Convergence and Stability of Estimated Error Variances derived from Assimilation Residuals in Observation Space

June 16, 2008

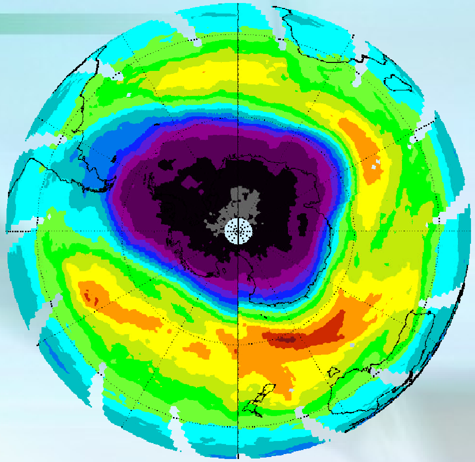
Richard Ménard
Yan Yang
Yves Rochon

Air Quality Research Division,
Atmospheric Science & Technology Directorate,
Environment Canada, Canada

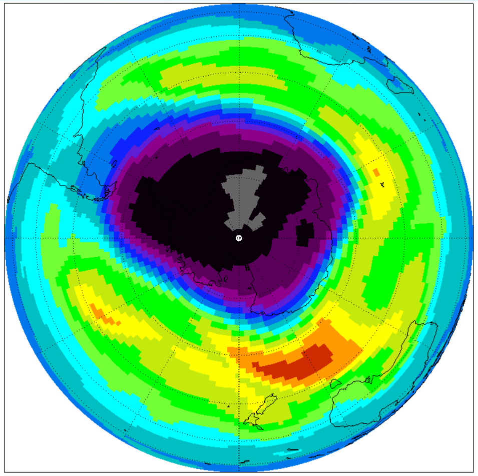
Context

- Environment Canada NWP
  + online stratospheric chemistry
  - BIRA 57 advected species
  - LINOZ (aka Cariolle)
  - LINOZ2 (O3, N2O, CH4, tendecies + parametrization of heterogeneous chem
  - FASTOC (High Dimensional Model Representation)
- (advected) ozone and water vapor radiation interaction
- 3D and 4DVar assimilation with possibilities of cross dynamics-chemistry coupling with balanced operators
- Extending the chemistry into the troposphere

TOMS



GEM-BACH no chem assim



30 September 2003

- Assimilation of all meteorological data + limb sounding observations of MIPAS/ENVISAT (T, H_2O, O_3, CH_4, HNO_3, H_2O, N_2O, $ClONO_2$)

GOAL

- Online estimation of observation and background error variance as a function of height (global mean or three regions, height) using O-F, O-A, A-F statistics
  - no chemical-radiation interaction, no coupling between meteorological and chemical error statistics

Convergence – scalar case

A - Iteration on observation error

$$\left\langle (O-A)(O-F)^T \right\rangle = \overline{\mathbf{R}}(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}})^{-1}(\mathbf{HBH} + \mathbf{R})$$

where $\left\langle (O-F)(O-F)^T \right\rangle = \mathbf{HBH}^T + \mathbf{R}$ is obtained from assimilation residuals and overbar denotes *prescribed* error covariances

i)- Correctly prescribed forecast error variance

$\overline{\mathbf{B}} = \mathbf{B} = \sigma_f^2$ $\overline{\mathbf{R}} = \alpha\mathbf{R} = \alpha\sigma_o^2$ optimal value $\alpha = 1$

$$\left\langle (O-A)(O-F) \right\rangle = \frac{\alpha\,\sigma_o^2}{\alpha\,\sigma_o^2 + \sigma_f^2}(\sigma_o^2 + \sigma_f^2) = \alpha\,\sigma_o^2\left(\frac{\gamma + 1}{\alpha\,\gamma + 1}\right)$$

where $\gamma = \dfrac{\sigma_o^2}{\sigma_f^2}$

Convergence – scalar case

let $\langle (O-A)(O-F) \rangle = \alpha_{n+1}\sigma_o^2$ be the next iterate

so the iteration on α_n takes the form

$$\alpha_{n+1} = \alpha_n \left(\frac{\gamma+1}{\alpha_n\gamma+1} \right) = G(\alpha_n)$$

Define a *mapping G*

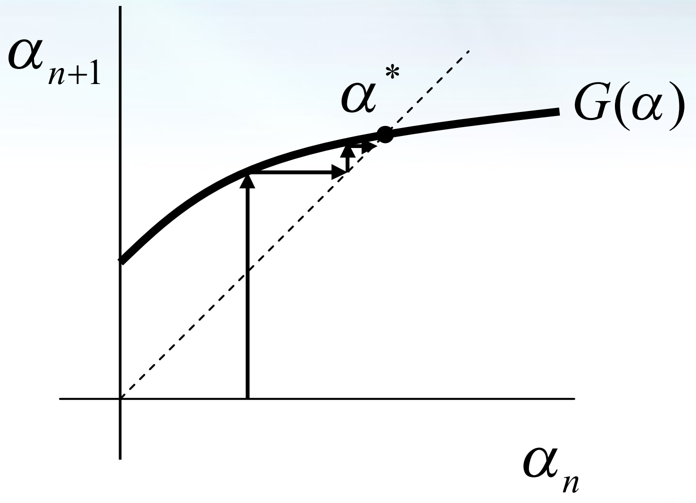
$$G(\alpha) = \alpha \left(\frac{\gamma+1}{\alpha\gamma+1} \right)$$

The fixed-point is

$$\alpha^* = G(\alpha^*)$$

condition for convergence

$$|G'(\alpha^*)| < 1$$



Convergence – scalar case

and so for this case we get $\alpha^* = 1$

$$G'(\alpha^*) = \frac{1}{\gamma+1} = \frac{\sigma_f^2}{\sigma_o^2 + \sigma_f^2} = K \le 1$$

the scheme is always convergent and converges to the true value, $\alpha = 1$

ii)- Incorrectly prescribed forecast error variance

$$\overline{\mathbf{B}} = \beta \mathbf{B} = \beta \sigma_f^2 \qquad \overline{\mathbf{R}} = \alpha \mathbf{R} = \alpha \sigma_o^2$$

the mapping is now different

$$\alpha_{n+1} = \alpha_n \left(\frac{\gamma+1}{\alpha_n \gamma + \beta} \right) = G(\alpha_n)$$

Convergence – scalar case

The fixed-point is

$$\alpha^* = 1 + \frac{1-\beta}{\gamma} = 1 + \frac{\left(\sigma_f^2 - \overline{\sigma}_f^2\right)}{\sigma_o^2}$$

that is *not the true observation error* value.

- If forecast error variance is underestimated, obs error is overestimated
- If forecast error variance is overestimated, obs error is underestimated

$$G'(\alpha^*) = \frac{\beta}{\gamma + 1} = \frac{\beta\sigma_f^2}{\sigma_o^2 + \sigma_f^2}$$

Will not converge if: $\beta\sigma_f^2 = \overline{\sigma}_f^2 > \sigma_o^2 + \sigma_f^2$

In practice the estimated forecast error variance will never be larger than the innovation error variance, so for all practical cases the scheme converges.

Convergence – scalar case

B - Iteration on forecast error

$$\left\langle (A-F)(O-F)^T \right\rangle = \mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T+\overline{\mathbf{R}})^{-1}(\mathbf{HBH}+\mathbf{R})$$

i)- Correctly prescribed observation error variance

$$\overline{\mathbf{R}} = \mathbf{R} = \sigma_o^2 \qquad \overline{\mathbf{B}} = \beta\,\mathbf{B} = \beta\,\sigma_f^2 \qquad \beta_{n+1} = \beta_n\left(\frac{\gamma+1}{\gamma+\beta_n}\right) = F(\beta_n)$$

converges to the true forecast error variance $\beta^* = 1$

ii)- Incorrectly prescribed observation error variance

$$\overline{\mathbf{R}} = \alpha\mathbf{R} = \alpha\,\sigma_o^2 \qquad \overline{\mathbf{B}} = \beta\,\mathbf{B} = \beta\,\sigma_f^2 \qquad \beta_{n+1} = \beta_n\left(\frac{\gamma+1}{\alpha\gamma+\beta_n}\right) = F(\beta_n)$$

for most practical cases will converge, but to the wrong value

$$\beta^* = 1+\gamma\,(1-\alpha) = 1+\frac{\left(\sigma_0^2 - \overline{\sigma}_o^2\right)}{\sigma_f^2}$$

If observation error variance is underestimated,
→ forecast error is overestimated
if observation error variance is overestimated,
→ forecast error is underestimated

Lagged-innovation covariance

An optimal KF has a lagged-innovation covariance equal to zero (Daley 1992)

$$\mathbf{C}_{k+1}^{k} = \left\langle (O-F)_k (O-F)_{k+1}^T \right\rangle$$
$$= \mathbf{H}_{k+1}\mathbf{M}_k \left[\mathbf{B}_k \mathbf{H}_k^T - \overline{\mathbf{B}}_k \mathbf{H}_k^T (\mathbf{H}_k \overline{\mathbf{B}}_k \mathbf{H}_k^T + \overline{\mathbf{R}}_k)^{-1} (\mathbf{H}_k \mathbf{B}_k \mathbf{H}_k^T + \mathbf{R}_k) \right]$$
$$\approx \mathbf{H}_k \mathbf{B}_k \mathbf{H}_k^T - \mathbf{H}_k \overline{\mathbf{B}}_k \mathbf{H}_k^T (\mathbf{H}_k \overline{\mathbf{B}}_k \mathbf{H}_k^T + \overline{\mathbf{R}}_k)^{-1} (\mathbf{H}_k \mathbf{B}_k \mathbf{H}_k^T + \mathbf{R}_k)$$

for a very small time step $\mathbf{M}_k = \mathbf{I} + \Delta t\, \Phi_k \approx \mathbf{I}$

The mean lagged-innovation in the *n*-th assimilation cycle

$$\left\langle \mathbf{C}_{k+1}^{k} \right\rangle \approx \mathbf{HBH}^T - \mathbf{H}\hat{\mathbf{B}}_n \mathbf{H}^T$$

Even if $\hat{\mathbf{B}}_n$ converges, it may not converge to the truth, and then the the lagged-innovation covariance converges to a non zero value

Remark

A - Iteration on observation error variance

- If forecast error variance is underestimated, obs error is overestimated
- If forecast error variance is overestimated, obs error is underestimated

B - Iteration on forecast error variance

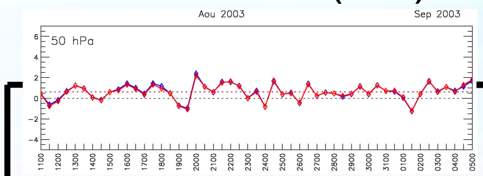
If observation error variance is underestimated,
→ forecast error is overestimated
if observation error variance is overestimated,
→ forecast error is underestimated

Are we getting anywhere if we iterate both, observation and forecast errors ?

Stability of the scheme

- Cycling the assimilation

Assimilation (*n-1*)



new set

A_{n-1}

F_{n-1}

Estimate the error statistics

$$\left\langle (O-A_{n-1})(O-F_{n-1})^T \right\rangle = \overline{\mathbf{R}}_n$$

$$\left\langle (A_{n-1}-F_{n-1})(O-F_{n-1})^T \right\rangle = \mathbf{H}\overline{\mathbf{B}}_n\mathbf{H}^T$$

reconduct an assimilation over a period of a month with new error stats

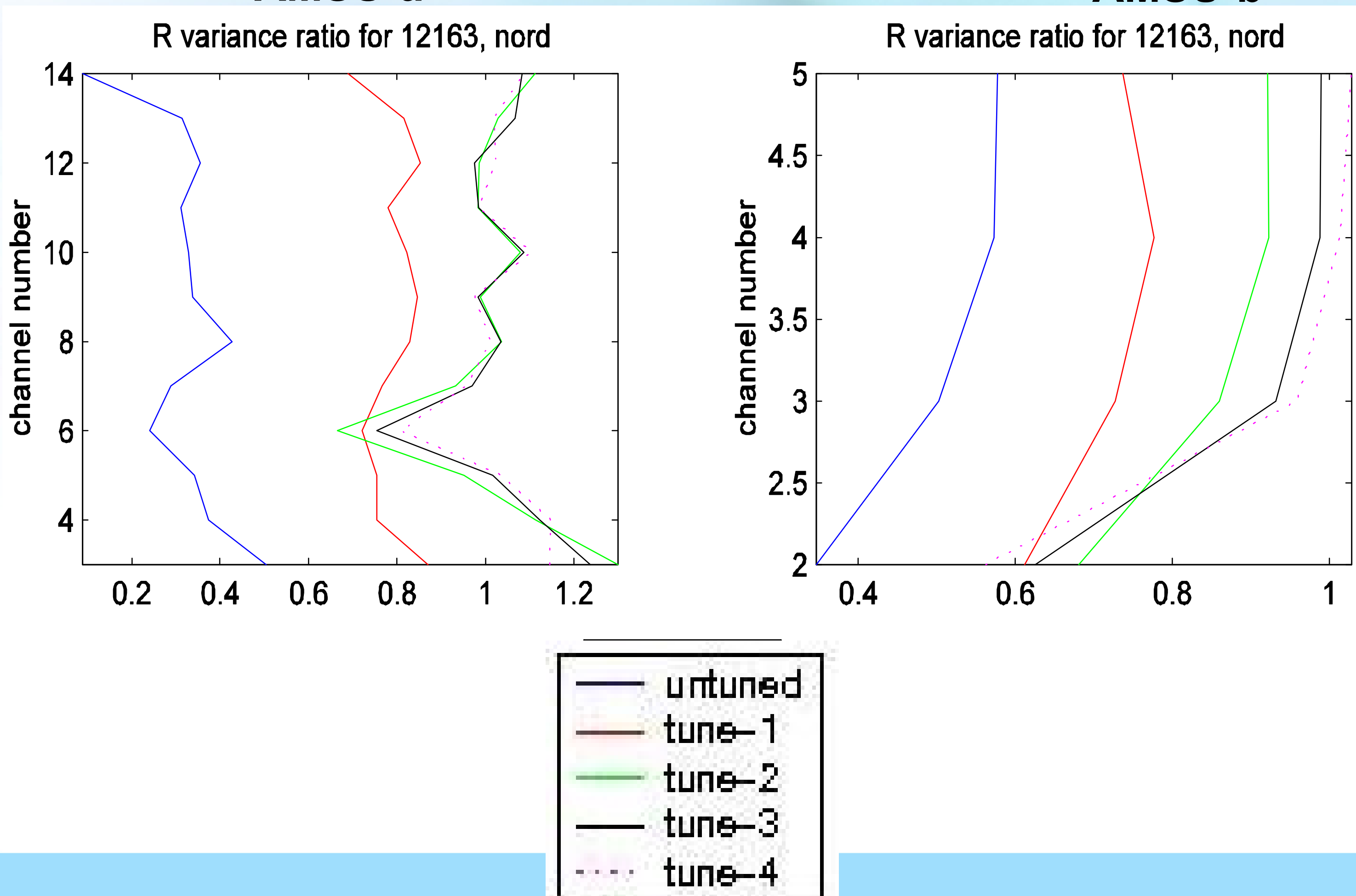
- All the statistics O-A, O-F, A-F are derived from assimilation residuals, and iterated on the whole assimilation cycle, so the O-A, O-F, A-F are diagnostics of the assimilation system

- With this scheme, we also have a *testbed for online estimation of error statistics* (in perpertual mode)

Iteration on observation error variance

Error variance (n) / reference error variance

AMSU-a **AMSU-b**

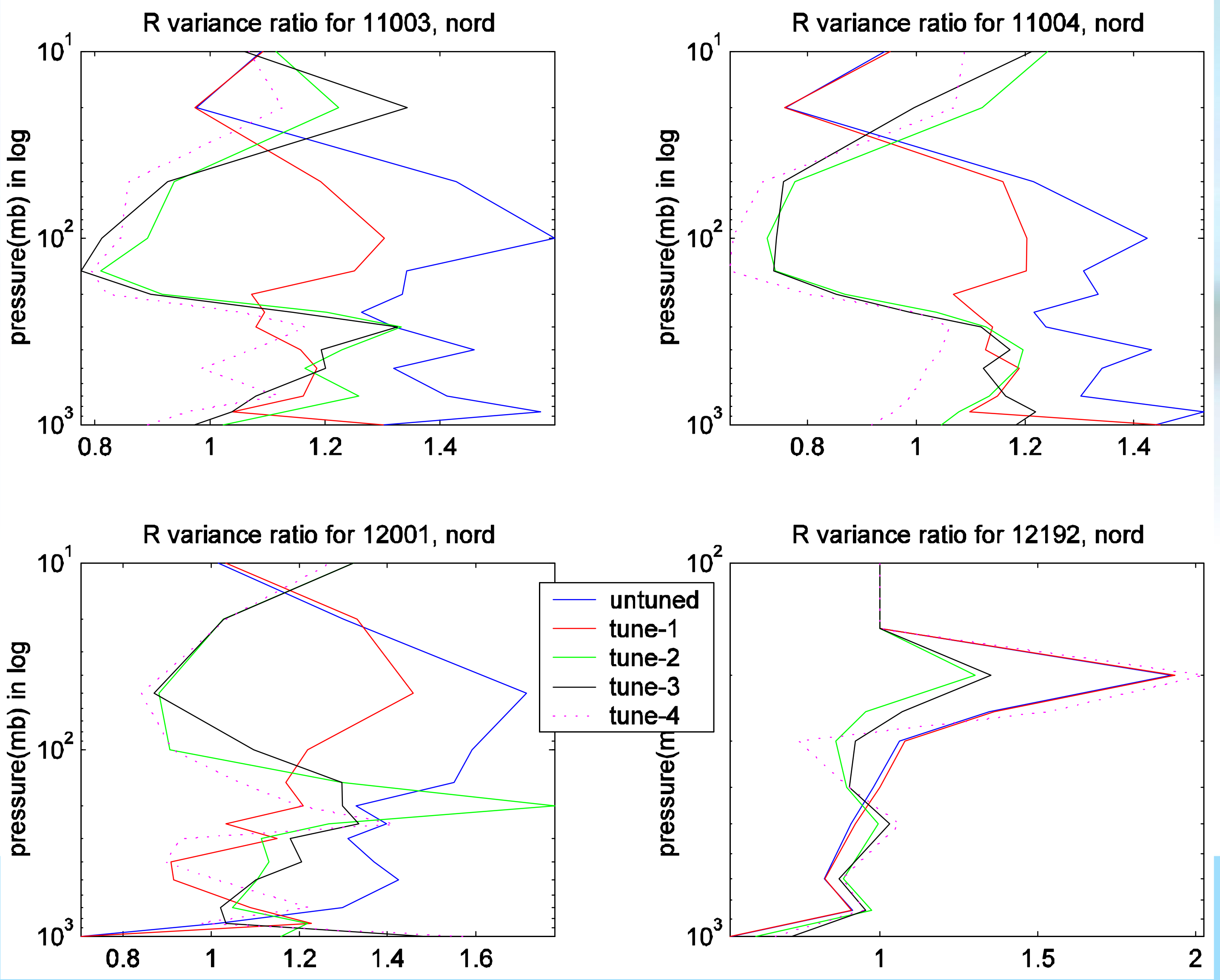


Iteration on observation error variance

Error variance (n) / reference error variance

RAOBS U

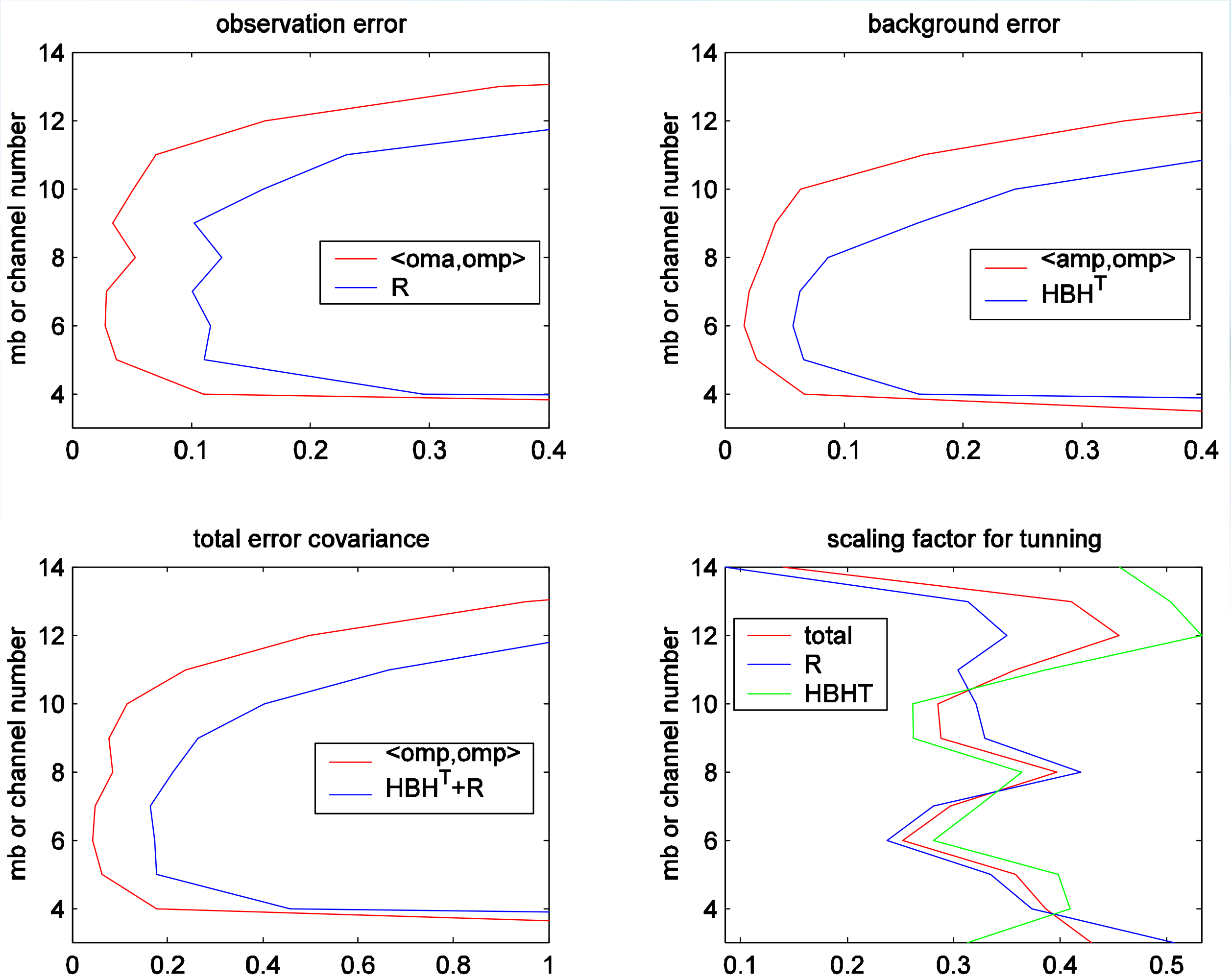
RAOBS V



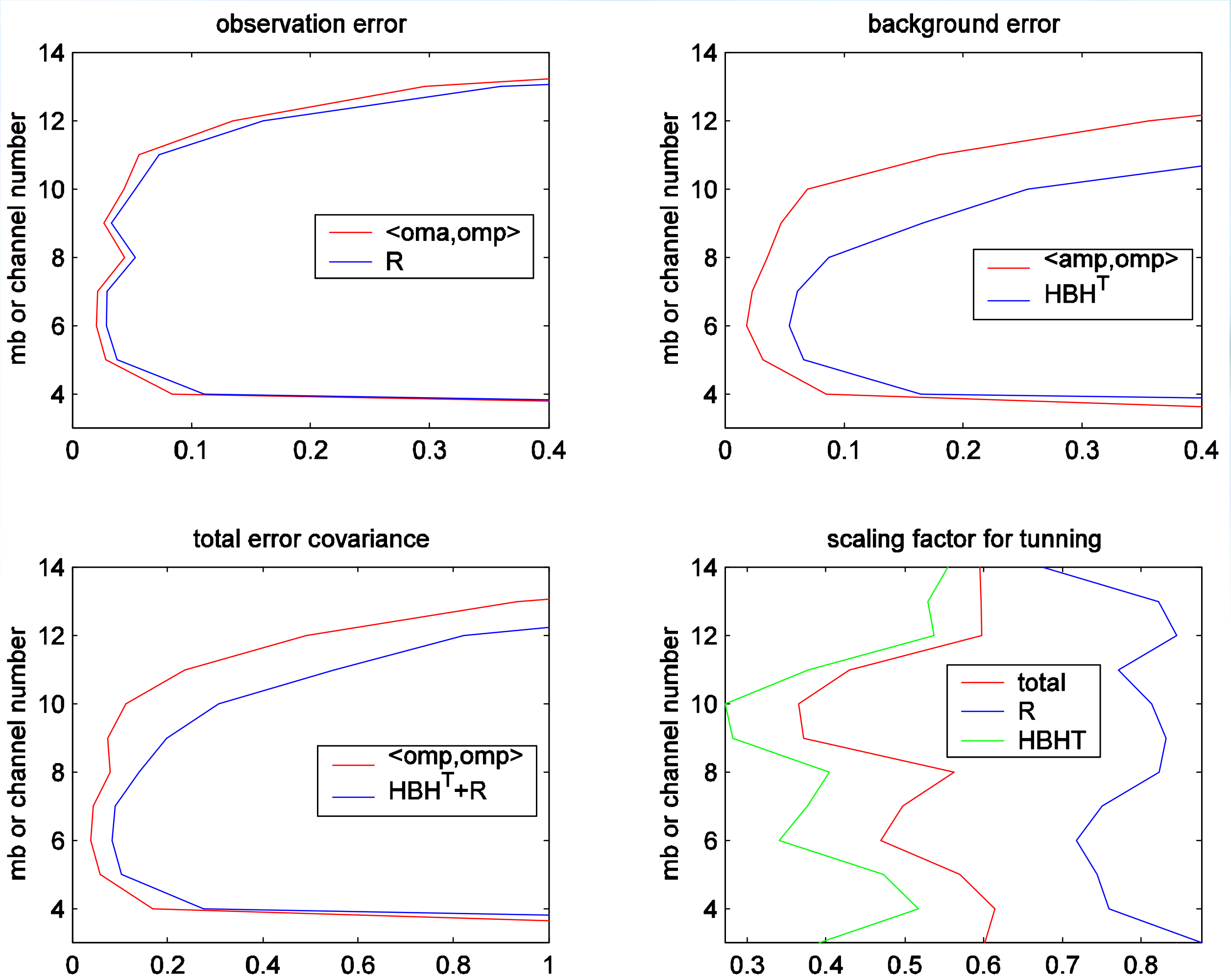
RAOBS T

RAOBS T-Td

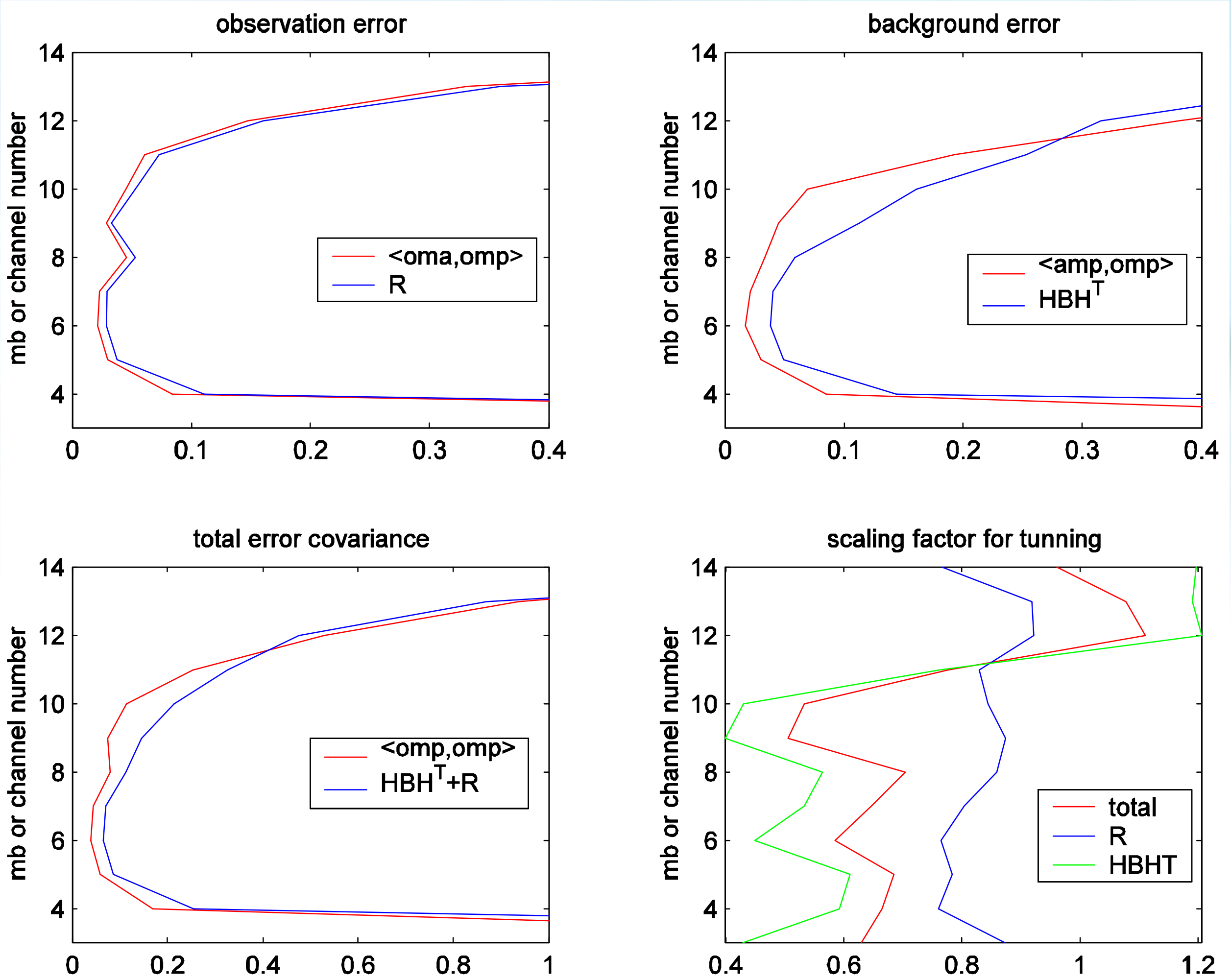
AMSU-a
Before any iteration



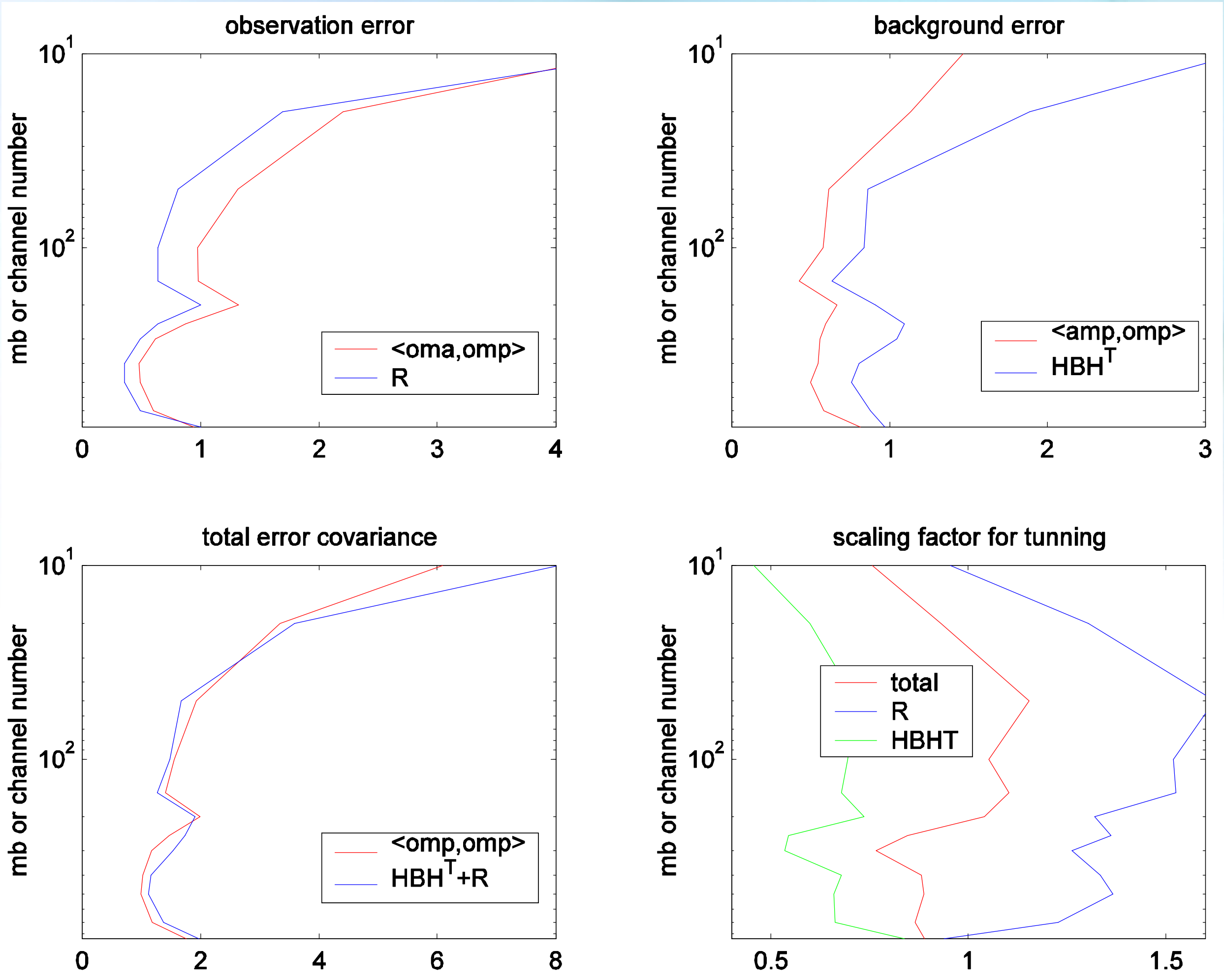
AMSU-a
one iteration on observation error



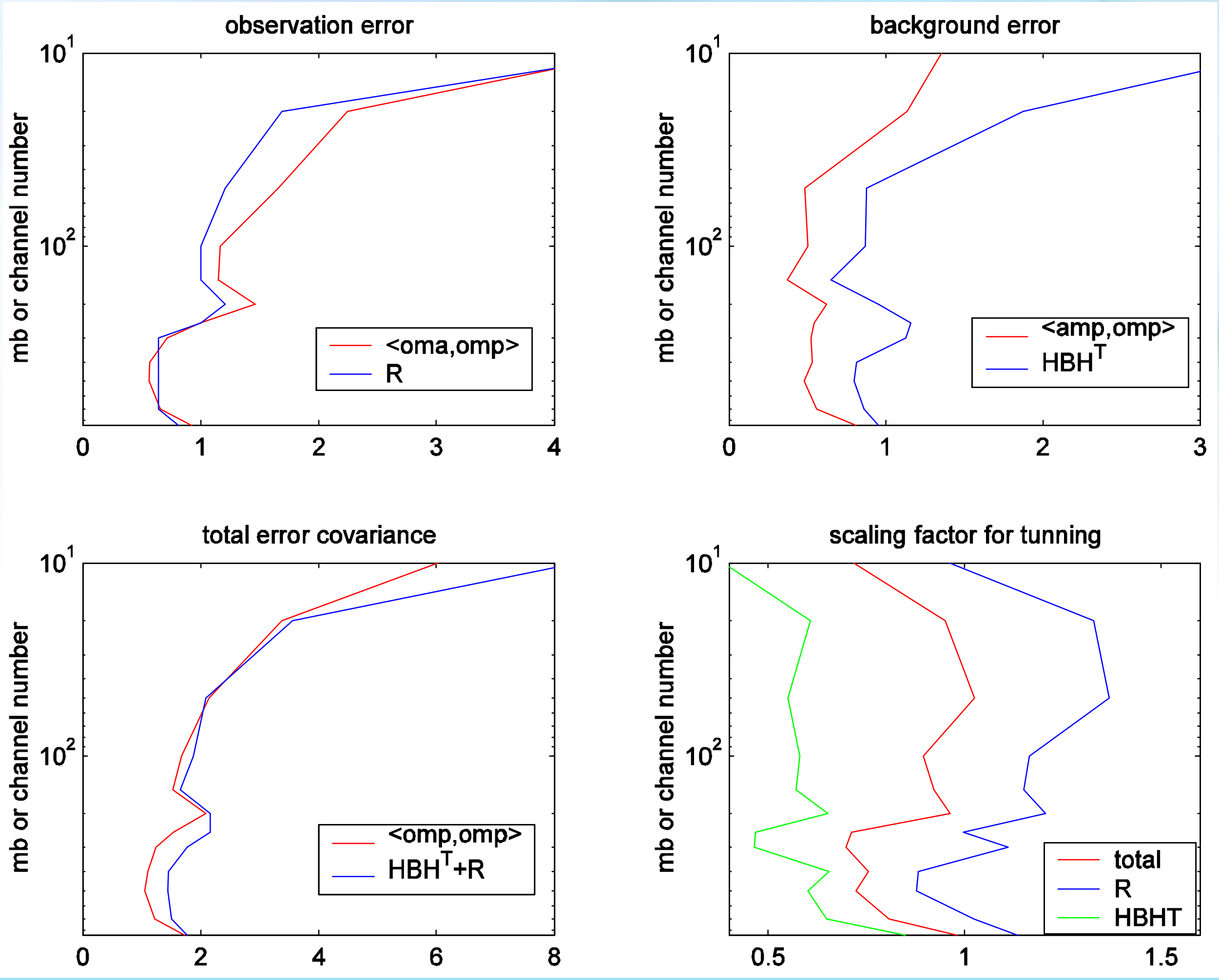
AMSU-a
one iteration on observation and background error



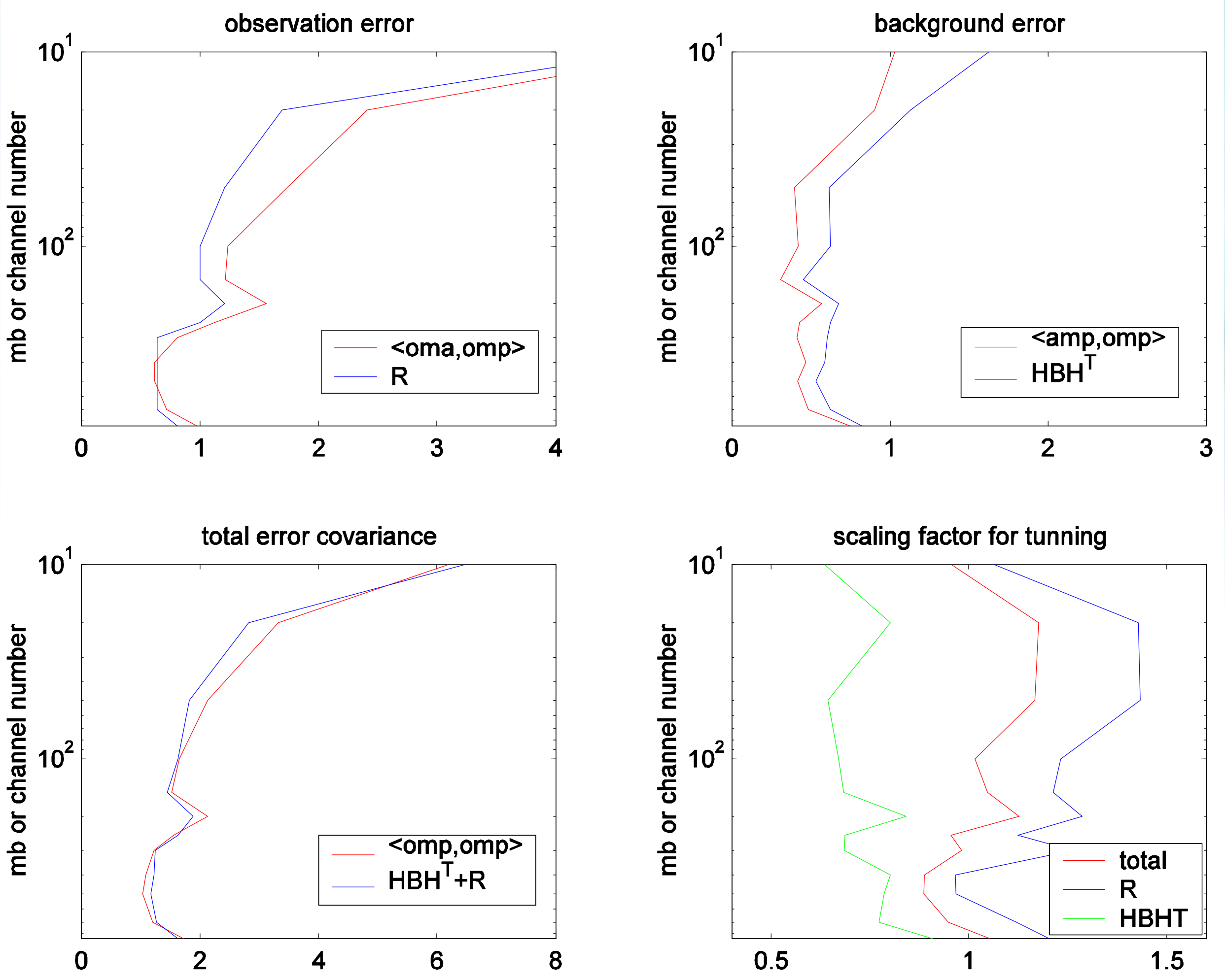
RAOBS T
Before any iteration



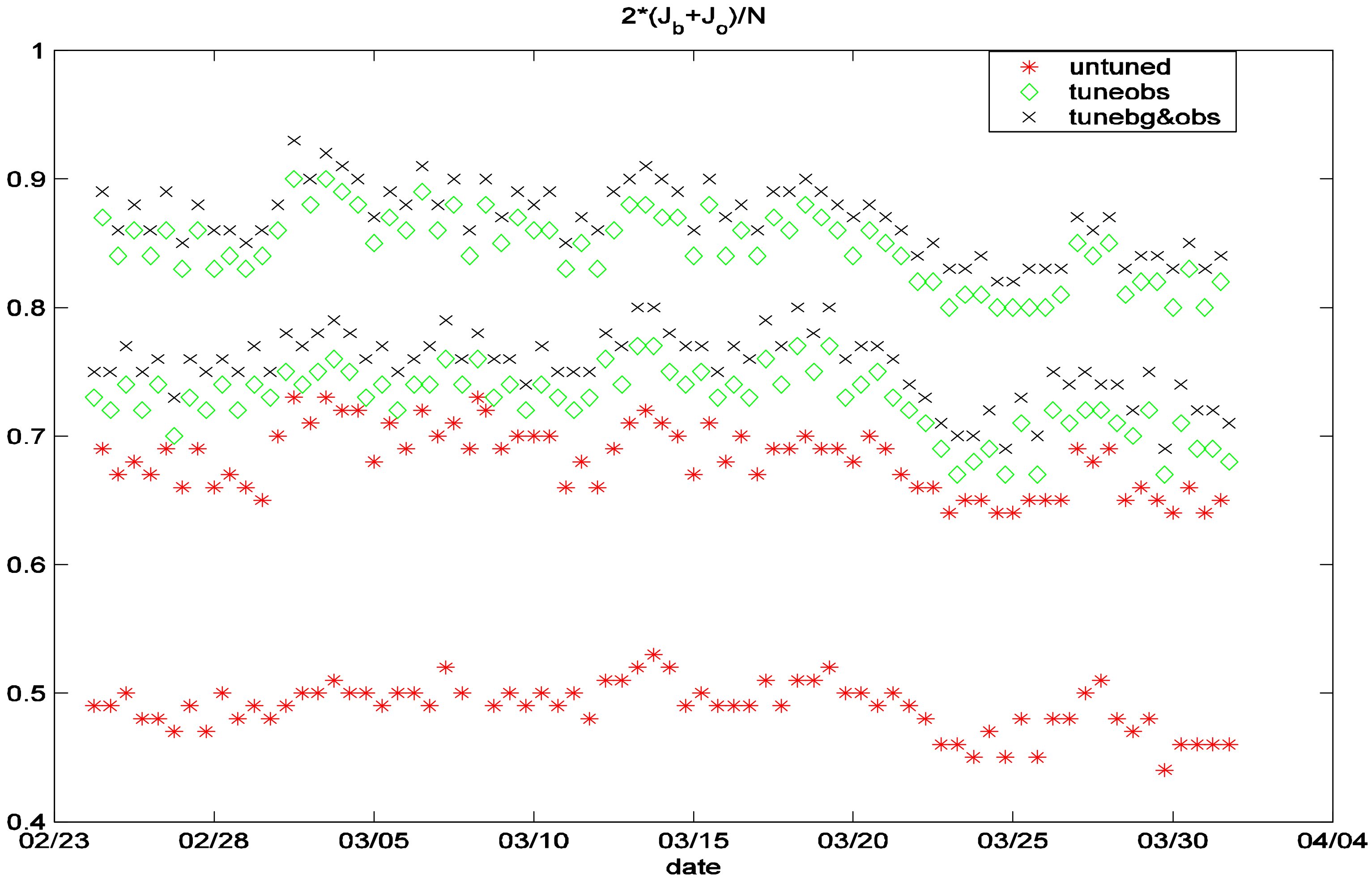
RAOBS T
one iteration on observation error



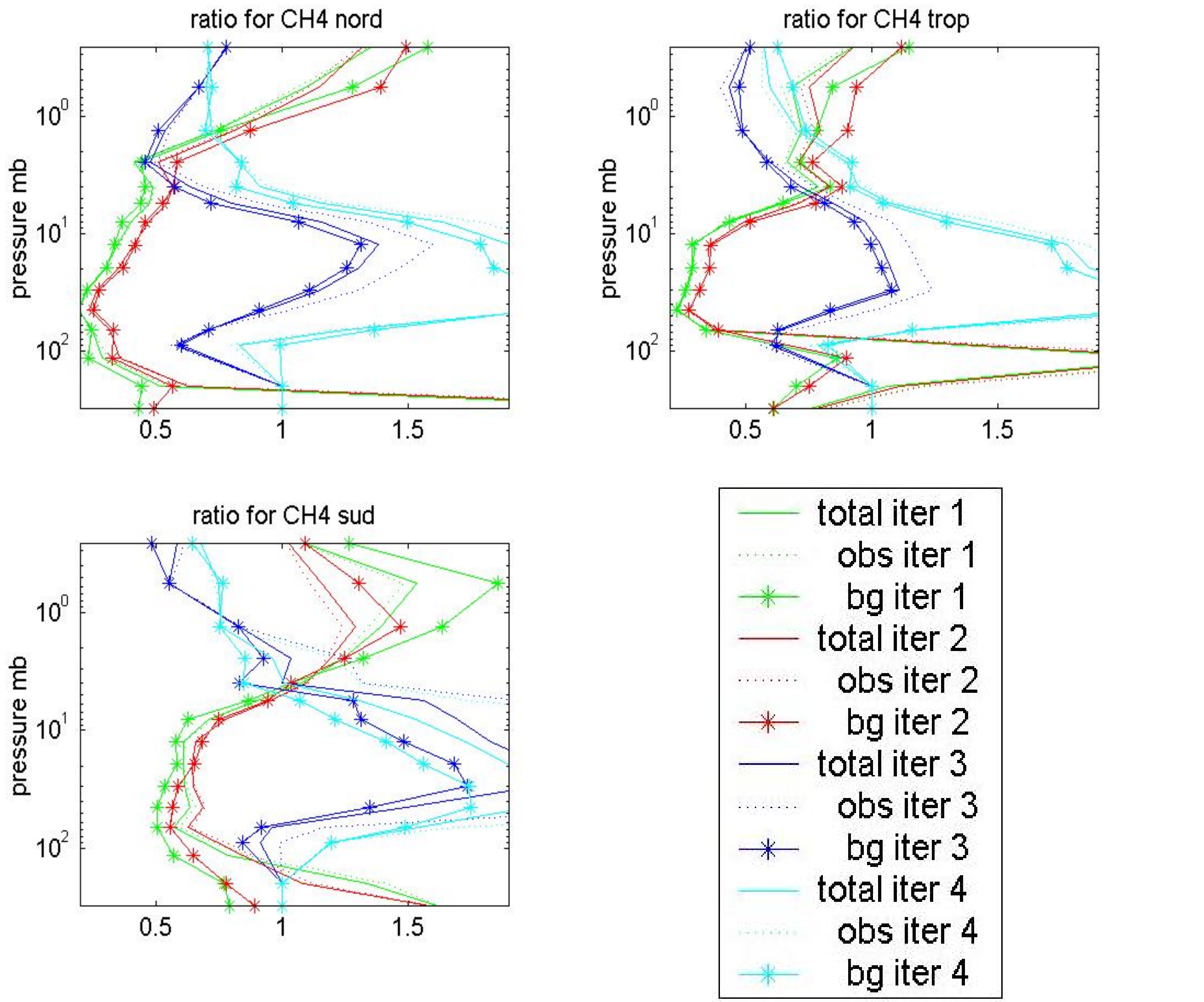
RAOBS T
then one iteration on background error



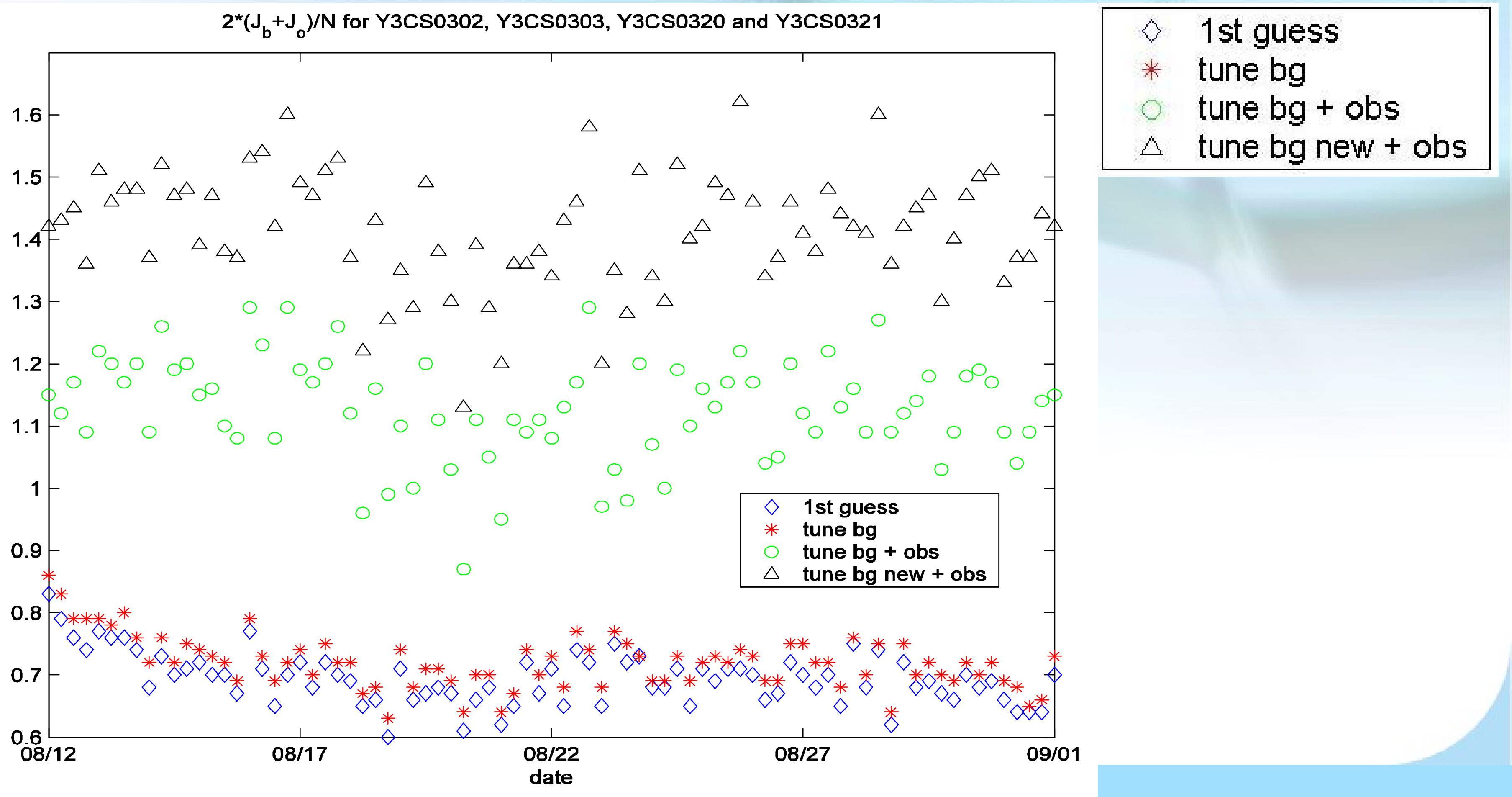
RAOBS T
one iteration on observation and background error



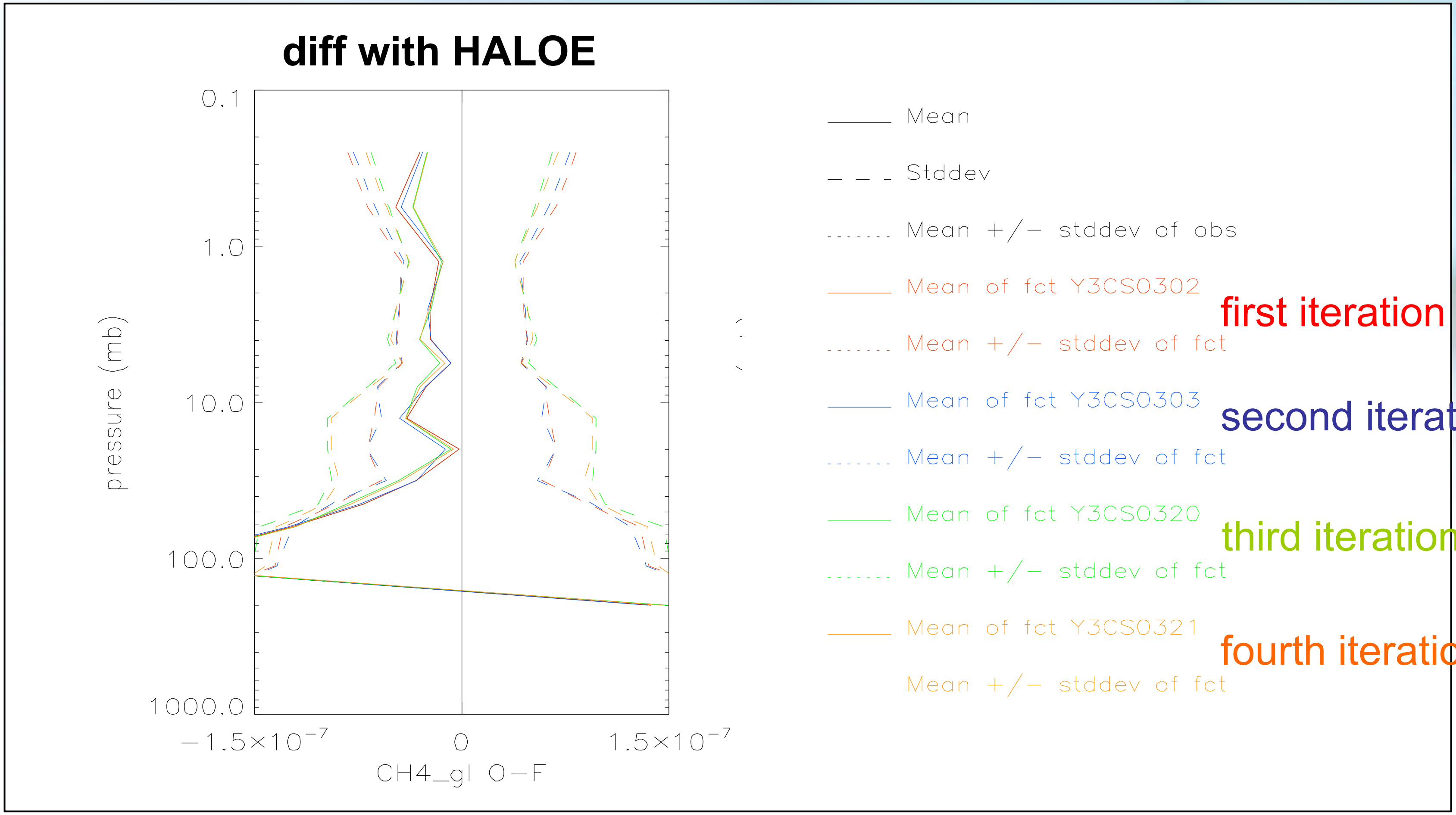
CH4 assimilation
combined iterations of observation and background error



CH4 assimilation
combined iterations of observation and background error



CH4 assimilation
combined iterations of observation and background error



Iteration on both observation and background error variances

Consider the case of tuning together α and β in each iteration

$$\alpha_{n+1} = \alpha_n \left(\frac{\gamma + 1}{\alpha_n \gamma + \beta_n} \right) = G(\alpha_n, \beta_n)$$

$$\beta_{n+1} = \beta_n \left(\frac{\gamma + 1}{\alpha_n \gamma + \beta_n} \right) = F(\alpha_n, \beta_n)$$

then the ratio

$$\mu_{n+1} = \frac{\alpha_{n+1}}{\beta_{n+1}} = \frac{\alpha_n}{\beta_n} = \mu_n = \cdots = \mu_0$$

is constant.

The mapping $(\alpha_n, \beta_n) \leftrightarrow (\alpha_{n+1}, \beta_{n+1})$ is in fact ill-defined, since the Jacobian

$$\frac{\partial(G,F)}{\partial(\alpha_n, \beta_n)} = \frac{\gamma + 1}{(\alpha_n \gamma + \beta_n)} \begin{pmatrix} \beta_n & -\alpha_n \\ -\beta_n \gamma & \alpha_n \gamma \end{pmatrix}$$

is rank deficient !

Iteration on both observation and background error variances

In fact the full system

$$\left\langle (O-A)(O-F)^T \right\rangle = \overline{\mathbf{R}}(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}})^{-1} \left\langle (O-F)(O-F)^T \right\rangle$$

$$\left\langle (A-F)(O-F)^T \right\rangle = \mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T (\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}})^{-1} \left\langle (O-F)(O-F)^T \right\rangle$$

is "rank deficient" !

Define $\left\langle (O-F)(O-F)^T \right\rangle = \mathbf{O}$

$$\left\langle (O-A)(O-F)^T \right\rangle = \mathbf{A}$$

$$\left\langle (A-F)(O-F)^T \right\rangle = \mathbf{P}$$

The system can be rewritten as

$$\mathbf{A}\mathbf{O}^{-1}(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}}) = \overline{\mathbf{R}} \qquad \text{.......................................} \ (1)$$

$$\mathbf{P}\mathbf{O}^{-1}(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}}) = \mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T$$

Iteration on both observation and background error variances

Since

$$\mathbf{A} + \mathbf{P} = \mathbf{O}$$

then

$$\mathbf{AO}^{-1} = (\mathbf{O} - \mathbf{P})\mathbf{O}^{-1} = \mathbf{I} - \mathbf{PO}^{-1}$$

So the first equation of the system (1)

$$\mathbf{AO}^{-1}(\mathbf{H\overline{B}H}^T + \overline{\mathbf{R}}) = \overline{\mathbf{R}}$$

$$(\mathbf{I} - \mathbf{PO}^{-1})(\mathbf{H\overline{B}H}^T + \overline{\mathbf{R}}) = \overline{\mathbf{R}}$$

$$\mathbf{H\overline{B}H}^T + \overline{\mathbf{R}} - \mathbf{PO}^{-1}\ (\mathbf{H\overline{B}H}^T + \overline{\mathbf{R}}) = \overline{\mathbf{R}}$$

$$\mathbf{H\overline{B}H}^T = \mathbf{PO}^{-1}\ (\mathbf{H\overline{B}H}^T + \overline{\mathbf{R}})$$

as the same information content than in the second equation !

The scalar equations applies as well for the spectral variances

Case where the background error covariance is *spatially correlated* and the observation error covariance is *spatially uncorrelated*

Assume an homogeneous **B** in a 1D periodic domain with observations at each grid points, **H = I** .
We can write the Fourier transform as a matrix **F**, and its inverse as $\mathbf{F}^T$

Then in the system

$$\mathbf{R}_{n+1} = \mathbf{R}_n (\mathbf{B}_n + \mathbf{R}_n)^{-1}\ \mathbf{O}$$

$$\mathbf{B}_{n+1} = \mathbf{B}_n (\mathbf{B}_n + \mathbf{R}_n)^{-1}\ \mathbf{O}$$

All matrices can be simultaneously diagonalized giving a N systems of scalar (variance) equations (one for each wavenumber *k*)

$$\hat{\mathbf{R}}_{n+1} = \hat{\mathbf{R}}_n (\hat{\mathbf{B}}_n + \hat{\mathbf{R}}_n)^{-1}\ \hat{\mathbf{O}}$$

$$\hat{\mathbf{B}}_{n+1} = \hat{\mathbf{B}}_n (\hat{\mathbf{B}}_n + \hat{\mathbf{R}}_n)^{-1}\ \hat{\mathbf{O}}$$

⇒

$$\hat{\alpha}_{n+1} = \hat{\alpha}_n \left(\frac{\gamma + 1}{\hat{\alpha}_n \gamma + \hat{\beta}_n} \right) = G(\hat{\alpha}_n, \hat{\beta}_n)$$

$$\hat{\beta}_{n+1} = \hat{\beta}_n \left(\frac{\gamma + 1}{\hat{\alpha}_n \gamma + \hat{\beta}_n} \right) = F(\hat{\alpha}_n, \hat{\beta}_n)$$

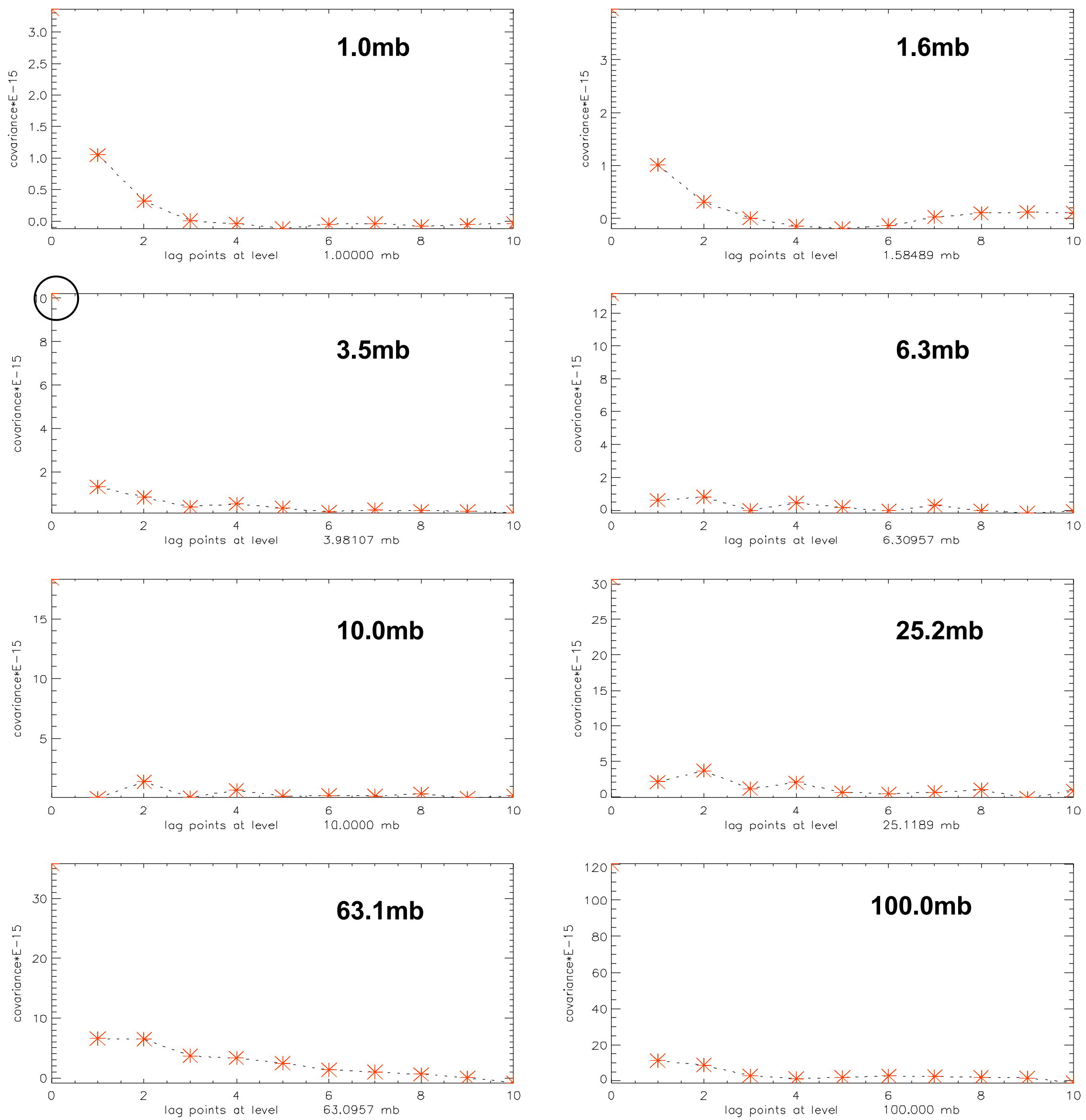
Summary and Conclusions

- The convergence of the Desrosiers' et al (2005) scheme has been investigated in the context of cycling assimilation
- Iteration on either observation error variance or background error variance generally converges, but will converge to an overestimate if the counterpart in underestimated, and vice versa
- Iteration on both observation and background error variance is in principle non convergent because the system of equation is rank deficient – the same information is contained in the O-A and A-F equations
- Consideration about the correlation length scales (different for obs and background) seems not to influence the converge as shown by a spectral analysis on a simplified system
- Divergence of the scheme is clearly demonstrated in the case of assimilation of a long-lived specie from a single instrument, but is unclear for meteorological variables, perhaps because of multivariate coupling and multiple source of observations that may restrain the feedback in assimilation cycles

Merci

Thank you

CH_4



1.0mb
1.6mb
3.5mb
6.3mb
10.0mb
25.2mb
63.1mb
100.0mb
covariance*E-15
lag points at level
1.00000 mb
1.58489 mb
3.98107 mb
6.30957 mb
10.0000 mb
25.1189 mb
63.0957 mb
100.000 mb

Tuning in alternance –CH4