Accounting for an imperfect model in 4D-Var

Yannick TRÉMOLET

ECMWF - NASA/GMAO

July 2007

- Introduction: Incremental 4D-Var
- Model Error in 4D-Var
- Weak constraint 4D-Var

Introduction to 4D-Var

4D Variational Data Assimilation

- Model: $\mathcal{M}$,
- Observations: y,
- Background: x_b,
- Observation operator: $\mathcal{H}$,
- Cost function to minimise:

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}[y - \mathcal{H}(x)]^T R^{-1}[y - \mathcal{H}(x)] + J_c$$

- In discrete form:

$$\begin{aligned} J(x) &= \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) \\ &+ \frac{1}{2}\sum_{i=0}^{n}[y_i - \mathcal{H}_i(\mathcal{M}_i(x))]^T R_i^{-1}[y_i - \mathcal{H}_i(\mathcal{M}_i(x))] + J_c \end{aligned}$$

Incremental 4D-Var

- The cost function is expressed in terms of increments with respect to the background state $\delta x = x - x_b$.
- $\mathcal{H}$ and $\mathcal{M}$ are linearised around $x_i = \mathcal{M}_i(x_0)$.

$$J(\delta x) = \frac{1}{2}\delta x^T B^{-1} \delta x + \frac{1}{2}\sum_{i=0}^{n}(H_i M_i \delta x - d_i)^T R_i^{-1}(H_i M_i \delta x - d_i)$$

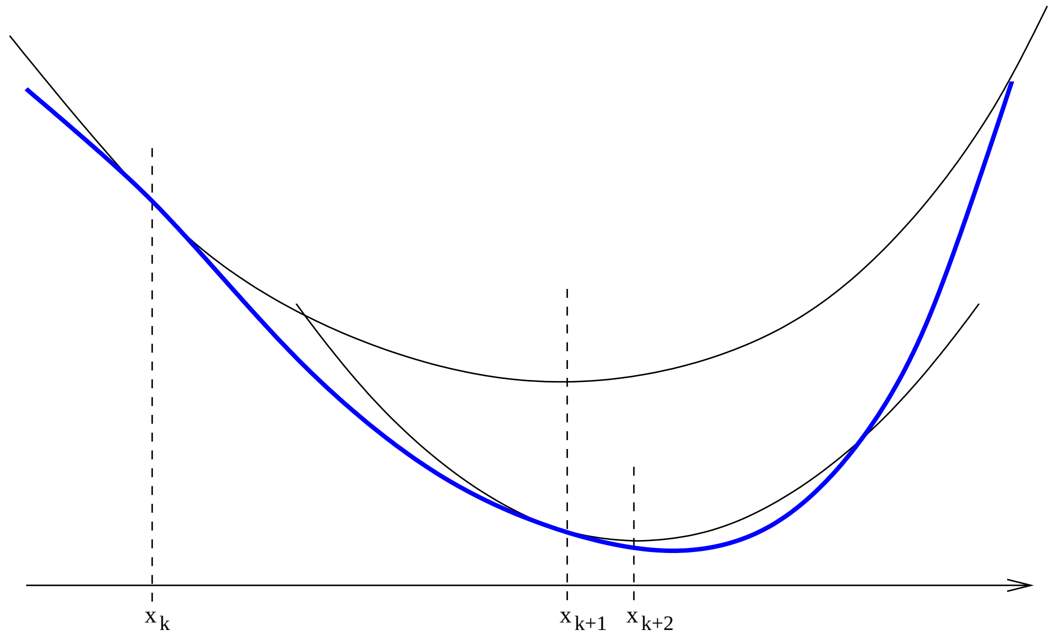
where

  - H_i and M_i are the linearised observation operator and model,
  - $d_i = y_i - \mathcal{H}_i(\mathcal{M}_i(x_b))$ are the innovations.

- The innovations, which are the primary input to the assimilation, are always computed using the full observation operator and model, at full resolution, to ensure the highest possible accuracy.

Incremental 4D-Var

- One complex problem is replaced by a series of (slightly) easier problems.
- Gauss-Newton algorithm.



The Outer Iterations

After each minimisation at inner level:

- x is updated: $x_a^j = x_b + \delta x_j$,
- H_i and M_i are re-linearised around $x_i = \mathcal{M}_i(x_a^j)$,
- Innovations are re-calculated using the full nonlinear observation operator $\mathcal{H}$ and model $\mathcal{M}$:

$$d_i^j = y_i - \mathcal{H}_i(\mathcal{M}_i(x_a^j))$$

- Superscript j represents the **outer iteration**.
- The nonlinear model $\mathcal{M}$ remains at the highest resolution (T799) throughout.

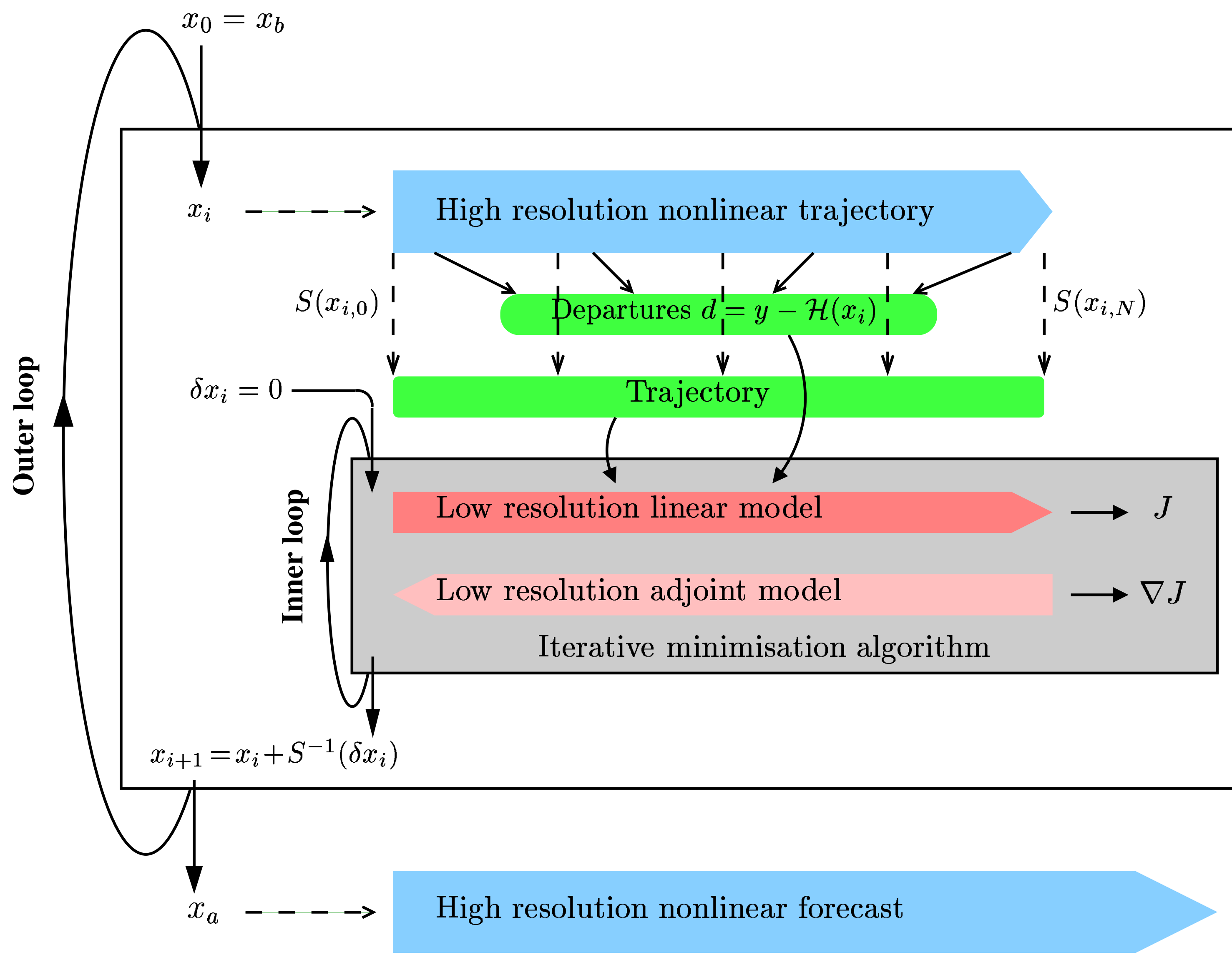
The Inner Iterations

- Tangent Linear approximation:

$$\mathcal{M}(x + \delta x) \approx \mathcal{M}(x) + M\delta x \quad \text{and} \quad \mathcal{H}(x + \delta x) \approx \mathcal{H}(x) + H\delta x$$

- Approximations to reduce cost: the tangent linear model (and its adjoint) is degraded with respect to the full model $\mathcal{M}$:
  - Lower resolution (T95/T159/T255 ($\approx$ 200, 120, 80 km) instead of T799 ($\approx$ 25 km)),
  - Simplified physics (some processes are ignored),
  - Simpler dynamics (spectral instead of grid-point humidity).
- The results in **shorter control vector** and **cheaper TL and AD models** during the minimisation.

Incremental 4D-Var



Model Error in 4D-Var

Weak constraint 4D-Var

- The model $\mathcal{M}$ is verified exactly although it is not perfect...
- The model can be imposed as a constraint in the cost function, in the same way as other sources of information:

$$\mathcal{F}_i(x) = x_i - \mathcal{M}_i(x_{i-1})$$

- Model error η is defined as: $\eta_i = x_i - \mathcal{M}_i(x_{i-1})$
- The cost function becomes:

$$\begin{aligned} J(x) \quad = \quad & \frac{1}{2}\sum_{i=0}^{n}[\mathcal{H}(x_i) - y_i]^T R_i^{-1}[\mathcal{H}(x_i) - y_i] \\ + \quad & \frac{1}{2}(x_0 - x_b)^T B^{-1}(x_0 - x_b) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\eta_i^T Q_{i,j}^{-1}\eta_j \end{aligned}$$

- The model error covariance matrix Q is required.

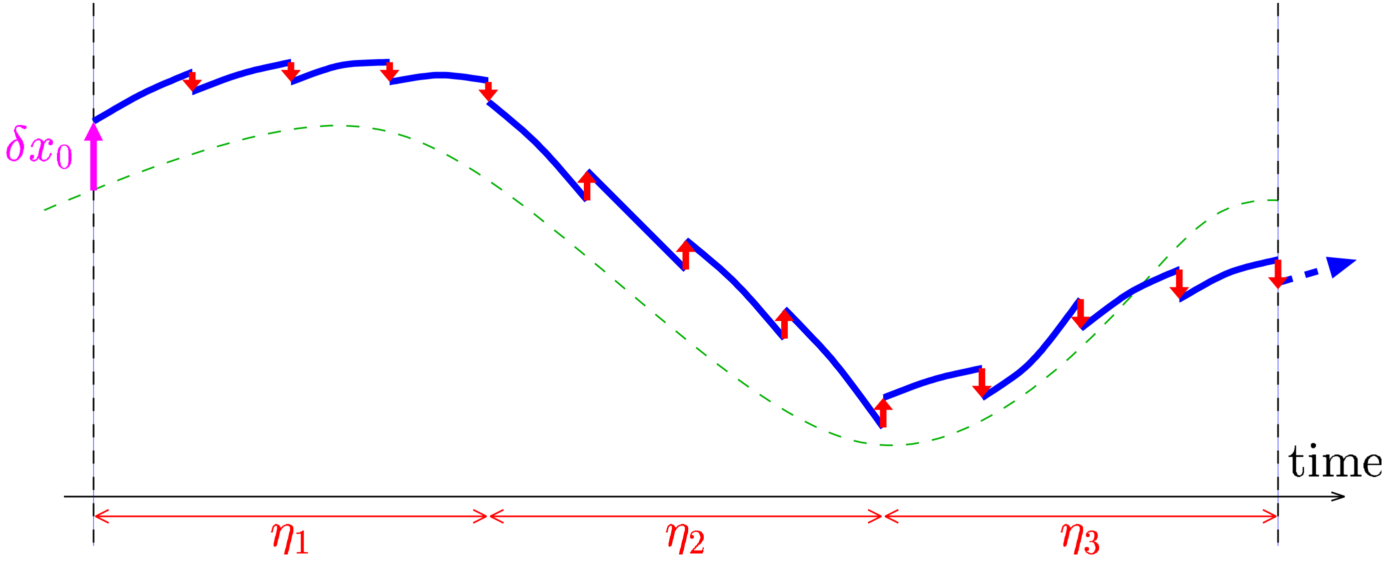
4D-Var with Model Error Forcing

- The cost function is written as a function of (x_0, η):

$$
\begin{aligned}
J(x_0, \eta) &= \frac{1}{2}\sum_{i=0}^{n}[\mathcal{H}(x_i) - y_i]^T R_i^{-1}[\mathcal{H}(x_i) - y_i] \\
&+ \frac{1}{2}(x_0 - x_b)^T B^{-1}(x_0 - x_b) + \eta^T Q^{-1}\eta \\
&\text{with } x_i = \mathcal{M}_i(x_{i-1}) + \eta_i.
\end{aligned}
$$

- The *usual* model error term in 4D-Var.
- η_i has the dimension of a 3D atmospheric state,
- η_i represents the instantaneous model error,
- η is constrained by the fact that it is propagated by the model.

4D-Var with Model Error Forcing



- TL and AD models can be used with little modification,
- Information is propagated between obervations and IC control variable by TL and AD models.

Model Error Covariance Matrix

ECMWF

Model error covariance matrix

- The *usual* choice is $Q = \alpha B$.
- Linearisation in incremental formulation gives:

$$\delta x_n = M_n \dots M_1 \delta x_0 + \sum_{i=1}^{n} M_n \dots M_{i+1} \eta_i$$

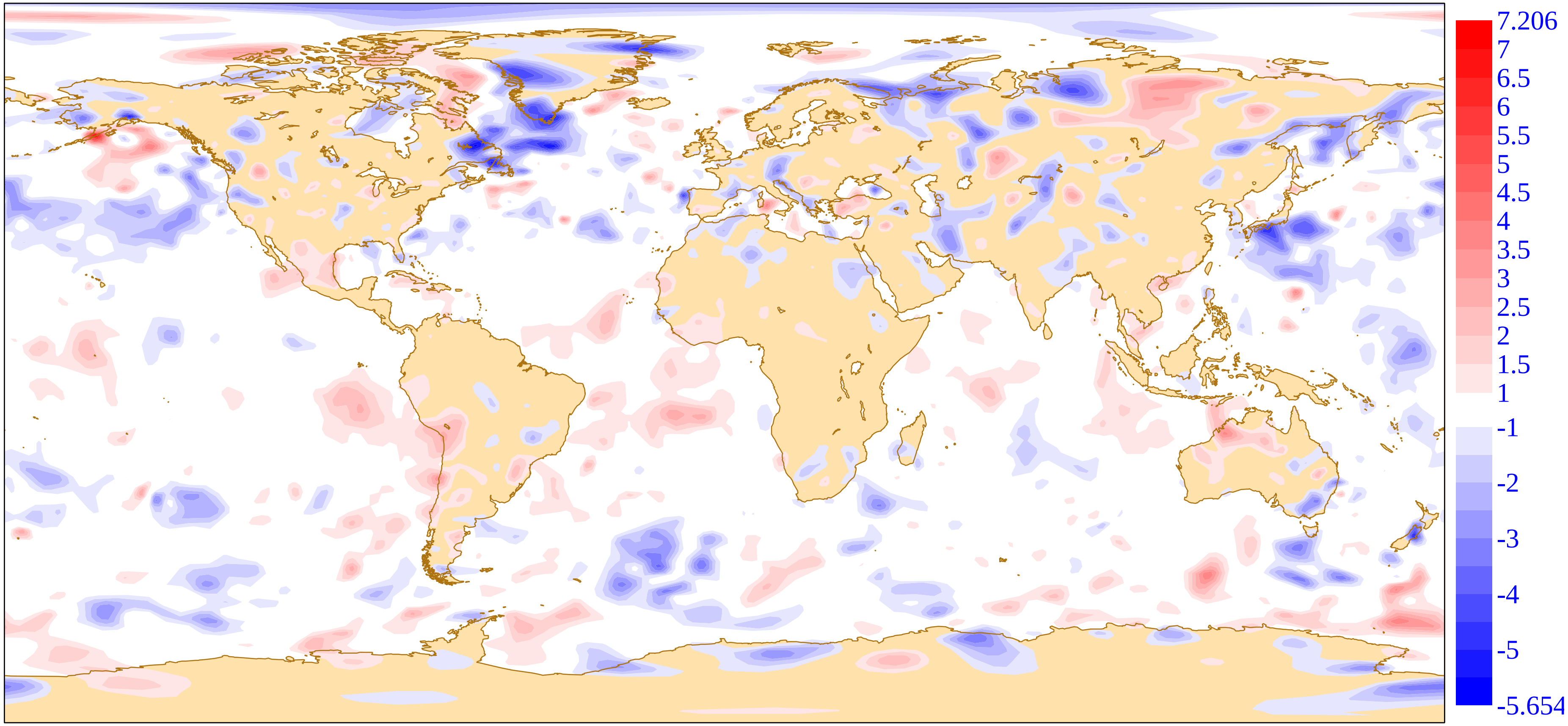
- δx_0 can be identified with η_0.
- The solution of the analysis equation satisfies:

$$\delta x_0 = BH^T (R + HBH^T)^{-1} (y - \mathcal{H}(x_b))$$

$$\eta = QH^T (R + HQH^T)^{-1} (y - \mathcal{H}(x_b))$$

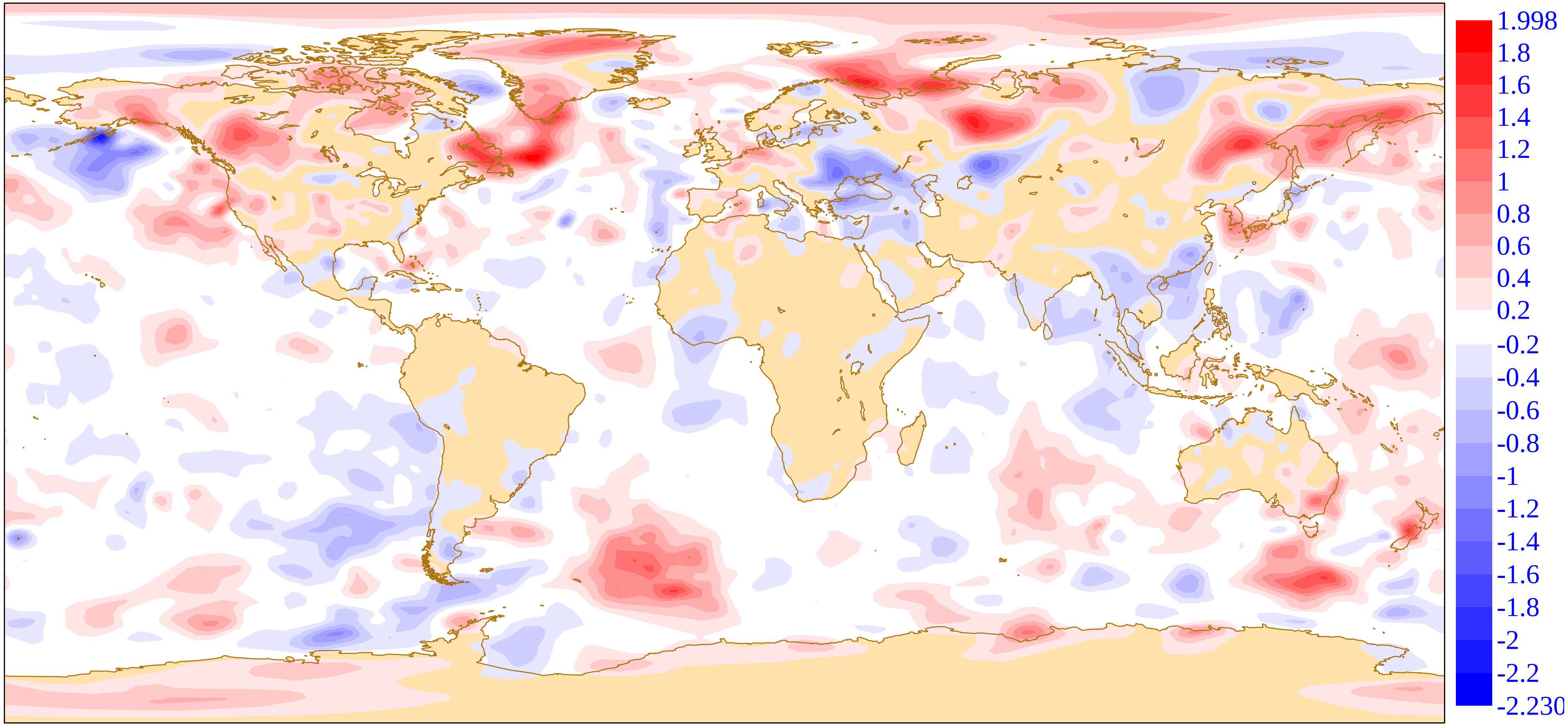
- If Q and B are proportional, δx_0 and η are constrained in the same directions, may be with different relative amplitudes.
- They both predominantly retrieve the same information: $Q = \alpha B$ is too limiting.

Model error covariance matrix



Model Error Forcing

Model error covariance matrix



Initial Condition Difference

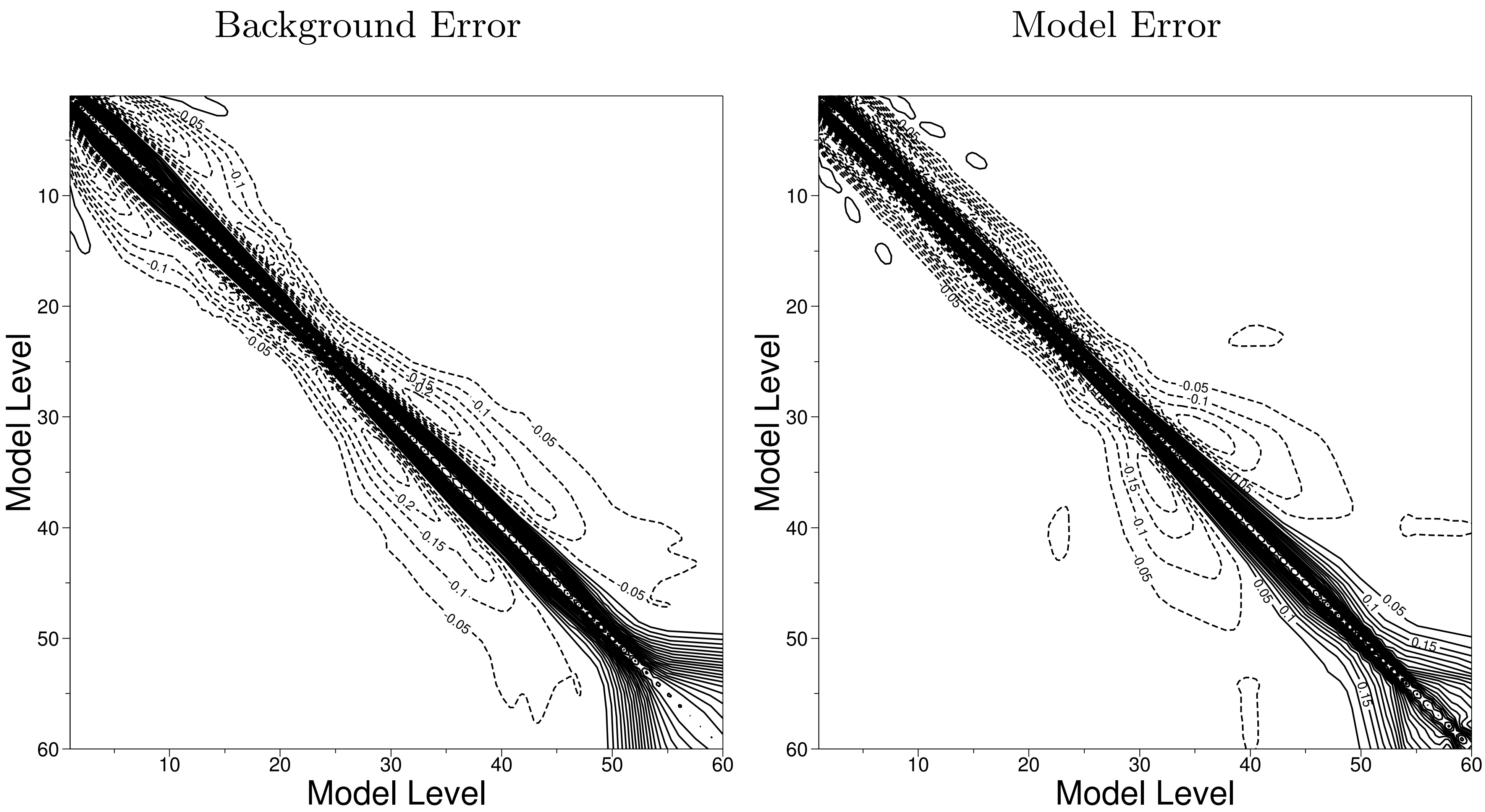
Generating a Model Error Covariance Matrix

- B is estimated from an ensemble of 4D-Var assimilations.
- Considering the forecasts run from the 4D-Var members:
  - At a given step, each model state is supposed to represent the same *true* atmospheric state,
  - The tendencies from each of these model states should represent possible evolutions of the atmosphere from that same *true* atmospheric state,
  - The differences between these tendencies can be interpreted as possible uncertainties in the model or realisations of *model error*.
- Q can be estimated by applying the statistical model used for B to tendencies instead of analysis increments.

Model Error Covariance Matrix

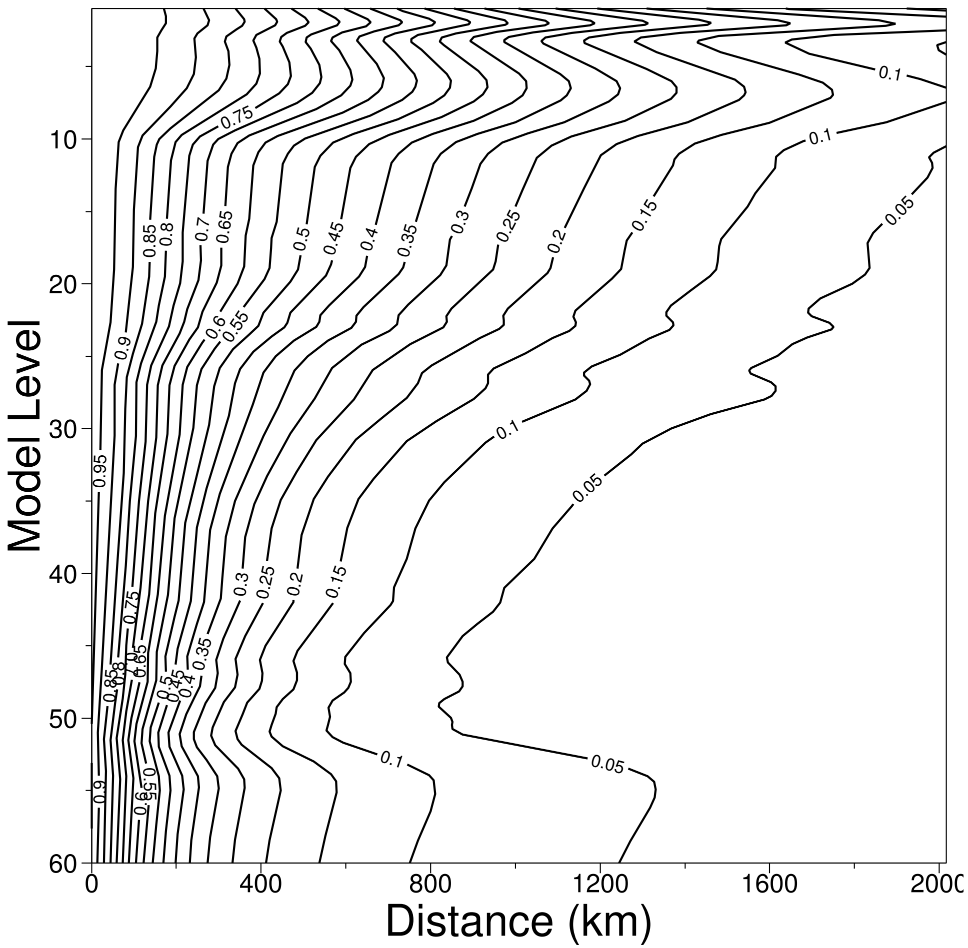
- Run forecasts from 4D-Var ensemble (T319),
- Save tendencies after 12h, 18h, 24h, 30h,
- Normalise tendencies to 1 hour time-steps,
- 10 members, one month (26 days),
- 936 realisations of *model errors*,
- Run statistical model.

Average Temperature Vertical Correlations

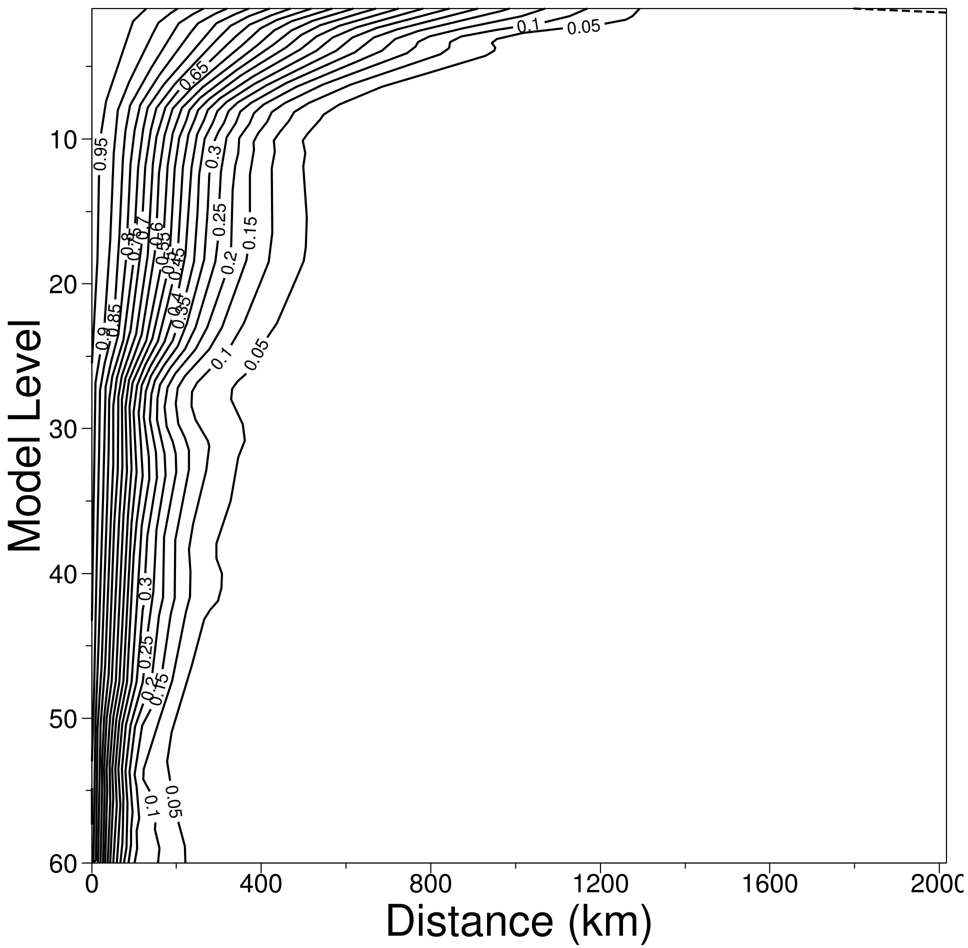


Temperature Horizontal Correlations

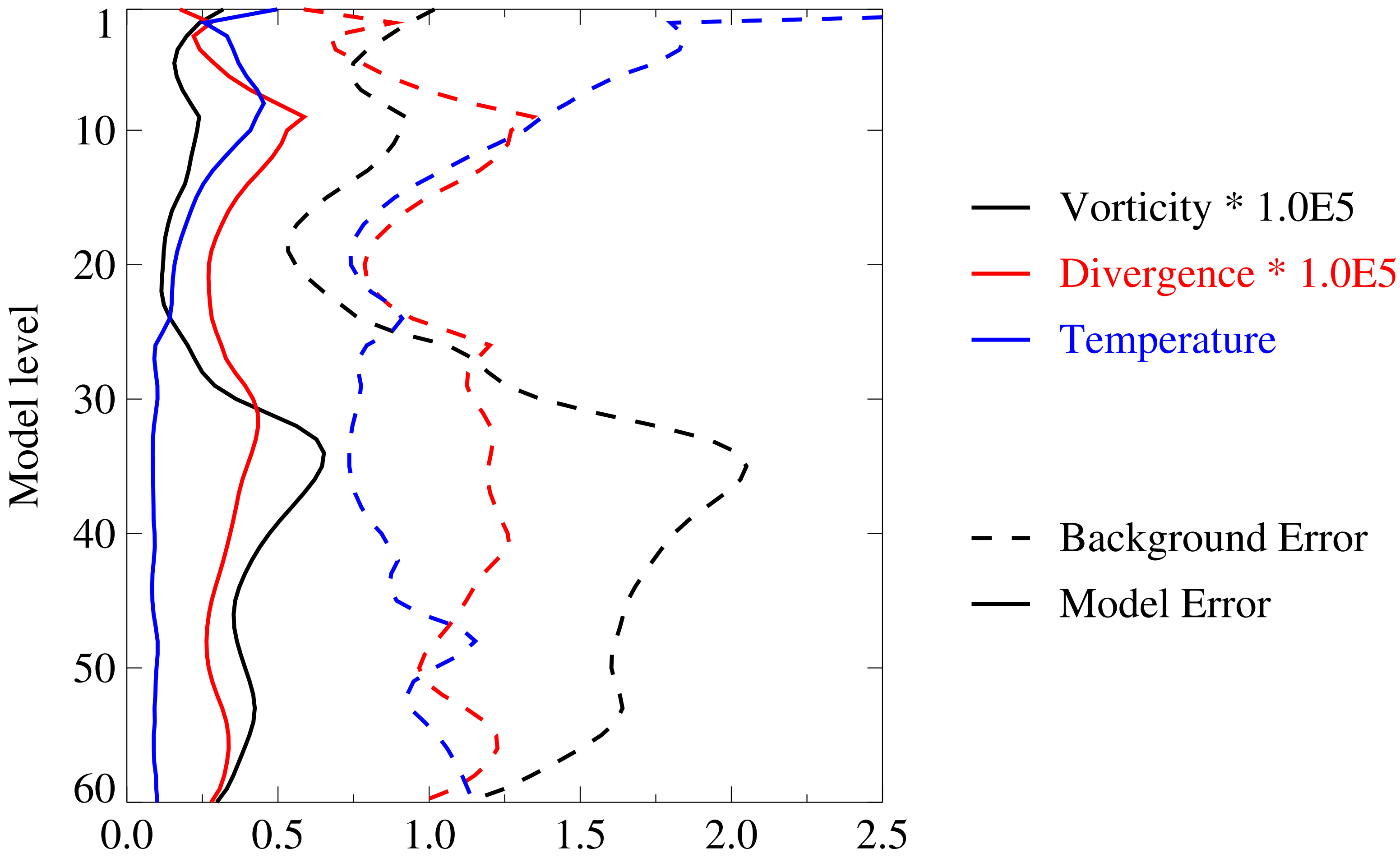
Background Error



Model Error



Average Standard Deviation Profiles

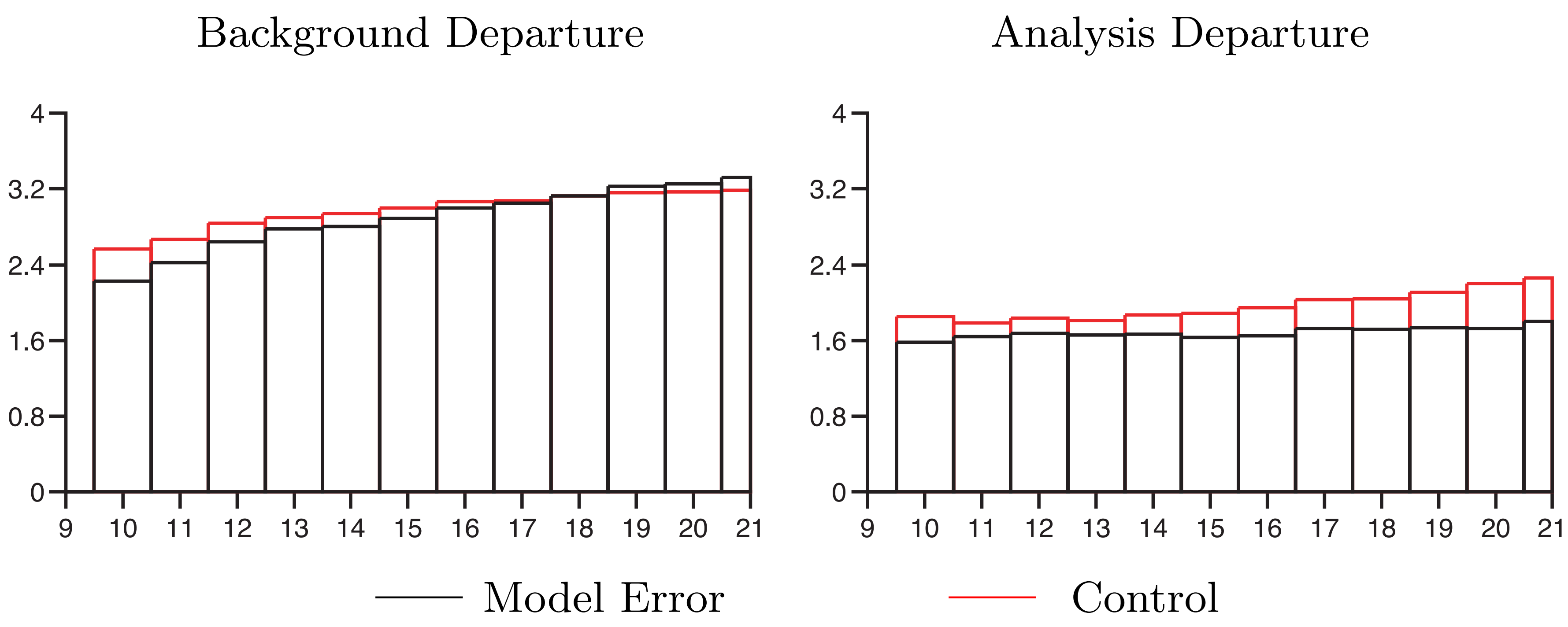


Q has narrower correlations and smaller amplitudes than B

Results: Constant Model Error Forcing

Results: Fit to observations

`AMprofiler-windspeed Std Dev N.Amer`



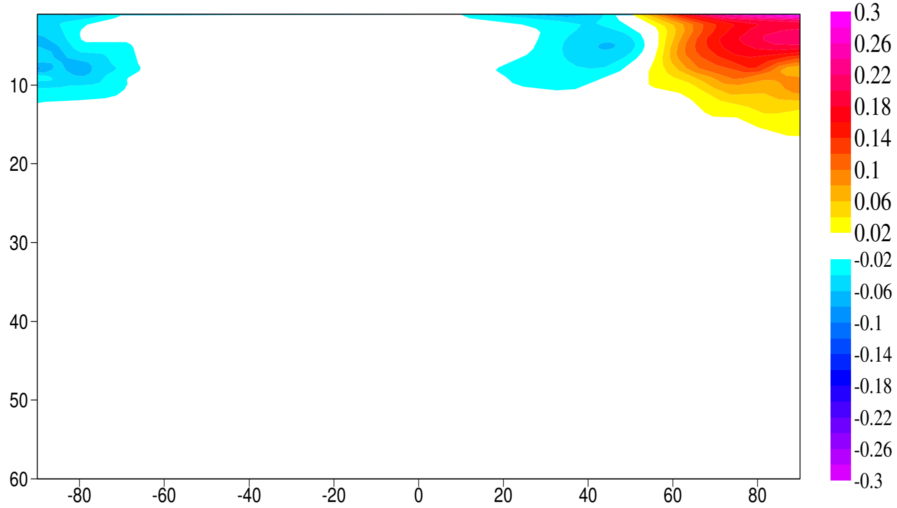
- Fit to observations is more uniform over the assimilation window.
- Background fit improved only at the start: error varies in time ?

Model Error Forcing

Zonal Mean Temperature

July 2004

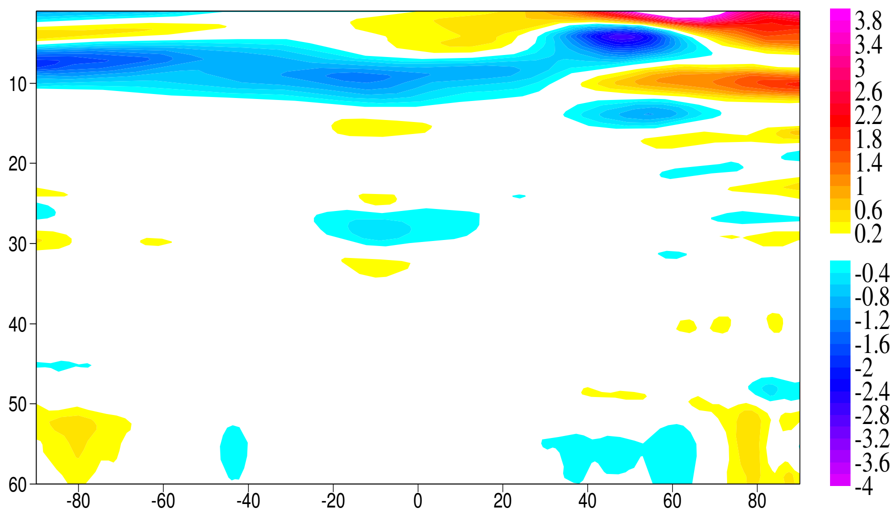
M.E. Forcing →

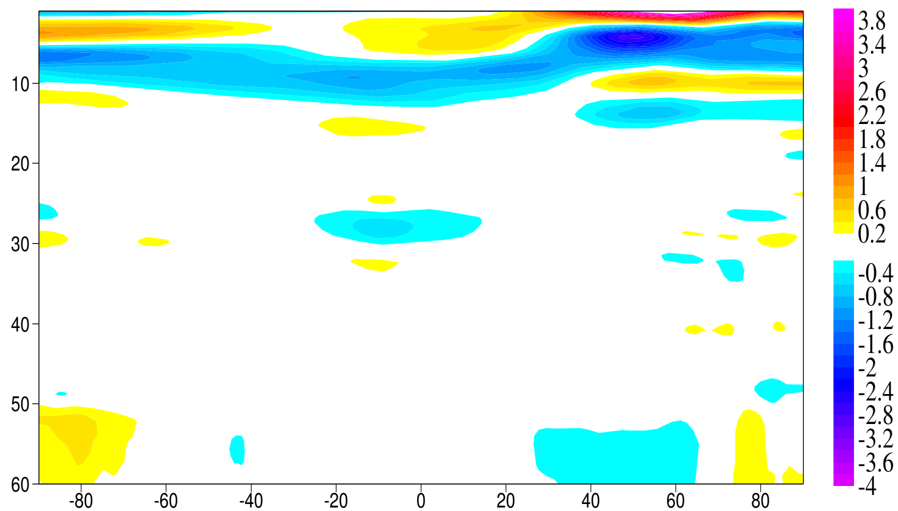


M.E. Mean Increment ↘

Control Mean Increment

↓



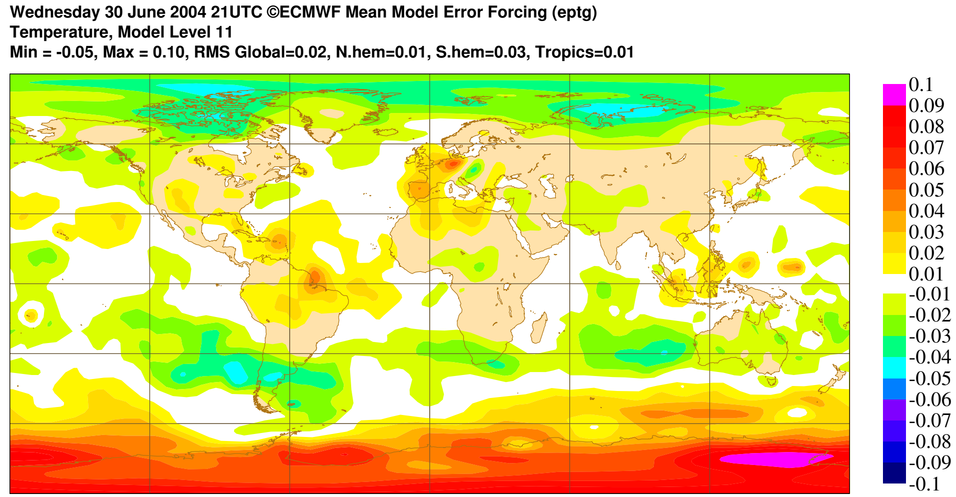


Mean Model Error Forcing

Temperature

Model level 11 (≈5hPa)

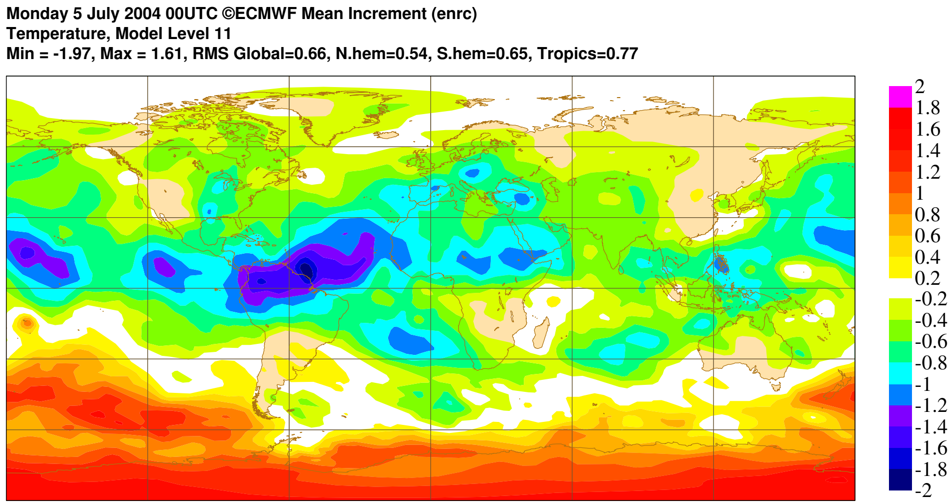
Mean M.E. Forcing ⟶

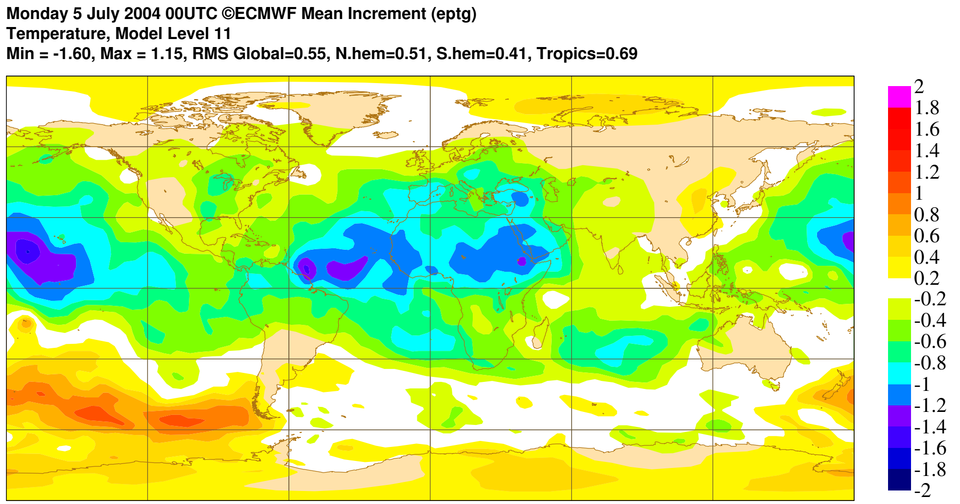


M.E. Mean Increment ↘

Control Mean Increment

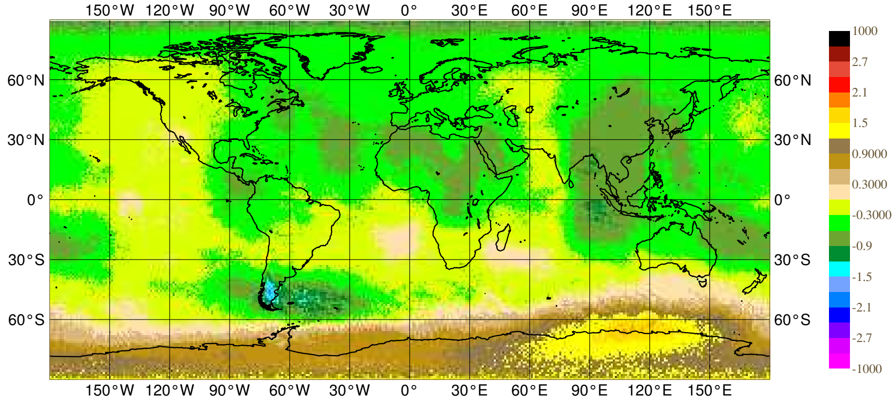
↓



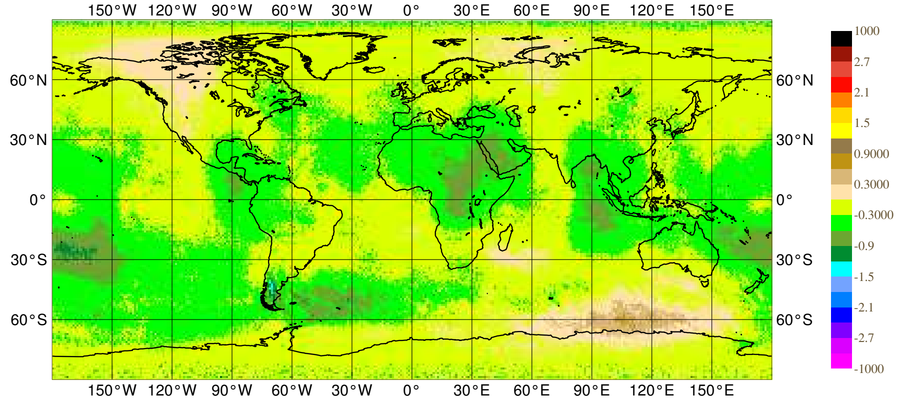


AMSU-A Departures

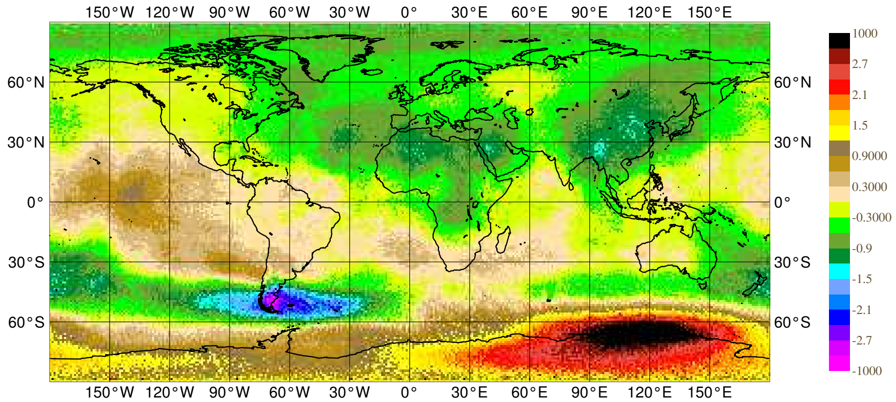
STATISTICS FOR RADIANCES FROM NOAA-16 / AMSU-A - 13
MEAN FIRST GUESS DEPARTURE (OBS-FG) (BCORR.) (CLEAR)
DATA PERIOD = 2004070200 - 2004073118 , HOUR = ALL
EXP = ENRC
Min: -1.9618 Max: 2.7 Mean: -0.169506



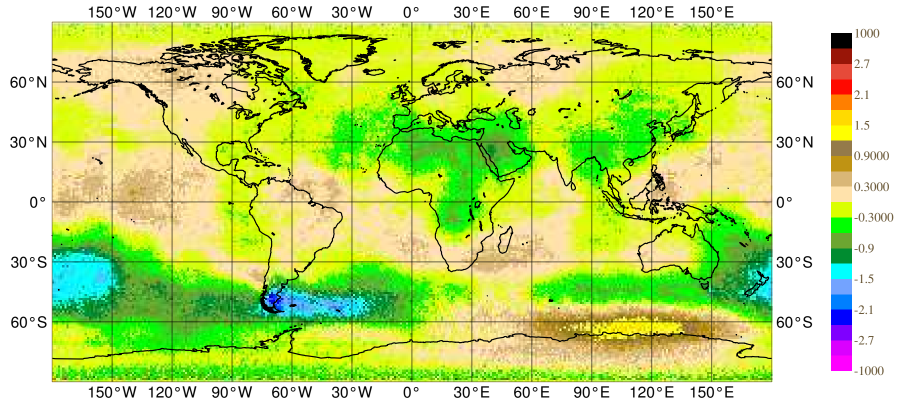
STATISTICS FOR RADIANCES FROM NOAA-16 / AMSU-A - 13
MEAN FIRST GUESS DEPARTURE (OBS-FG) (BCORR.) (CLEAR)
DATA PERIOD = 2004070100 - 2004073118 , HOUR = ALL
EXP = EPTG
Min: -1.6688 Max: 0.8 Mean: -0.231773



STATISTICS FOR RADIANCES FROM NOAA-16 / AMSU-A - 14
MEAN FIRST GUESS DEPARTURE (OBS-FG) (BCORR.) (CLEAR)
DATA PERIOD = 2004070200 - 2004073118 , HOUR = ALL
EXP = ENRC
Min: -3.3564 Max: 5.46 Mean: 0.006309



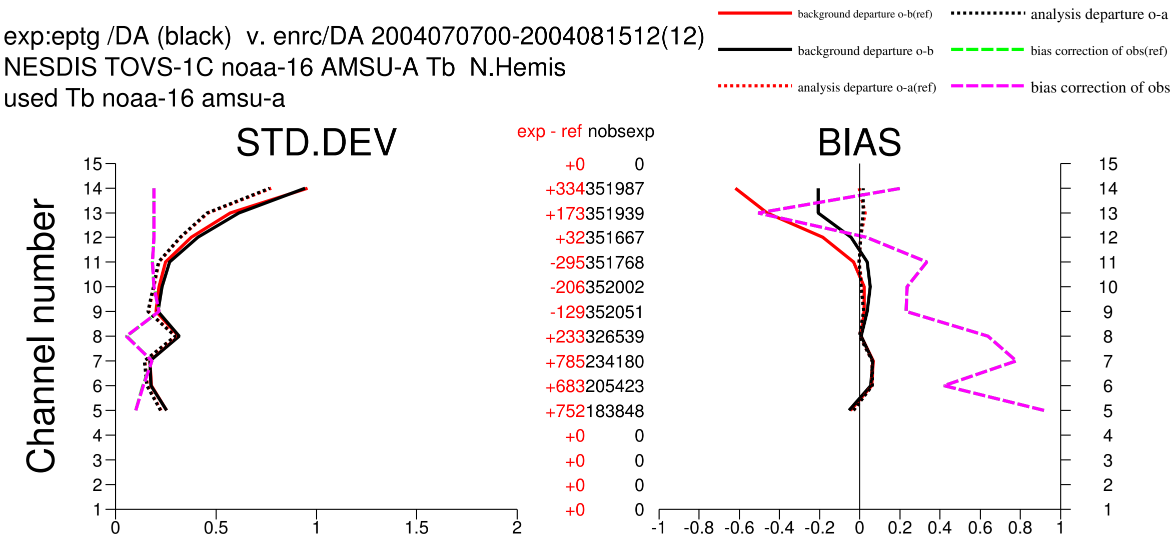
STATISTICS FOR RADIANCES FROM NOAA-16 / AMSU-A - 14
MEAN FIRST GUESS DEPARTURE (OBS-FG) (BCORR.) (CLEAR)
DATA PERIOD = 2004070100 - 2004073118 , HOUR = ALL
EXP = EPTG
Min: -2.6 Max: 2.16 Mean: -0.111883

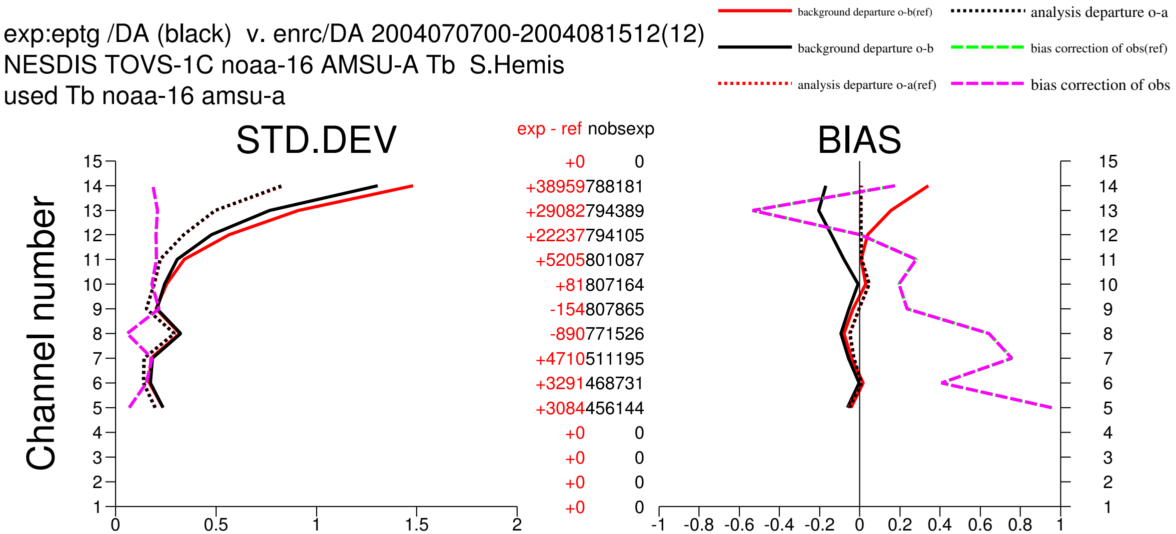


Control

Model Error

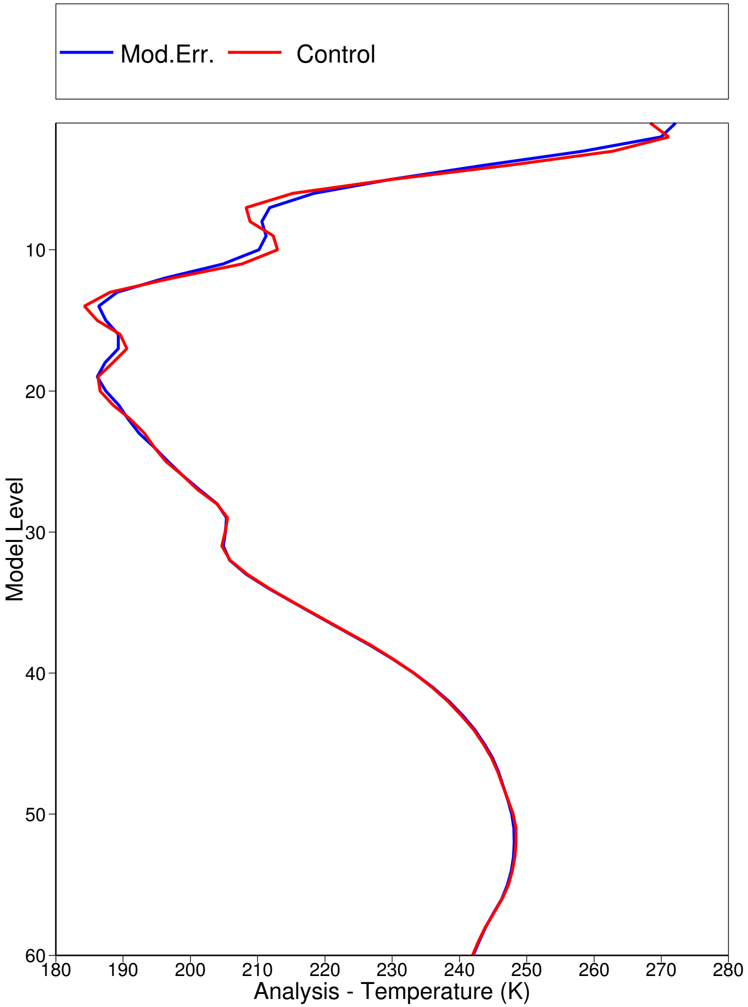
AMSU-A Statistics



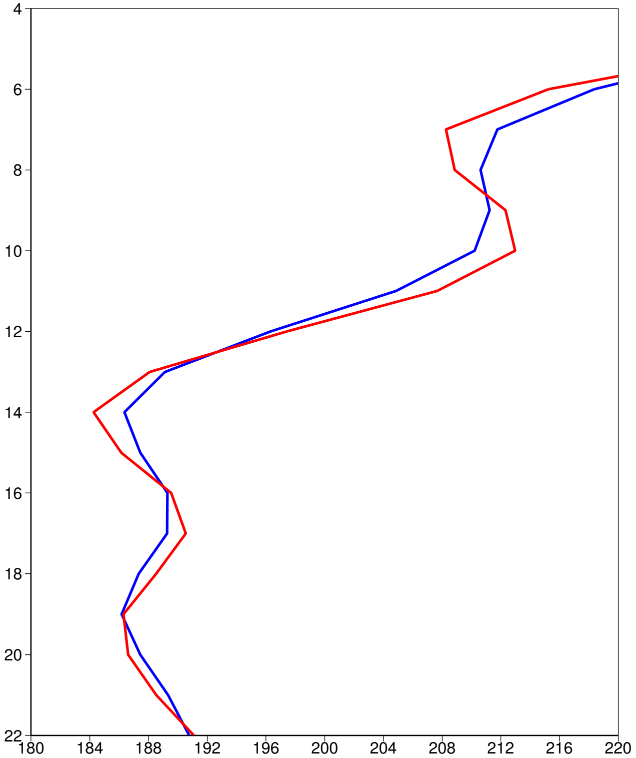


- Bias is more uniform,
- BG std. dev. is reduced in SH,
- More data is used.

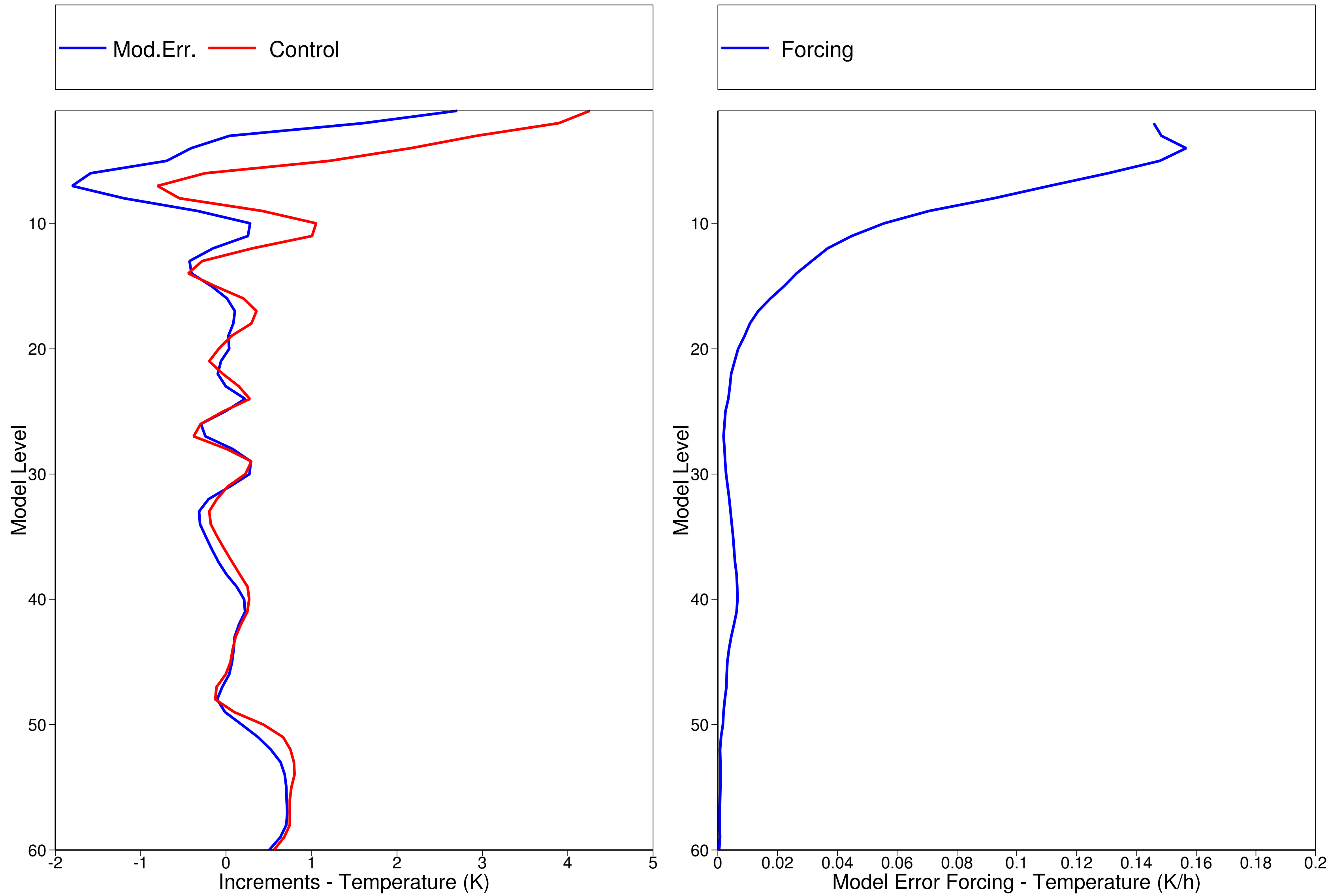
Temperature Analysis Profile



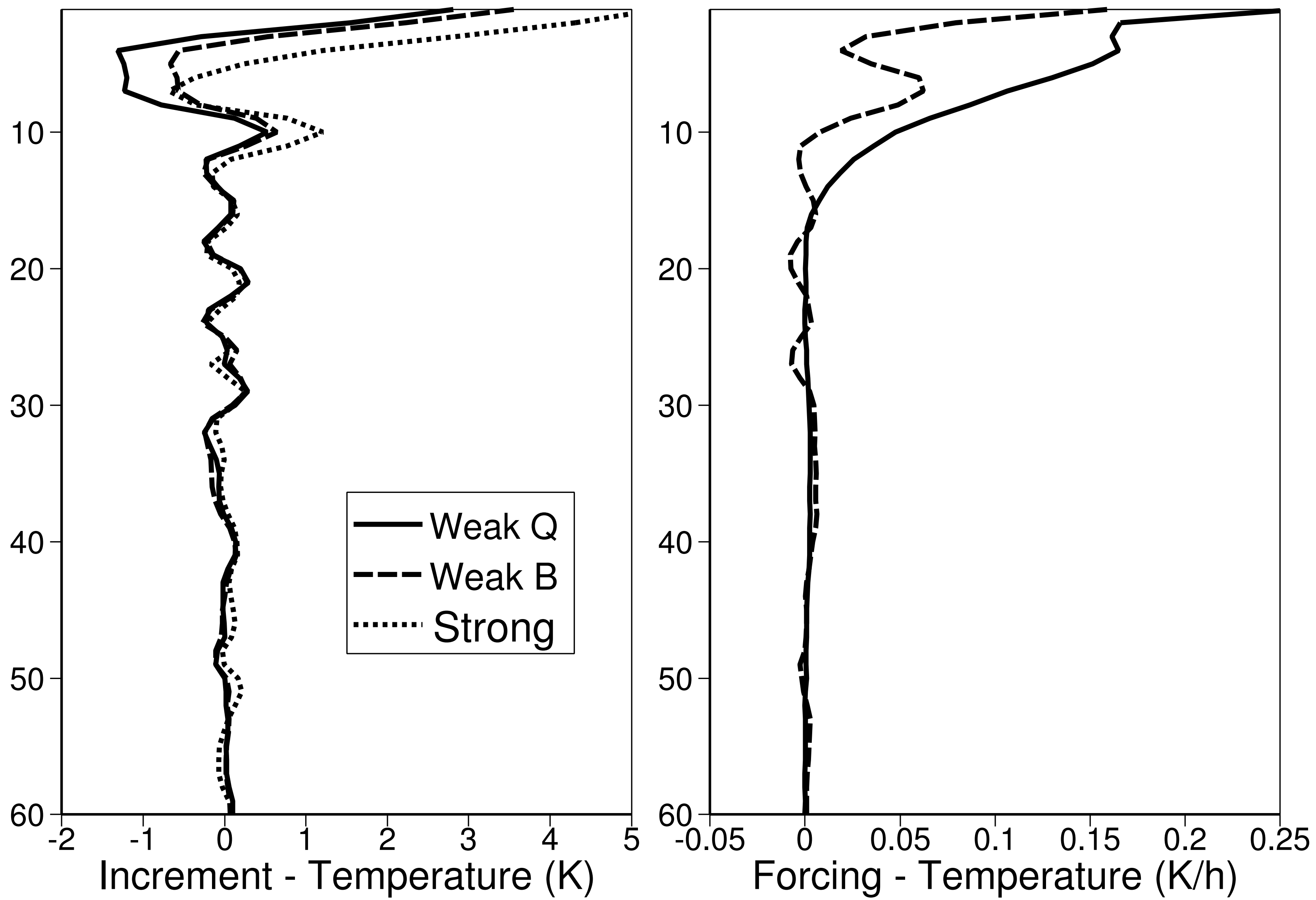
⟹



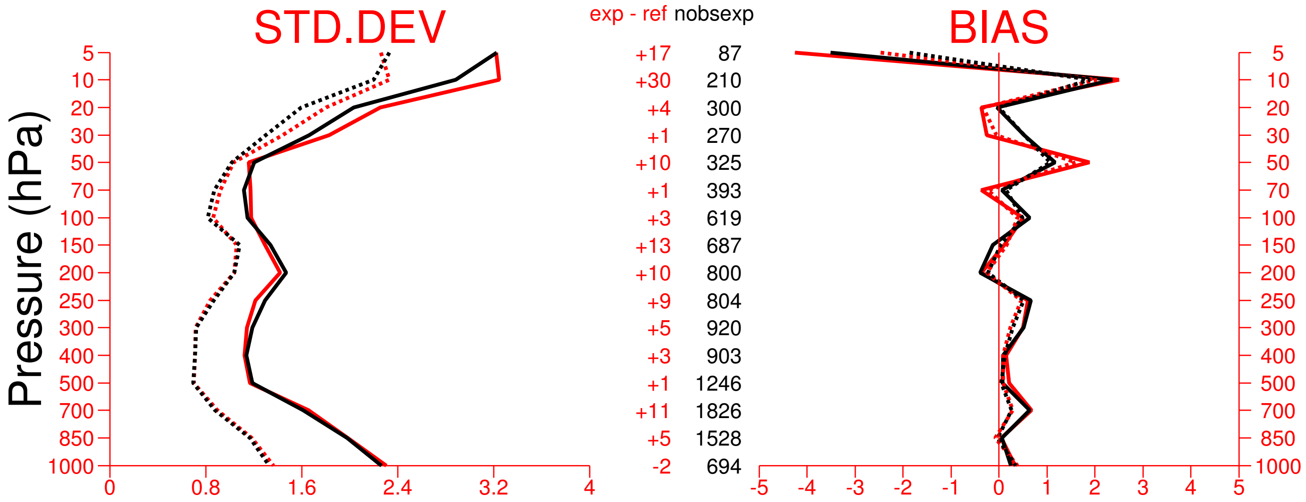
Temperature Increments Profile



Temperature Increments Profile



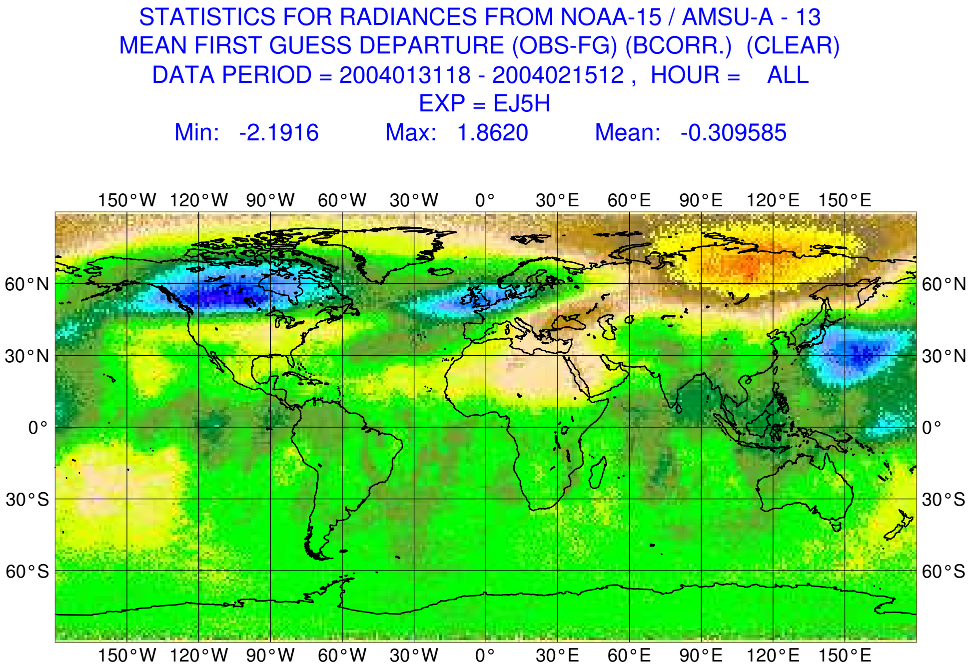
Fit to radiosonde data

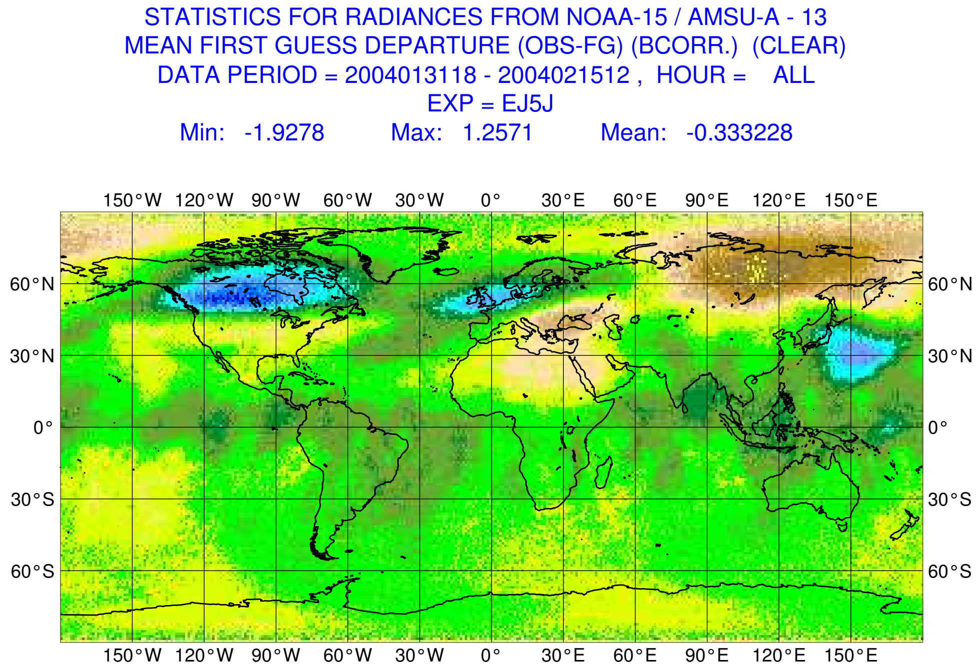


- Oscillations in bias are reduced,
- Std. deviation is reduced above 50 hPa (bg and an).

One source of model error was corrected by the forcing term.

Model Error Forcing



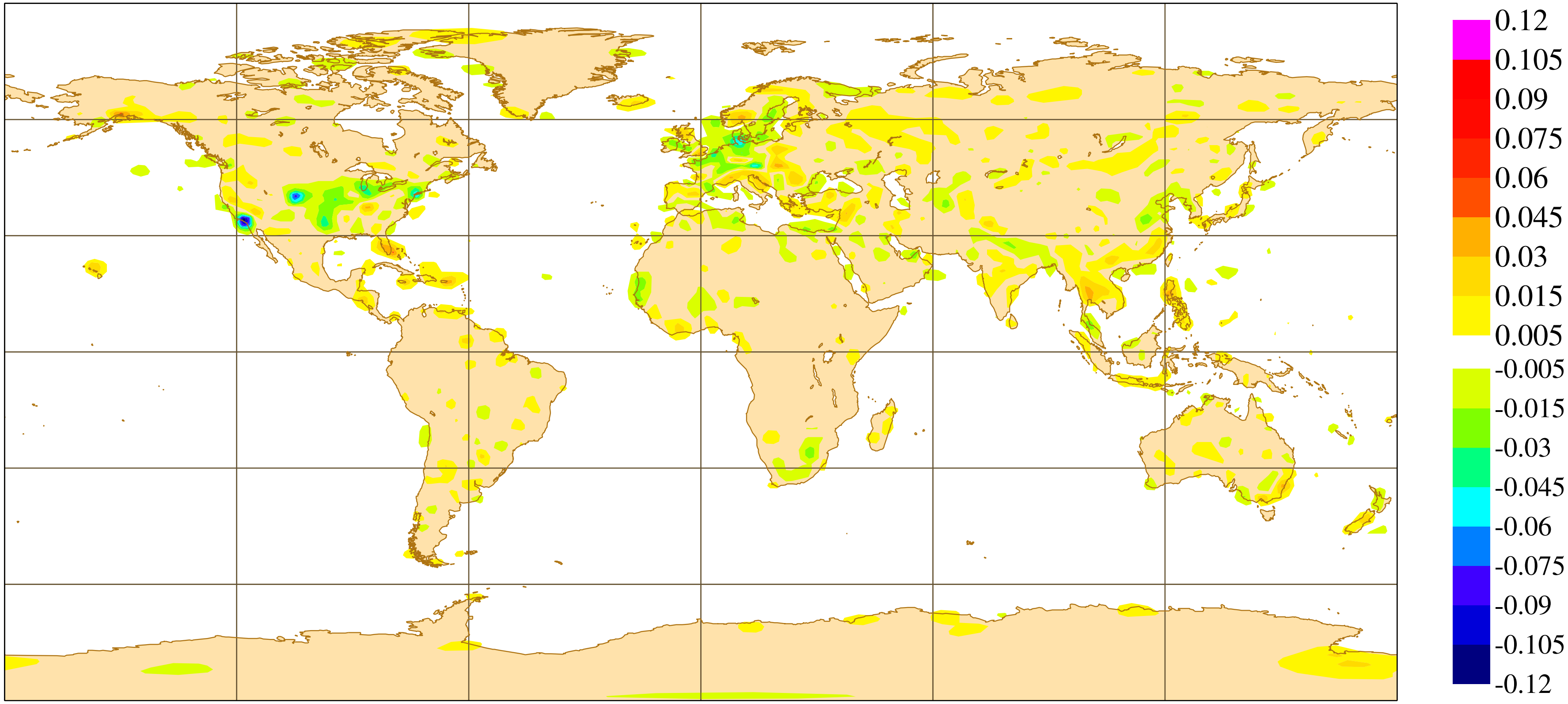


Northern hemisphere winter case (February 2004)

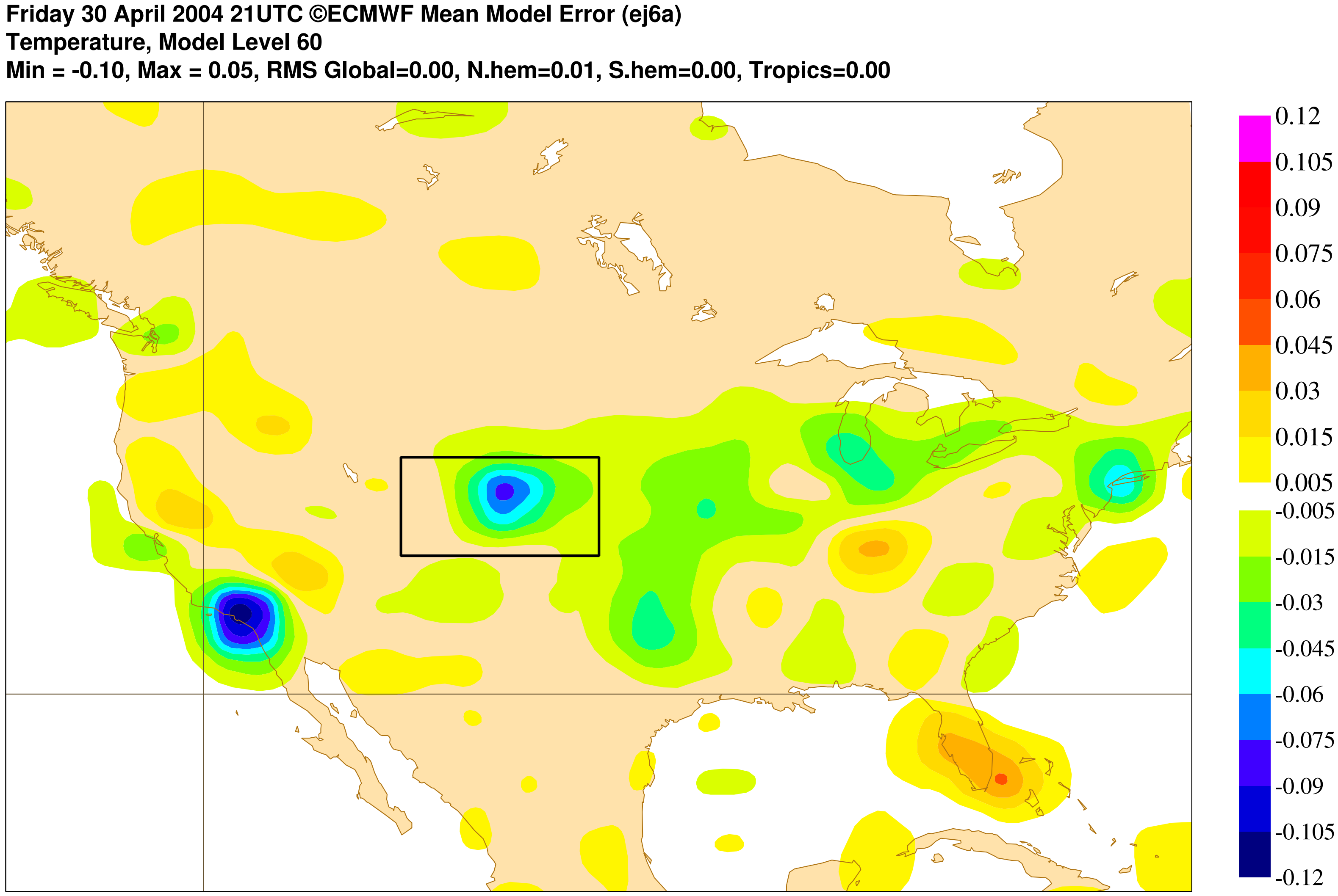
Model bias in the stratosphere is reduced.

Low Level Mean Model Error Forcing

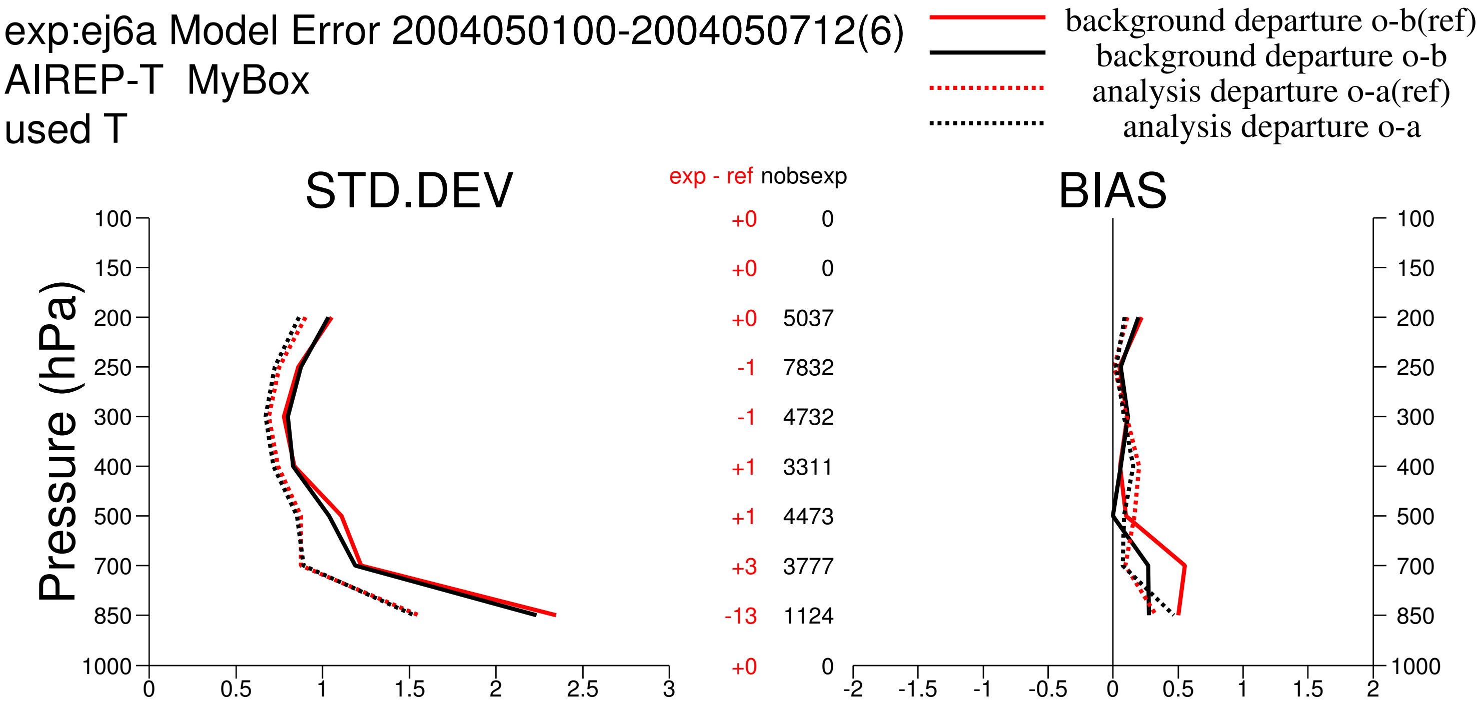
Friday 30 April 2004 21UTC ©ECMWF Mean Model Error (ej6a)
Temperature, Model Level 60
Min = -0.10, Max = 0.05, RMS Global=0.00, N.hem=0.01, S.hem=0.00, Tropics=0.00



Low Level Mean Model Error Forcing

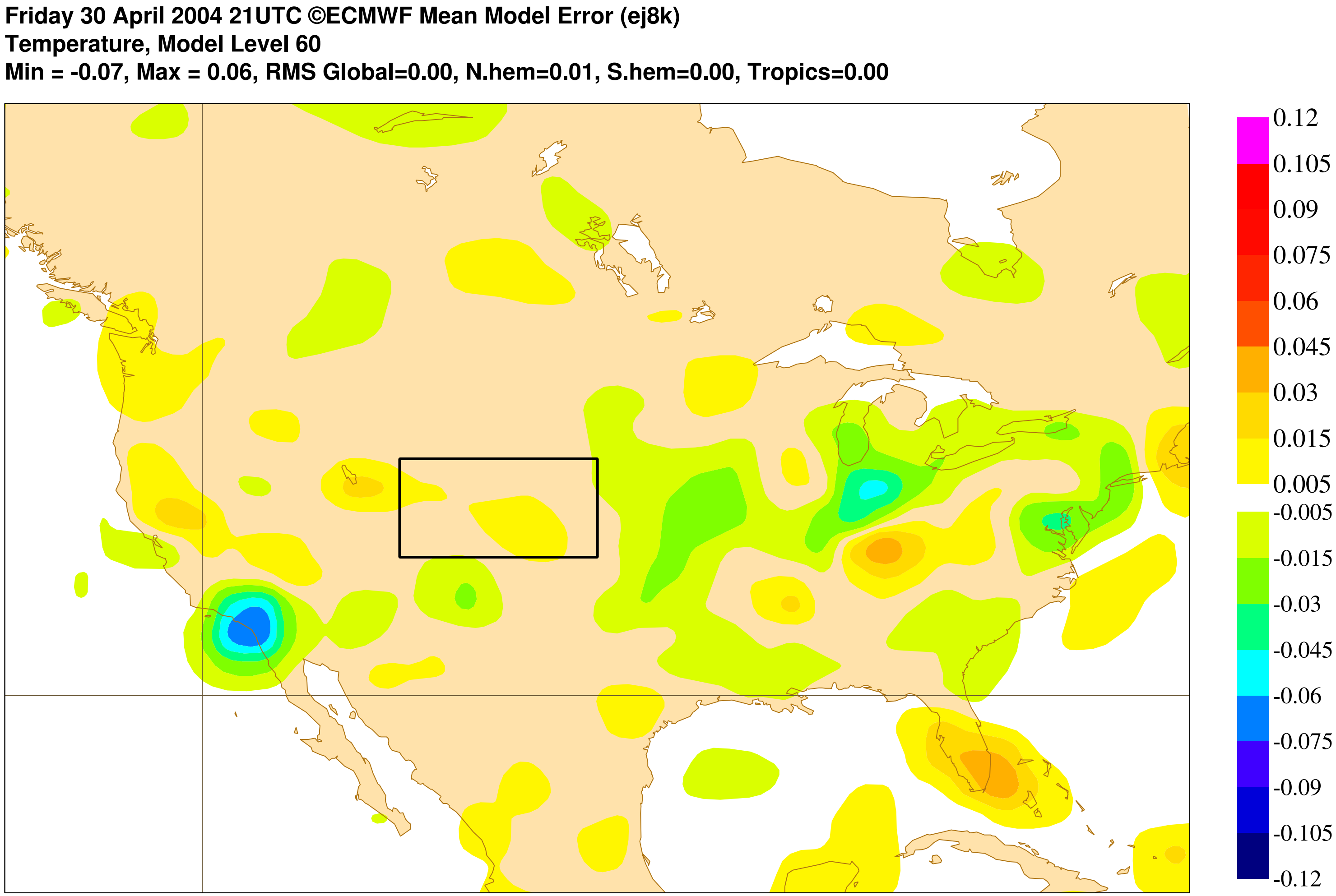


Fit to Observation with Model Error



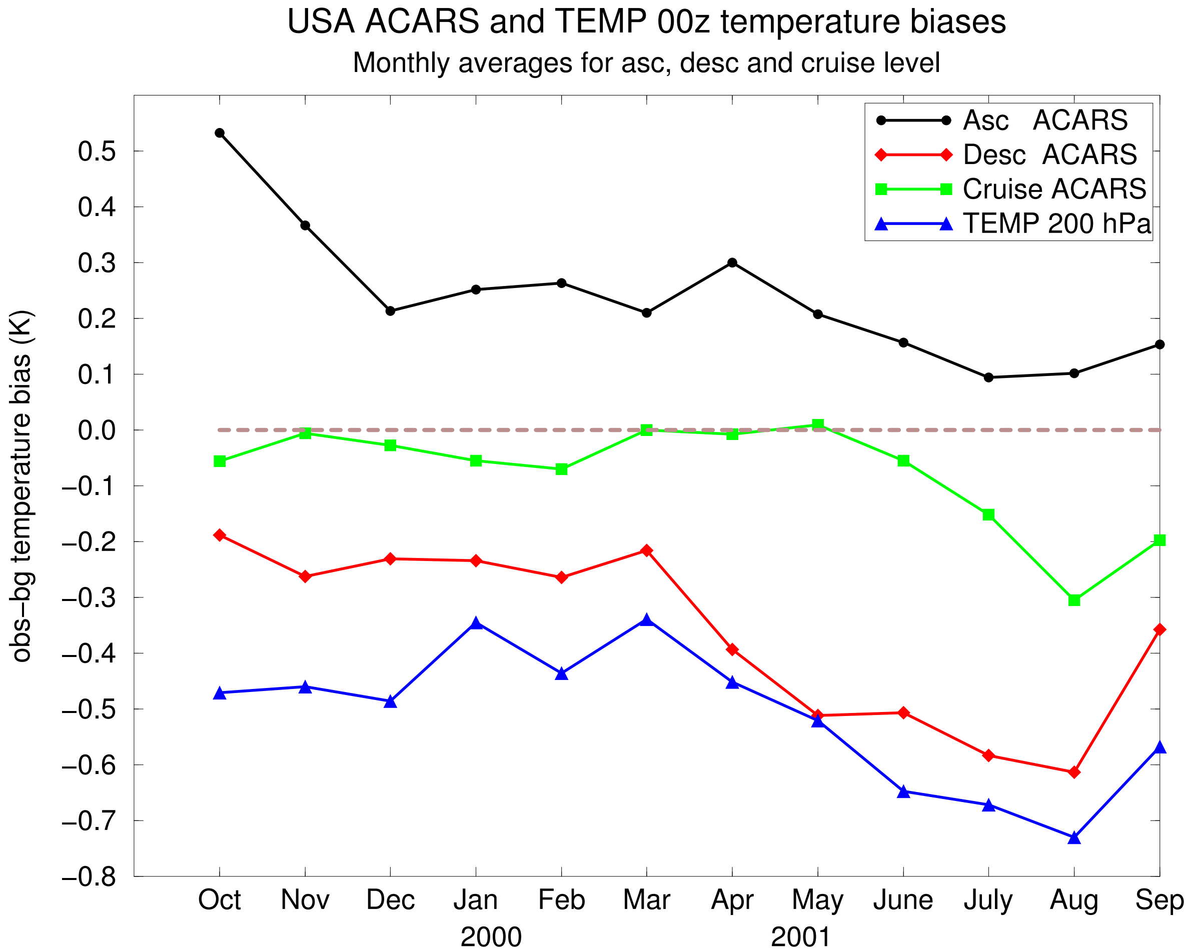
- The only significant source of data in the box is aircraft data (Denver airport).
- The bias for aircraft low level temperature observations was reduced.

Low Level Mean Model Error Forcing



Removing aircraft data in the box eliminates the spurious forcing.

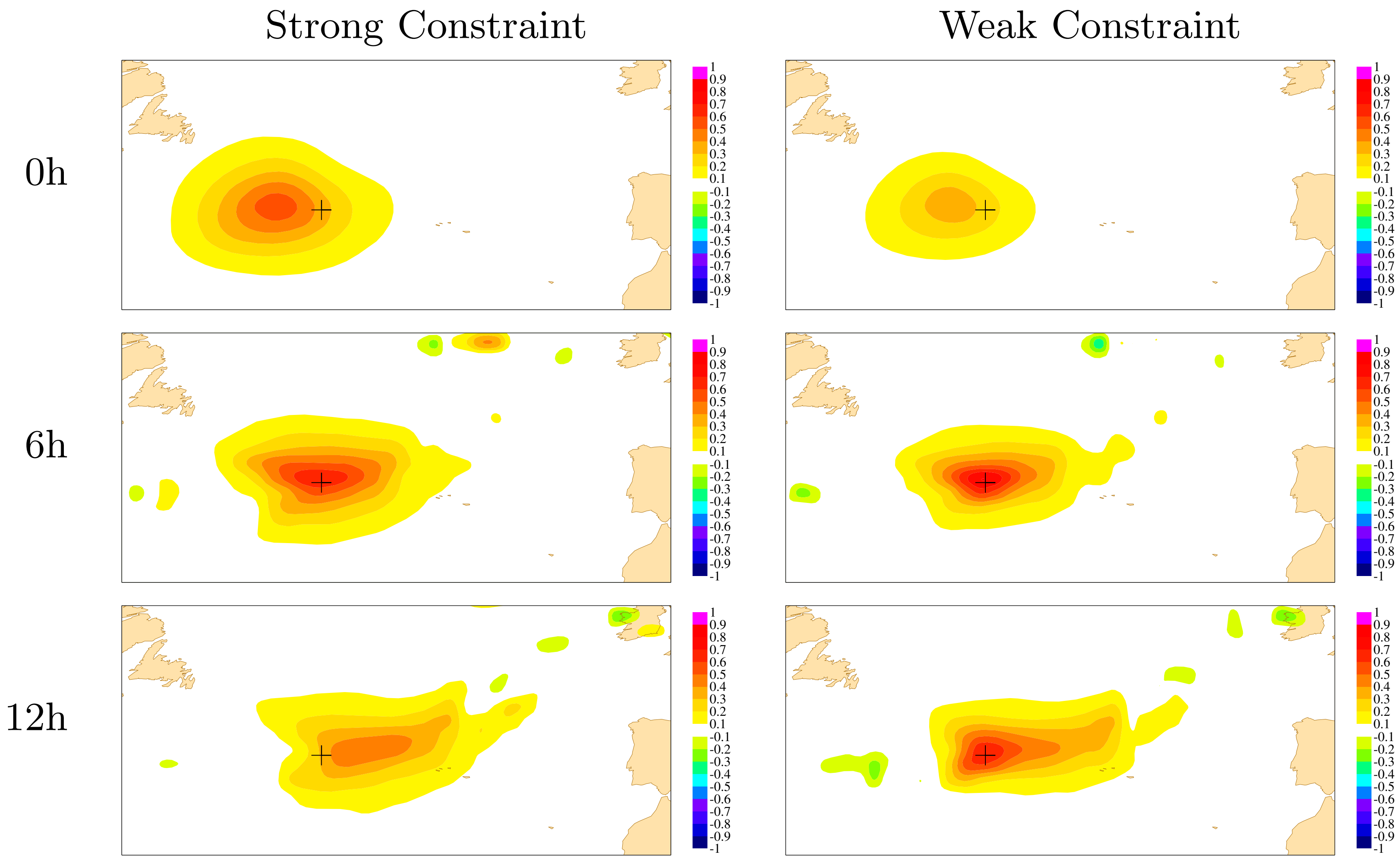
Aircraft Temperature Bias



Observations are biased.

Figure from Lars Isaksen

12 Observations Experiment



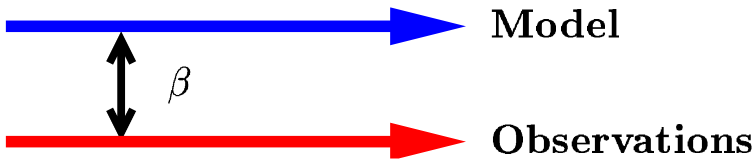
Systematic model error captures stationary errors (model or observations).

Model Bias Control Variable

4D-Var with Model Bias

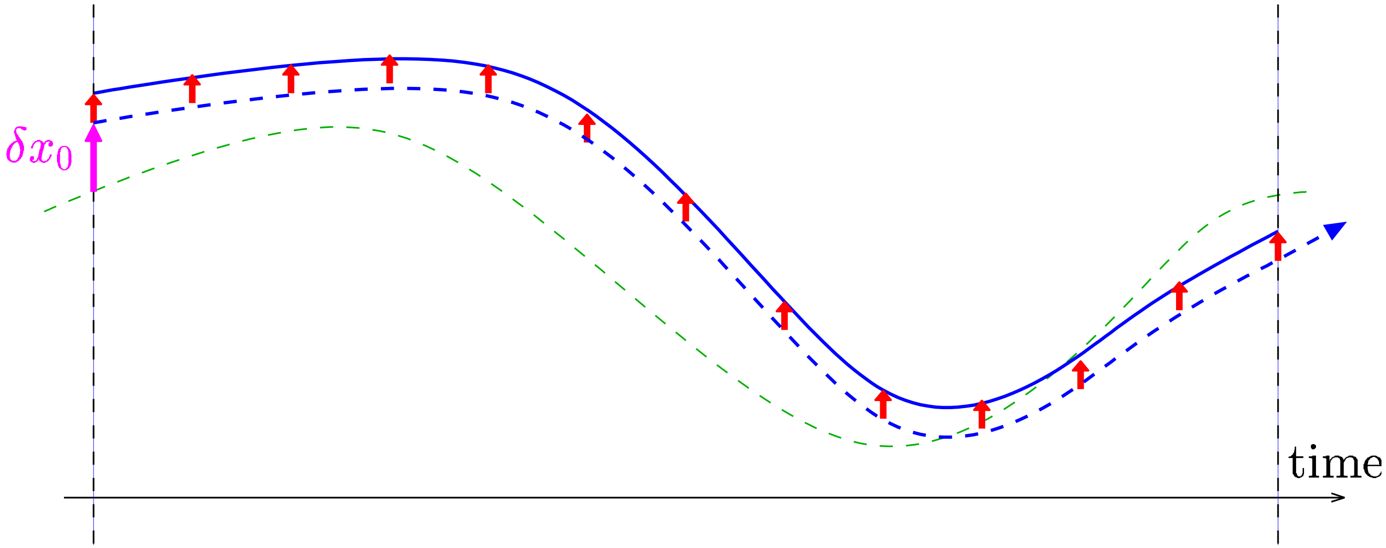
$$
\begin{aligned}
J(x_0, \beta) \quad &= \quad \frac{1}{2}\sum_{i=0}^{n}[\mathcal{H}(x_i^m + \beta_i) - y_i]^T R_i^{-1}[\mathcal{H}(x_i^m + \beta_i) - y_i] \\
&+ \quad \frac{1}{2}(x_0 - x_b)^T B^{-1}(x_0 - x_b) + \beta^T Q_\beta^{-1} \beta
\end{aligned}
$$

$$
\text{with } x_i^m = \mathcal{M}_{i,0}(x_0).
$$



- β_i is a 3D atmospheric state,
- The model is not perturbed,
- β sees global (model − all observations) bias,
- Does not correct for bias of one subset of observations against another subset of observations.

4D-Var with Model Bias



- Bias added to forecast at post-processing stage,
- Makes sense if β is slowly varying or constant ($\beta_i = \beta$),
- Information is propagated between obervations and IC control variable by TL and AD models (not modified),
- Model bias is represented by additional parameters, without entering the model equations,
- Optimisation problem is very similar to strong constraint 4D-Var.

Weak Constraint 4D-Var

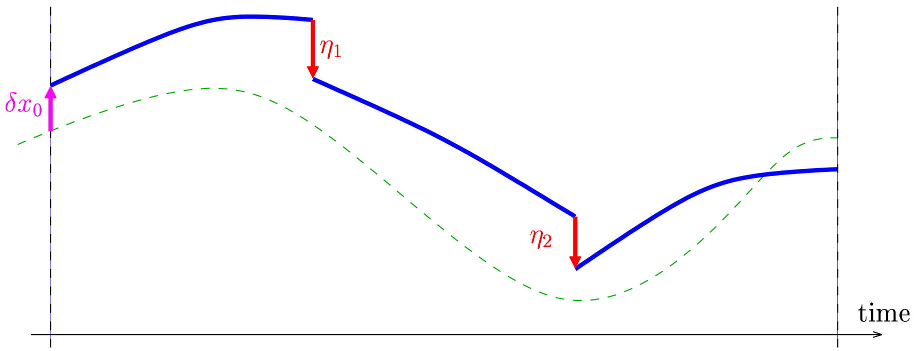
Model State Control Variable

- Use $\{x_i\}_{i=0,\ldots,n}$ as the control variable.
- Incremental cost function:

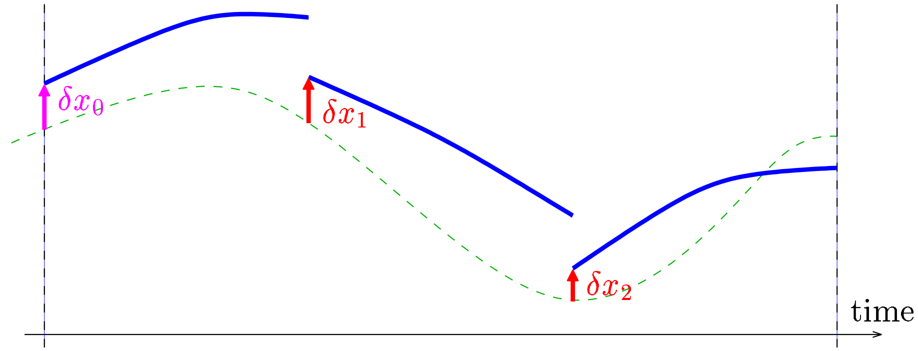
$$
\begin{aligned}
J(\delta x) &= \frac{1}{2}(\delta x_0 - b)^T B^{-1}(\delta x_0 - b) + \frac{1}{2}\sum_{i=0}^{n}(H\delta x_i - d_i)^T R_i^{-1}(H\delta x_i - d_i) \\
&+ \frac{1}{2}\sum_{i=1}^{n}(q_i + M_{i-1}\delta x_{i-1} - \delta x_i)^T Q_i^{-1}(q_i + M_{i-1}\delta x_{i-1} - \delta x_i)
\end{aligned}
$$

where $b = x^g - x_b$, $d_i = \mathcal{H}(x_i^g) - y_i$ and $q_i = \mathcal{M}_{i-1}(x_{i-1}^g) - x_i^g$.

Model State Control Variable



Forcing Control Variable



Model State Control Variable

- Model integrations within each time-step (or sub-window) are independent:
  - Information is not propagated across sub-windows by TL/AD models,
  - Natural parallel implementation (in theory...).
- Tangent linear and adjoint models:
  - can be used without modification,
  - propagate information between observations and control variable within each sub-window.

Model State Control Variable: Properties

- Preconditioning with $B^{-1/2}, Q_1^{-1/2}, \cdots, Q_n^{-1/2}$
- The Hessian of the preconditioned cost function is:

$$\hat{J}" = \begin{pmatrix} I + M_1^T M_1 & -M_1^T & & & & 0 \\ -M_1 & I + M_2^T M_2 & -M_2^T & & & \\ & -M_2 & I + M_3^T M_3 & & & \\ & & & & & \\ & & & & I + M_n^T M_n & -M_n^T \\ 0 & & & & -M_n & I \end{pmatrix} + \hat{J_o}"$$

- The off-diagonal terms propagate the information between the sub-windows.

Model State Control Variable: Properties

- Over one time step, $M_i \approx I$:

$$\hat{J}" \approx \begin{pmatrix} 2I & -I & & & 0 \\ -I & 2I & -I & & \\ & -I & \ddots & \ddots & \\ & & \ddots & 2I & -I \\ 0 & & & -I & I \end{pmatrix} + \hat{J_o}"$$

- The largest eigenvalue is:

$$\lambda_{max} \approx 4 + 2n_{obs}/n \max\left[(\sigma_b/\sigma_o)^2, (\sigma_q/\sigma_o)^2\right]$$

- Approximately the same as the maximum eigenvalue of strong constraint 4D-Var for the sub-windows.
- But...

Model State Control Variable: Properties

- Smallest eigenvalue:

$$\lambda_{min} \propto 1/n^2$$

- Condition number:

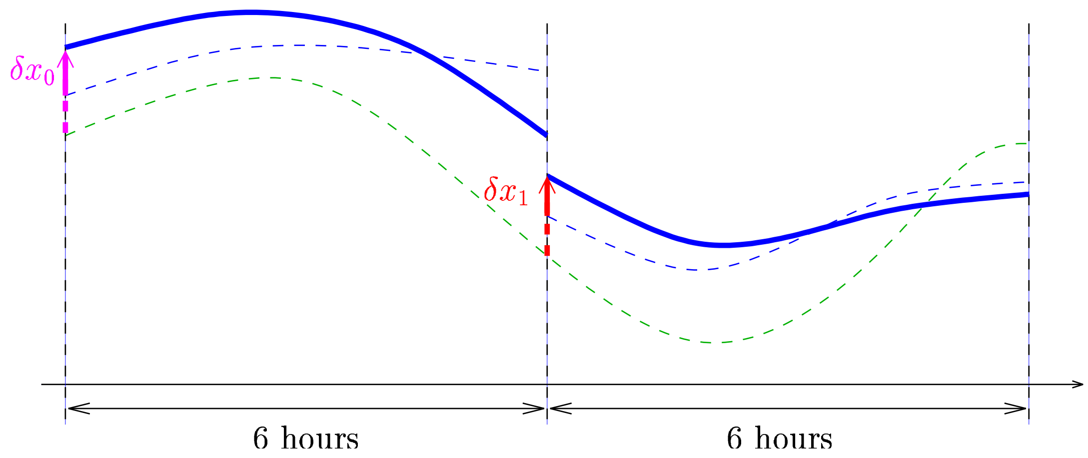
$$\kappa \approx 2\, n\, n_{obs}\, (\sigma_q/\sigma_o)^2$$

  - Larger than for strong constraint 4D-Var,
  - Increases with the number of sub-windows (it takes n iterations to propagate information).

- Simplified Hessian of the cost function
  - can be used as preconditioner,
  - $\approx$ Laplacian operator: small eigenvalues are obtained for constant perturbations which might be well observed and project onto eigenvectors of J_o" associated with large eigenvalues.

Weak Constraint 4D-Var: Examples

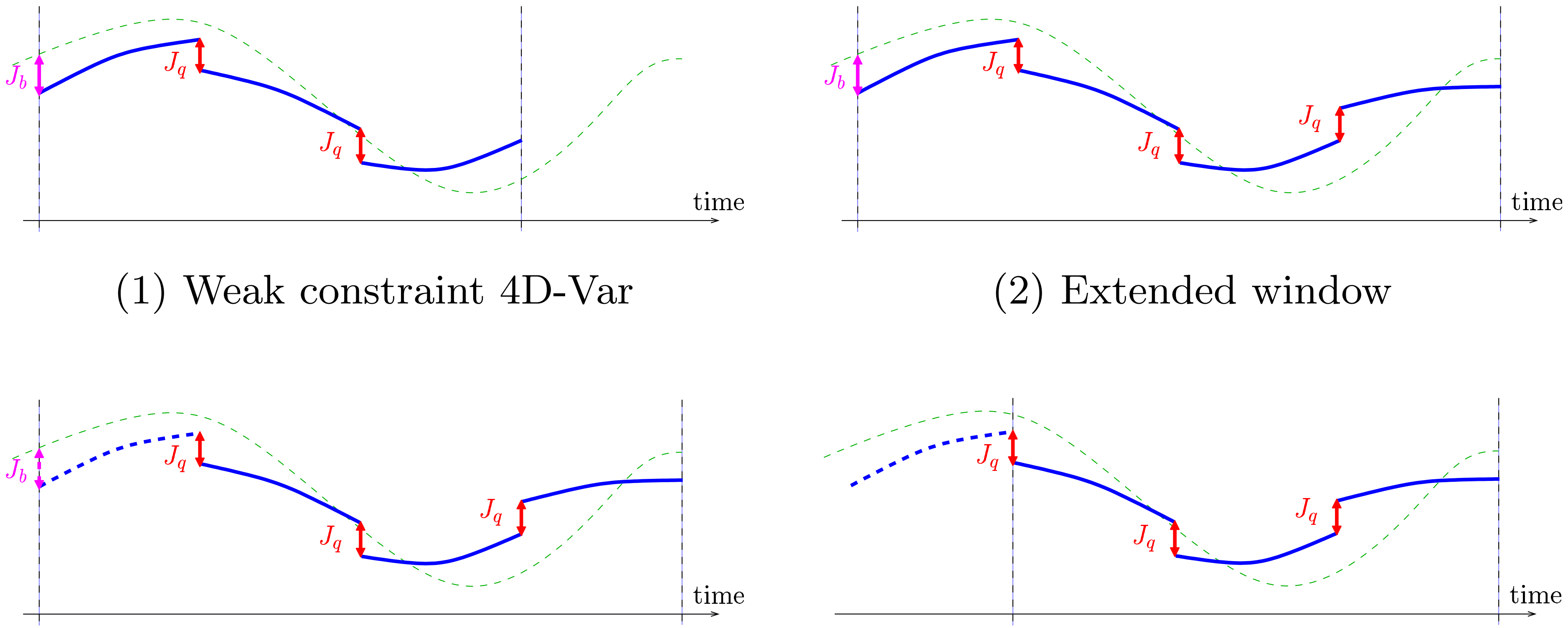
- 6-hour sub-windows:



  - Better than 6-hour 4D-Var: two cycles are coupled through J_q,
  - Better than 12-hour 4D-Var: more information (imperfect model), more control,
  - It makes more sense to use $Q = \alpha B$ in that case.

- One time step sub-windows:
  - Each assimilation problem is instantaneous = 3D-Var,
  - Equivalent to a string of 3D-Var problems coupled together and solved as a single minimisation problem,
  - Approximation can be extended to non instantaneous sub-windows.

Weak Constraint 4D-Var: Sliding Window



(1) Weak constraint 4D-Var

(2) Extended window

(3) Initial term has converged

(4) Assimilation window is moved forward

- This implementation is an approximation of weak contraint 4D-Var with an assimilation window that extends indefinitely in the past...
- ...which is equivalent to a Kalman filter that has been running indefinitely.

ECMWF

Conclusions

Control Variable in 4D-Var

4D-Var	4D-Var$_x$	4D-Var$_\eta$	4D-Var$_\beta$
x_0	x	$x_0,\ \eta$	$x_0,\ \beta$
$x_i = \mathcal{M}_i(x_{i-1})$	$x_i \approx \mathcal{M}_i(x_{i-1})$	$x_i = \mathcal{M}_i(x_{i-1}) + \eta_i$	$x_i = \mathcal{M}_{i,0}(x_0) + \beta_i$
$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$
3D Initial Condition	4D Model Trajectory	3D I.C. + Model Error Forcing	3D I.C. + Model Bias

Weak Constraint 4D-Var

- *Weak constraint 4D-Var* with model bias or forcing model error is essentially an initial value problem with parameter estimation (parameters happen to represent model error).
- *Weak constraint 4D-Var* with model state control variable is a 4D problem.
- J_q acts as coupling between sub-windows:
  - propagates information across sub-windows,
  - forgetting factor because model equations are not strictly verified.
- Takes into account the fact that the model is imperfect without directly estimating model error.
- *Long window weak constraint 4D-Var* is equivalent to a full rank Kalman smoother: the sliding window is an efficient algorithm to implement it.

Results

- Weak constraint 4D-Var is technically implemented in the IFS (x coming...),
- With constant model error forcing:
  - Fits the data more uniformly over the assimilation window,
  - Capture some model errors (winter stratosphere),
  - Captures some observation bias (ascending/descending aircrafts),
- Systematic differences between model and observations:
  - Distinction between model bias and observation bias?
  - Interactions with variational observation bias correction?
- Two tools to define the errors we wish to capture:
  - Model error covariance matrix (scales, ...),
  - Model for model error in η and β formulations (constant 3D state).

Future Work

- Compare various formulations of weak constraint 4D-Var (η, β, x),
- What type of model error (random vs. systematic vs. bias)?
- Determine appropriate model error covariance matrix,
- Interactions between model error and observation bias,
- Long window 4D-Var,
- Operational analysis and Reanalysis.