Use of analysis ensembles in estimating flow-dependent background error variances

Lars Isaksen

ECMWF

Acknowledgements: Mike Fisher, Jean-Noël Thépaut ,

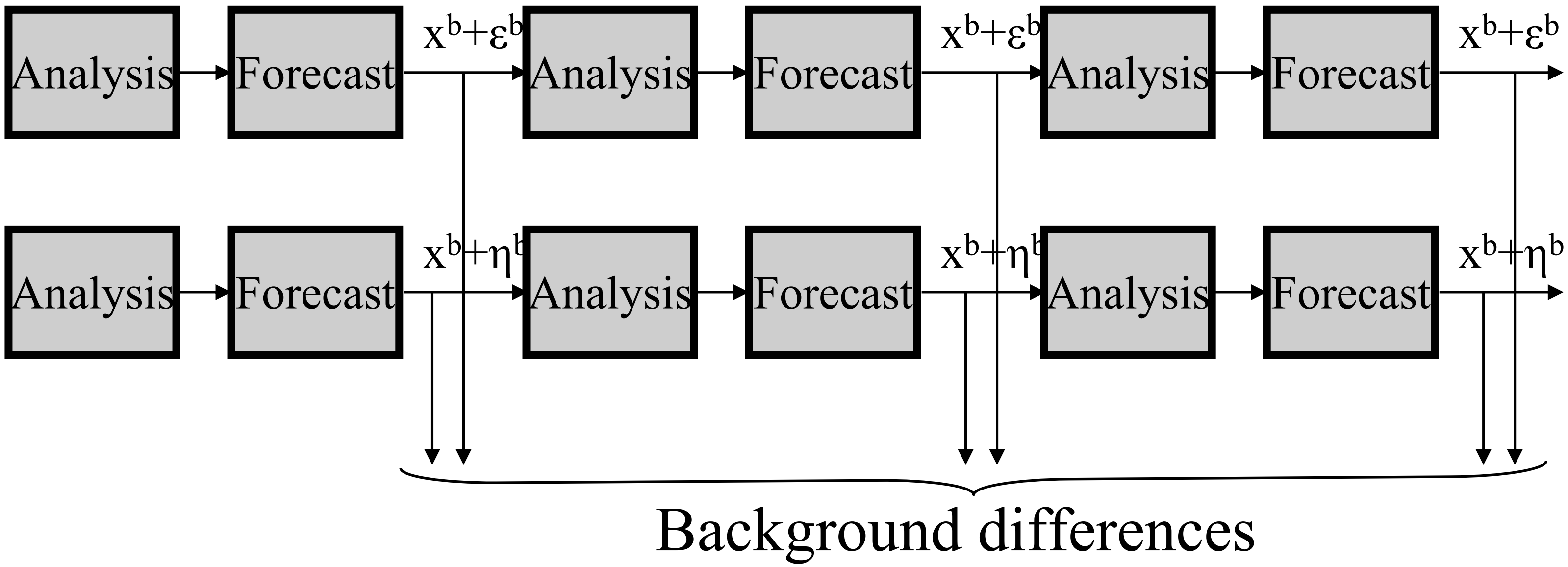
Judith Berner, Dick Dee and Edit Hagel

Content of talk

- **Ensemble Data Assimilation at ECMWF**
  - **Present status of ensemble DA at ECMWF**
  - **Estimating analysis error and short range forecast error**
  - **How can be increase the ensemble spread to become more realistic**
  - **Investigate impact of flow dependent Background Error Covariances – can we benefit from "Errors of The Day" modifications of variances?**
  - **Why is the impact of the present flow dependent background errors so small on the general forecast scores**

Estimating Flow dependent Background Error from an ensemble DA method

- **Run an ensemble of analyses with random observation and SST perturbations, and form differences between pairs of background fields.**
- **These differences will have the statistical characteristics of background error (but twice the variance).**



Estimating analysis error from an ensemble

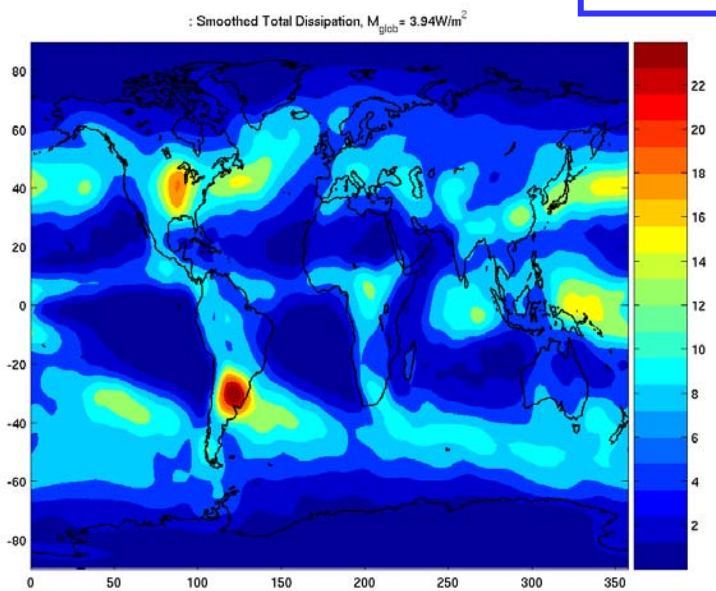
- **We used a 10 member ensembles of 4D-Var assimilations**
- **Additional low-resolution 50 member ensemble also run**
- **New spectral backscatter (SPBS) method used to represent model error (developed by Judith Berner, ECMWF)**
- **Perturbed observations and SST as explained earlier. Only feature tracking winds (SATOB) are correlated**
- **45 days of T255/T159 ensemble data assimilations performed**
- **Each day at each grid point we calculated the standard deviation of ensemble spread among the 10 members**
- **Standard deviation values averaged over 40 days to give an estimate of average analysis error**

Judith Berner's Spectral Backscatter Scheme

Rationale: A fraction of the dissipated energy is scattered upscale and acts as streamfunction forcing for the resolved-scale flow (LES, CASBS: Shutts and Palmer 2004, Shutts 2005); New: spectral pattern generator

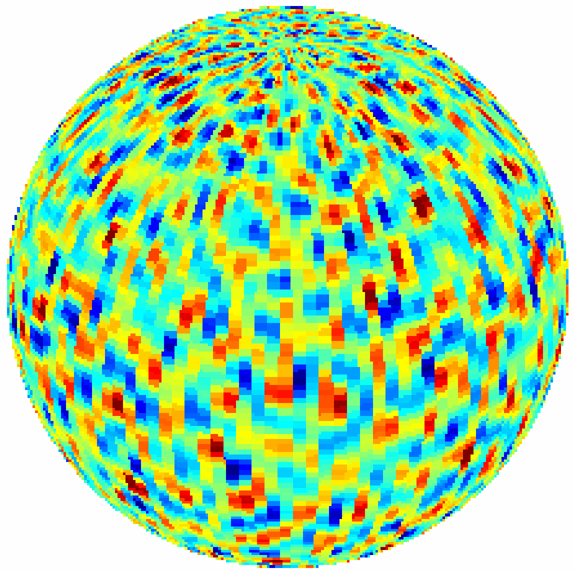
$$\boxed{\Delta \psi^* \propto \sqrt{D}\psi'}$$

$\boxed{D}$



Total Dissipation rate from numerical dissipation, convection, gravity/mountain wave drag.

$\boxed{\psi'}$



Spectral Markov chain: temporal and spatial correlations prescribed

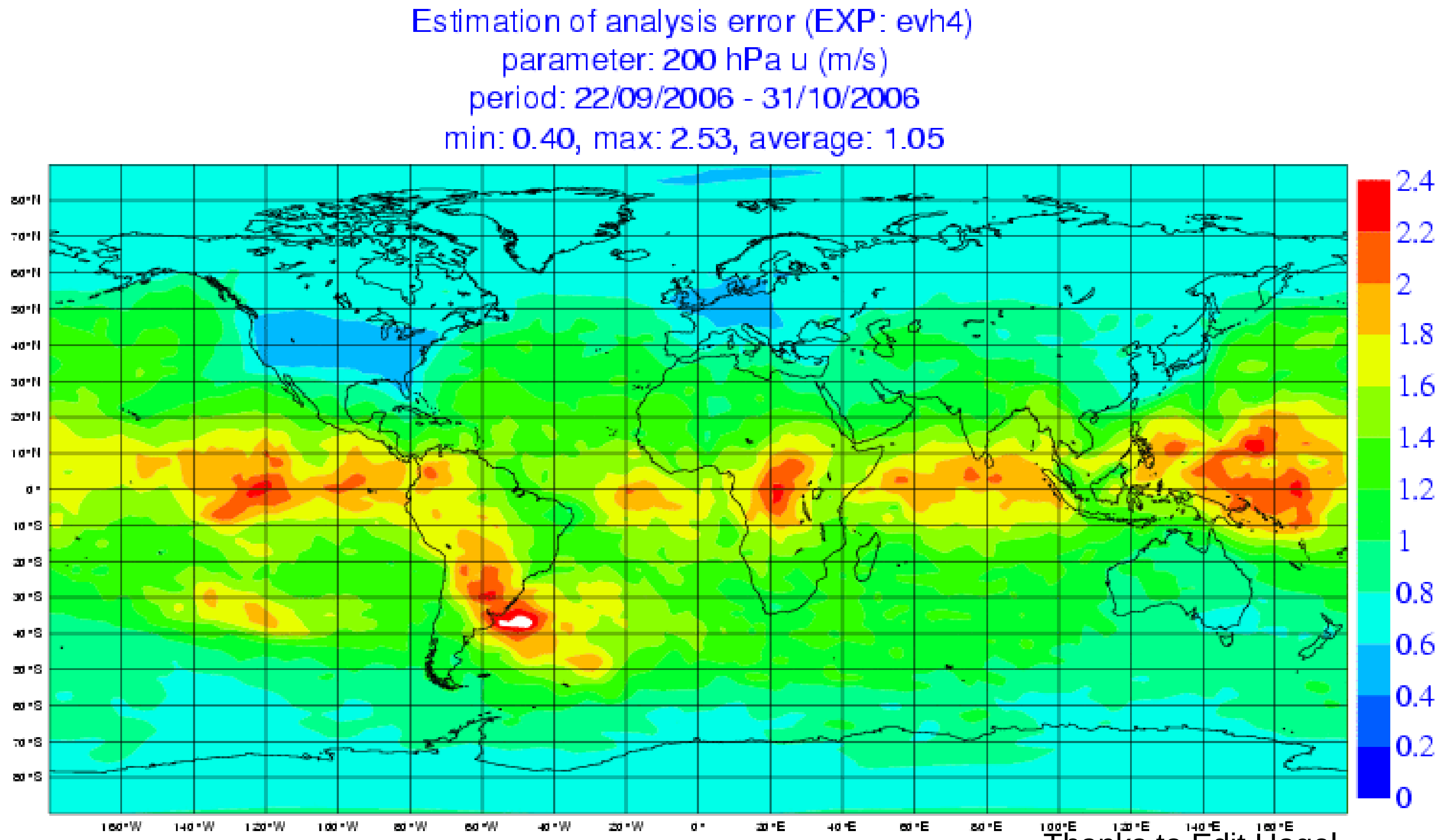
Characteristics of the Spectral Backscatter Scheme

- ***Stochastic***
- **Spatial and temporal correlations**
- **Control over wavenumber forcing allows *scale selection***
- ***Flow-dependent* (weighting with dissipation rates)**
- ***Spectral* (consistent with spectral dynamical core); isotropic pattern in sphere;**

➢ ***Injects energy in regions of large dissipation, which are the regions of large model errors***

ECMWF

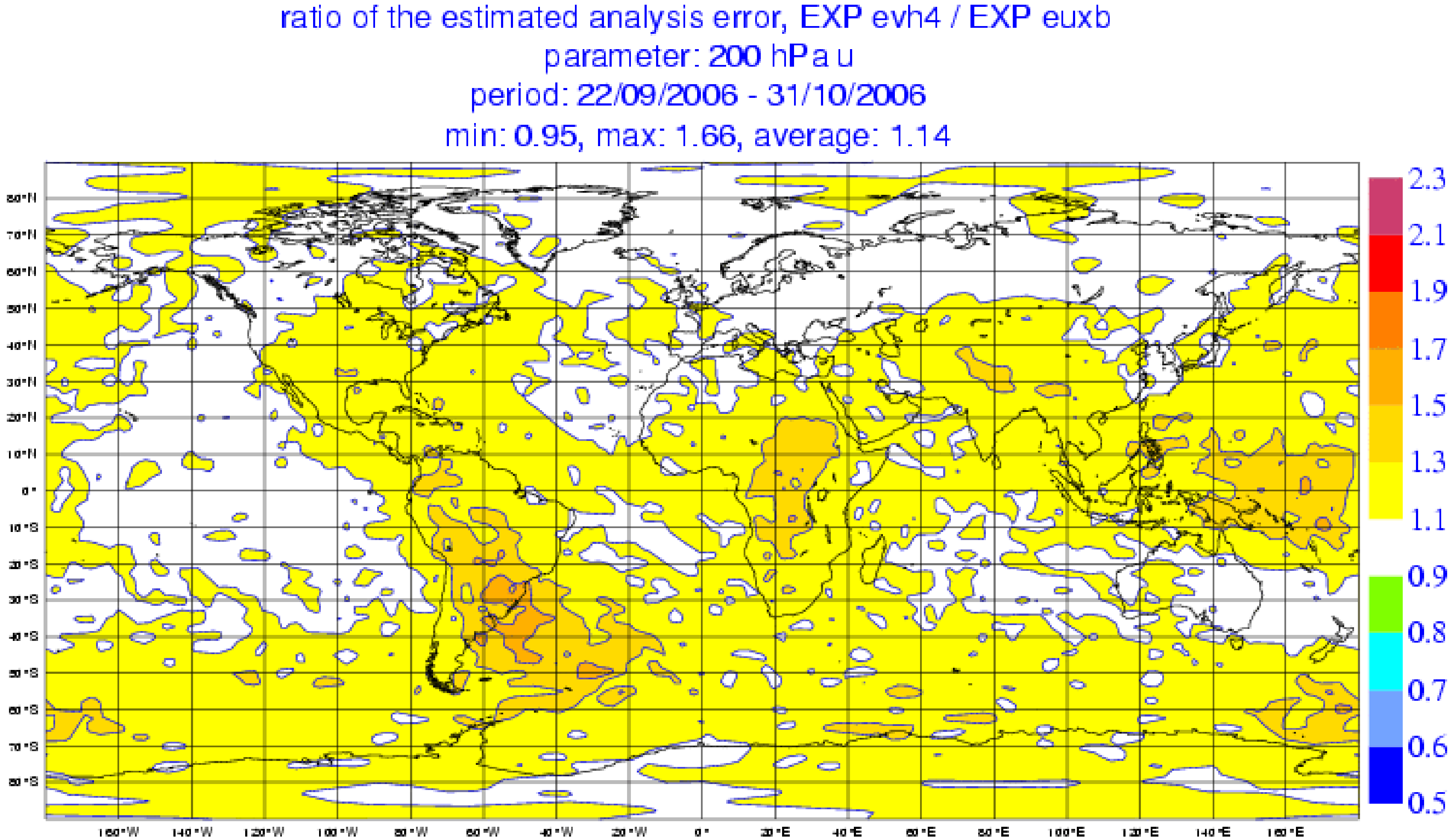
Static estimation of analysis error for 200hPa u-wind from the standard deviation of 40 days of 10 member ensemble DA with SPBS looks reasonable.
Scaling factor of 1.5-2 required due to underestimation of model errors in ensemble.



Thanks to Edit Hagel

ECMWF

Static estimation of analysis error for 200hPa u-wind from the standard deviation of 40 days of 10 member ensemble DA. Measure of the increased spread due to the Spectral Backscatter (SPBS) scheme.



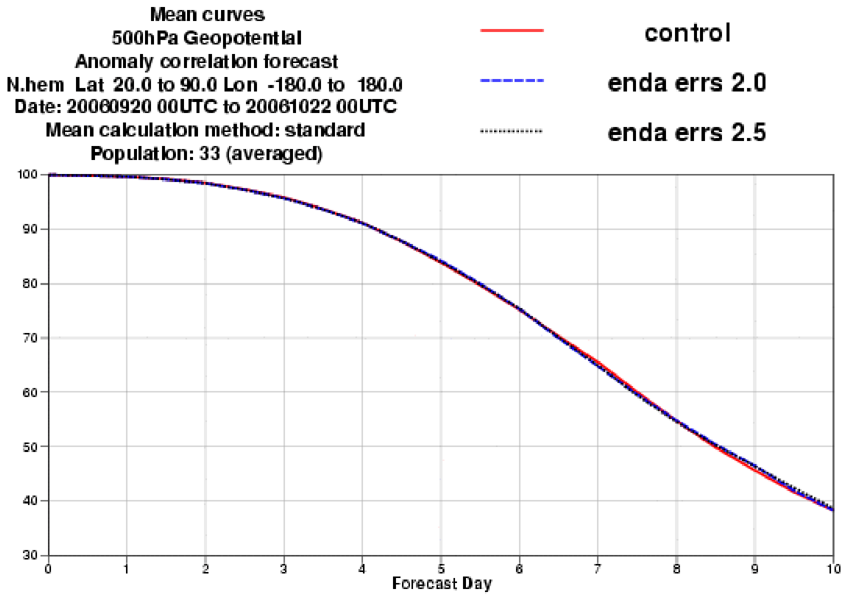
Thanks to Edit Hagel

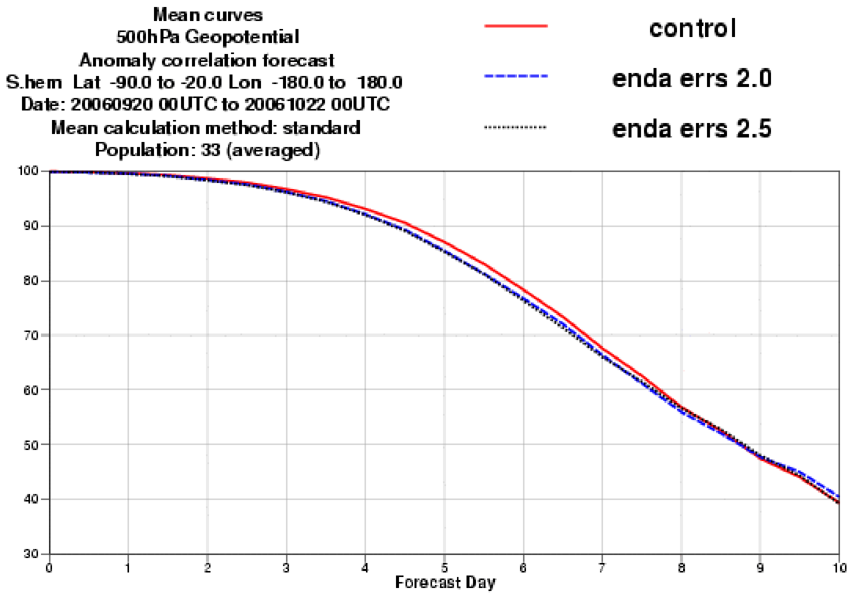
ECMWF

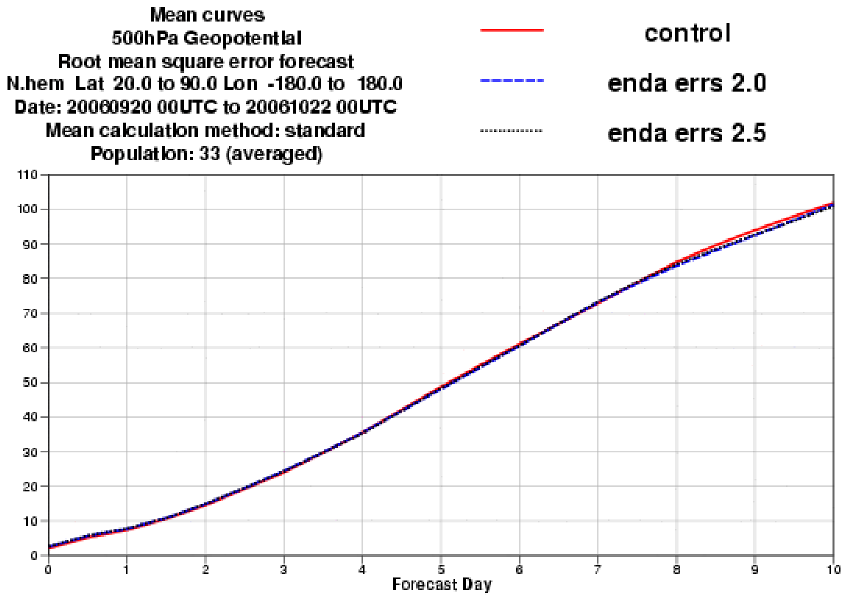
Flow dependent background error variance experimental set-up

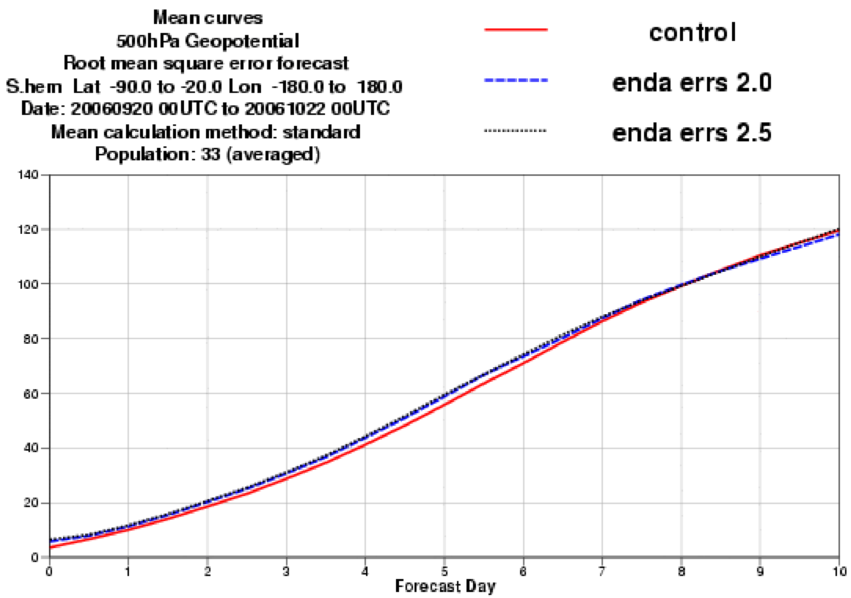
- **System used and period under investigation:**
  - **2006091500-2006103112**
  - **12h 4D-Var T319 outer loop, T95+T159+T255 inner loop**
  - **5 first days used for warm-up and excluded from statistics**
- **Three different assimilations:**
  - **4DVAR control (static Jb + randomization method applied for background error variance)**
  - **4DVAR Flow dependent background error variances. Defined as 2.0*StDev(10 member Ensemble 3h forecasts)**
  - **4DVAR Flow dependent background error variances. Defined as 2.5*StDev(10 member Ensemble 3h forecasts)**

The impact of using ENDA spread versus “static Jb” for background error variances



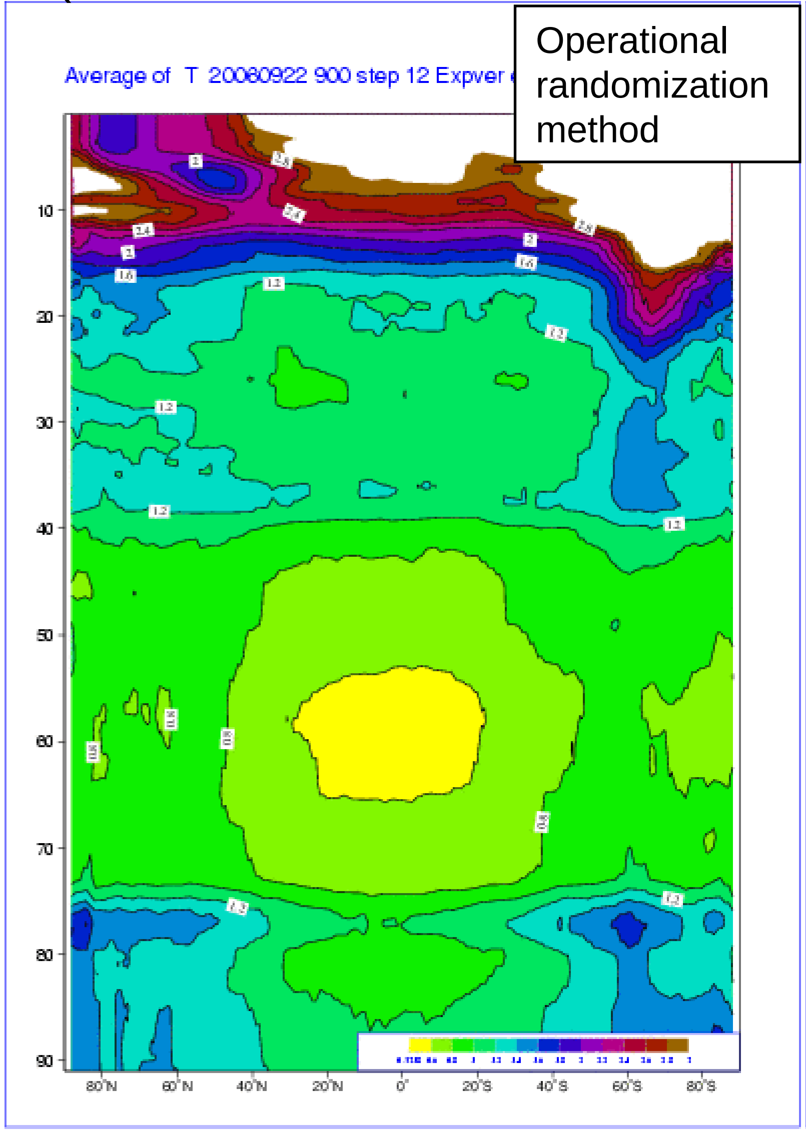




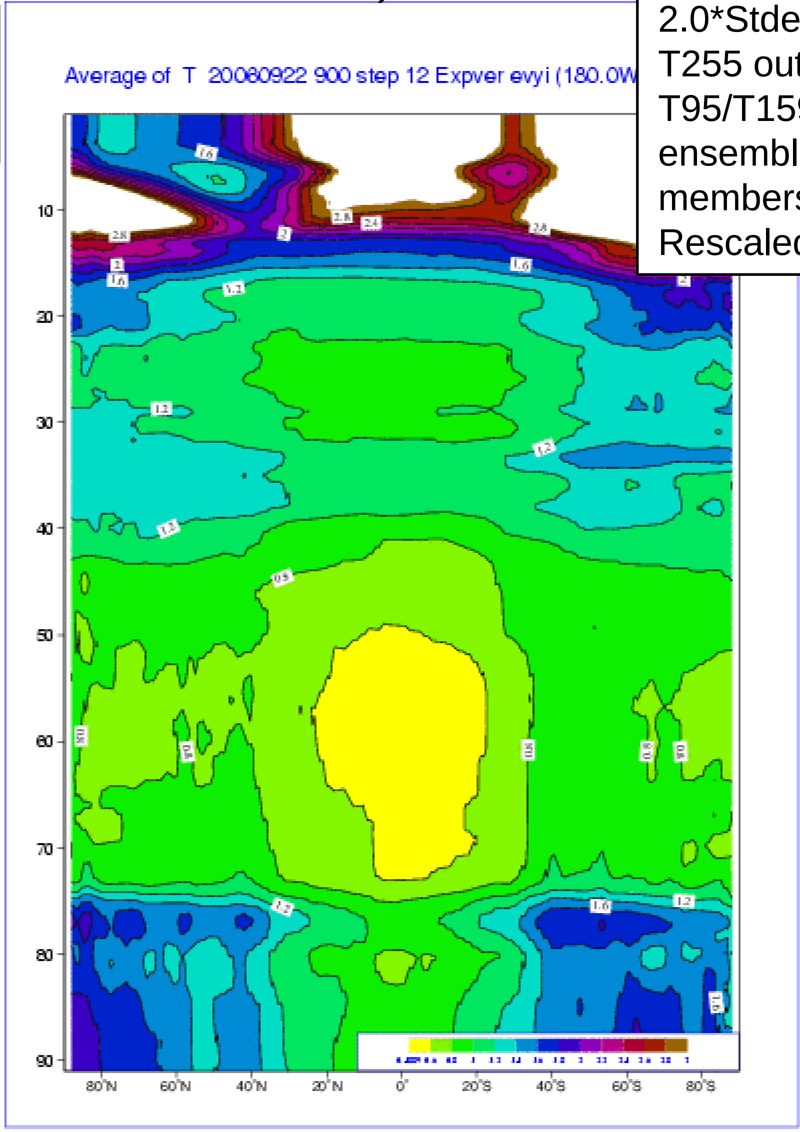


Day-to-day zonal mean background error estimates for **temperature** have fairly similar structure and amplitude, calculated from operational randomization method and from the 2.0*(standard deviation of the 10 ensemble DA members).

Operational randomization method

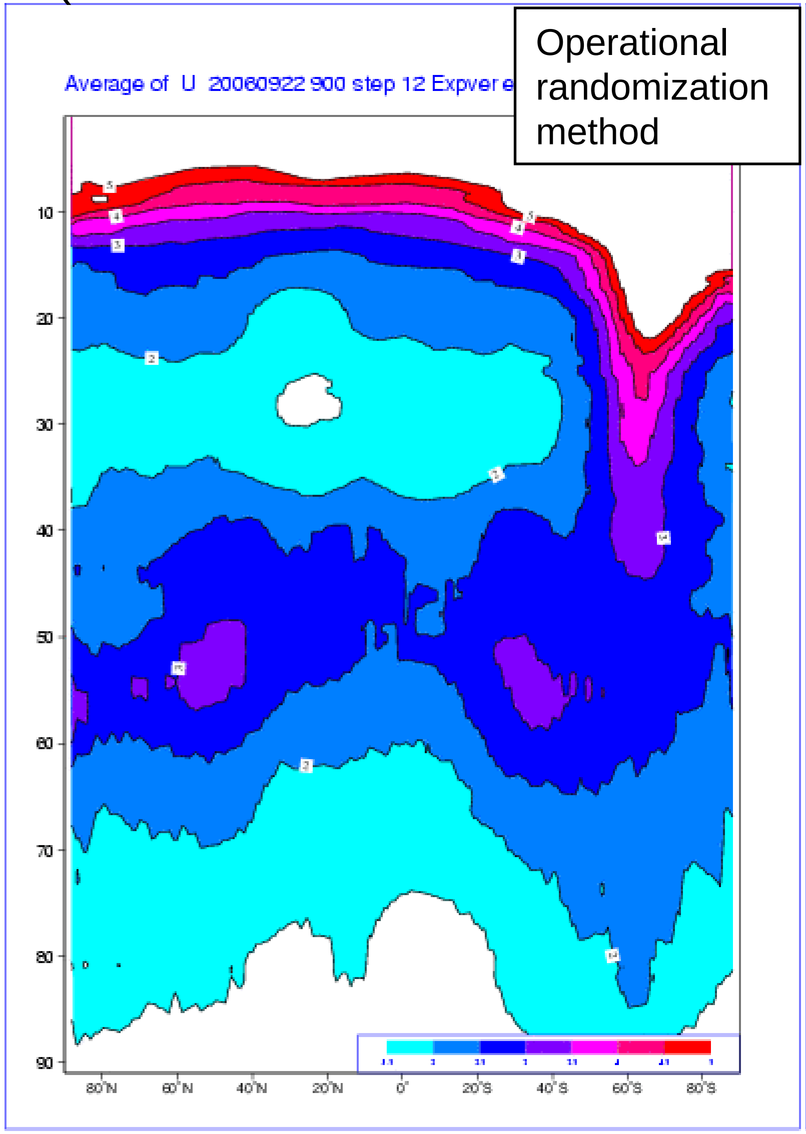


2.0*Stdev of 10 T255 outer loop+ T95/T159 inner ensemble members. Rescaled by 0.7

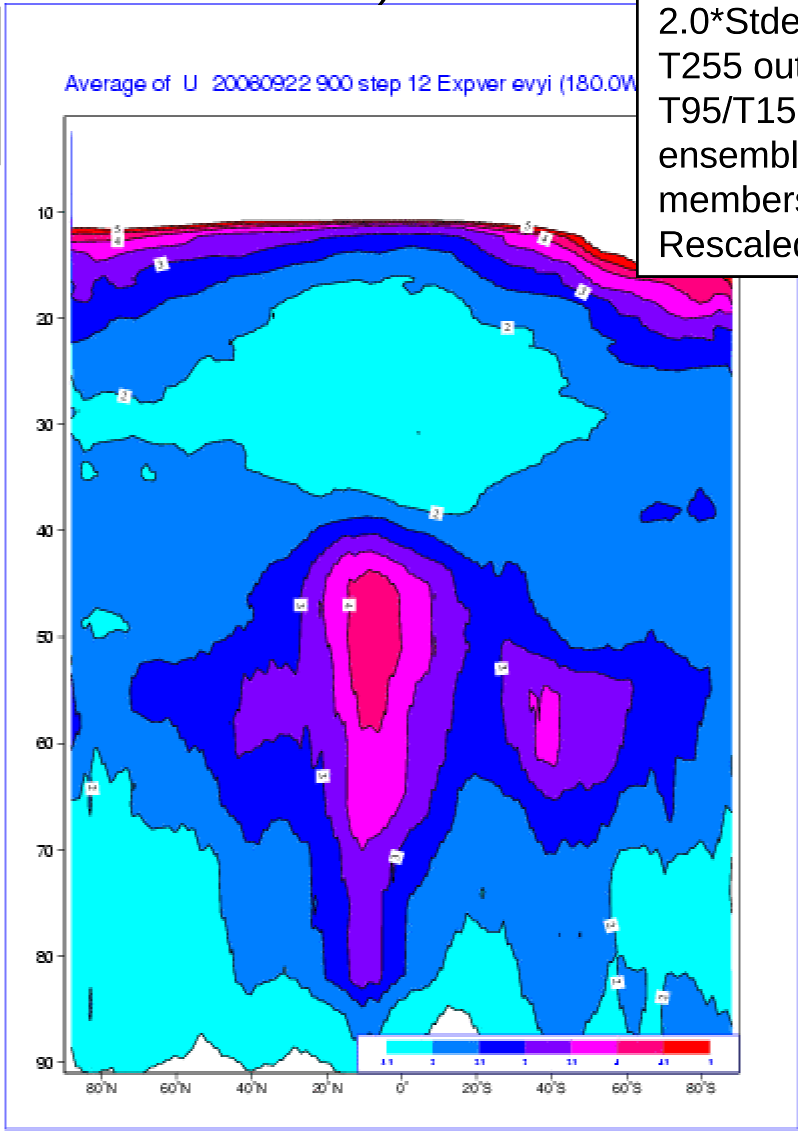


Day-to-day zonal mean background error estimates for **U-wind** have fairly different structure and amplitude, calculated from operational randomization method and from the 2.0*(standard deviation of the 10 ensemble DA members).

Operational randomization method

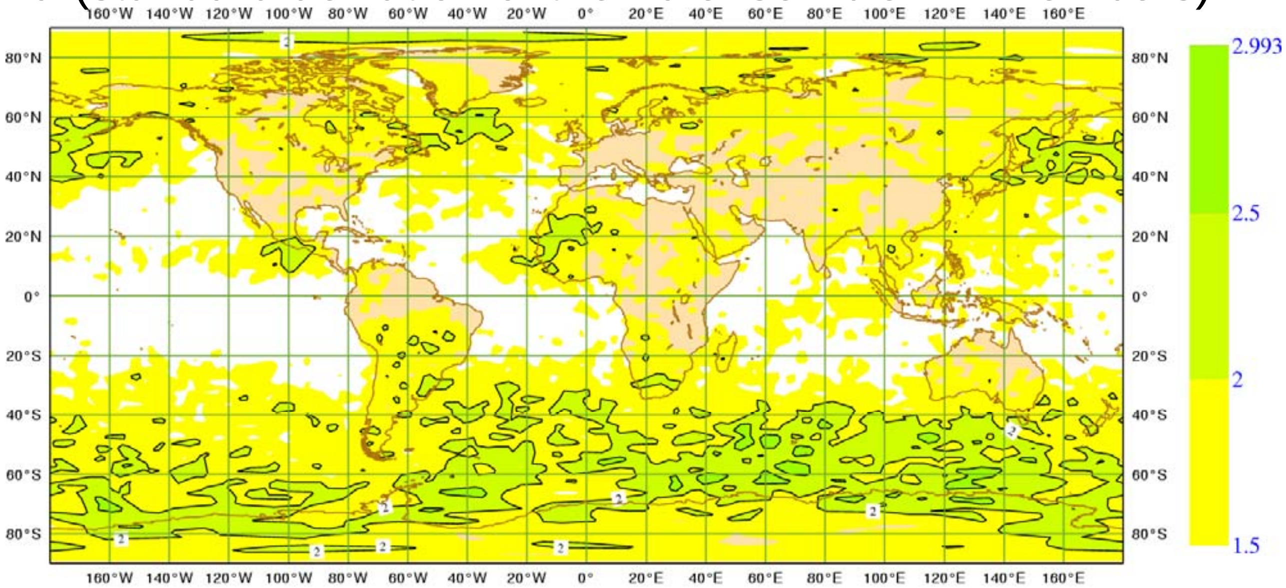


2.0*Stdev of 10 T255 outer loop+ T95/T159 inner ensemble members. Rescaled by 0.7



ECMWF

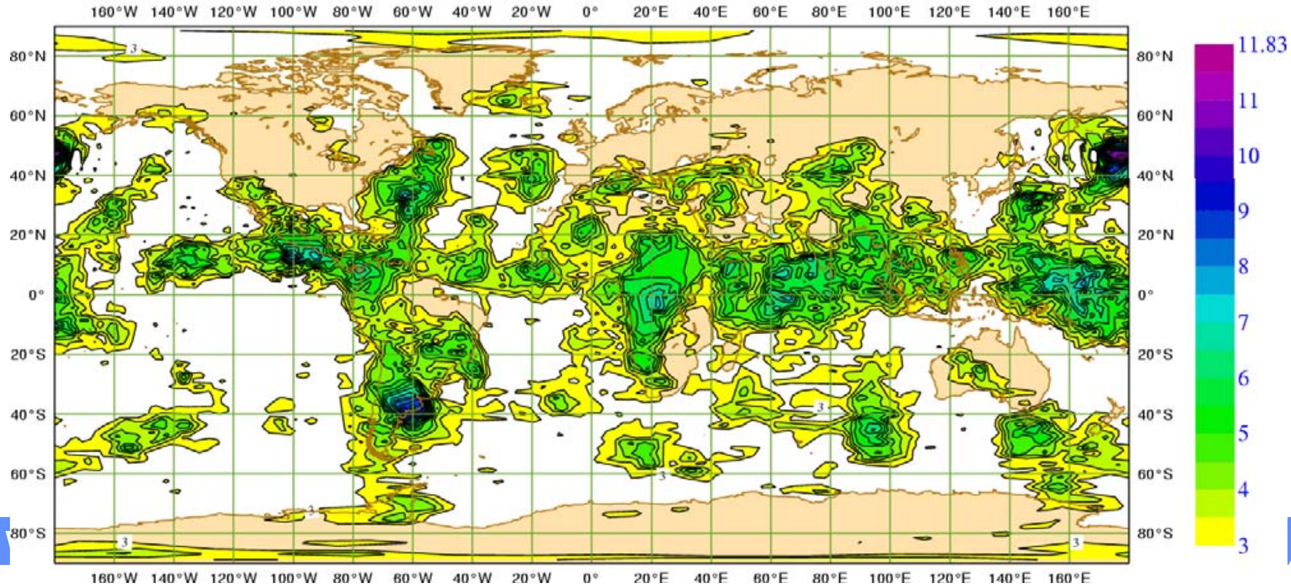
Day-to-day horizontal layer background error estimates for **U-wind** have very different structure and amplitude, calculated from operational randomization method and from the 2.0*(standard deviation of the 10 ensemble DA members).



Max=2.99m/s

Operational randomization method

Both ~850hPa



Max=11.83m/s

2.0*Stdev of 10 T255 outer loop+ T95/T159 inner ensemble members.

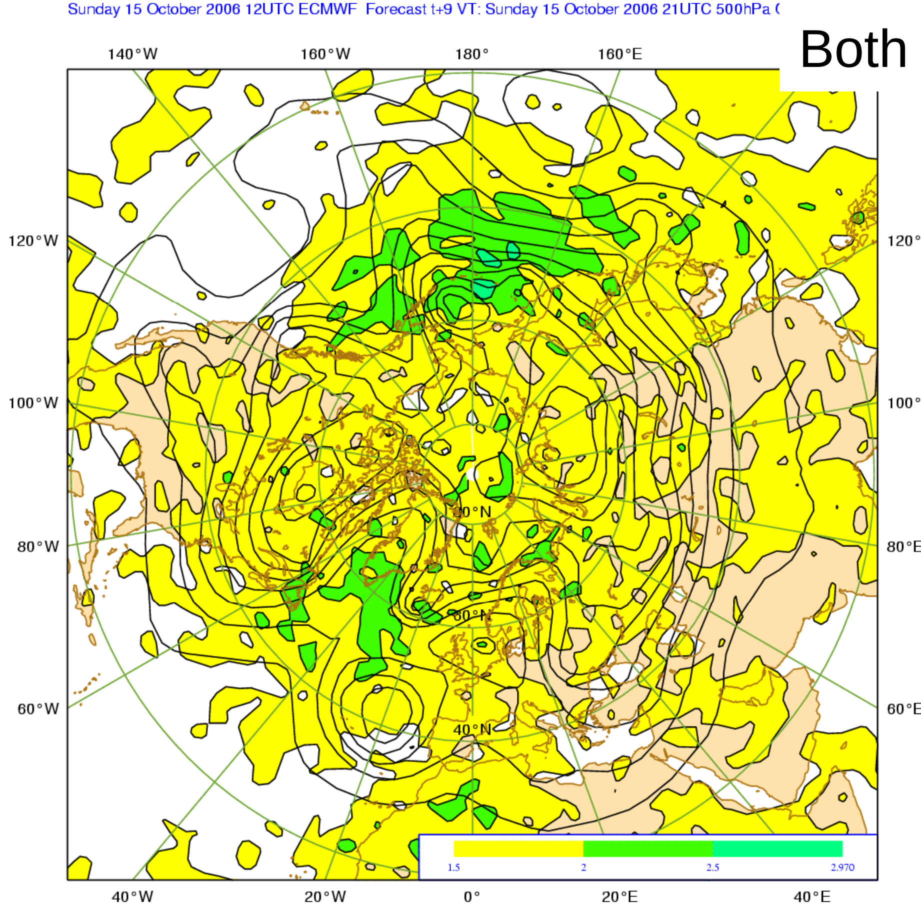
ECMWF

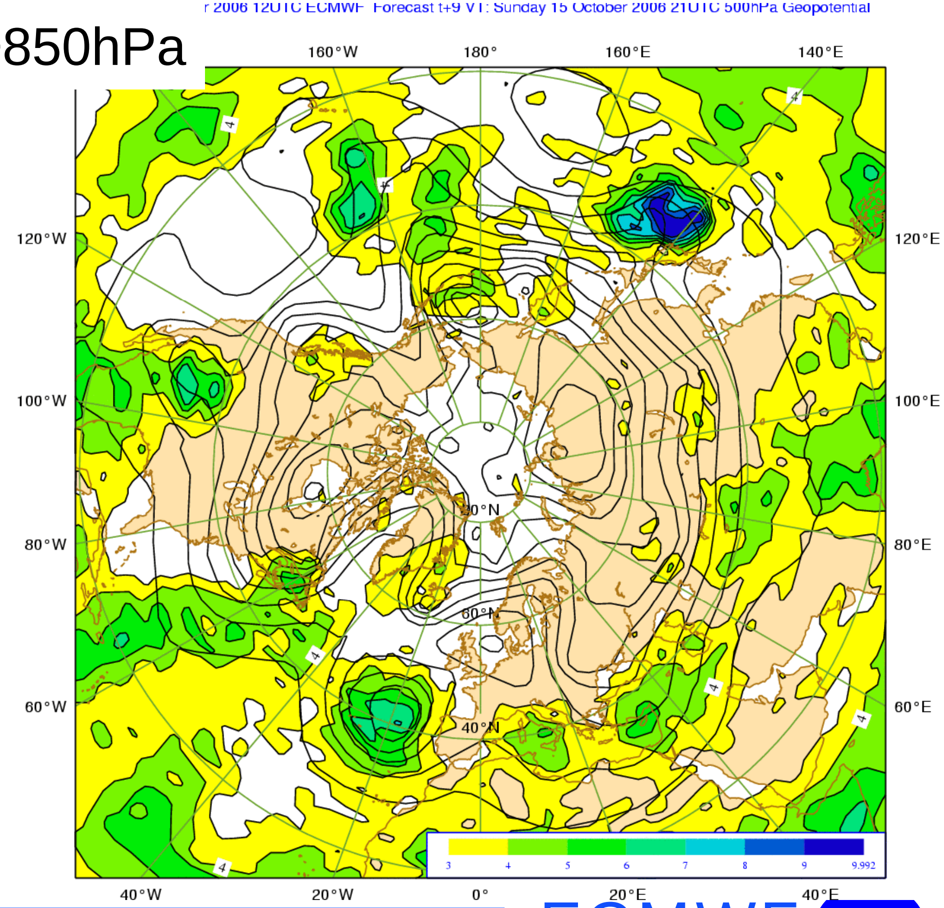
Day-to-day horizontal layer background error estimates for **U-wind** have very different structure and amplitude, calculated from operational randomization method and from the 2.0*(standard deviation of the 10 ensemble DA members).
The ensemble based version is much more flow-dependent.

Operational randomization method

2.0*Stdev of 10 T255 outer loop+ T95/T159 inner ensemble members.

Both ~850hPa

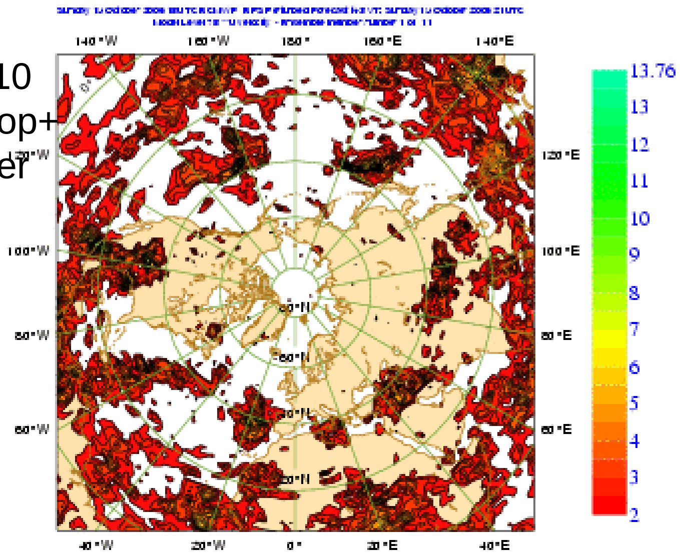




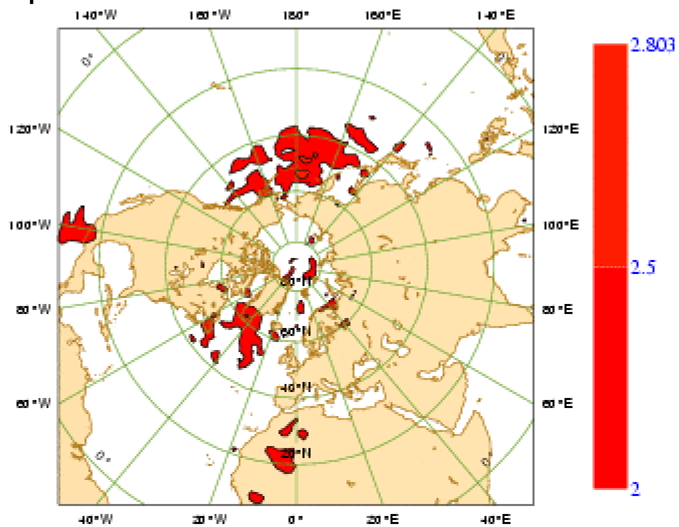
ECMWF

Background error variances for **U-wind** at 850hPa used for Quality Control estimated from ensemble spread *2.0 and oper. randomization method are **even more different**.

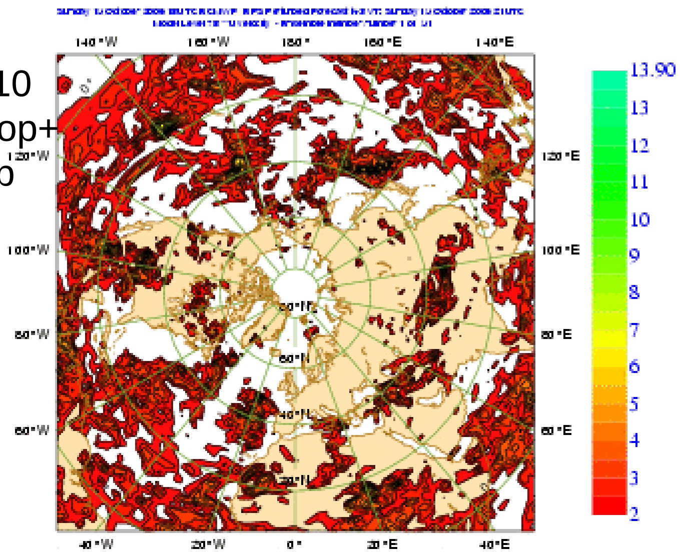
2.0*Stdev of 10 T255 outer loop+ T95/T159 inner ensemble members

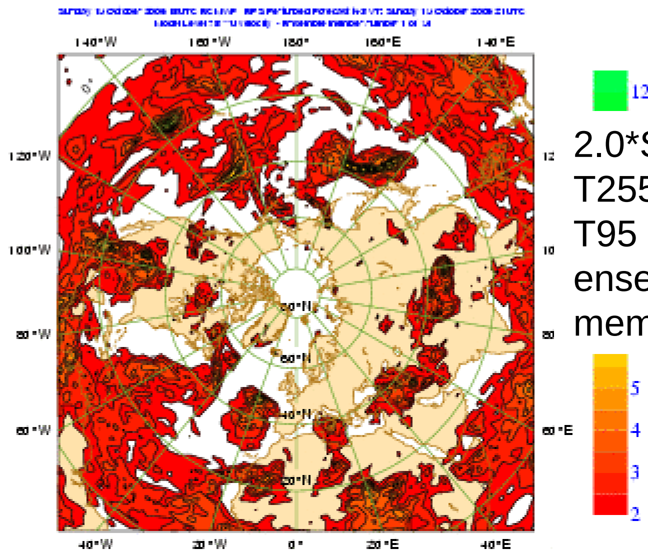


Operational EF fields from randomization method



2.0*Stdev of 10 T255 outer loop+ T95 inner loop ensemble members



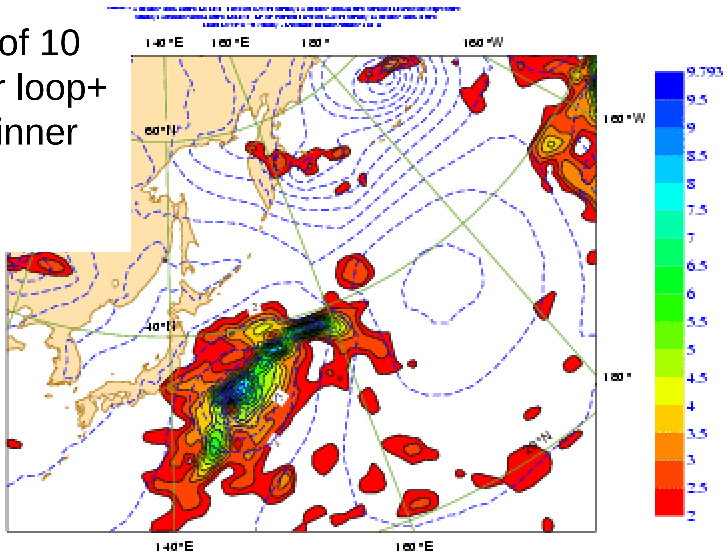


2.0*Stdev of 50 T255 outer loop+ T95 inner loop ensemble members

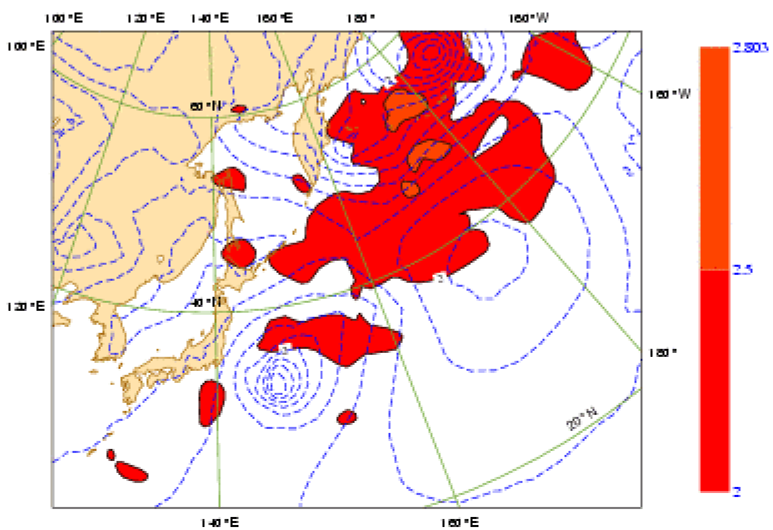
ECMWF

Background error variances for **U-wind** at 850hPa used for Quality Control estimated from ensemble spread *2.0 and oper. randomization method are **even more different**.

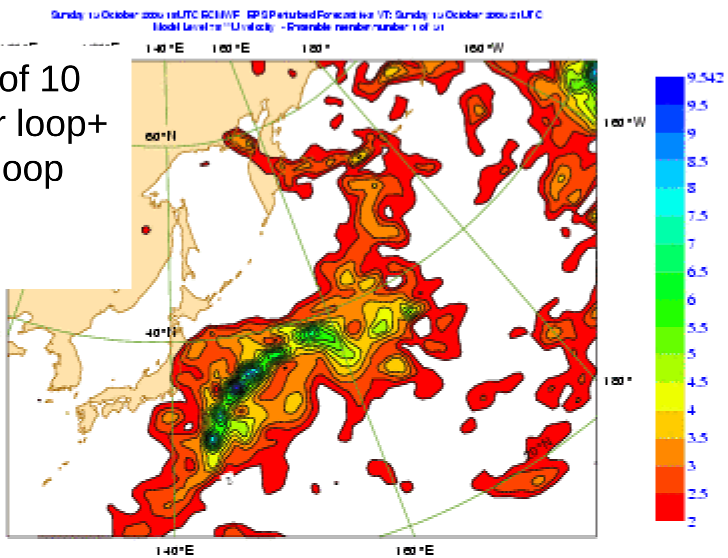
2.0*Stdev of 10 T255 outer loop+ T95/T159 inner ensemble members

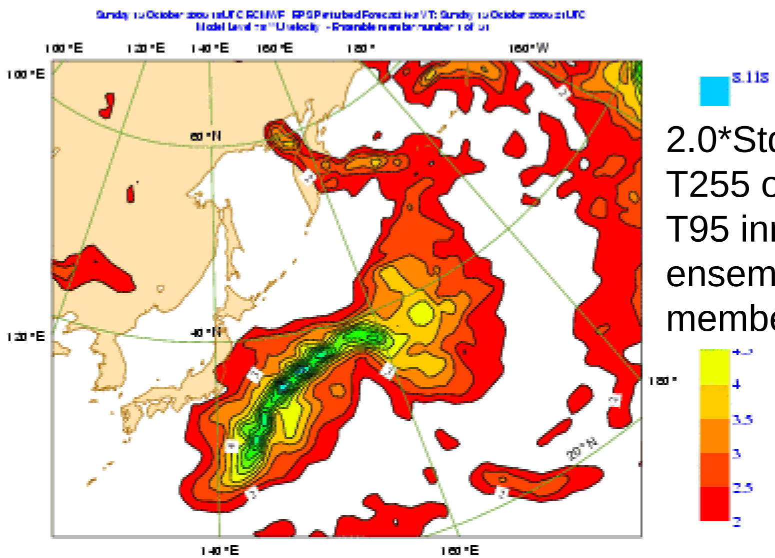


Operational EF fields from randomization method



2.0*Stdev of 10 T255 outer loop+ T95 inner loop ensemble members





2.0*Stdev of 50 T255 outer loop+ T95 inner loop ensemble members

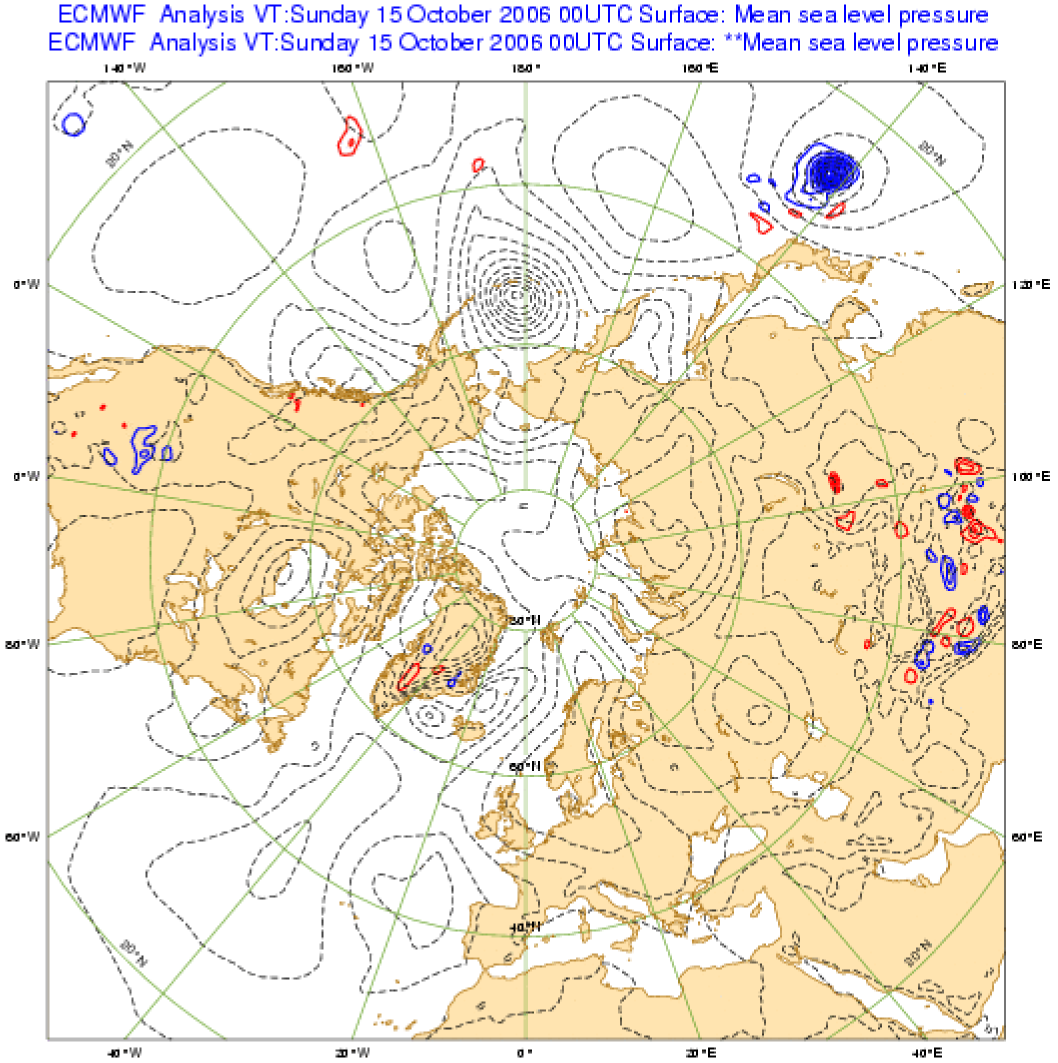
ECMWF

But analysis from assimilations with background error estimates based on ensemble spread *2.0 and operational randomization method, respectively are **very similar in general.**

Analyses only differ near tropical cyclones, troughs and extra-tropical lows.

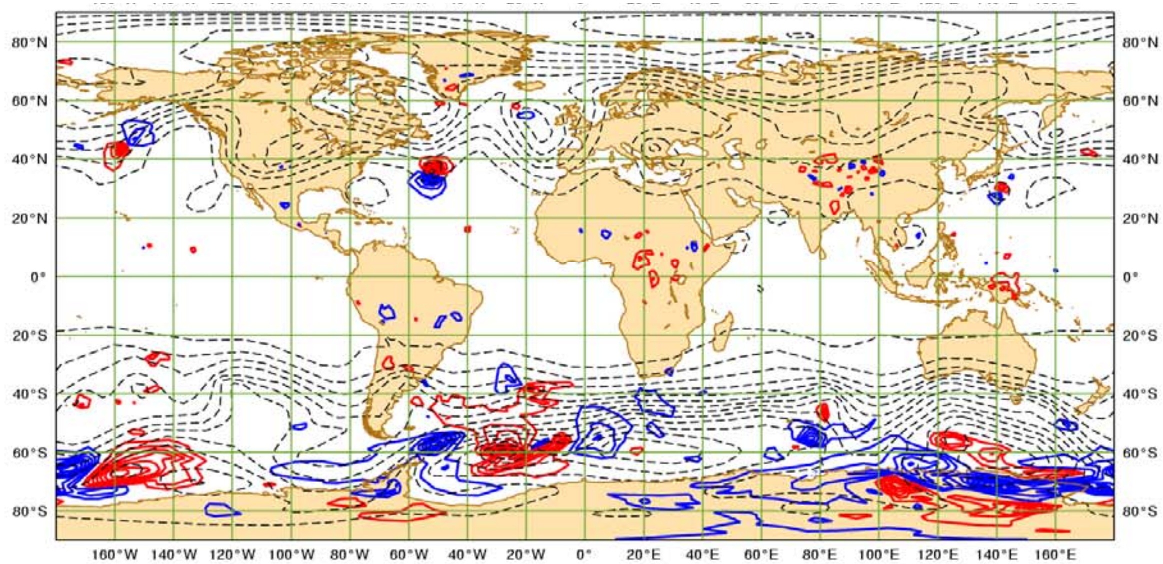
MSLP analysis difference of ENDA and static Jb runs (1hPa contours, red=positive, blue=negative).

Dashed contours MSLP field (5hPa contours)

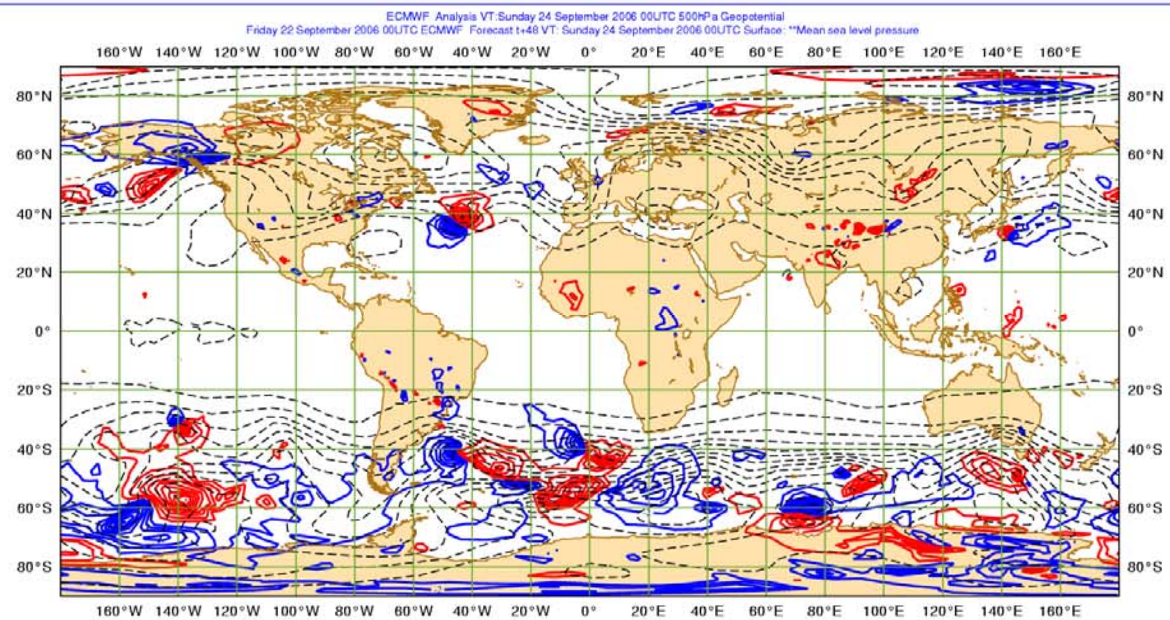


ECMWF

The 24h and 48h MSLP forecast differences for assimilations with background error estimates based on ensemble spread *2.0 and operational randomization method, respectively, show the same pattern

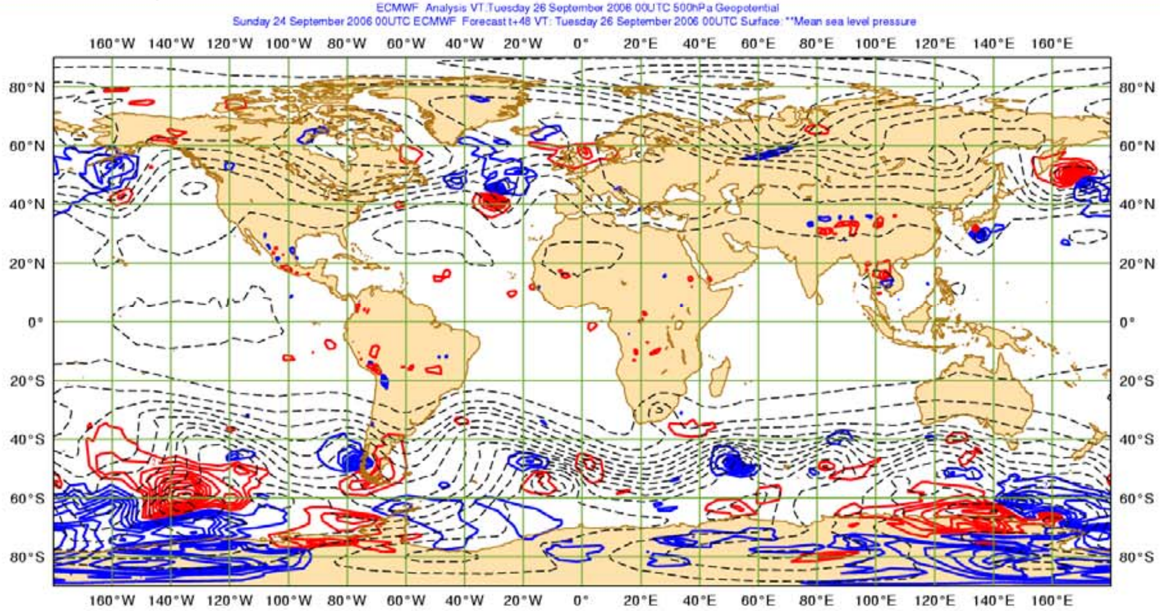


MSLP 24h fc difference of ENDA and static Jb runs (1hPa contours, red=positive, blue=negative). Dashed contours 500hPa height field (8dm contours)



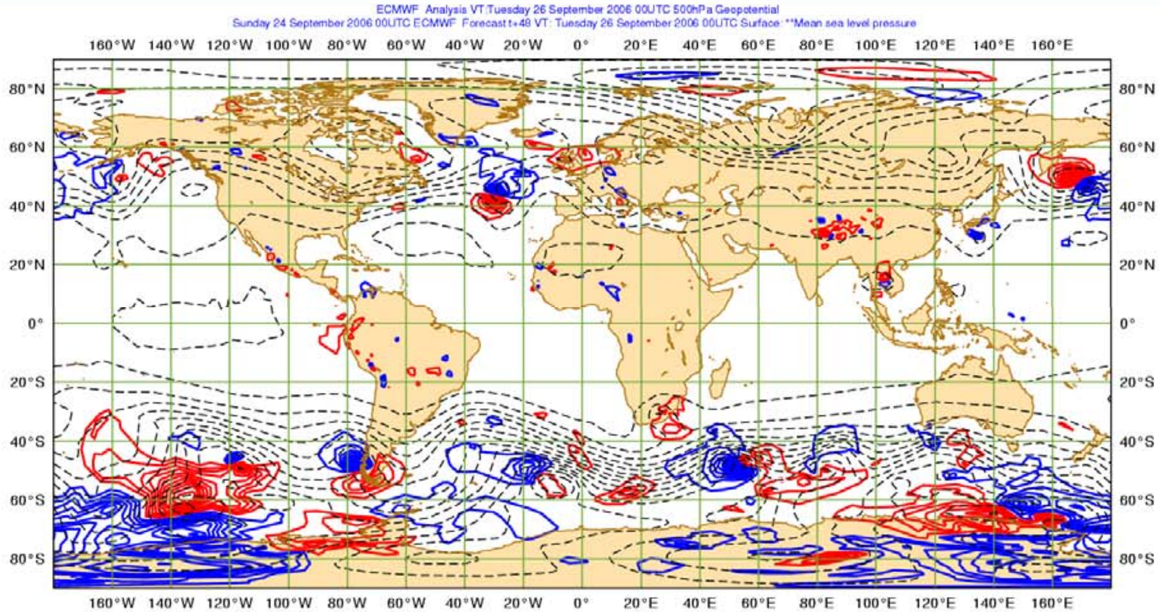
MSLP 48h fc difference of ENDA and static Jb runs (1hPa contours, red=positive, blue=negative). Dashed contours 500hPa height field (8dm contours)

The impact on MSLP 48h forecast of increasing ENDA based background error variances from a factor 2.0 to a factor 2.5 is small



2.0*Stdev(ENDA) run.

MSLP 24h fc difference of ENDA and static Jb runs (1hPa contours, red=positive, blue=negative). Dashed contours 500hPa height field (8dm contours)

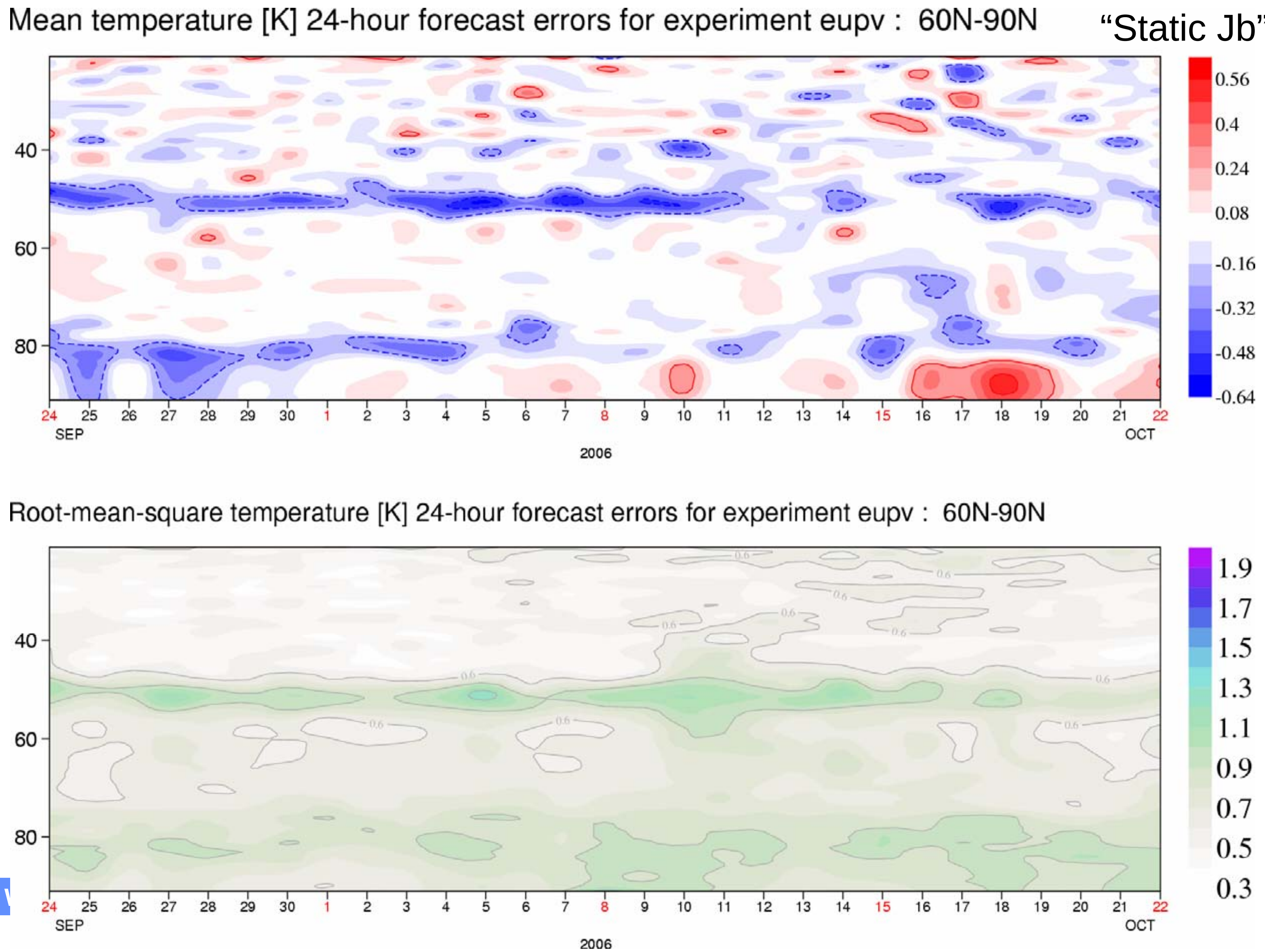


2.5*Stdev(ENDA) run.

MSLP 24h fc difference of ENDA and static Jb runs (1hPa contours, red=positive, blue=negative). Dashed contours 500hPa height field (8dm contours)

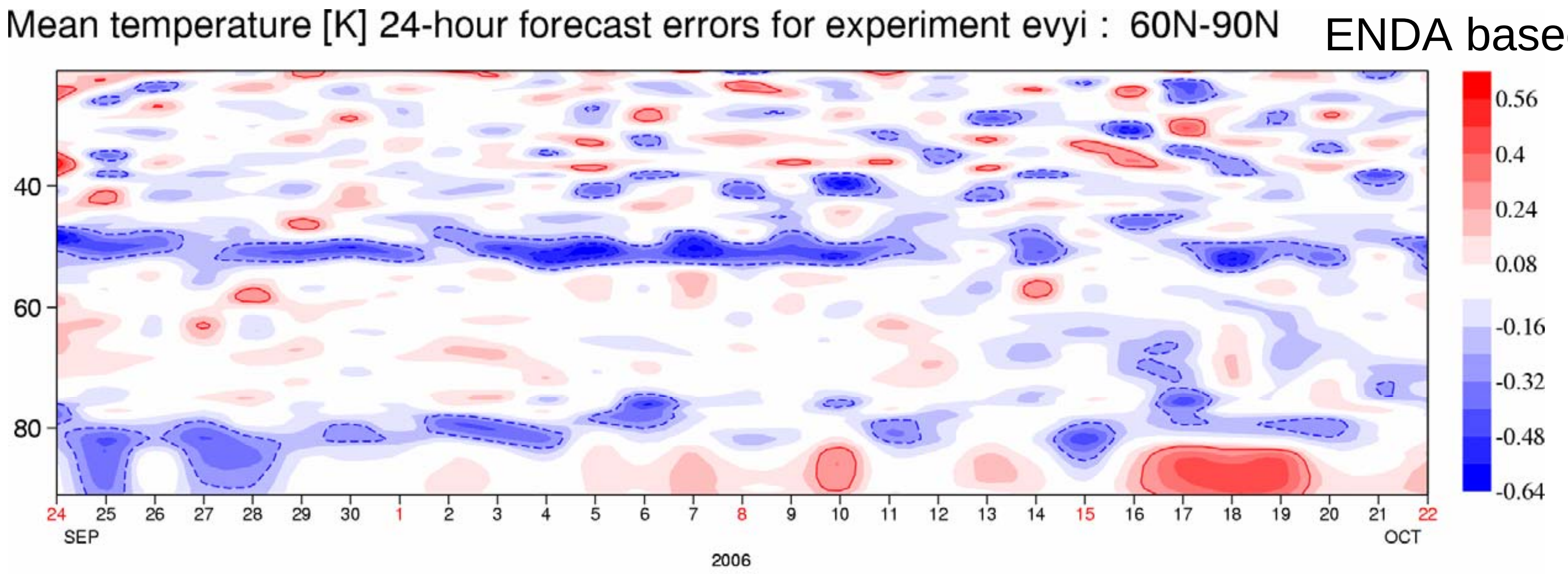
The zonal mean 24h forecast errors for temperature for assimilations with background error estimates based on ensemble spread *2.0 and operational randomization method, respectively, show the same pattern

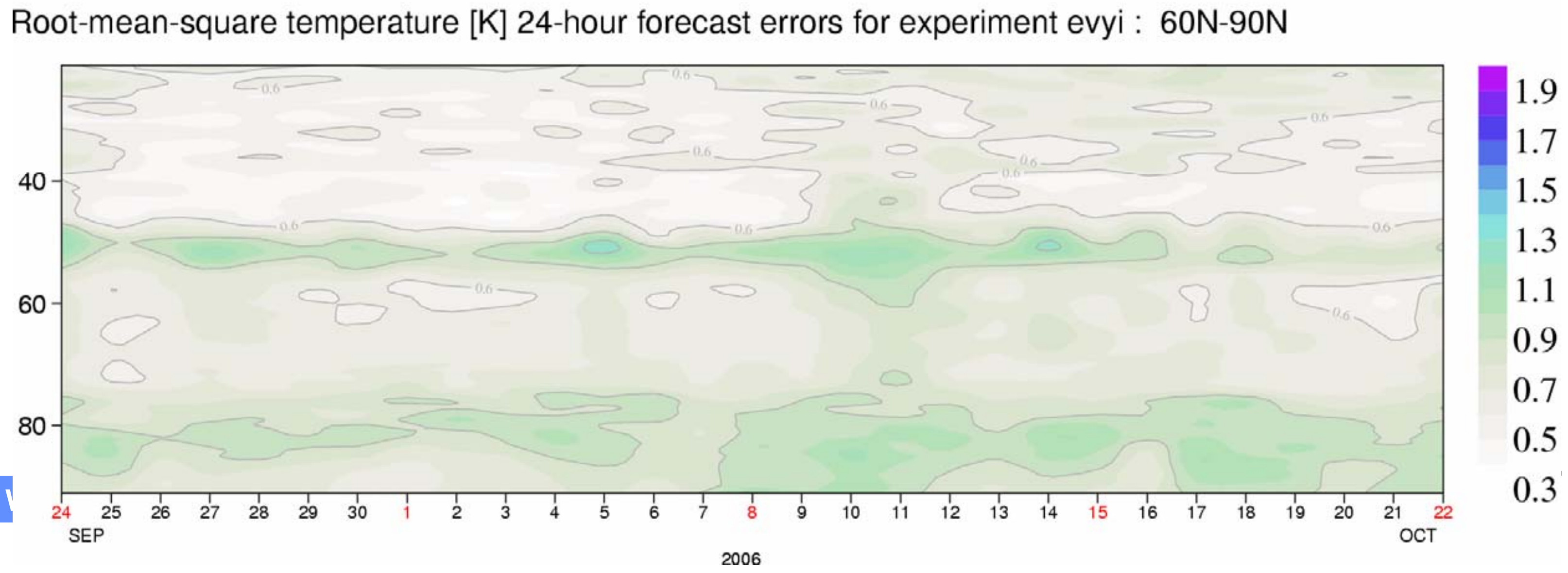
“Static Jb”



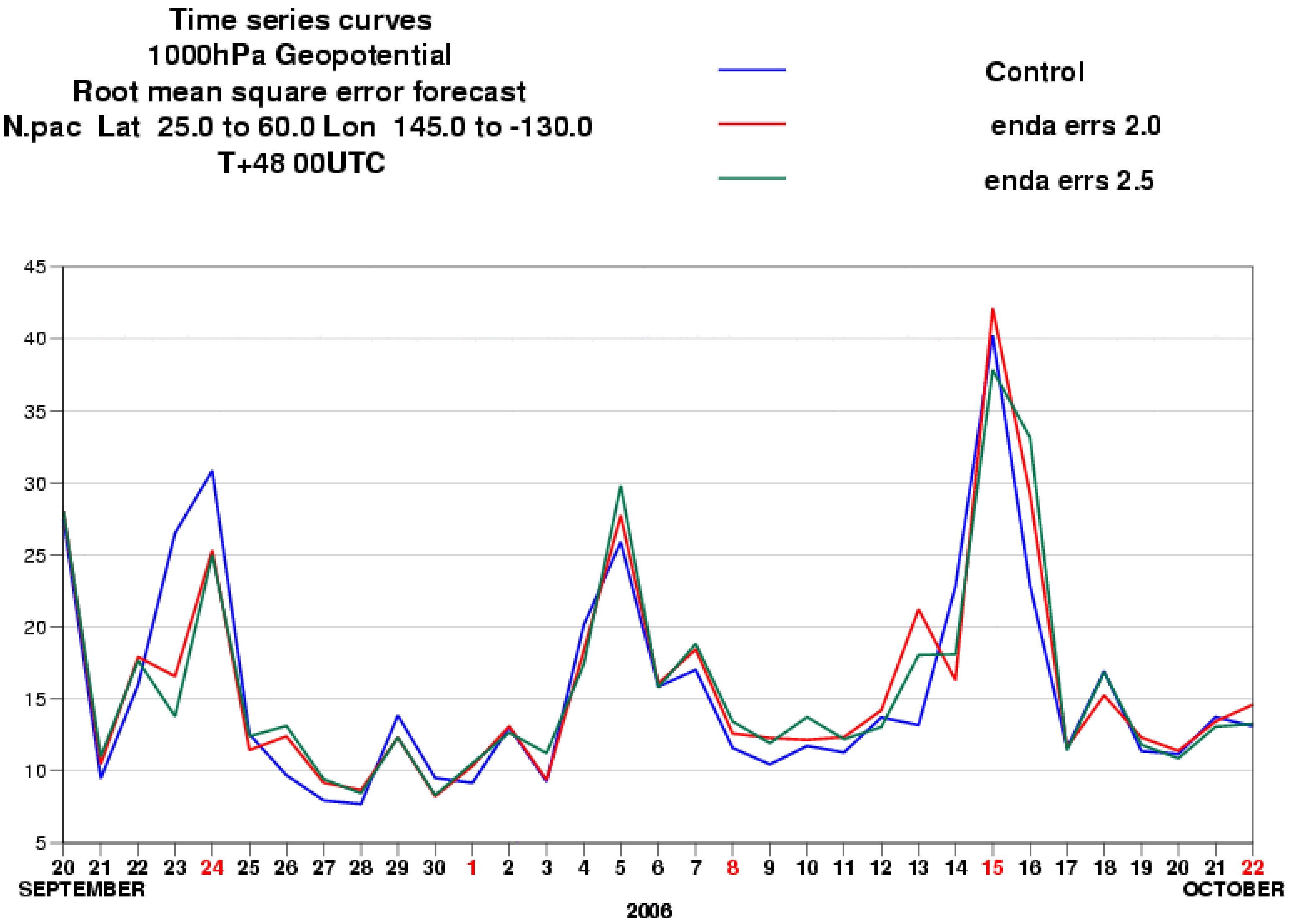
The zonal mean 24h forecast errors for temperature for assimilations with background error estimates based on ensemble spread *2.0 and operational randomization method, respectively, show the same pattern

ENDA based





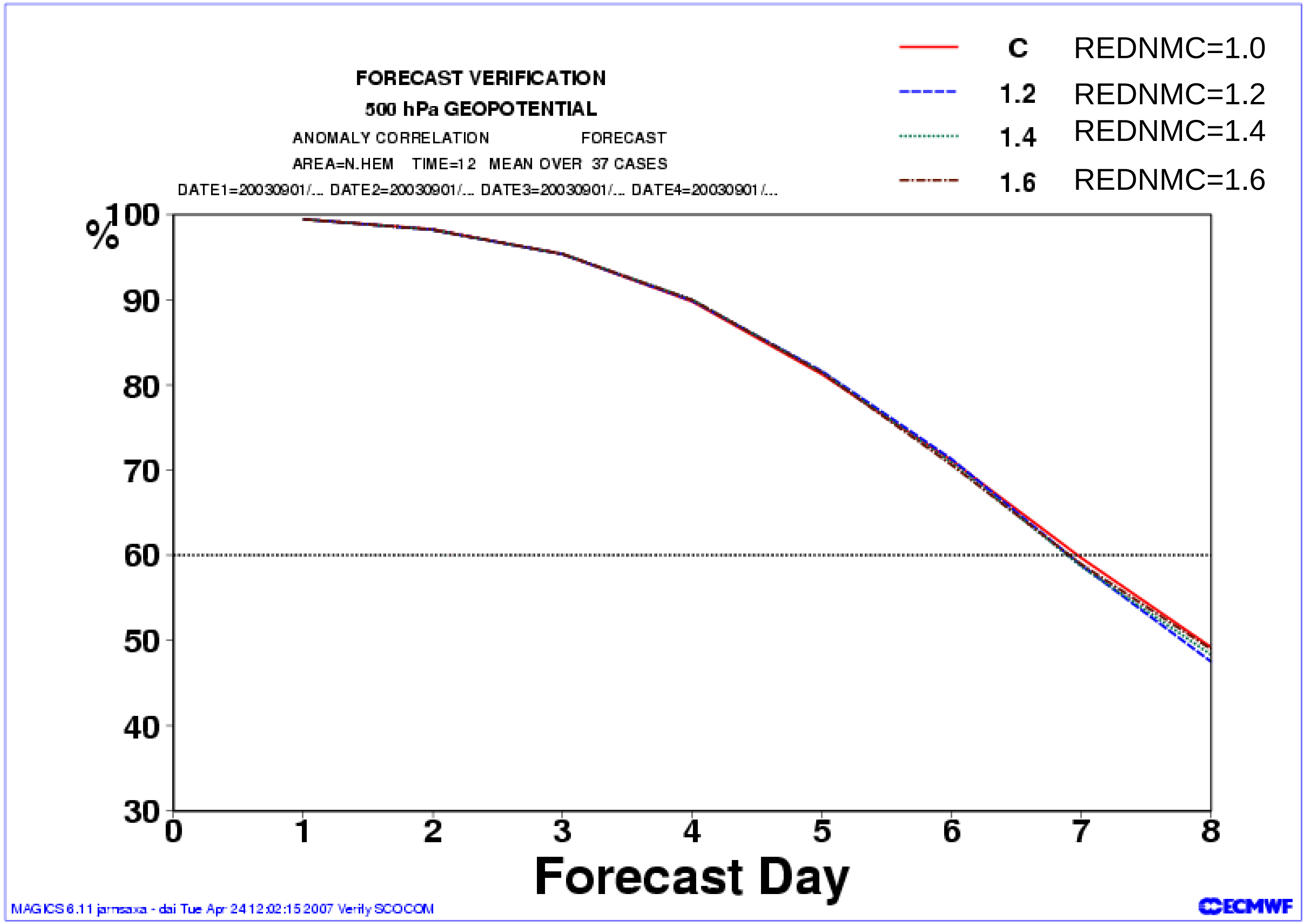
The impact of using ENDA spread versus “static Jb” for background error variances



ECMWF

Why does flow dependent background errors and REDNMC (background error variance scaling) not influence the general circulation and scores very much?

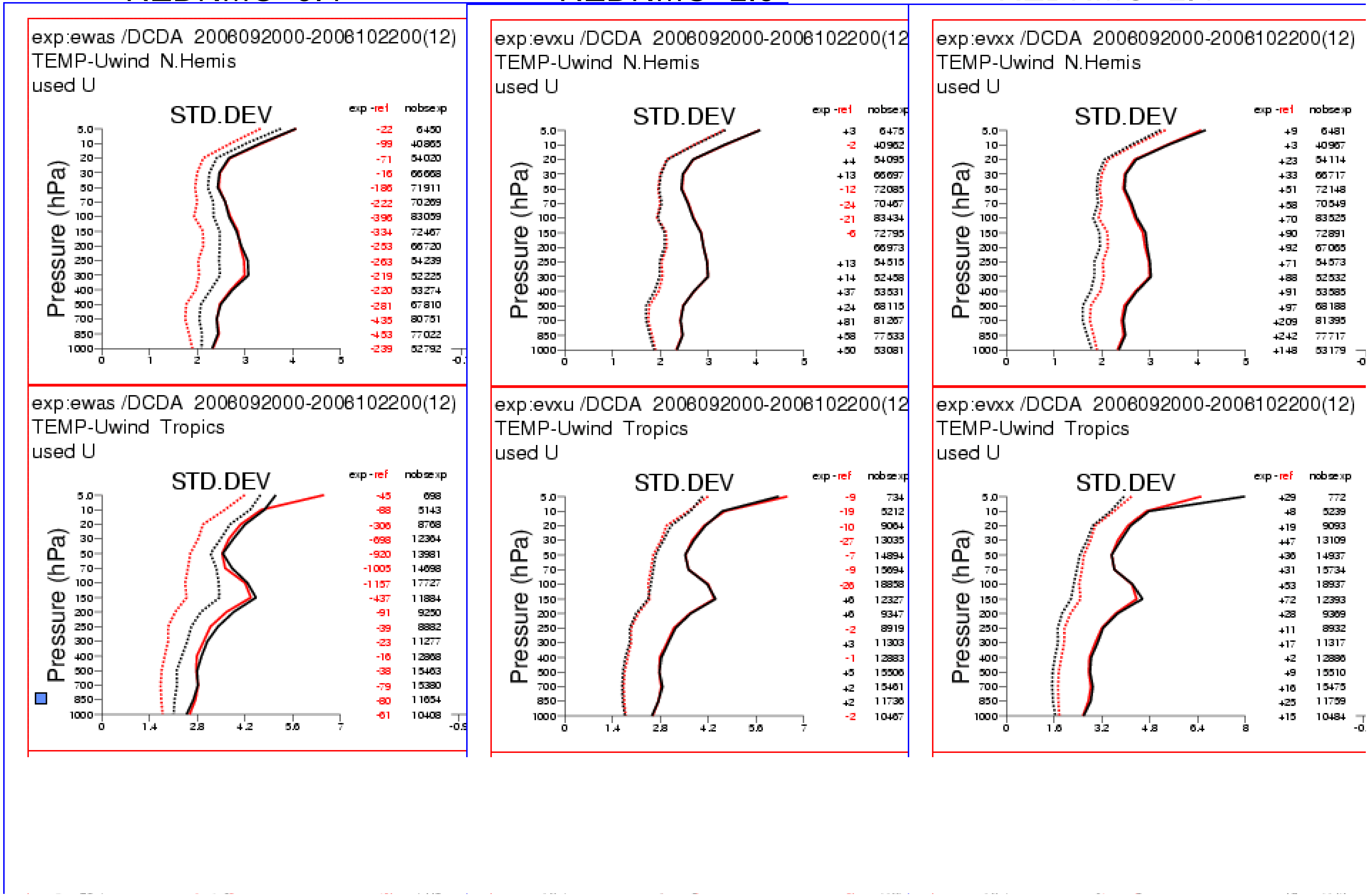
The impact of changing REDNMC on 500hPa height scores



ECMWF

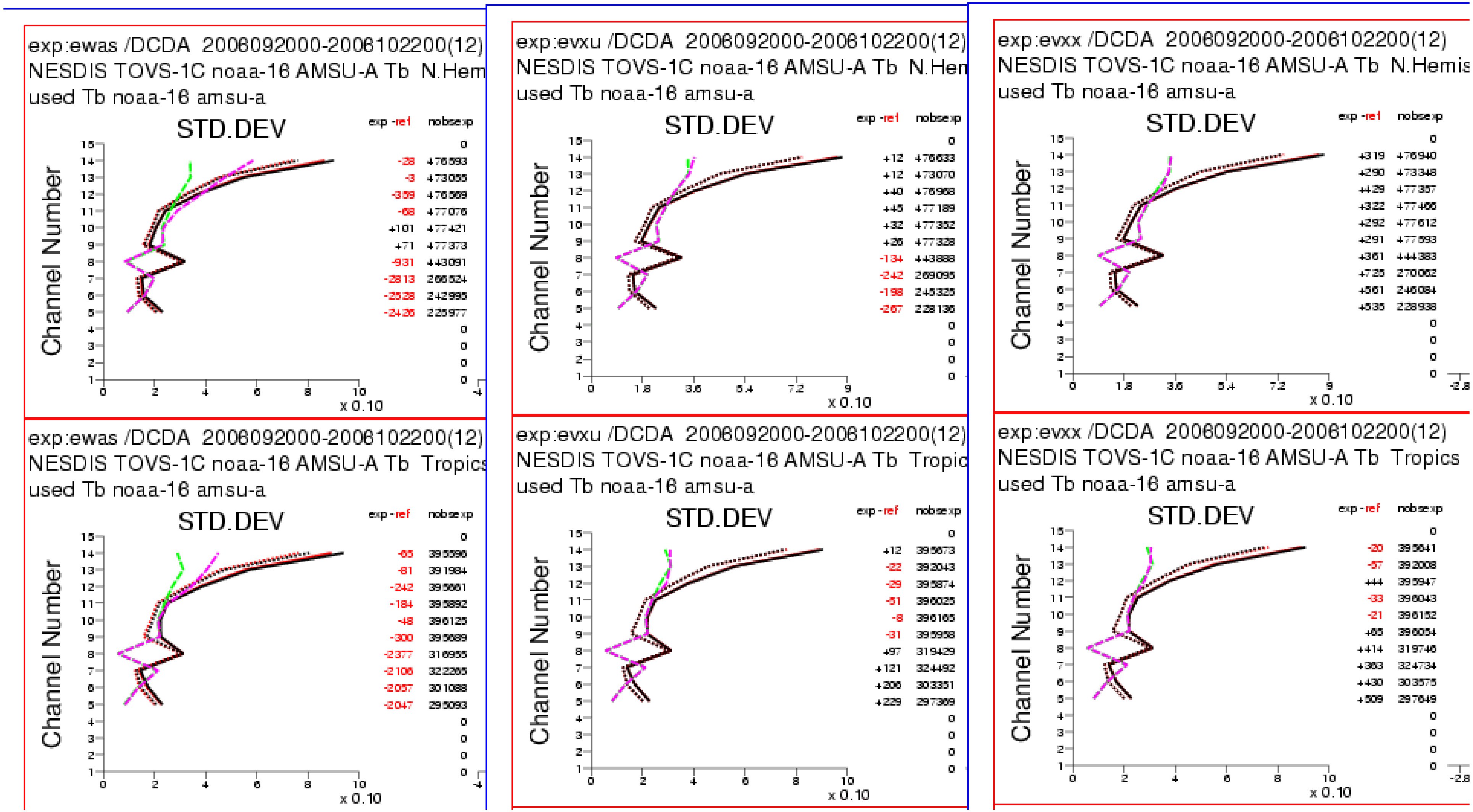
REDNMC impacts analysis fit for conventional data, but not the background fit

REDNMC=0.4 | REDNMC=1.0 | REDNMC=1.4



For almost all radiance data REDNMC does not influence the analysis or background fit

REDNMC=0.4 REDNMC=1.0 REDNMC=1.4

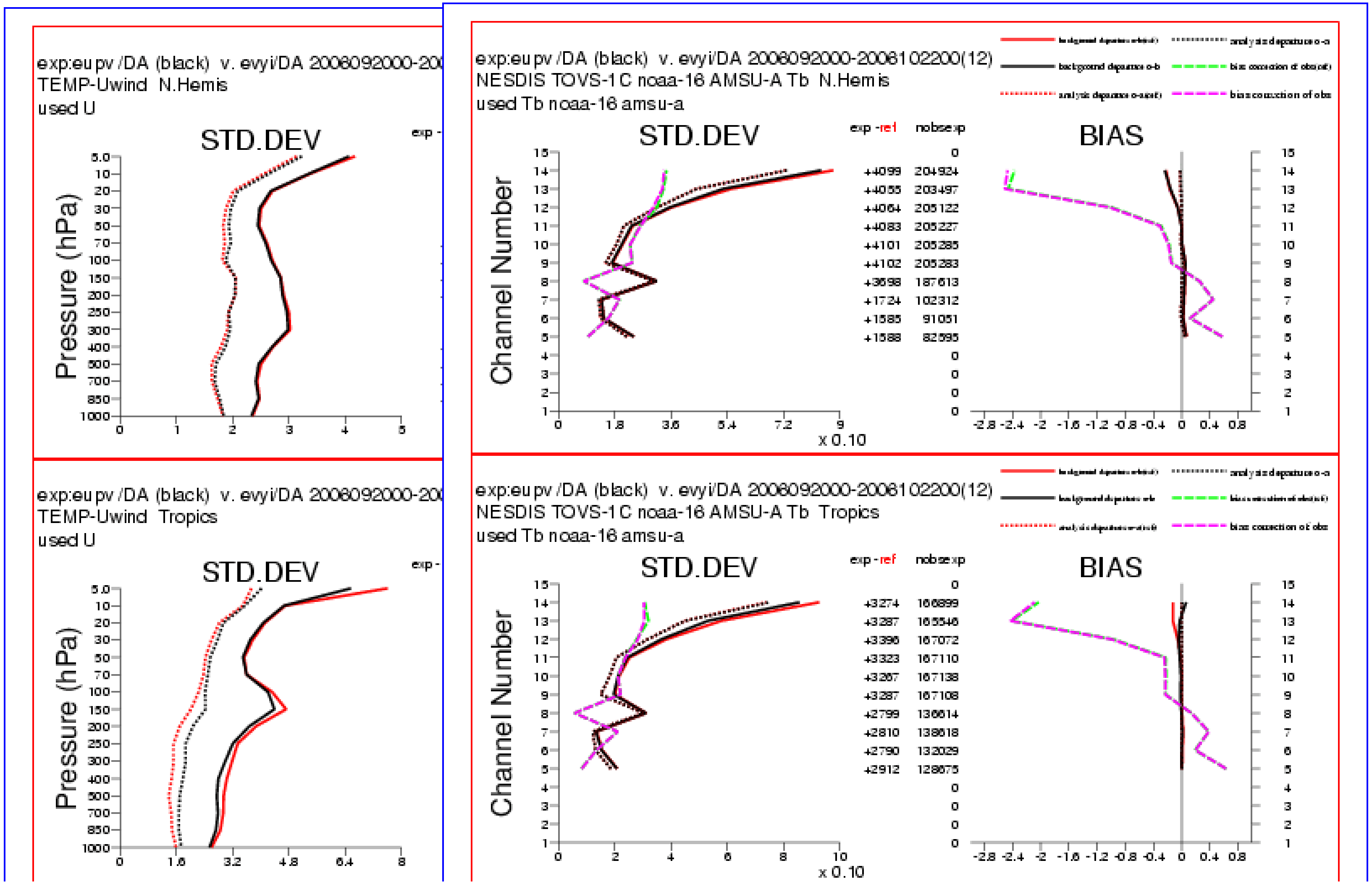


Flow dependent Jb has small impact on obs-bg fit and small impact on obs-an fit.
Larger ENDA spread in tropics results in a bigger effective REDNMC

Black curves: Static Jb   Red curves: Flow dependent Jb

N.Hem.
extra-tr.

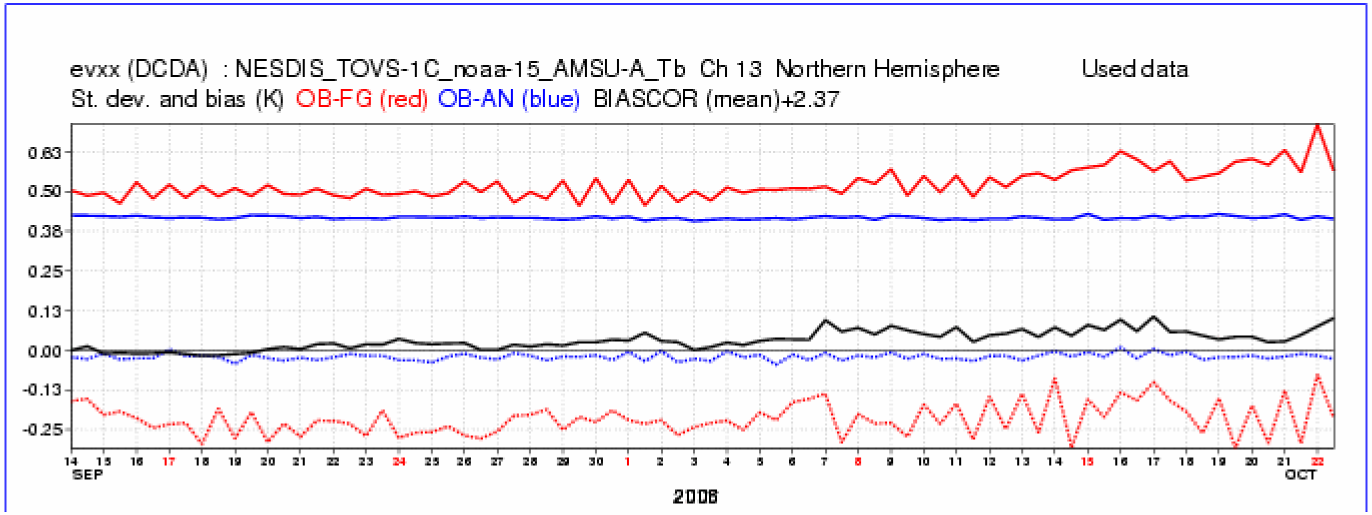
Tropics



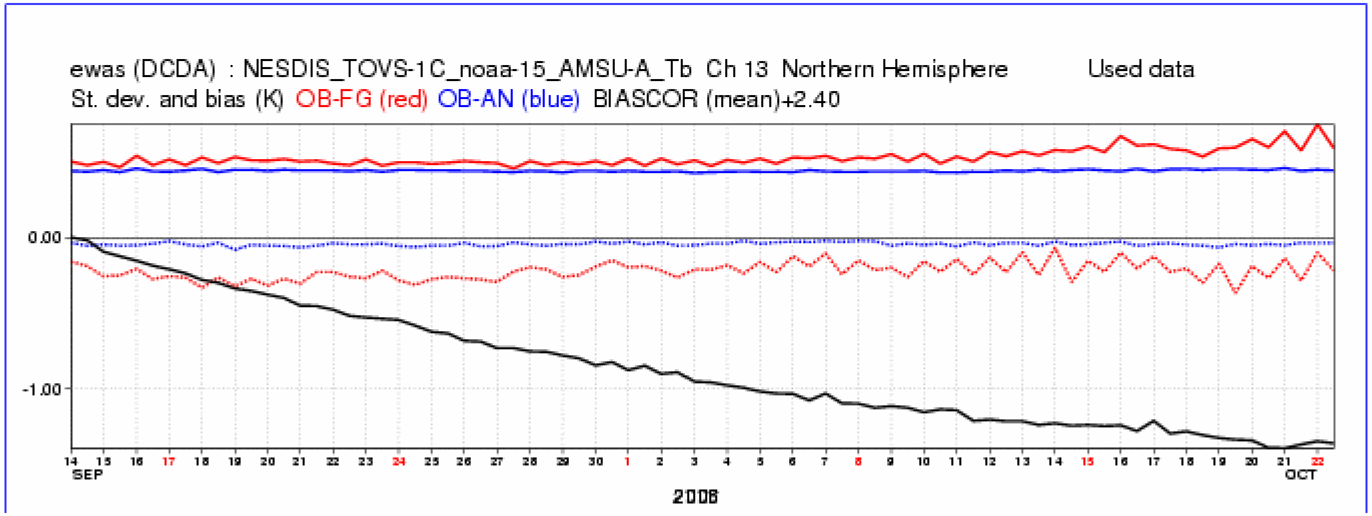
REDNMC influences the satellite biases

Time series of biases (black curve) for stratospheric AMSU-A channel 13 for assimilation experiment using REDNMC=1.4 and REDNMC=0.4, respectively.

REDNMC=1.4



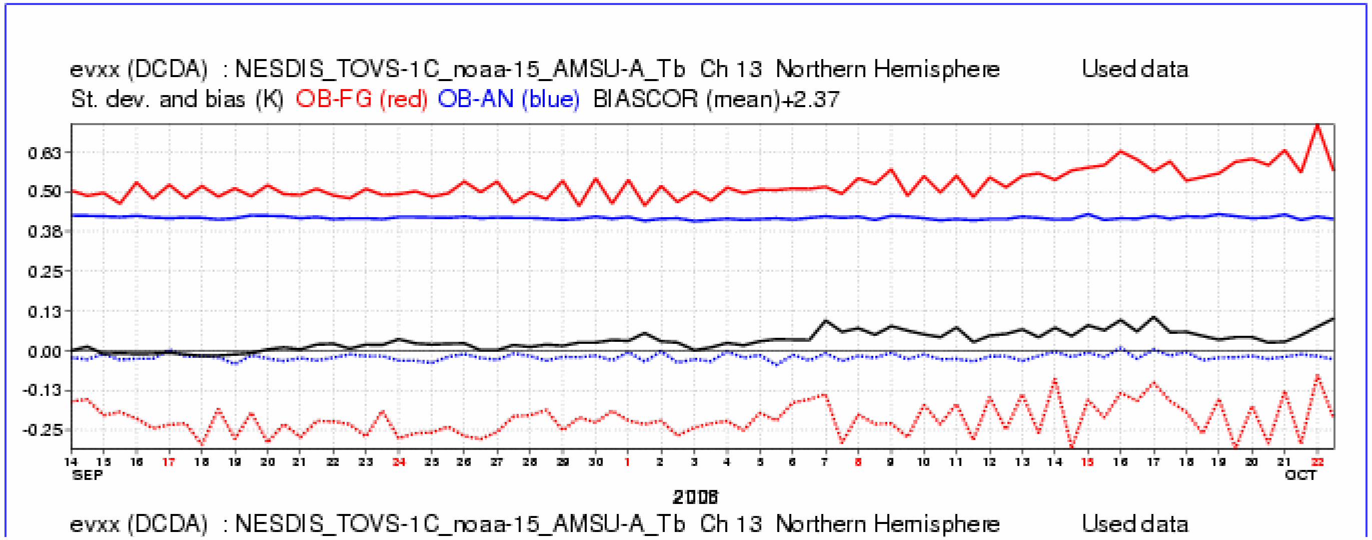
REDNMC=0.4



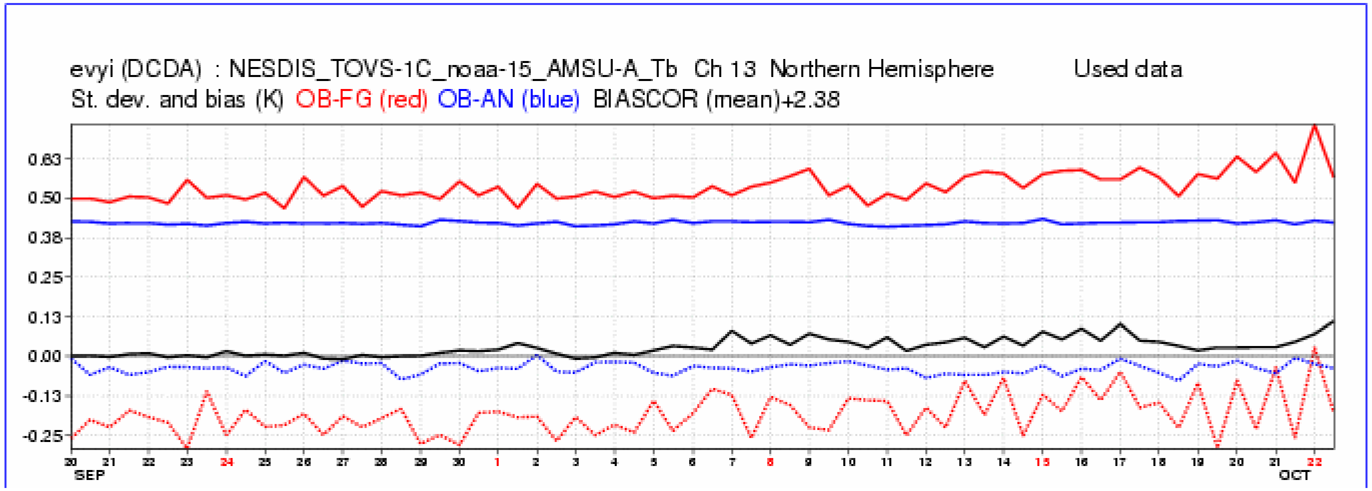
But flow dependence does not influences the satellite biases

Time series of biases (black curve) for stratospheric AMSU-A channel 13 for assimilation experiment using static Jb and flow dependent Jb, respectively.

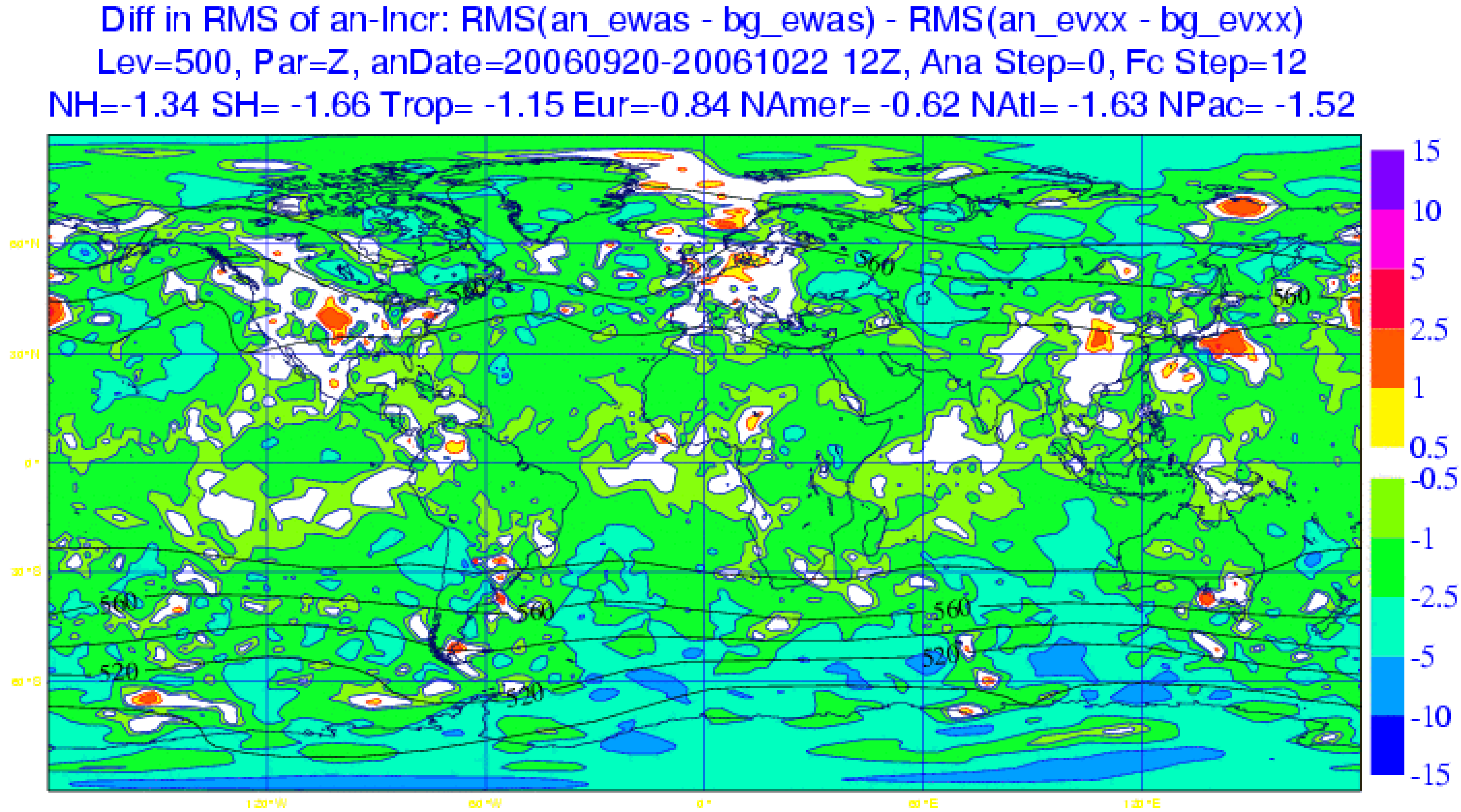
Static Jb



Flow dependent Jb based on 2.*STDEV(ENDA) variances

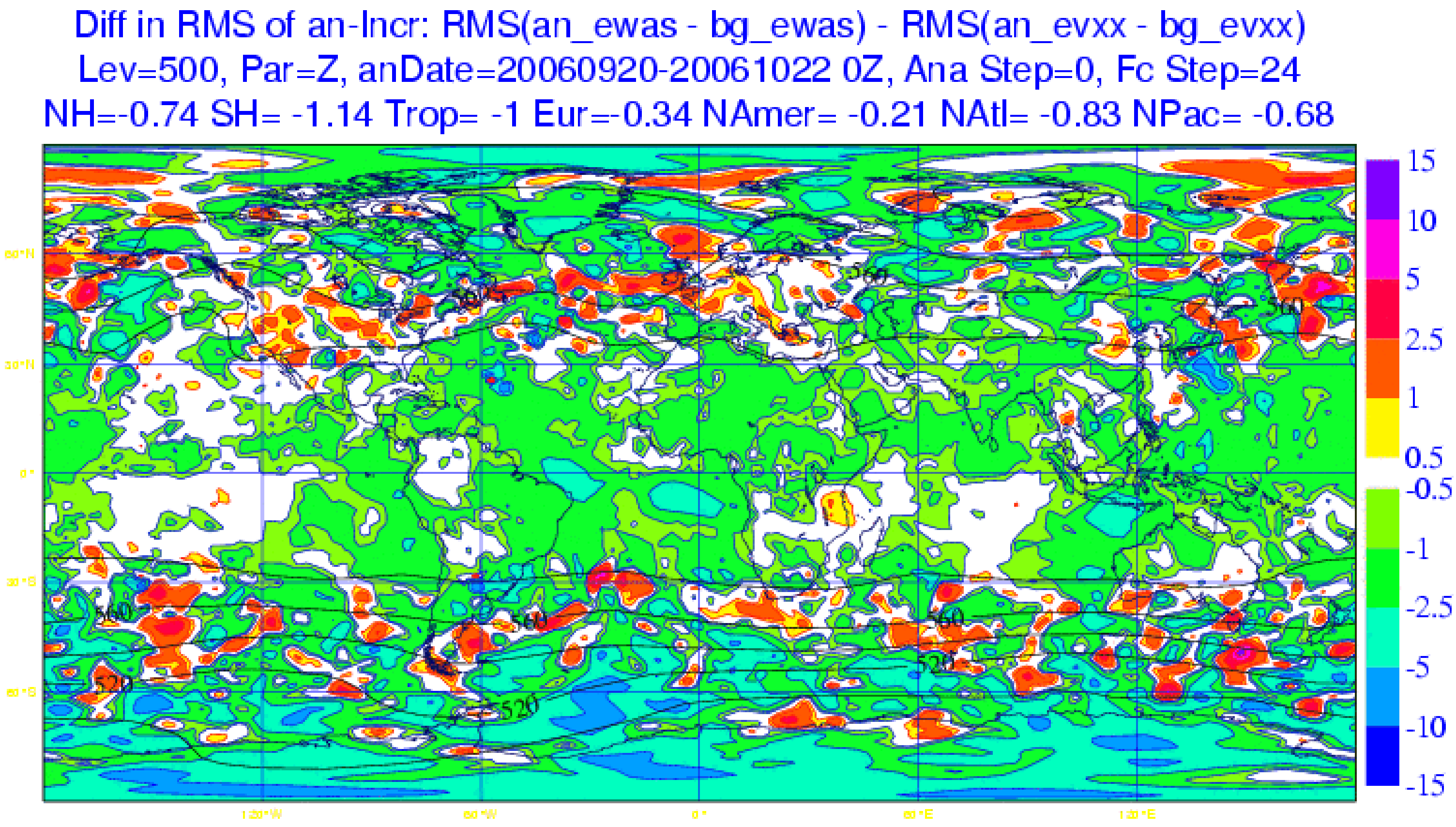


Analysis increment differences for assimilation experiments ewas (REDNMC=0.4) and evxx (REDNMC=1.4).



ECMWF

24h forecast error differences for assimilation experiments
ewas (REDNMC=0.4) and evxx (REDNMC=1.4).



ECMWF

Extracts from: Assimilating only surface pressure observations in 3D and 4DVAR

Jean-Noël Thépaut

ECMWF

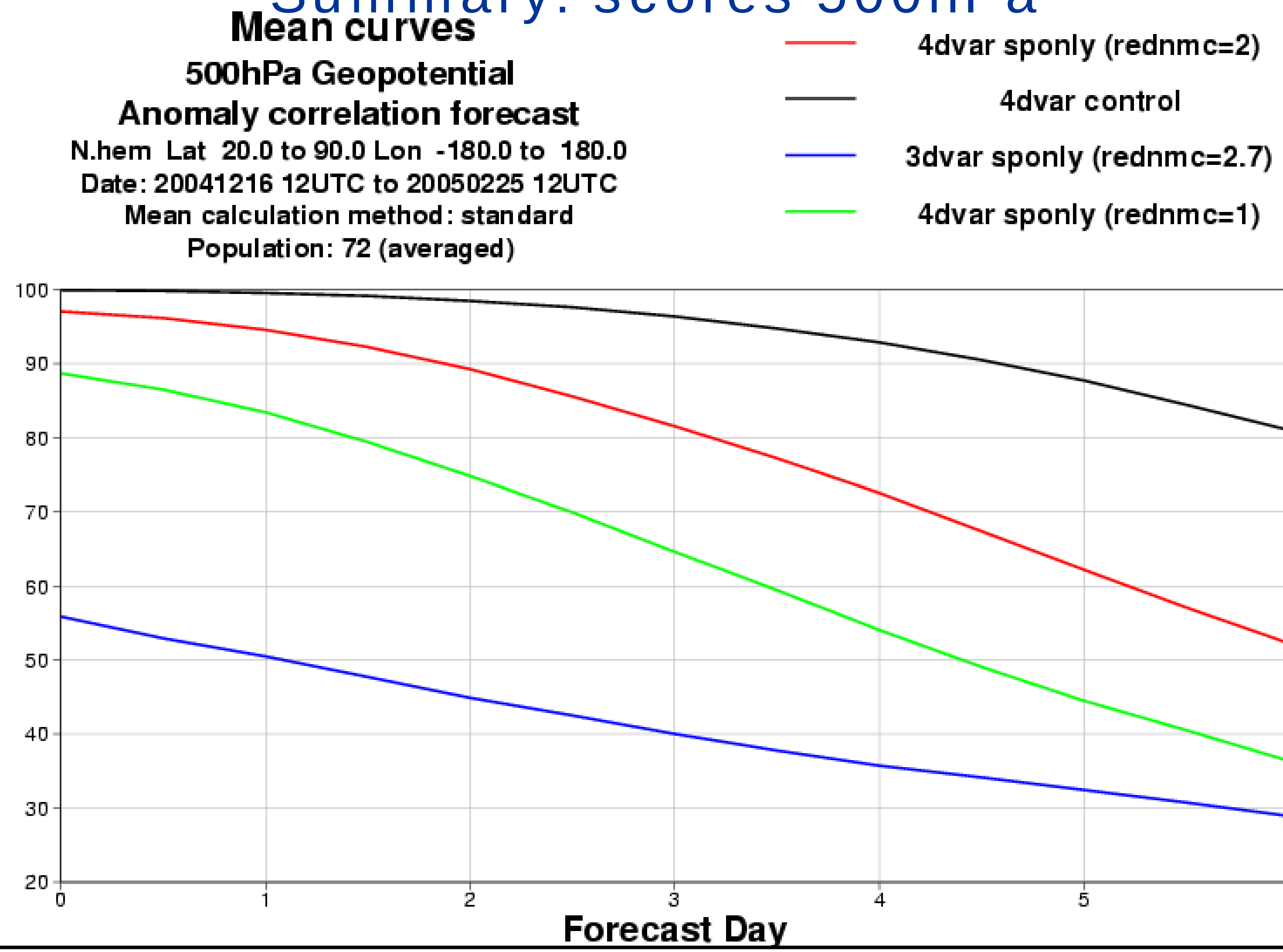
Acknowledgements: Graeme Kelly

Experimental set-up

- **Period under investigation:**
  - **2004120400-2005022512**
  - **12 first days used for warm-up and excluded from statistics**
- **Four different assimilations:**
  - **4DVAR control**
  - **3DVAR “surface pressure only”**
  - **4DVAR “surface pressure only” Default REDNMC=1.**
  - **4DVAR “surface pressure only” Retuned REDNMC=2.7**

ECMWF

Summary: scores 500hPa



Tuning the statistics of the assimilation system when the Observing System is substantially degraded is essential

Need for an adaptive covariance model

Conclusions (1)

- **We have developed a streamlined ensemble DA system**
- **We have included an improved representation of model error to get more realistic spread from the ensemble DA**
- **The use of flow dependent background error variances based on the ensemble DA spread does not improve the general scores; but impact near tropical cyclones, troughs and extra-tropical cyclones looks promising**
- **The general scores are at ECMWF determined by the broad temperature/wind structures that are well described by the high volume satellite data. Flow dependence does not matter for these structures. The forecast model can generate and evolve cyclones on its own from these accurate but broad initial conditions.**

Future work

- **We will try to take account of correlated radiance errors and improve representation of model error in the ensemble DA**
- **It would be beneficial to run a research mode ensemble DA system in real time mode to learn from daily monitoring and to calculate seasonal variance estimates**
  - **Resolution? TBD (likely T399 outer loop/T159 inner loop)**
  - **Number of members? TBD (likely 10)**
- **The wavelet J_b formulation used at ECMWF will ease introduction of flow dependent variances and structures**
- **It may be beneficial to use ensemble data assimilation based estimates of short range forecast errors in the EPS system (see Martin Leutbecher's talk tomorrow)**

Possible working group topics

- **Should we expect improved general scores from using flow dependent background error variances based on an ensemble DA?**
- **Can we utilize the ensembles for more than just background error variance estimation, i.e., to specify correlation lengths and non-isotropic structures?**