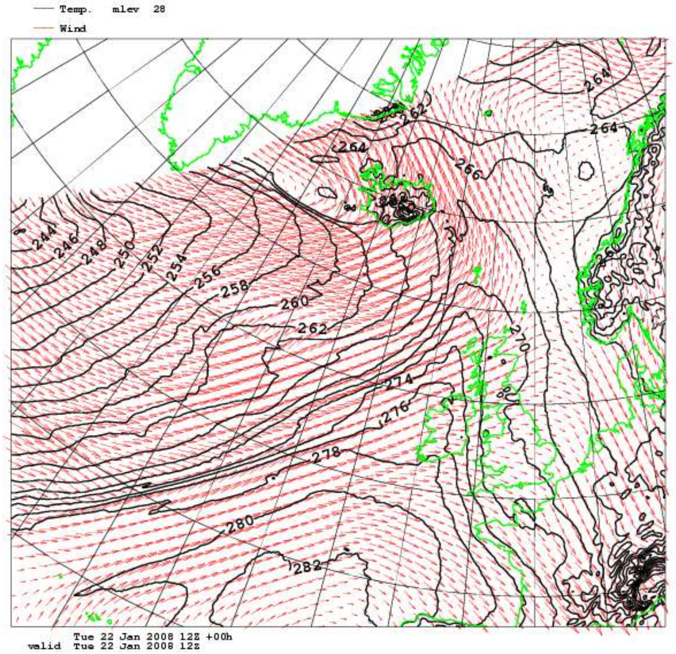
3DVAR



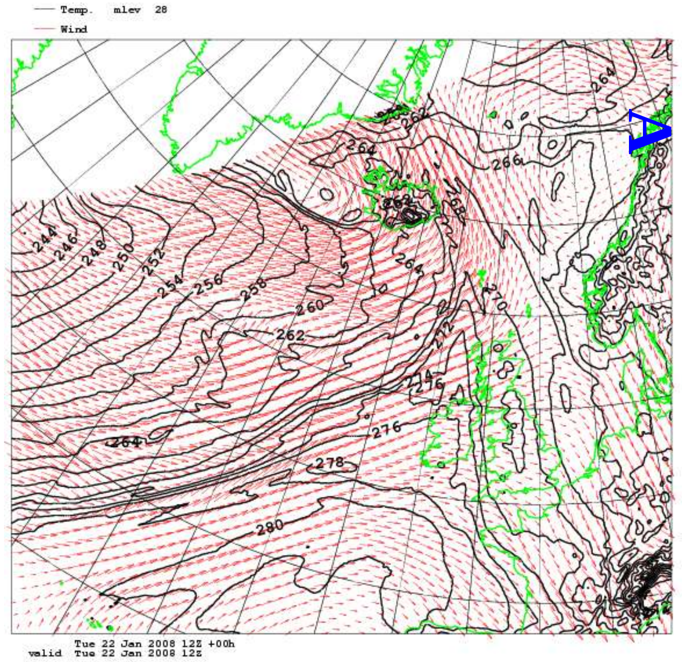
3DVAR + TEPS



3DVAR + ETKF

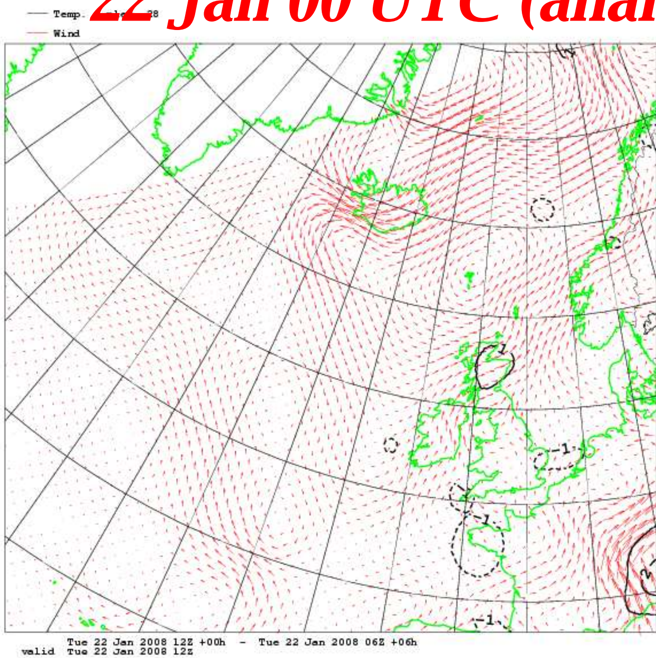


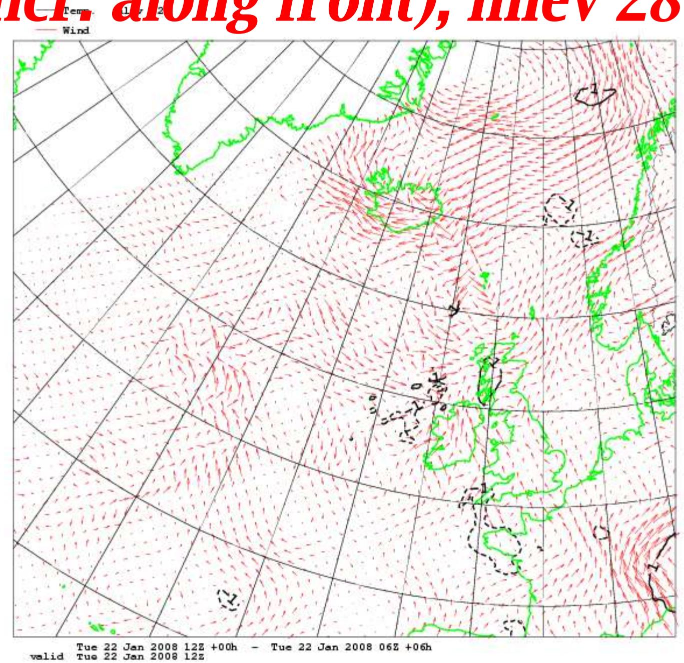
3DVAR+ EnsDA



22 Jan 00 UTC (analysis incr. along front), mlev 28

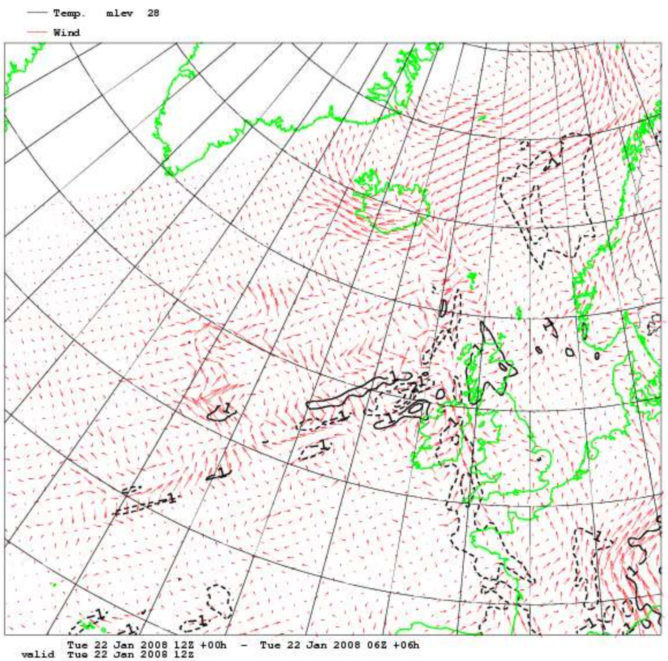
3DVAR

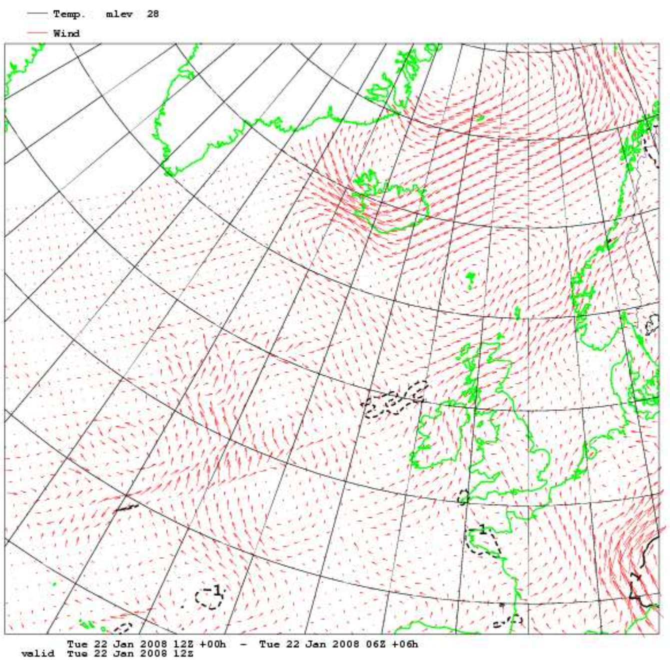




3DVAR+TEPS

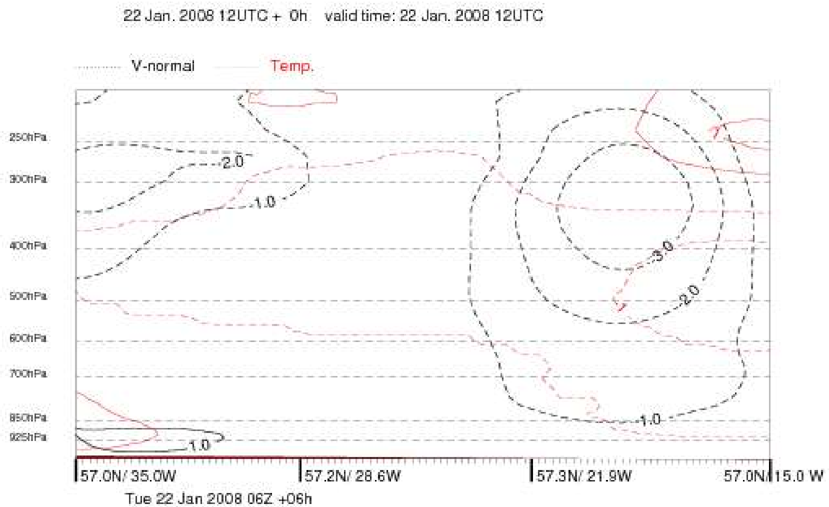
3DVAR+ETKF



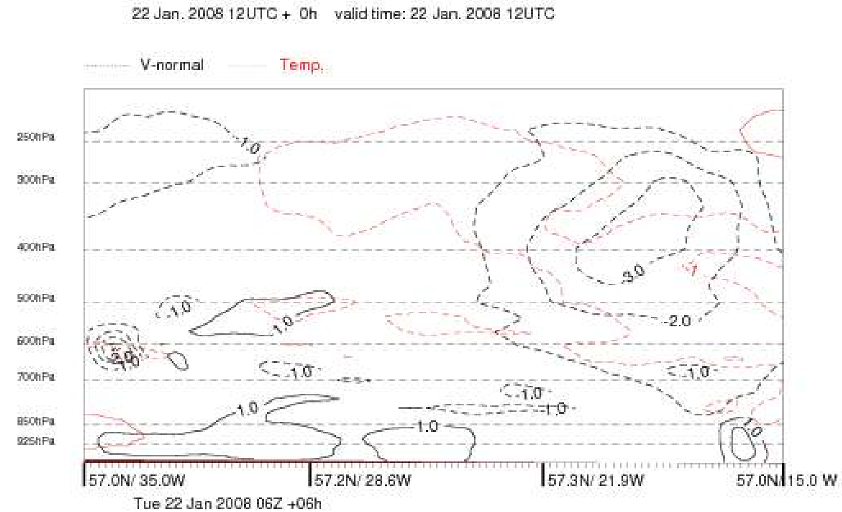


3DVAR+EnsDA

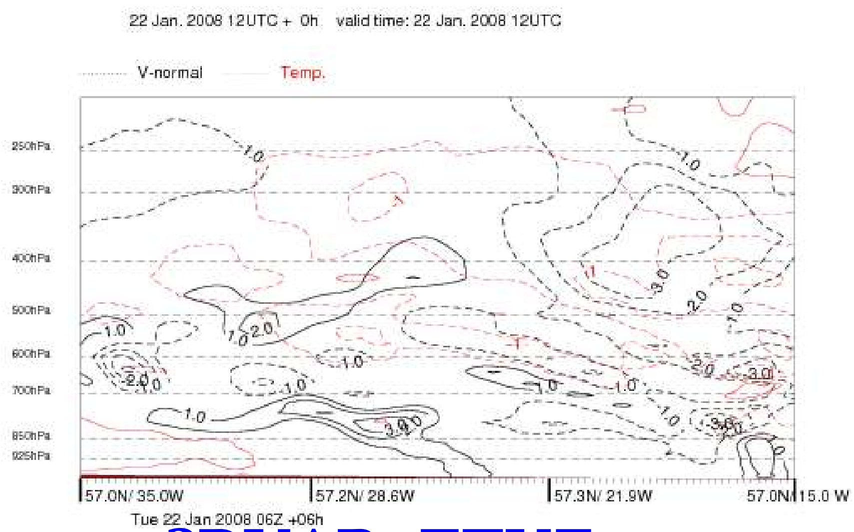
22 Jan 00 UTC (cross-section of across front of analysis incr.)



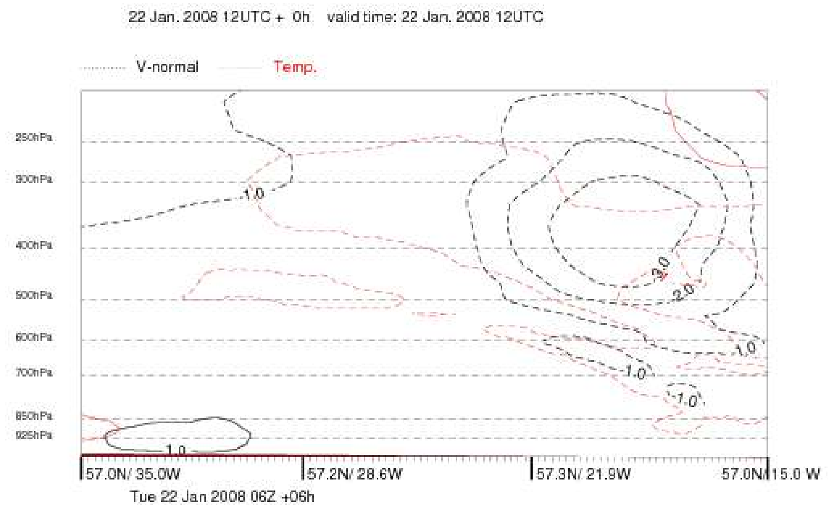
3DVAR



3DVAR+TEPS

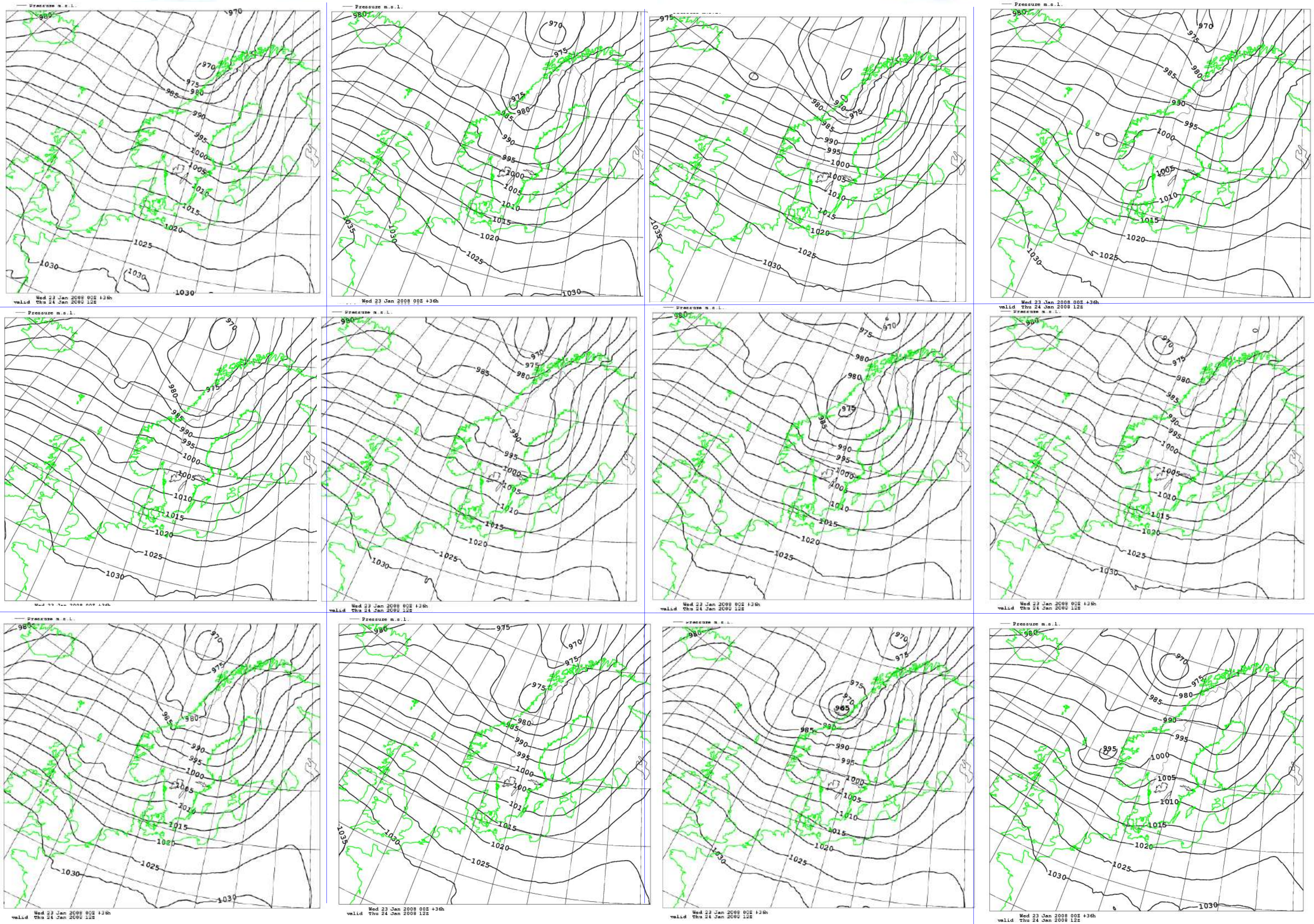


3DVAR+ETKF

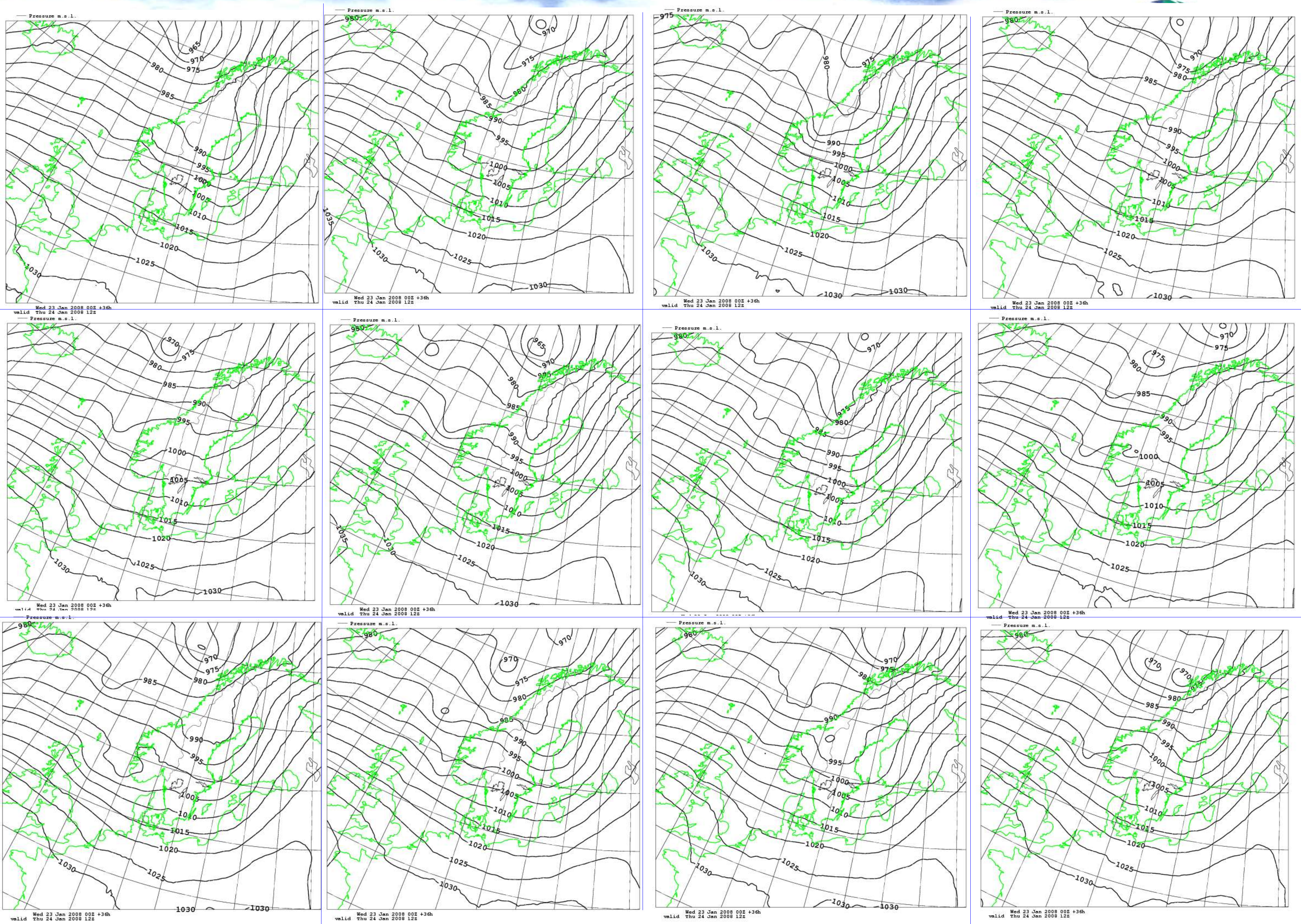


3DVAR+EnsDA

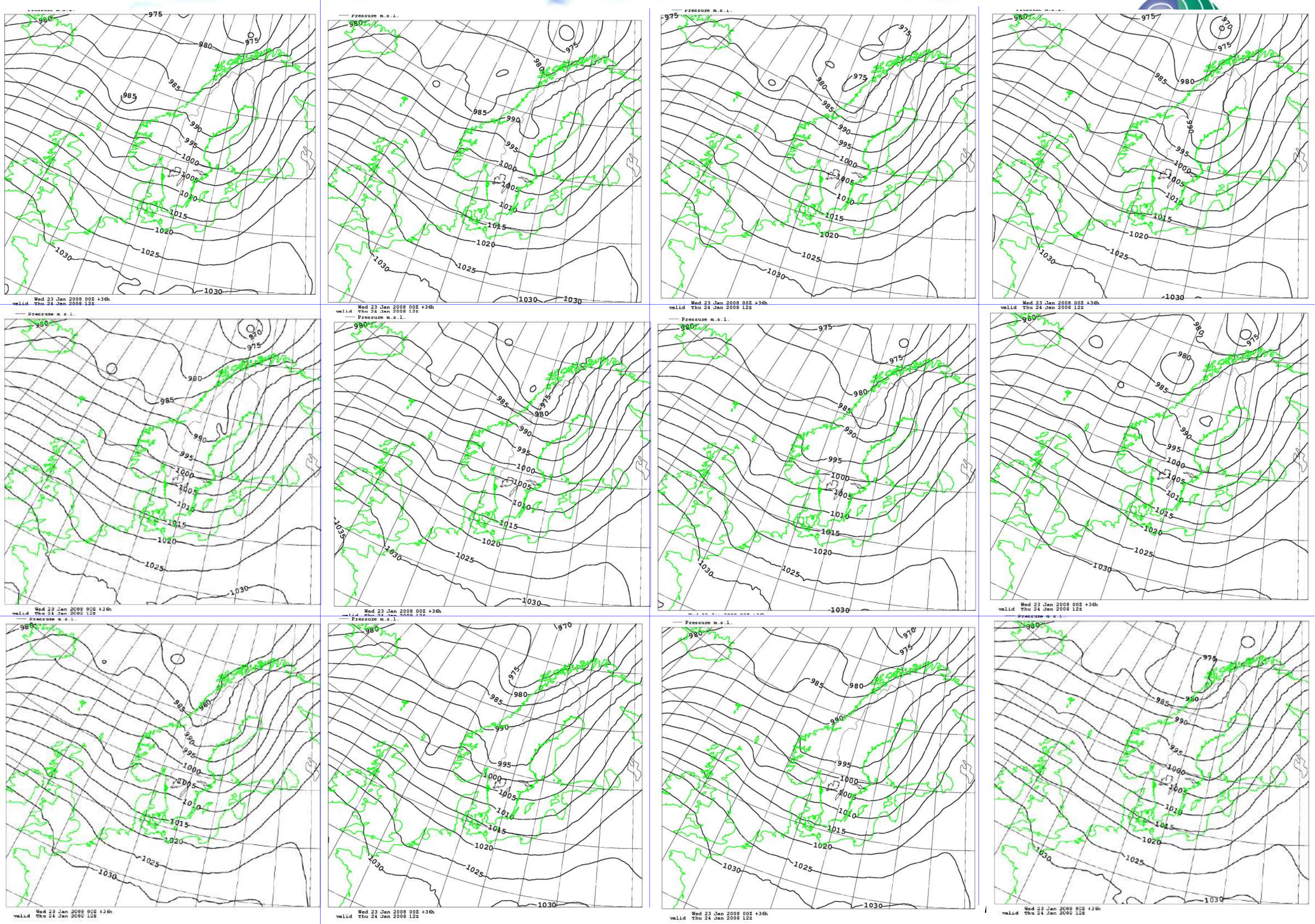
+36 h EuroTEPS valid at 24 Jan 12 UTC



+36 h ETKF EPS valid at 24Jan 12 UTC



+36h EnsDA EPS valid at 24 Jan 12 UTC



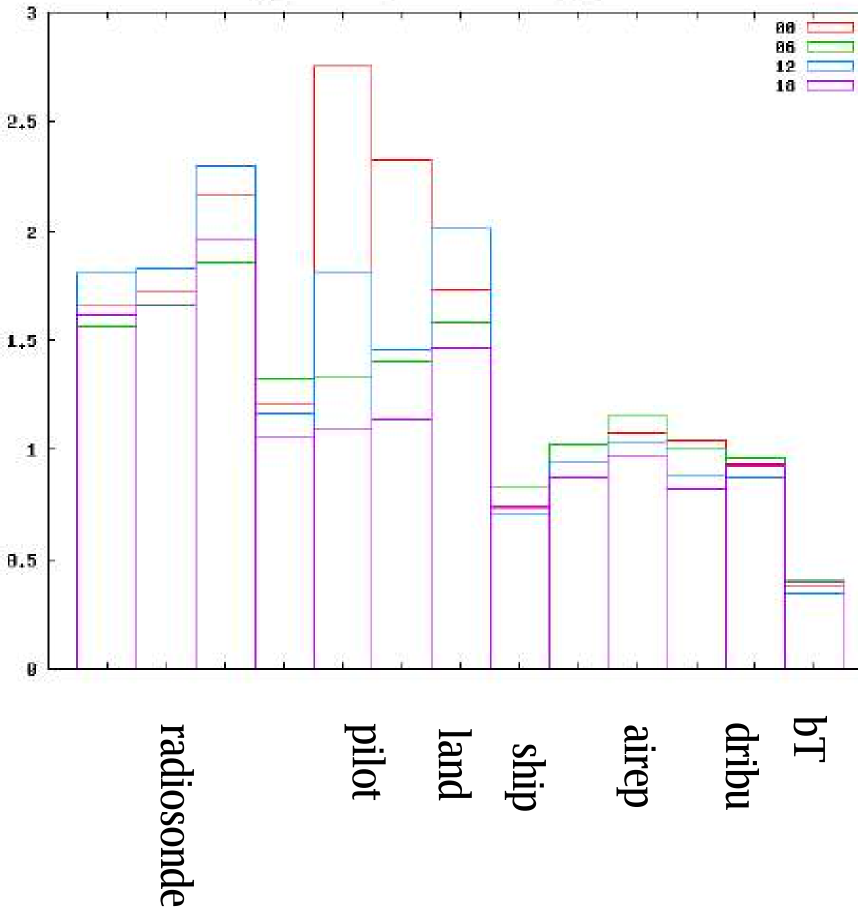
Conclusions

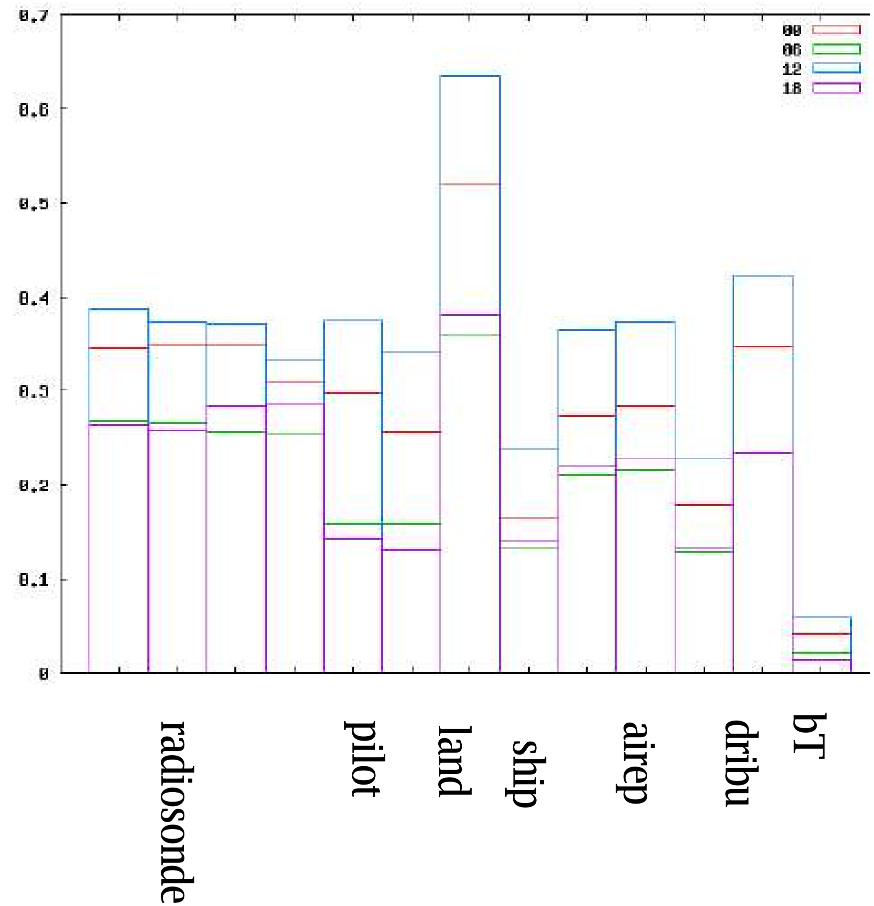
- The flow-dependent structures reflected in ensemble based forecast error covariances improves 3D-Var performance.
- All three ensemble generation strategies (EuroTEPS, ETKF and EnsDA) are useful for this purpose.
- Improvement of the standard verification scores for the hybrid ensemble variational data assimilation scheme over pure 3D-Var is of the same magnitude as of the 4D-Var over the 3D-Var.
- Further tuning is needed to improve perforance of EuroTEPS and in particular EnsDA based hybrid scheme.
- The impact of the increased ensemble size on the EPS performance and the usefulness of the hybrid4 D-Var ensemble data assimilation sc heme will be investigated.

Relative squared innovation variance versus relative spread

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

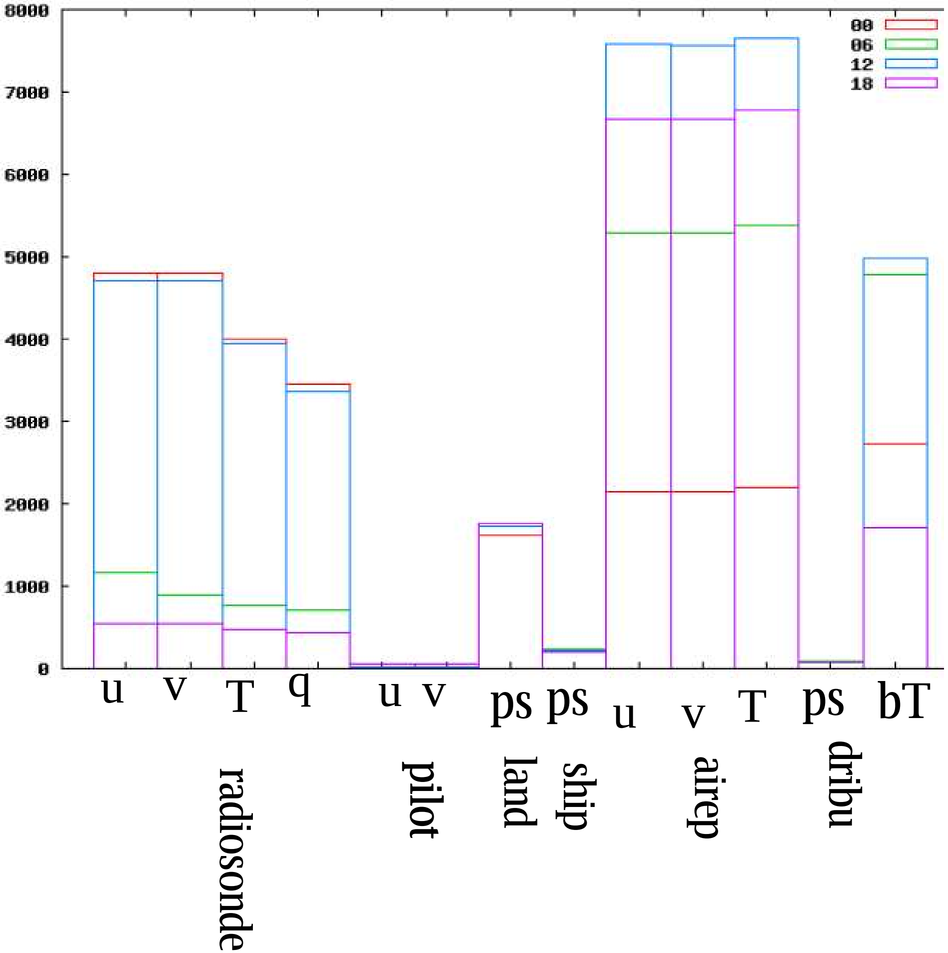
$$trace(B^f_{ens}) = p\frac{1}{p}\sum_{j=1}^{p} \frac{Var(\mathcal{H}_j(\delta X^f))}{\sigma_{o,j}^2}$$



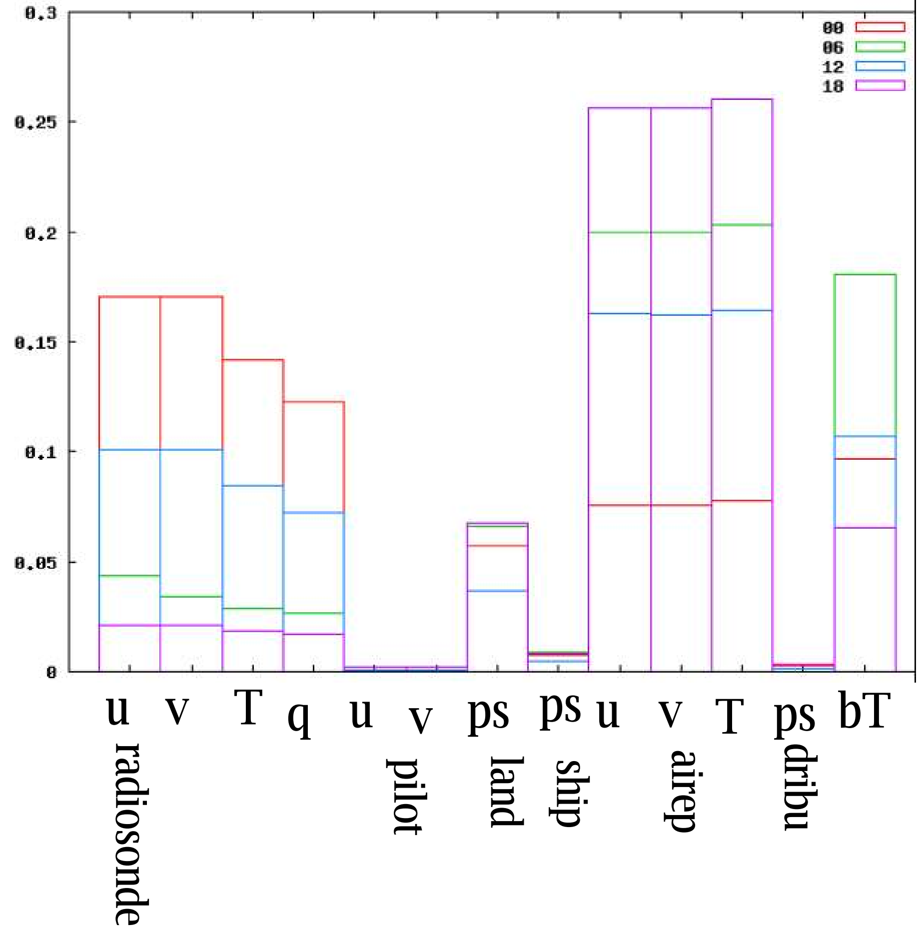


Earth observing system

absolute amount of observations



relative amount of observations



Thank You

for

Attention !..