Ensemble derived background error covariances for CMAM-DA

Yulia Nezlin[1] , Yves Rochon[2], Suzhan Ren[1], Saroja Polavarapu[2]

1 - University of Toronto

2 - Environment Canada

Error covariances in CMAM

- Stationary background error covariances
- Based on 6-hour differences scheme
  - ♣ Proposed by Yves Rochon
  - ♣ To account for propagation of errors by dynamics ???
  - ♣ no impact of observations

How does it relate to statistics of 6 hour differences?

Formulation of the method: 6-hour differences scheme

by Yves Rochon

Final variances taken as minimum of 6-hour differences and day-to-day variability (climate) variances:

$$\sigma^2_{\text{6hr-diff}} = \frac{1}{4}\sum_{i=1}^{4}\left\langle \left(\mathbf{x}_{t+6} - \mathbf{x}_t - \left\langle \mathbf{x}_{t+6} - \mathbf{x}_t \right\rangle\right)\left(\mathbf{x}_{t+6} - \mathbf{x}_t - \left\langle \mathbf{x}_{t+6} - \mathbf{x}_t \right\rangle\right)^{\mathrm{T}} \right\rangle$$

$$\sigma^2_{\text{climate}} = \frac{1}{4}\sum_{i=1}^{4}\left\langle \left(\mathbf{x}_t - \left\langle \mathbf{x}_t \right\rangle\right)\left(\mathbf{x}_t - \left\langle \mathbf{x}_t \right\rangle\right)^{\mathrm{T}} \right\rangle$$ ← above ~stratopause

where t = 0, 6, 12, 18 hrs,
< > - is the monthly mean (over a single month) for each synoptic period and each grid point.

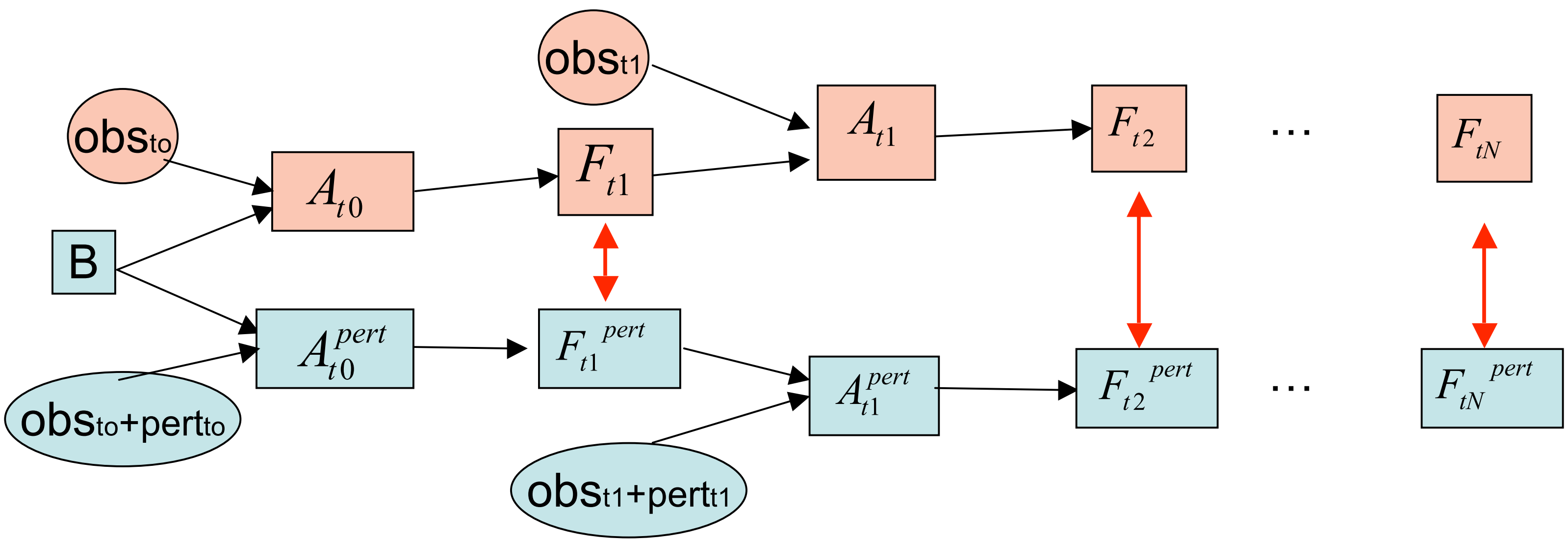
noting that

$$\frac{\sigma^2_{\text{6hr-diff}}}{\sigma^2_{\text{climate}}} \approx 2(1-\rho) \quad \begin{cases} < 1 & \text{below} \sim \text{stratopause} \\ > 1 & \text{above} \sim \text{stratopause} \end{cases}$$

where ρ is the coefficient of correlation between x at *t* and *t+6*.

Ensemble perturbation method

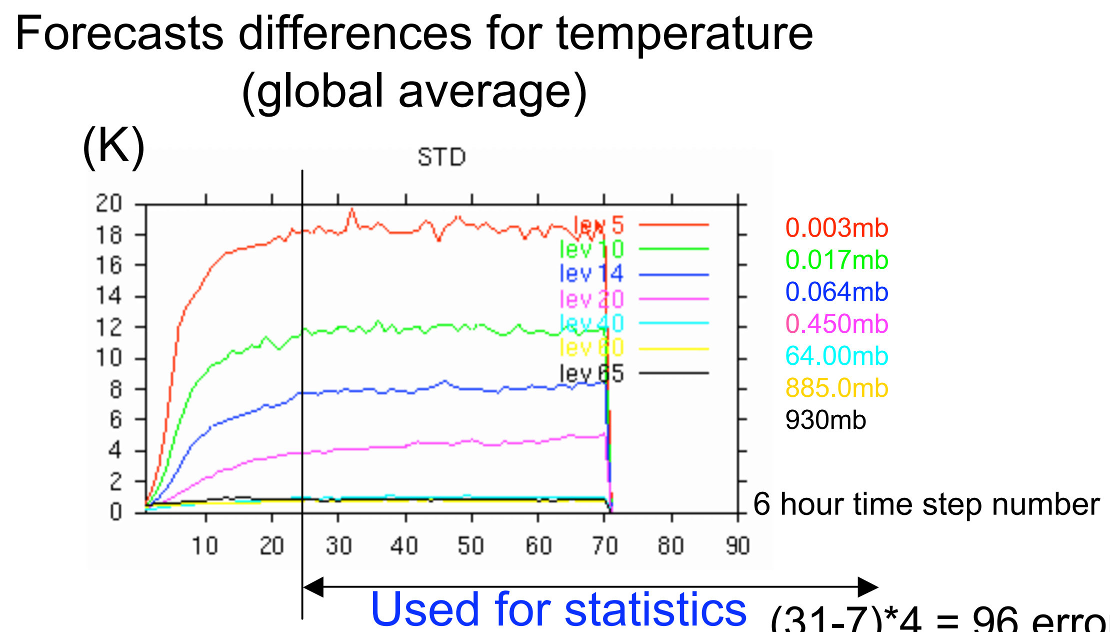
- Mark Buehner scheme (GEM) *[Q.J.R.Meteorol.Soc. 2004].*
  - ♣ Perform 2 different assimilations:
    - Control assimilation (one month)
    - Assimilation with perturbed observations (one month)



♣ Differences between the forecasts from perturbed and unperturbed assimilations are treated as random samples of background error distribution.

- Perfect model assumption → underestimated Errors

Perturbations of observations



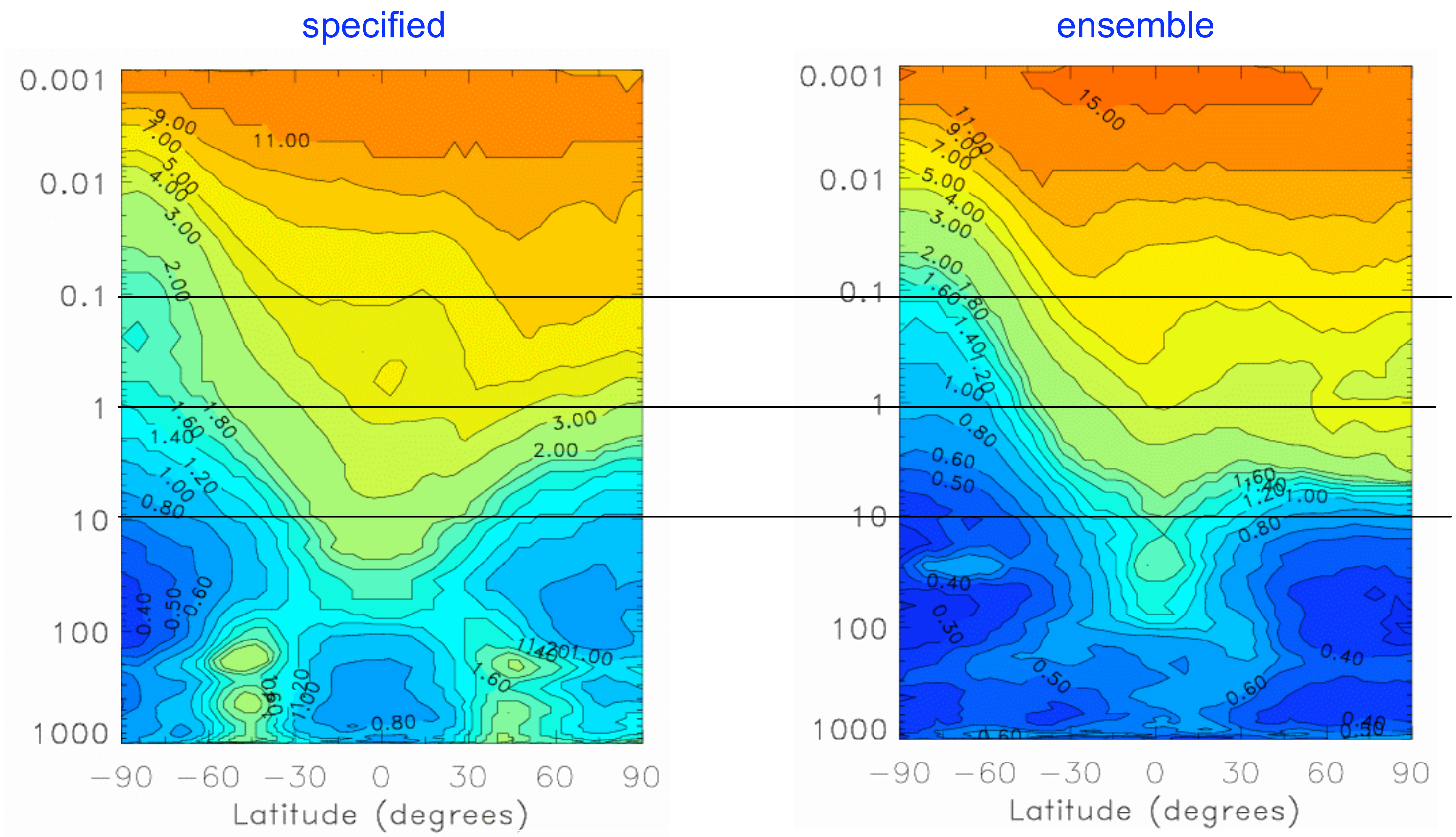
(31-7)*4 = 96 error samples

- Because of perturbations the differences between the forecasts grow
- After a few days of assimilation they stop growing

- No observations in mesosphere
- Perturbations of observations (below mesosphere) propagate to mesosphere

- In troposphere and stratosphere propagation of errors is controlled by analysis (observations) and dynamics
- In mesosphere – by dynamics only.

Ensemble derived error covariances: std for temperatures (January)



In troposphere:

- ♣ smaller variances
- ♣ Different pattern
- ♣ The pattern reflects
  - data rich/ data sparse regions
  - Dynamically active regions

In stratosphere:

♣ Similar pattern, Why?

♣The pattern reflects dynamically active regions

- Smaller variances in south pole region
- bigger values in north pole region (even bigger than specified)

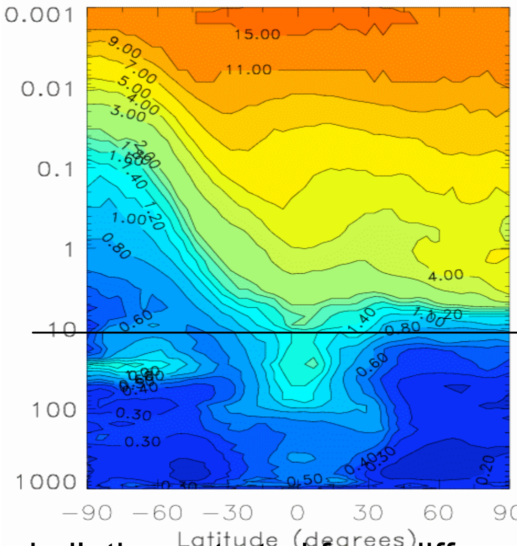
Satellite data ~ horizontally homogeneous

In mesosphere:

Climate variability

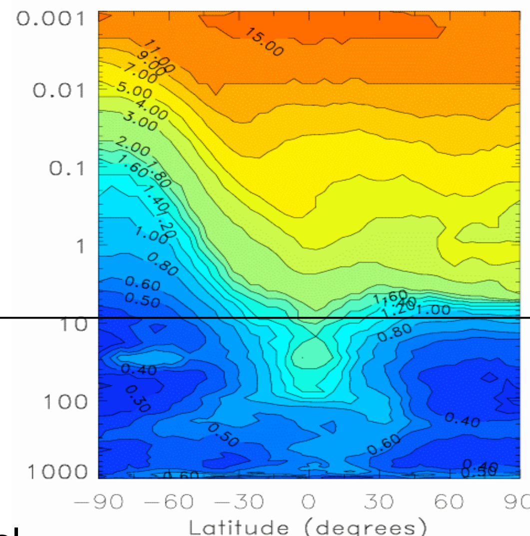
Modifications of ensemble method

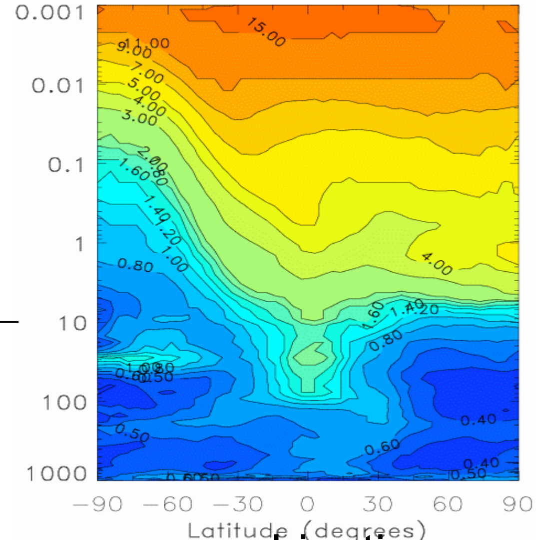
Derived error covariances: std for temperatures (January)



~=

<



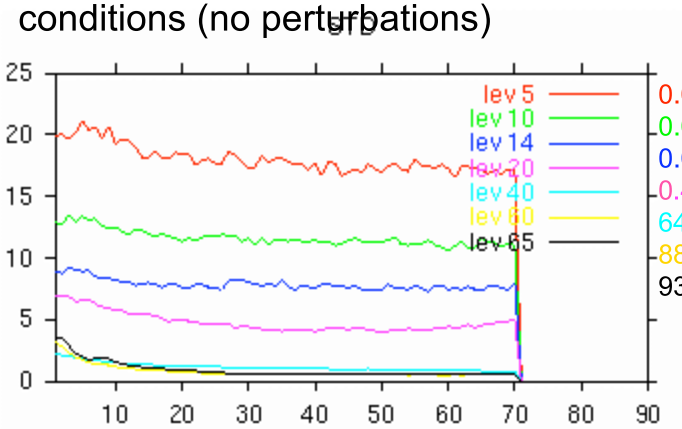


2 assimilations started from different initial conditions (no perturbations)

perturbed observations

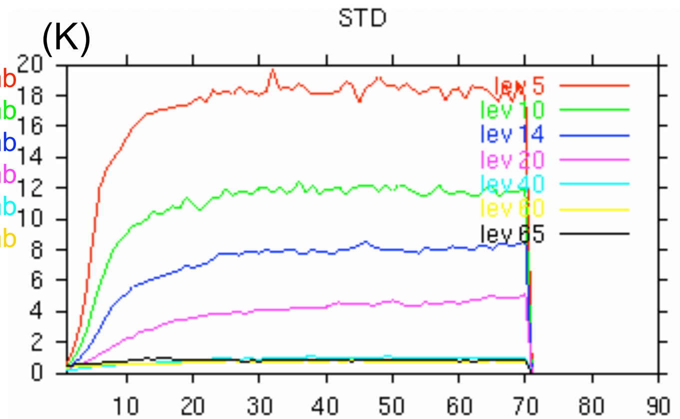
combination

Gives the maximum of the 2 methods



0.003mb
0.017mb
0.064mb
0.450mb
64.00mb
885.0mb
930mb

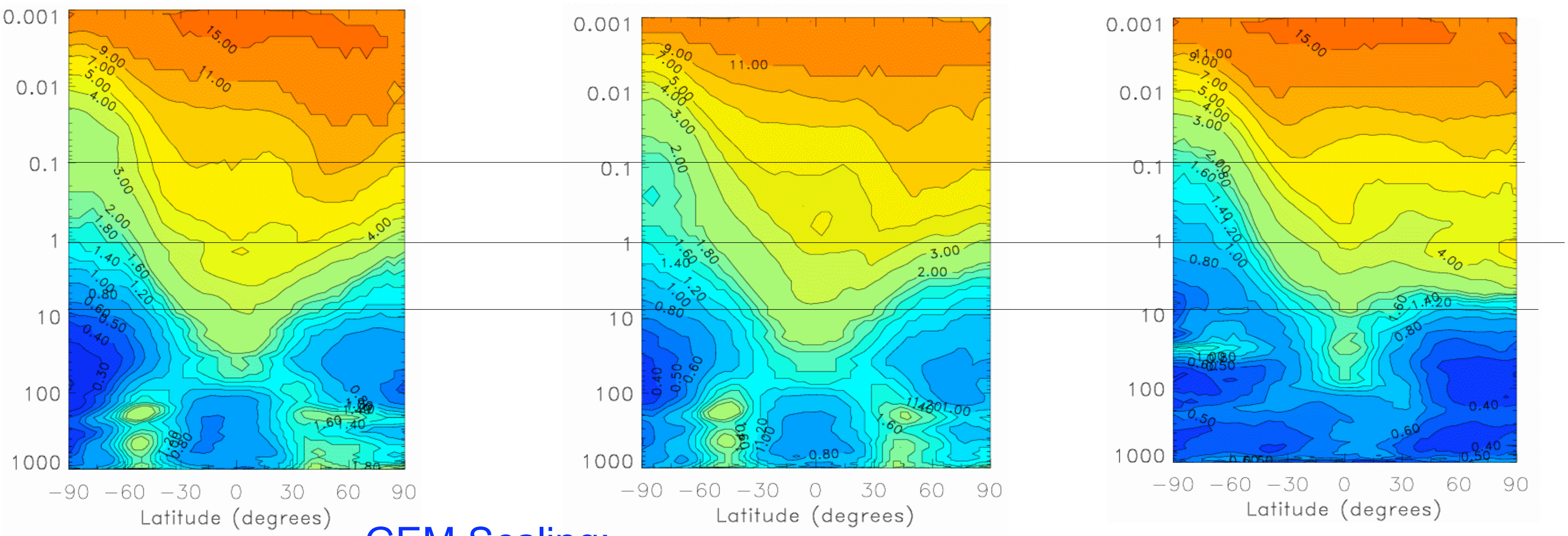
(K)



6 hour time step number

- In troposphere the differences decrease in few days
- In mesosphere the variances of differences almost do not change

Ensemble derived error covariances: std for temperatures (January)



6 hour diff → GEM Scaling: NMC to OmP → Specified

ensemble

- In most part of upper atmosphere (higher than 10mb) the pattern and the values of ensemble variances are similar to 6 hour diff. variances

→ 6 hour diff scheme is a very good proxy for this region

- In some regions in stratosphere ensemble variances are bigger than both specified and 6 hour diff. variances

→ In those regions 6 hour diff. variances were underestimated.

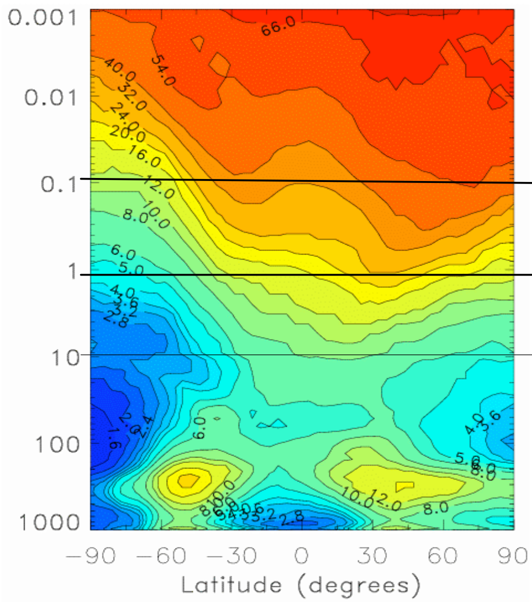
- In troposphere:
  - ♣ different pattern
  - ♣ Smaller values (on average)

What are real errors?

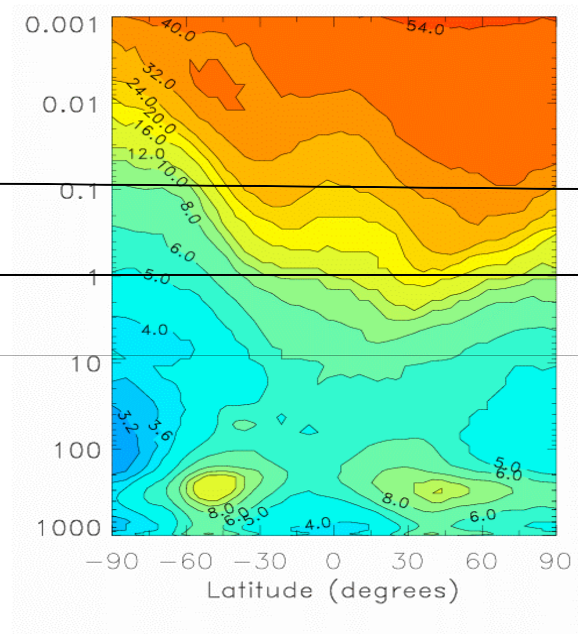
- Scale ensemble pattern to O-P
- Hybrid?

Ensemble derived error covariances: Std: VV (January)

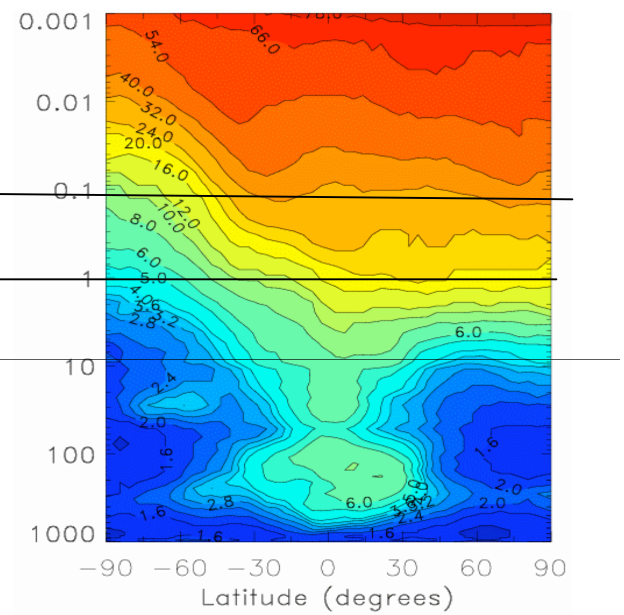
6 hour diff



Specified



ensemble



In mesosphere:
climate variability

In stratosphere:
The pattern reflects dynamically active regions

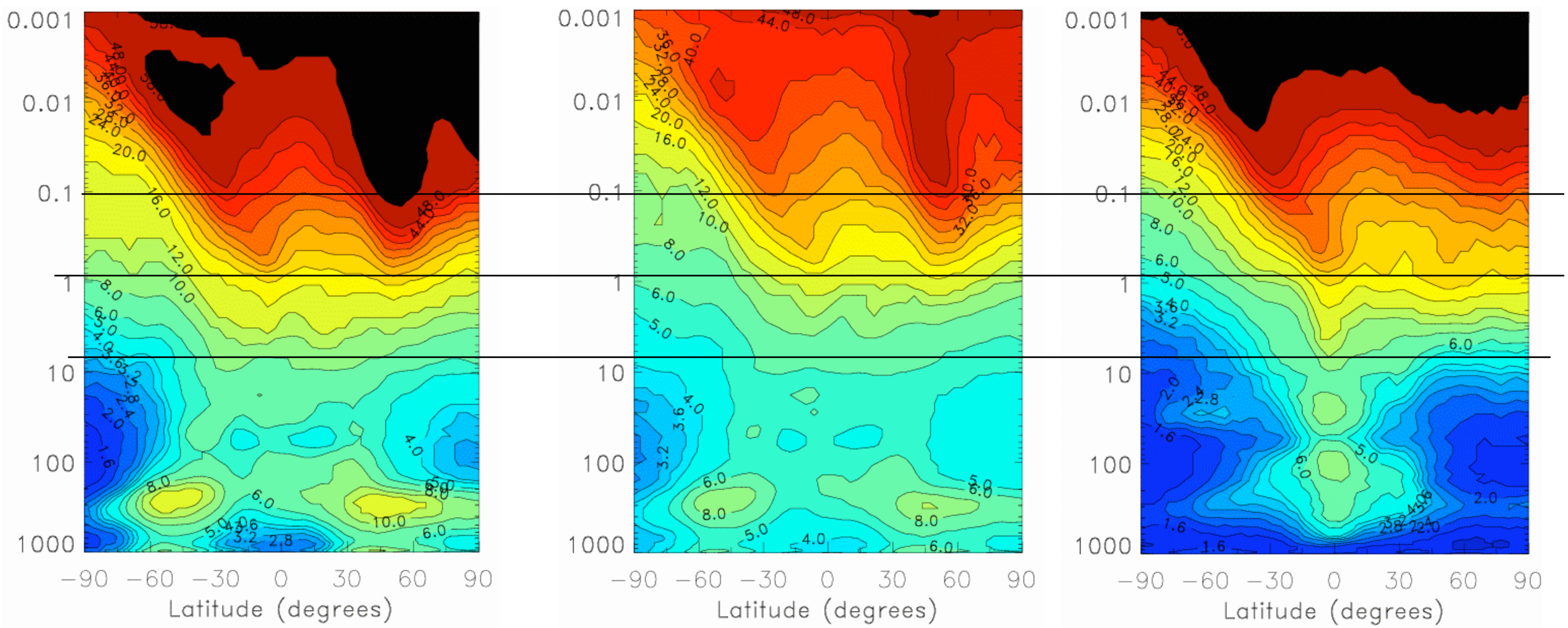
- bigger values in north pole region (near 1mb even bigger than specified)

In troposphere:

- smaller variances
- The pattern reflects
  - data rich/ data sparse regions
  - Dynamically active regions

Ensemble derived error covariances: UU std (January)

6 hour diff | Specified | ensemble



In mesosphere:
climate variability

In stratosphere:
The pattern reflects dynamically active regions

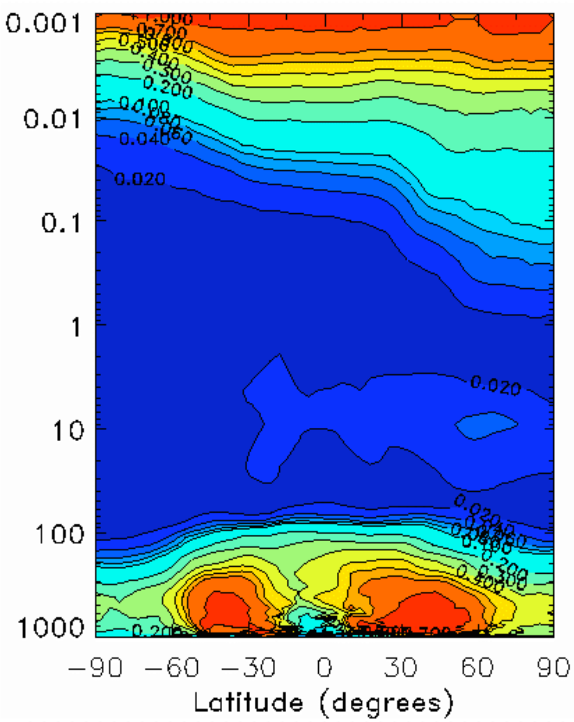
- Smaller variances in south pole region
- the maximum in equatorial region
- bigger values in north pole region (even bigger than specified)

In troposphere:

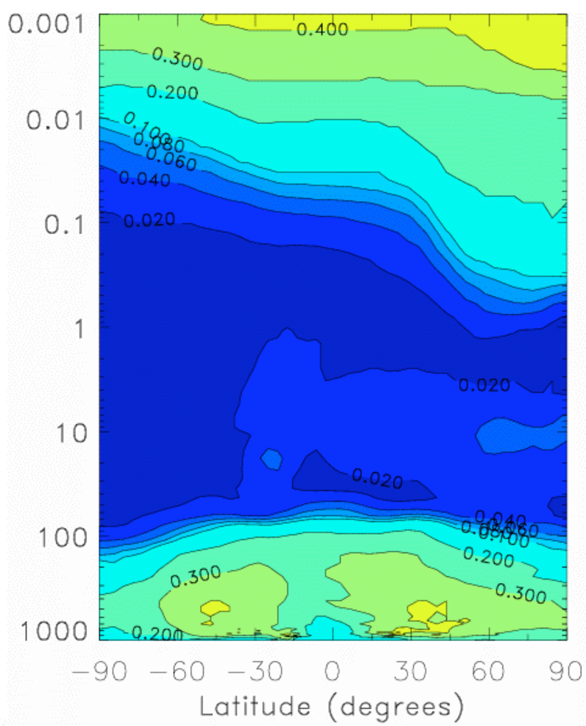
- ♣ smaller variances
- ♣ The pattern reflects
  - data rich/ data sparse regions
  - Dynamically active regions

Ensemble derived error covariances: Std: LQ (January)

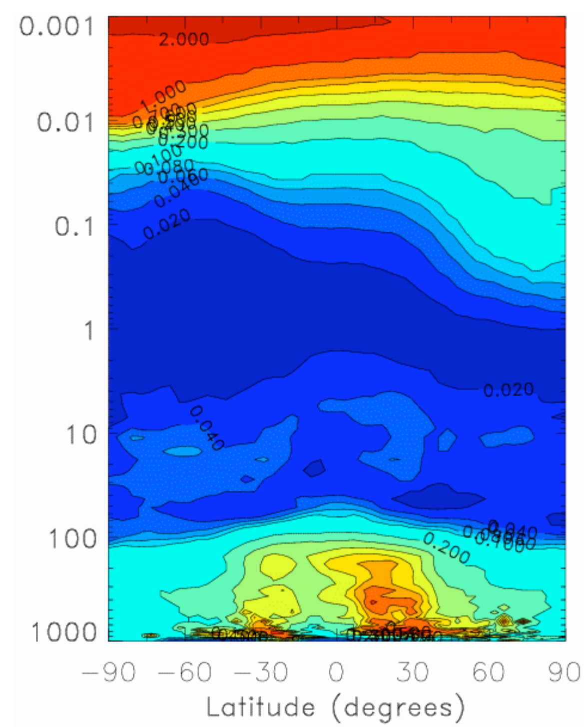
6 hour diff

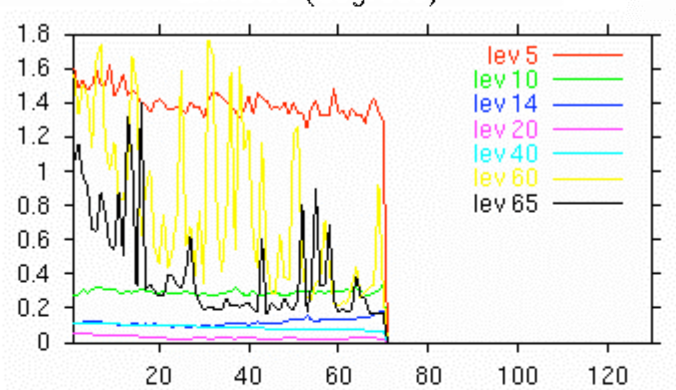


specified



ensemble

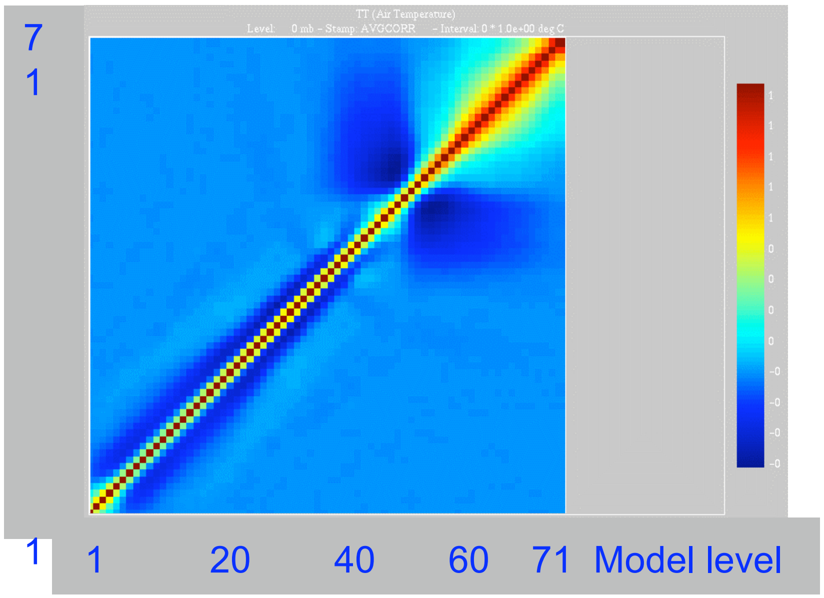




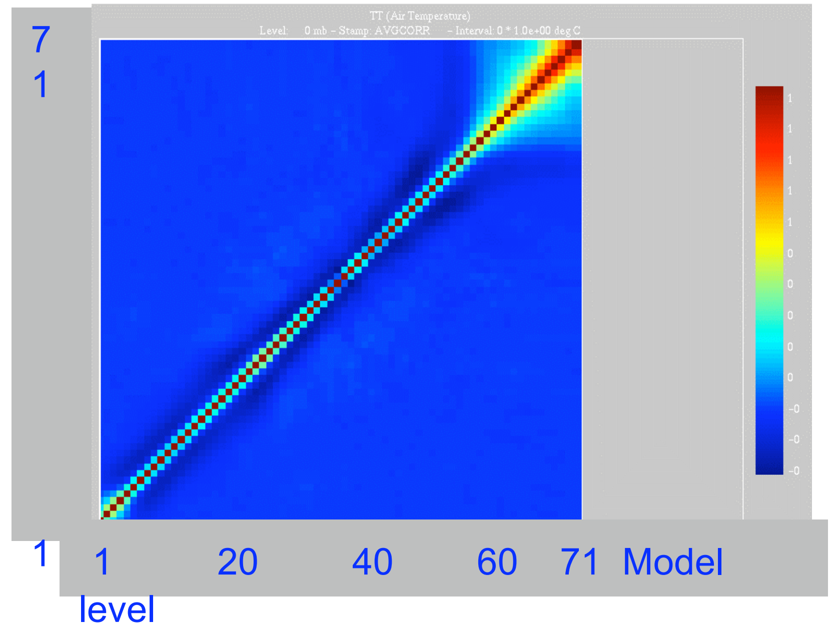
Ensemble variances are very different from specified but rather close to 6 hour diff. variances.

Average vertical correlations (temperature)

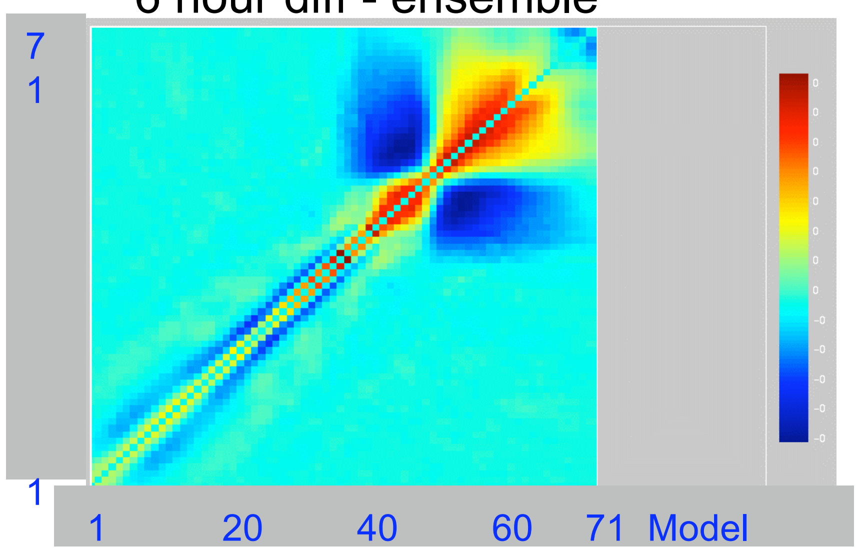
6 hour diff



ensemble



6 hour diff - ensemble



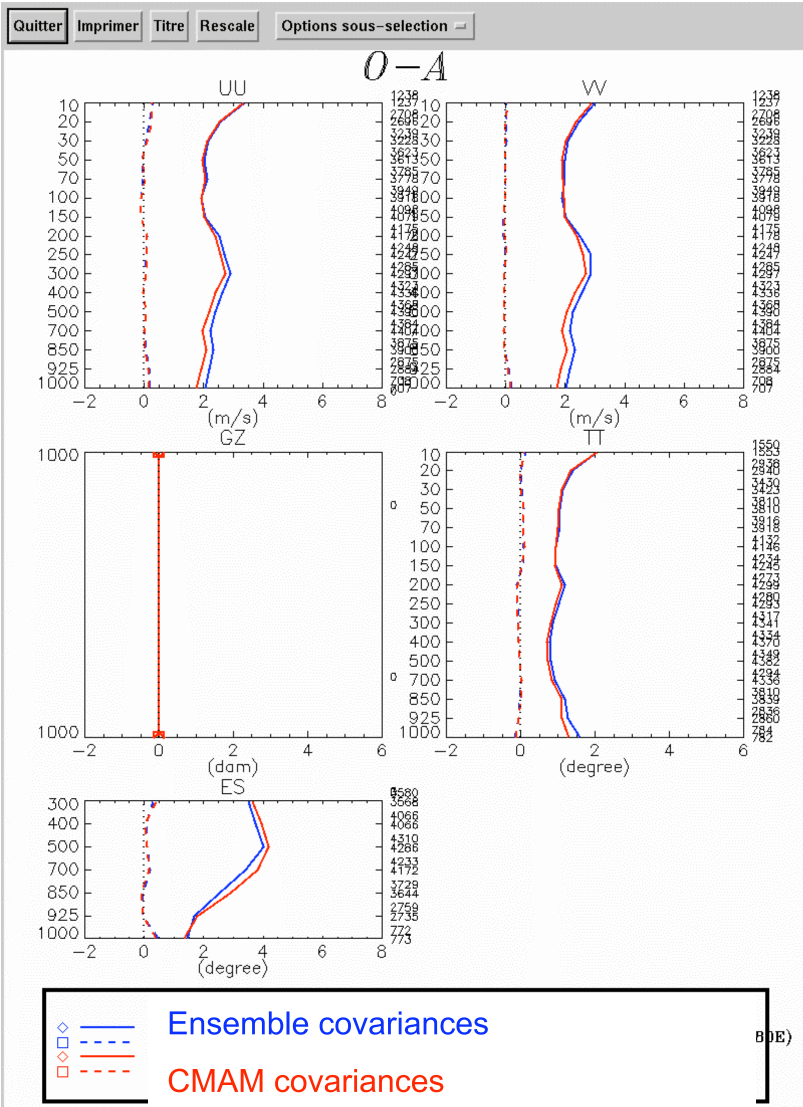
Ensemble correlations:

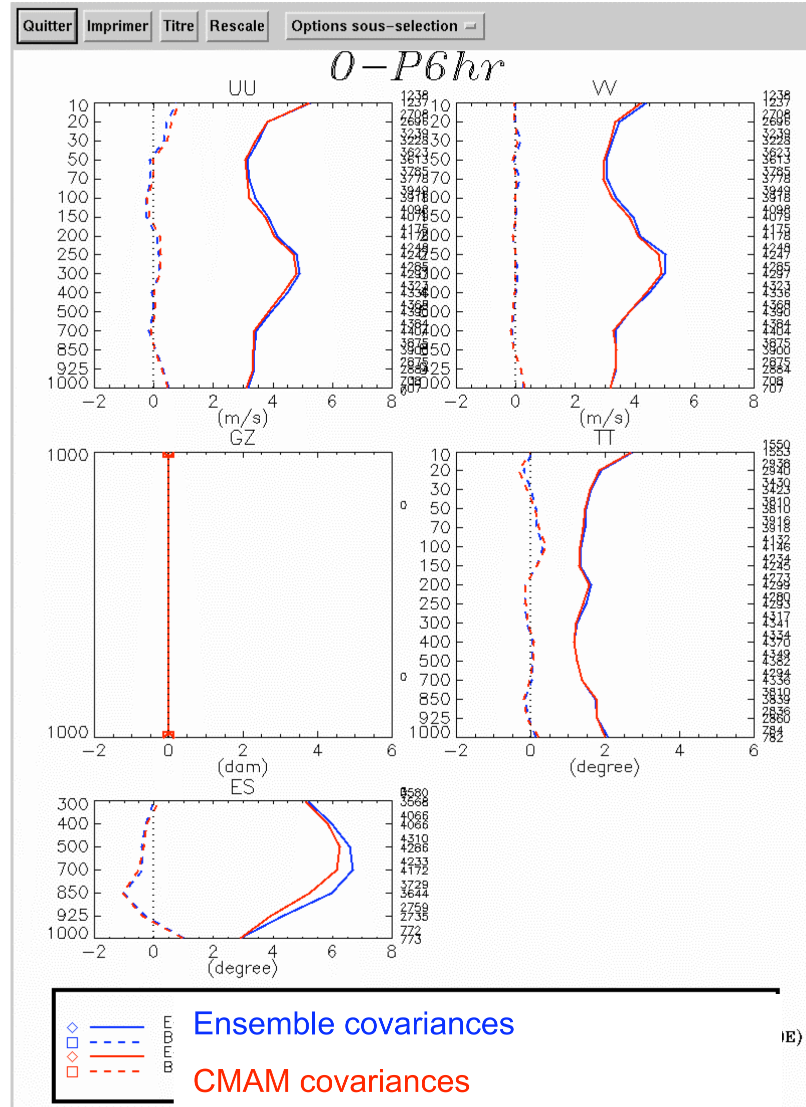
- narrower than specified
- Less negative

conclusions

- New background error covariances for CMAM were derived using ensemble method.
- As could be expected
  - The pattern of ensemble variances reflects the effects of observations density and dynamical activity.
  - Ensemble produced correlations are narrower and ‘less negative’.
- As could not be expected
  - In upper atmosphere the 6 hour differences scheme (Yves Rochon’s method) and Ensemble method result to rather close variances (by values and by pattern).
- For future:
  - To account for model error
  - To consider different ways of evaluation
  - To learn the impact (on scores) of variances and correlations separately
  - To increase ensemble size

Verification of Ensemble error covariances
compare against sonds, global average



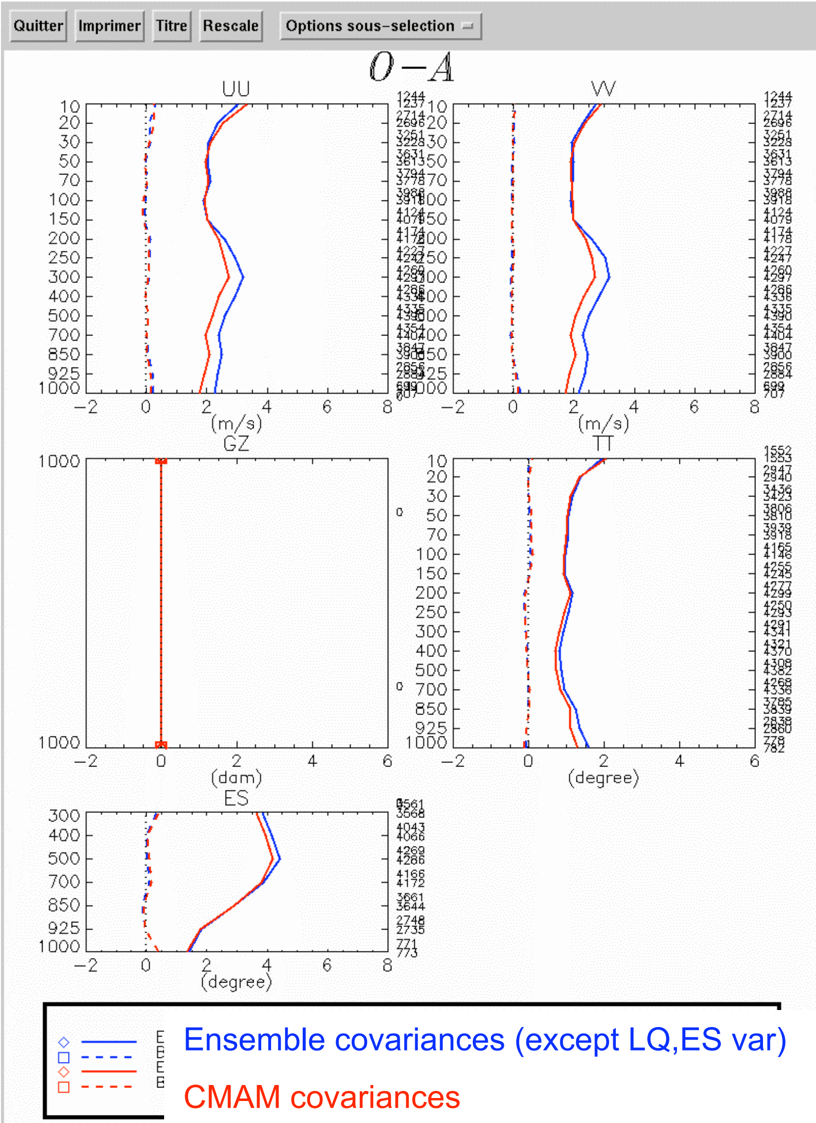


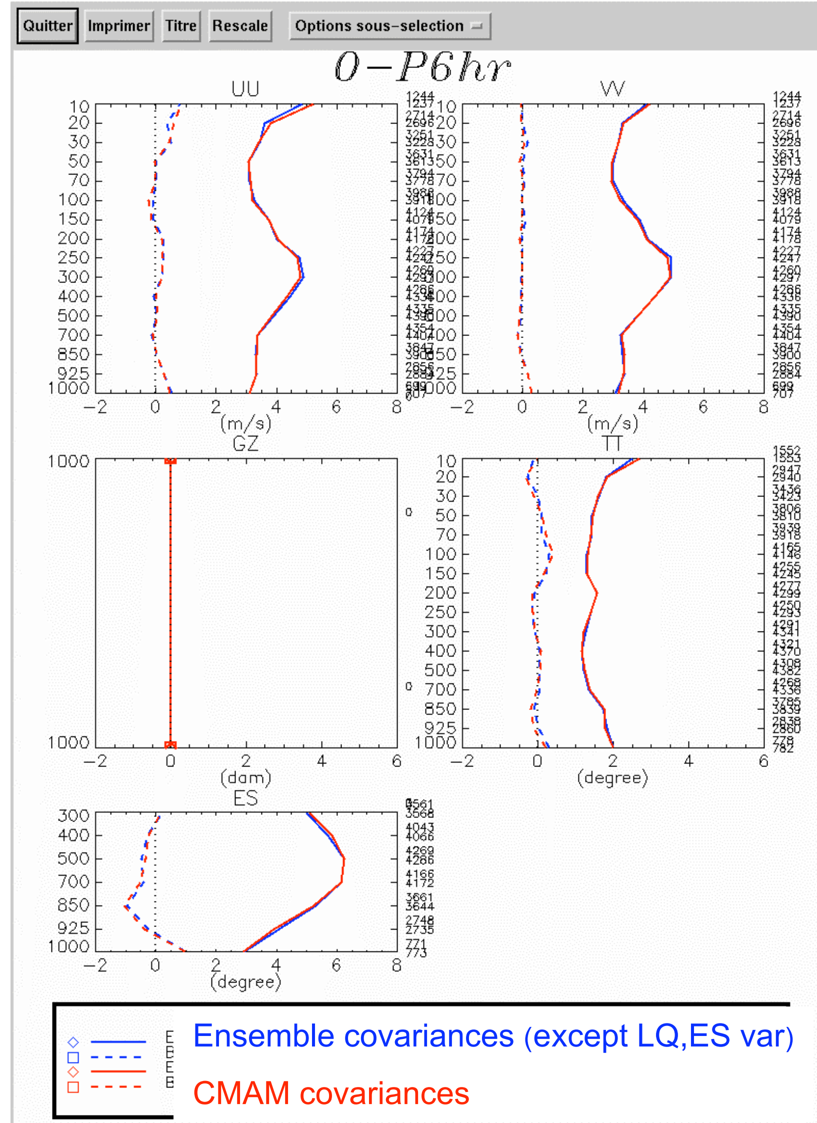
Since Ensemble TT,UU,VV variances are smaller in troposphere, analysis is farer from observations (blue curve) than in CMAM case (red curve)

O-P scores for ensemble covariances are worse than for CMAM

Verification of Ensemble error covariances: scores in respect to radiosond obs

New error covariances with old variances for LQ, ES (variances are scaled by 1.41)



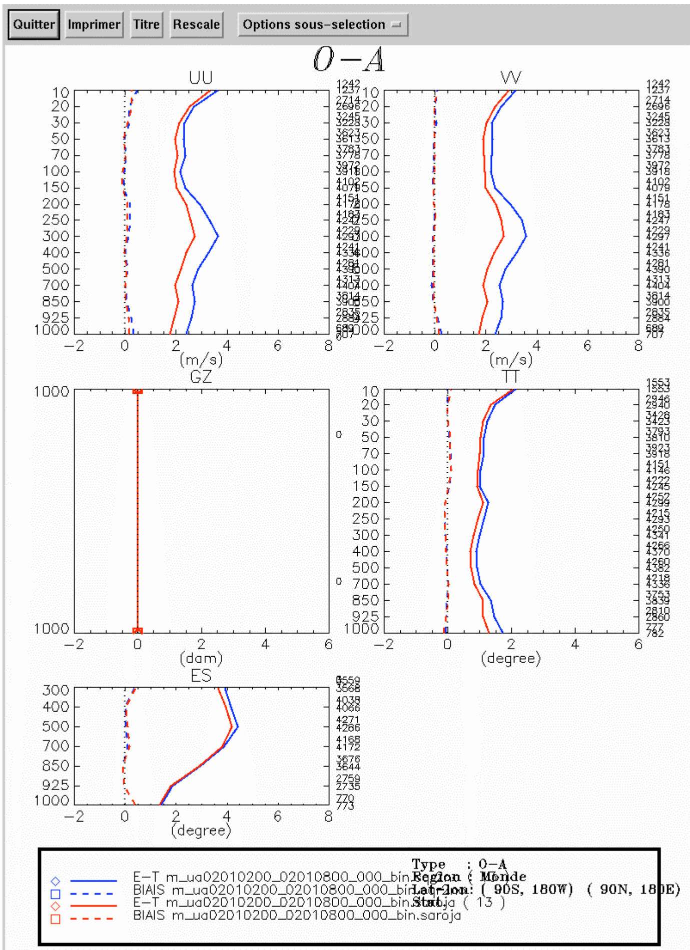


Now all ensemble covariances are smaller than specified in CMAM, thus the difference in O-A scores is bigger than in case with new LQ,ES var

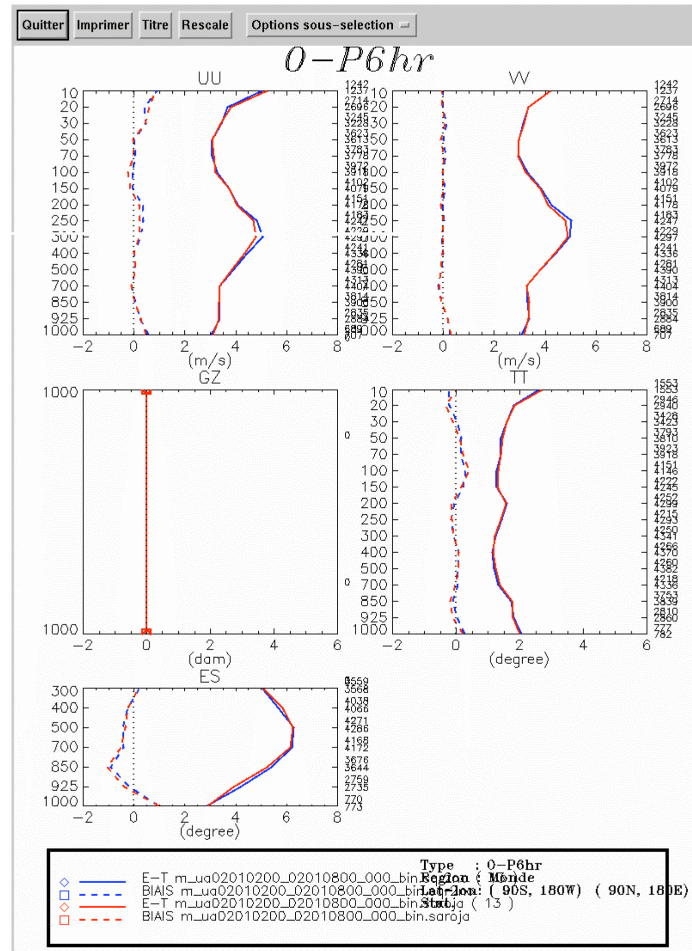
But O-P scores are the same

Verification of Ensemble error covariances: scores in respect to radiosond obs

New error covariances with old variances for LQ, ES



Bigger difference in O-A



Almost no difference in O-P

Why the scores are the same?

Does it mean that ensemble error covariances are as far from reality as currently specified?

Check the scores for new error covariances with different scaling factors (from 1 to 2) applied to variances.

Does it mean that current DA system is not sensitive to error covariances specification?

Do the scores matter?

Other ways of evaluation?