A new method for model error covariance estimation

Craig H. Bishop[1]

[1]Marine Meteorology Division, Naval Research Laboratory, Monterey, USA

Overview

- Todling's (2015a) KS+KF approach for **Q** estimation can be simplified using weak constraint 4DVAR.
- However, Todling's approach will fail if **R** estimates based on Desrozier et al. (2005).
- New Divide and Calibrate (DC) approach to **R** estimation
  - Accuracy of $\mathbf{R}_{\mathbf{DC}}$ with accurate **Q** in DA
  - Accuracy of $\mathbf{R}_{\mathbf{DC}}$ with inaccurate **Q** in DA
- Recovery of true **Q** even with very poor initial guesses of **R, P** and **Q**
- Concluding remarks

The "disrupted trajectories" form of weak constraint 4DVAR

- Consider the form of weak constraint 4DVAR that finds the 4D state $\underline{\mathbf{x}}^T = \left[\mathbf{x}_0^T, \mathbf{x}_1^T, \ldots, \mathbf{x}_n^T\right]$ that minimizes the cost function

$$-\ln\left[\rho_{post}\left(\underline{\mathbf{x}} \,|\, \underline{\mathbf{y}}\right)\right] \propto J\left(\underline{\mathbf{x}}\right) = J_b\left(\underline{\mathbf{x}}\right) + J_m\left(\underline{\mathbf{x}}\right) + J_o\left(\underline{\mathbf{x}}\right), \text{ where}$$

$$J_b\left(\underline{\mathbf{x}}\right) = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right)^T \mathbf{P}_0^{f\,-1}\left(\mathbf{x}_0 - \mathbf{x}_0^f\right),$$

$$J_m\left(\underline{\mathbf{x}}\right) = \frac{1}{2}\sum_{i=1}^{n}\left[M\left(\mathbf{x}_{i-1}\right) - \mathbf{x}_i\right]^T \mathbf{Q}^{-1}\left[M\left(\mathbf{x}_{i-1}\right) - \mathbf{x}_i\right], \text{ and}$$

$$J_o\left(\underline{\mathbf{x}}\right) = \frac{1}{2}\left[\underline{\mathbf{y}} - \underline{H}\left(\underline{\mathbf{x}}\right)\right]^T \mathbf{R}^{-1}\left[\underline{\mathbf{y}} - \underline{H}\left(\underline{\mathbf{x}}\right)\right]$$

$[M(\mathbf{x}_{i-1}) - \mathbf{x}_i]$ is a model error proxy.

However, unless everything is observed accurately, the differences $[M(\mathbf{x}_{i-1}) -$

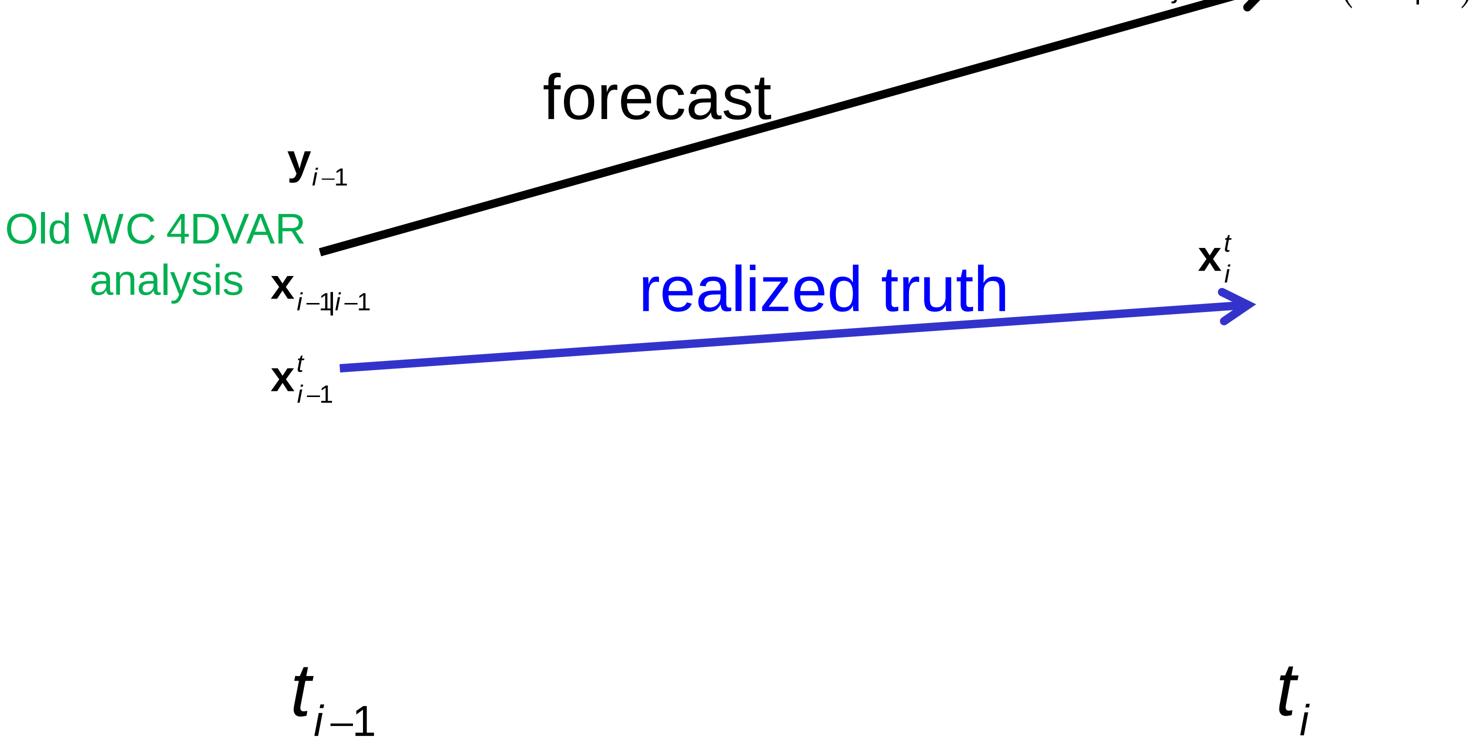
Todling's (2015) Q estimation approach using WC 4DVAR instead of KS + KF

$t_i =$ discrete time

$\mathbf{x}_i^t =$ truth at time t_i

$\mathbf{y}_i =$ observation of truth at time t_i

$\mathbf{x}_{i|j} =$ analysis at time t_i using all observations up to time t_j



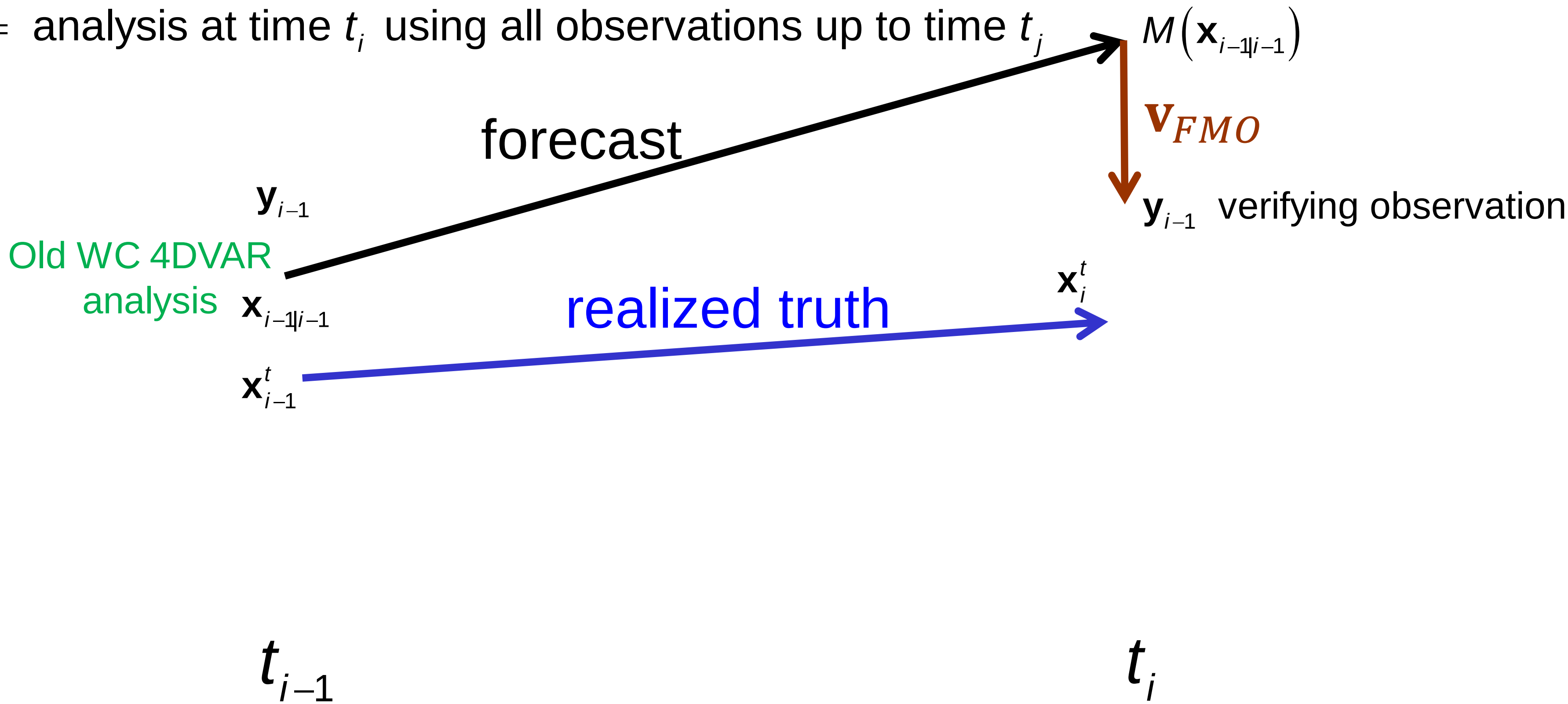
Todling's (2015) Q estimation approach using WC 4DVAR instead of KS + KF

t_i = discrete time

$\mathbf{x}_i^t$ = truth at time t_i

$\mathbf{y}_i$ = observation of truth at time t_i

$\mathbf{x}_{i|j}$ = analysis at time t_i using all observations up to time t_j



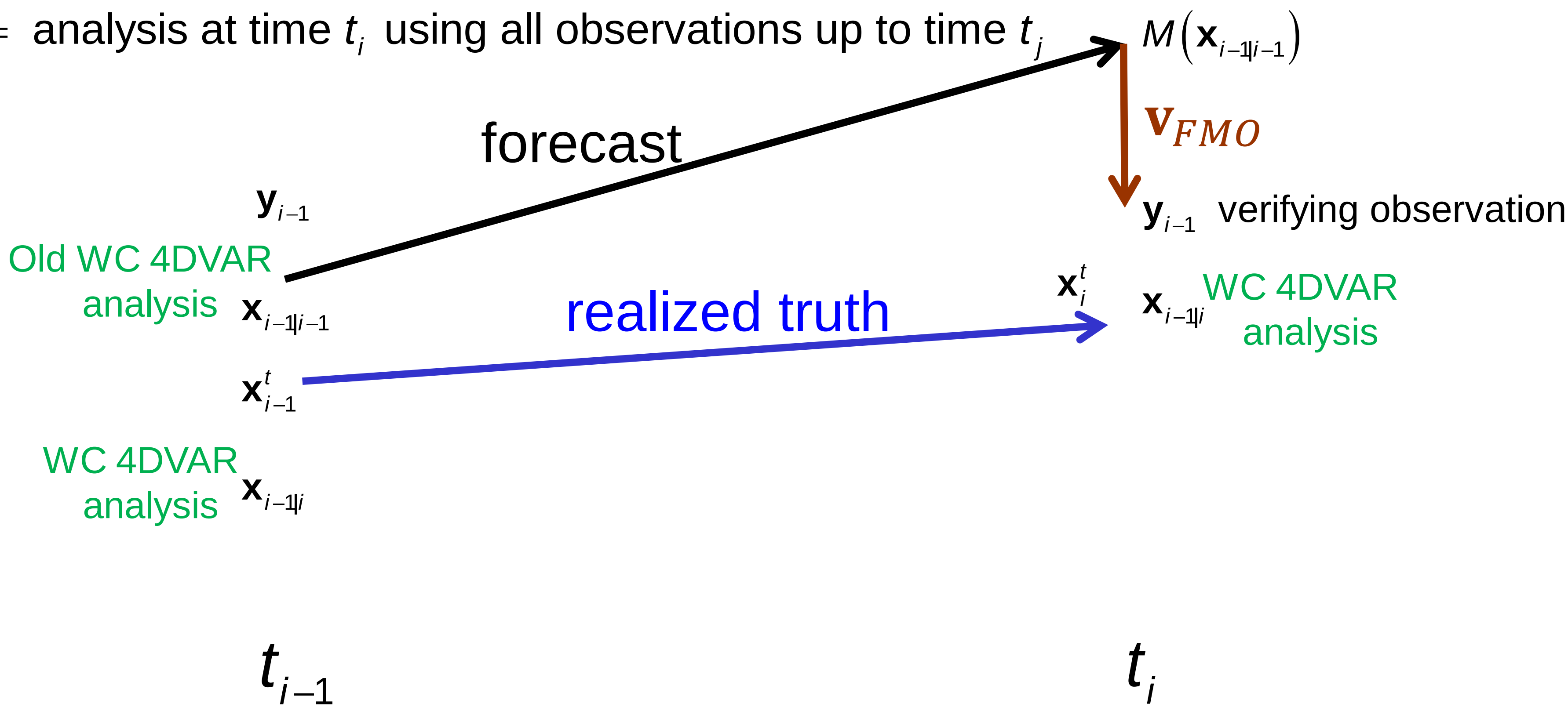
Todling's (2015) Q estimation approach using WC 4DVAR instead of KS + KF

t_i = discrete time

$\mathbf{x}_i^t$ = truth at time t_i

$\mathbf{y}_i$ = observation of truth at time t_i

$\mathbf{x}_{i|j}$ = analysis at time t_i using all observations up to time t_j



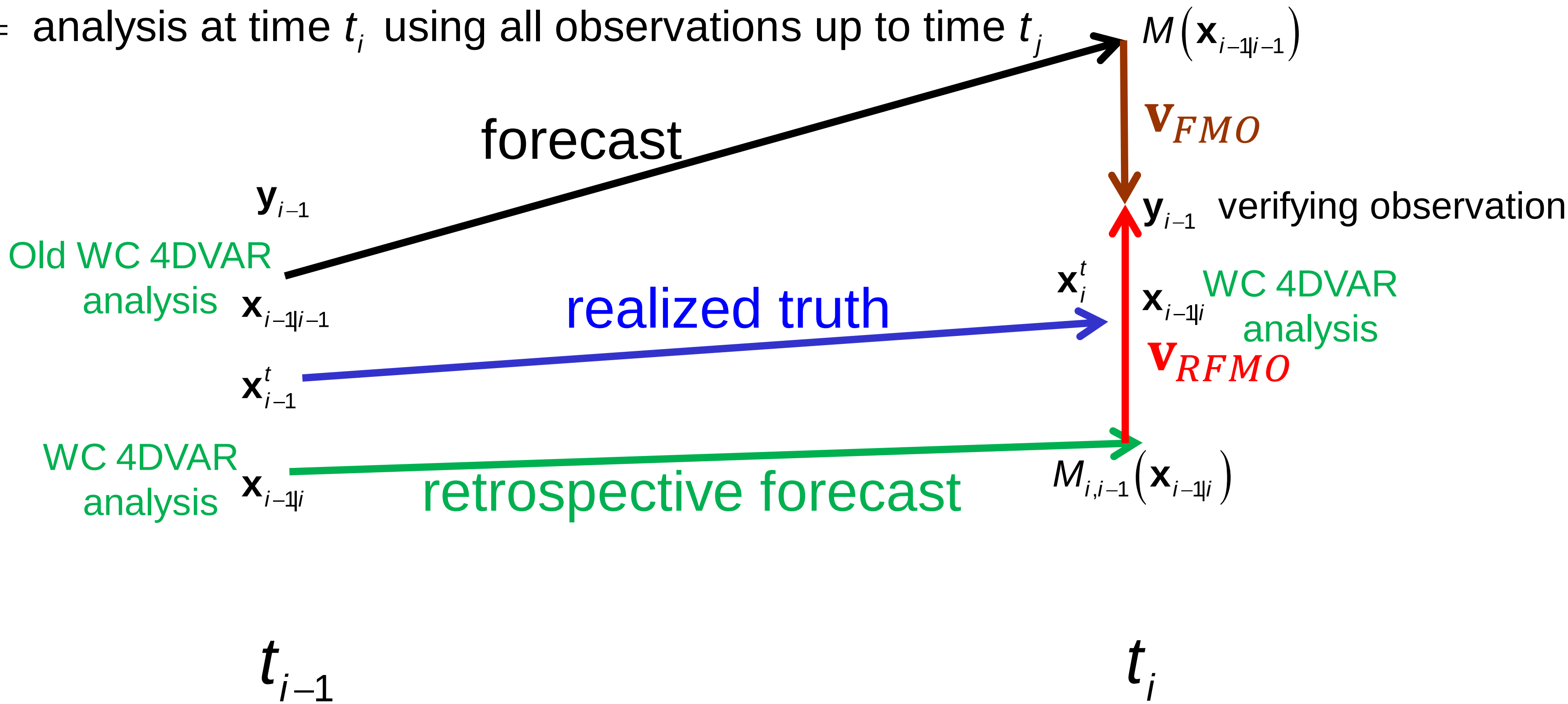
Todling's (2015) Q estimation approach using WC 4DVAR instead of KS + KF

t_i = discrete time

$\mathbf{x}_i^t$ = truth at time t_i

$\mathbf{y}_i$ = observation of truth at time t_i

$\mathbf{x}_{i|j}$ = analysis at time t_i using all observations up to time t_j



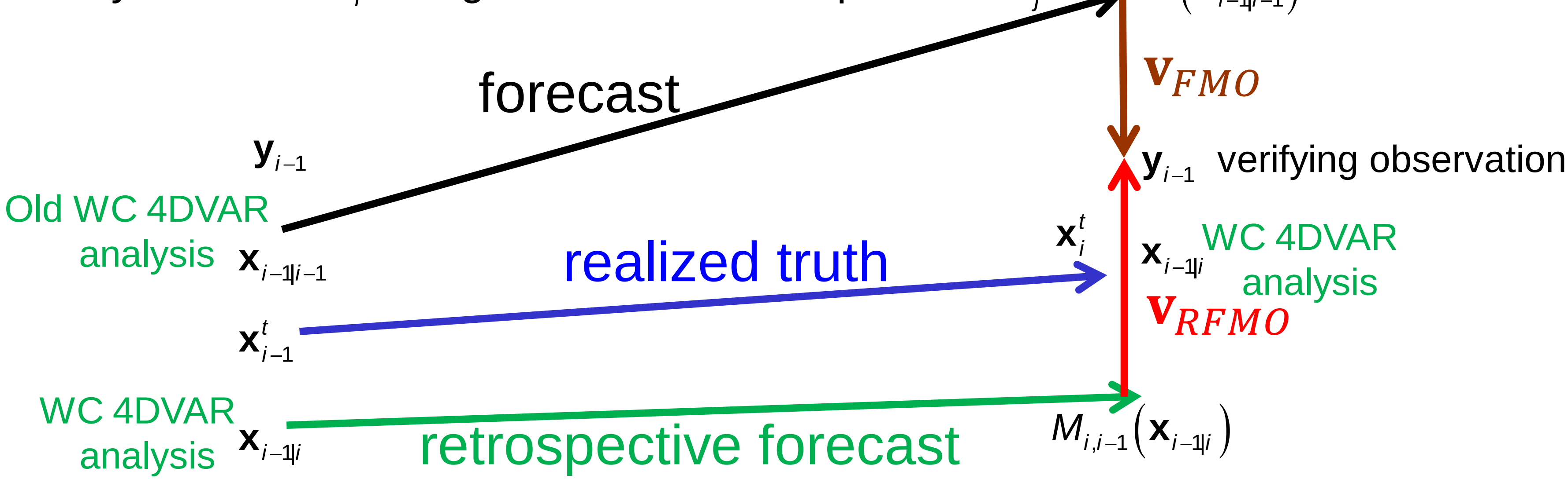
Todling's (2015) Q estimation approach using WC 4DVAR instead of KS + KF

$t_i =$ discrete time

$\mathbf{x}_i^t =$ truth at time t_i

$\mathbf{y}_i =$ observation of truth at time t_i

$\mathbf{x}_{i|j} =$ analysis at time t_i using all observations up to time t_j



Todling (2015a, July, QJ) shows that $\left\langle \mathbf{v}_{FMO}\mathbf{v}_{RFMO}^T \right\rangle = \mathbf{HQ}_k\mathbf{H}^T + \mathbf{R}_k$.

Hence, IF $\mathbf{R}_k$ is known $\mathbf{HQ}_k\mathbf{H}^T = \left\langle \mathbf{v}_{FMO}\mathbf{v}_{RFMO}^T \right\rangle - \mathbf{R}_k$.

However, …

Todling (2015b, Oct, QJ) shows that Desrozier et al. (2005) $\mathbf{R}_D$ is affected by unknown aspects of $\mathbf{Q}_k$ in a way that makes $\left\langle \mathbf{v}_{FMO}\mathbf{v}_{RFMO}^T \right\rangle - \mathbf{R}_D \neq \mathbf{H}\mathbf{Q}_k\mathbf{H}^T$

Divide and Calibrate method for R

- Step 1: Put every second observation into set A, and the remainder into set B
- Step 2: Make an analysis by assimilating all of the obs in set A using

$$\mathbf{x}_A^a = \mathbf{x}^f + \mathbf{P}^f\mathbf{H}_A^T\left(\mathbf{H}_A\mathbf{P}^f\mathbf{H}_A^T + \mathbf{R}_A\right)^{-1}\left[\mathbf{y}_A - H_A\left(\mathbf{x}^f\right)\right]$$

Divide and Calibrate method for R

$$H^B\left(\mathbf{x}_A^a\right) = H^B\left(\mathbf{x}^f\right) + \mathbf{H}^B\mathbf{P}\mathbf{H}^A\left(\mathbf{H}^A\mathbf{P}\mathbf{H}^A + \mathbf{R}^A\right)^{-1}\left[\mathbf{y}_A - H^A\left(\mathbf{x}^f\right)\right]$$

assimilated observation $\mathbf{y}_A$

unassimilated observation $\mathbf{y}_B$

analysis $H^A\left(\mathbf{x}_A^a\right)$

analysis $H^B\left(\mathbf{x}_A^a\right)$

forecast $H^A\left(\mathbf{x}^f\right)$

A

B

Divide and Calibrate method for R

$$H^B\left(\mathbf{x}_A^a\right) = H^B\left(\mathbf{x}^f\right) + \mathbf{H}^B\mathbf{P}\mathbf{H}^A\left(\mathbf{H}^A\mathbf{P}\mathbf{H}^A + \mathbf{R}^A + \delta\mathbf{R}^A\right)^{-1}\left[\mathbf{y}_A - H^A\left(\mathbf{x}^f\right)\right]$$

assimilated observation $\mathbf{y}_A$

$\mathbf{y}_B$ unassimilated observation

analysis $H^A\left(\mathbf{x}_A^a\right)$ —— $H^B\left(\mathbf{x}_A^a\right)$ analysis

forecast $H^A\left(\mathbf{x}^f\right)$

A B

This sensitivity is the gradient of $\mathbf{x}_A^a$ with respect to changes in $\mathbf{R}_A$.

Divide and Calibrate method for R

- Step 1: Put every second observation into set A, and the remainder into set B
- Step 2: Make an analysis by assimilating all of the obs in set A using

$$\mathbf{x}_A^a = \mathbf{x}^f + \mathbf{P}^f\mathbf{H}_A^T\left(\mathbf{H}_A\mathbf{P}^f\mathbf{H}_A^T + \mathbf{R}_A\right)^{-1}\left[\mathbf{y}_A - H_A\left(\mathbf{x}^f\right)\right]$$

This analysis is sensitive to infinitessimal changes $\delta\mathbf{R}_A$ in $\mathbf{R}_A$

$$\mathbf{x}_A^a + \delta\mathbf{x}_A = \mathbf{x}^p + \mathbf{P}^f\mathbf{H}_A^T\left[\mathbf{H}_A\mathbf{P}^f\mathbf{H}_A^T + \left(\mathbf{R}_A + \delta\mathbf{R}_A\right)\right]^{-1}\left[\mathbf{y}_A - H_A\left(\mathbf{x}^f\right)\right]$$

This sensitivity is the gradient of $\mathbf{x}_A^a$ with respect to changes in $\mathbf{R}_A$.

- Step 3: Simplify, e.g. by assuming that $\mathbf{R}_A = \mathrm{diag}\left(\mathbf{r}_A\right) = r_A\mathbf{I}$ for each observation type
- Step 4: Use variational method to find the value of $\mathbf{r}_A$ that minimizes

$$J_{ob}^u\left[\mathbf{r}_A\right] = \underbrace{\frac{1}{2}\left[\mathbf{y}^B - H^B\left(\mathbf{x}_A^a\right)\right]^T\mathbf{F}^{-1}\left[\mathbf{y}^B - H^B\left(\mathbf{x}_A^a\right)\right]}_{\text{Distance from unassimilated observations}} + \underbrace{\frac{1}{2}\left(\mathbf{r}_A - \mathbf{r}_A^f\right)^T\mathbf{G}^{-1}\left(\mathbf{r}_A - \mathbf{r}_A^f\right)}_{\text{Distance from prior guess of } \mathbf{r}_A}$$

- and compute the average of $diag\left[\left(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{P}^f\mathbf{H}^T\mathbf{R}^{-1/2}\right)\right]$ corresponding to this minimum
- Step 5: Repeat steps 1-4 but this time observations in set B are assimilated while observations in set A are left unassimilated.
- Step 6: Take the average of the ratios of obtained from Steps 4 and 5.

Ensure that the estimated R and P are consistent with innovation statistics

- Step 7: Recognize that the distance of an analysis from unassimilated observations solely depends solely on the *ratio* of **R** to **P** given by $\mathbf{R}^{-1/2}\mathbf{H}\mathbf{P}^{f}\mathbf{H}^{T}\mathbf{R}^{-1/2}$
- Step 8: Ensure that **R** and **P** are consistent with innovation statistics while preserving the ratio of **R** and **P** obtained from Steps 1-6 by multiplying them both by the same factor α that ensures that

$$J_{\min} = \frac{\left[\mathbf{y} - H\left(\mathbf{x}^{f}\right)\right]^{T}\left(\mathbf{H}\alpha\mathbf{P}^{f}\mathbf{H}^{T} + \alpha\mathbf{R}\right)^{-1}\left[\mathbf{y} - H\left(\mathbf{x}^{f}\right)\right]}{p} = 1,$$

then set $\mathbf{P}^{f}_{new} = \alpha\mathbf{P}^{f}$ and $\mathbf{R}_{new} = \alpha\mathbf{R}$

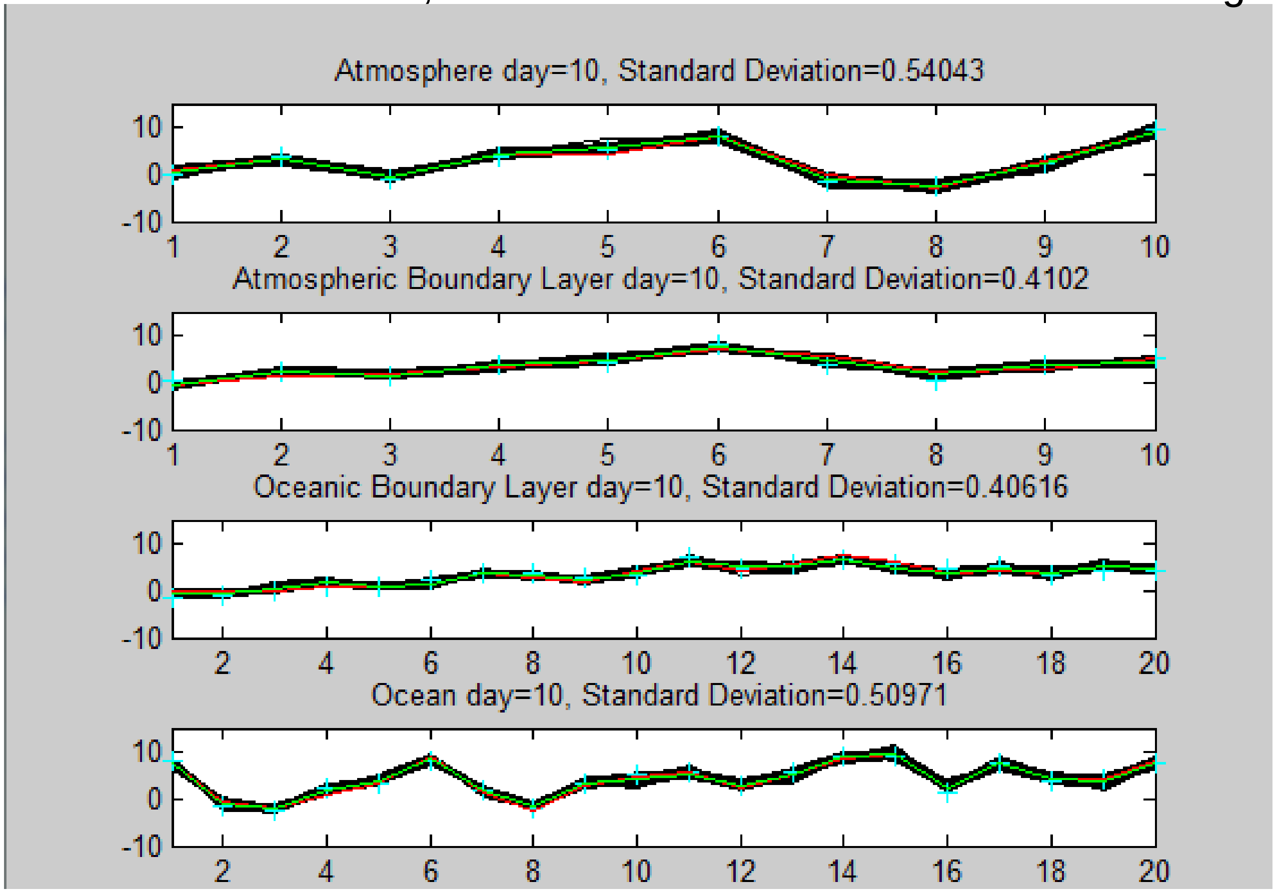
- where **y** contains all the observations (sets A and B) and *p* is the corresponding total number of observations.
- Step 9: Using the updated values of **R** and **P**, go ahead and assimilate all the observations using Weak constraint 4DVAR.
- Step 10: Repeat until you have a moderately stable **R** value and enough realizations to obtain another estimate of **Q** using Todling's (2015) equations.
- Incorporate this value in your DA/ensemble forecasting scheme and have another go at estimating **R** and **Q** by going back to step 1.

System to be used for tests

- 4-level idealized coupled model based on a stochastic version of Lorenz 95 model (model 1 of Lorenz 2005).
- At each time step, WC 4DVAR is used to create both a current analysis and a retrospective analysis – both of which will later be needed to apply Todling's equation.
- The square root Extended Kalman Filter (EKF) is used for propagating and updating error covariances. (An EnKF could have been used for this).

Weak constraint 4DVAR and model error detection

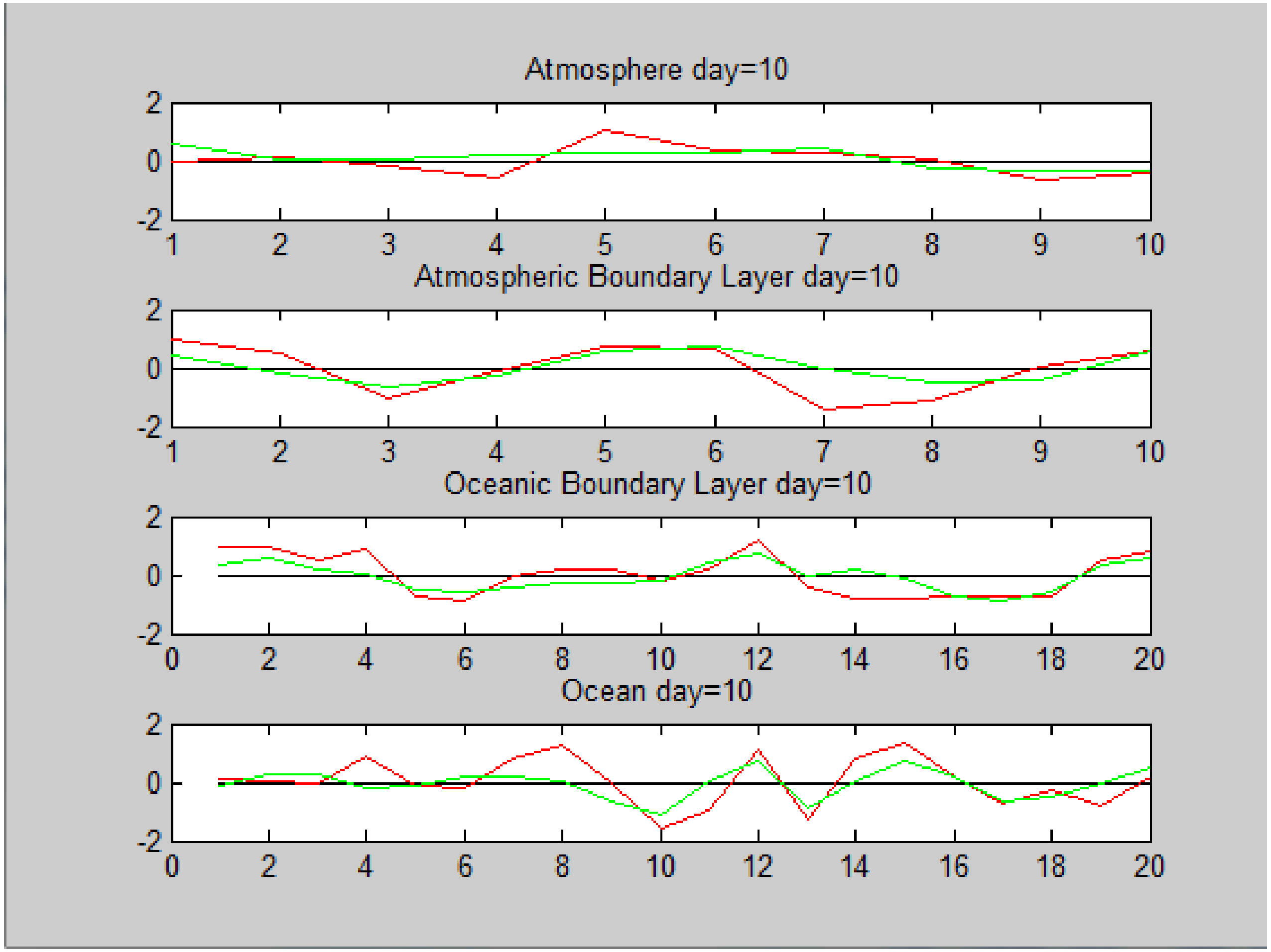
- Std of model error = 0.6, model error correlation same as model climatology.
- Std of observation error = 0.6, observation error correlation matrix is diagonal.



Red line is true state:
Cyan +s are observations: Truth plus observational noise
Black lines are ensemble members that have the same covariance as EKF P^a
Green line is the WC 4DVAR analysis

Weak constraint 4DVAR and model error detection

- Std of model error = 0.6, model error correlation same as model climatology.
- Std of observation error = 0.6, observation error correlation matrix is diagonal.

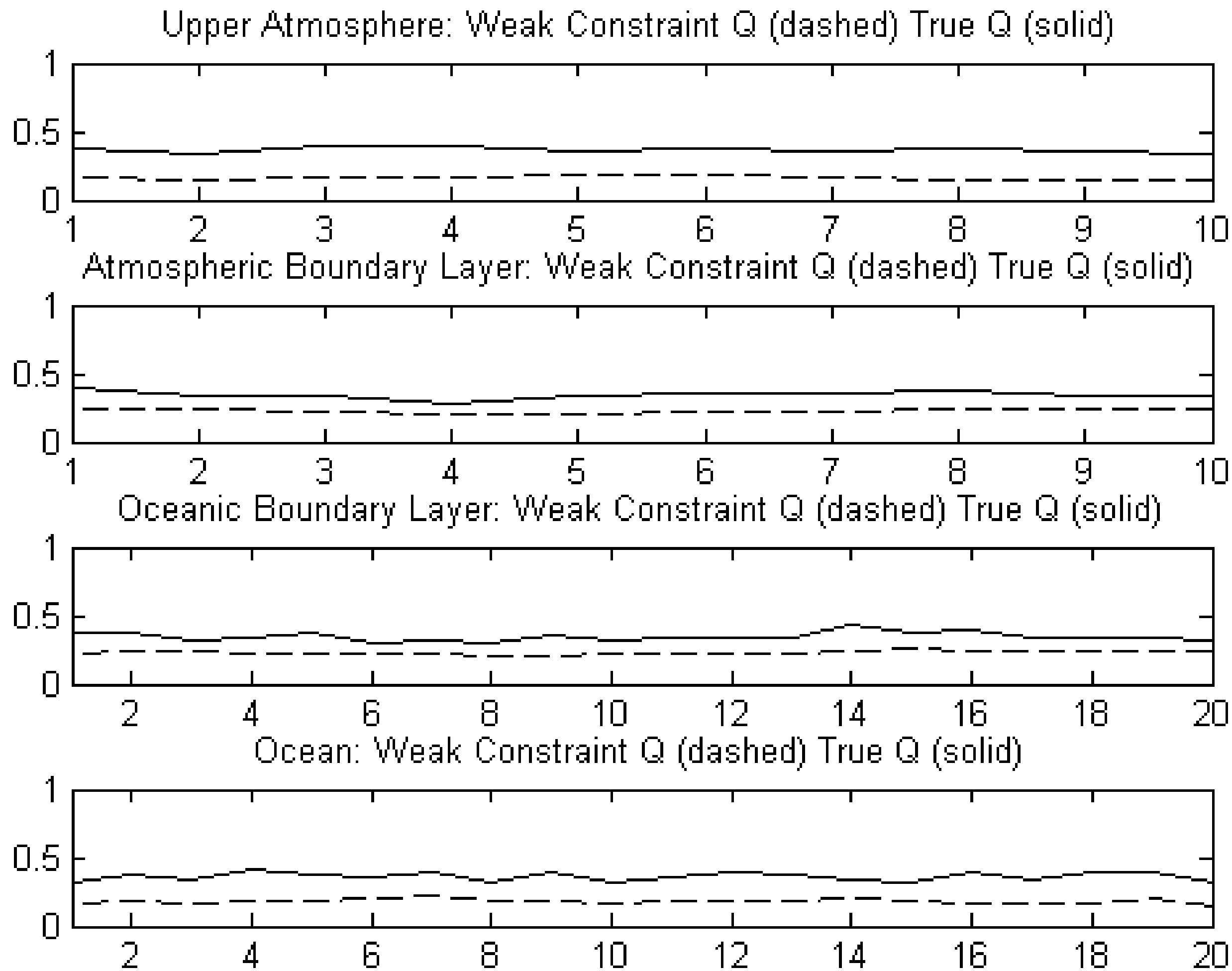


Red line is true model error: Obtained from a random draw

Green line is 4DVAR estimate of model error

Variance of estimates of model error obtained with WC 4DVAR

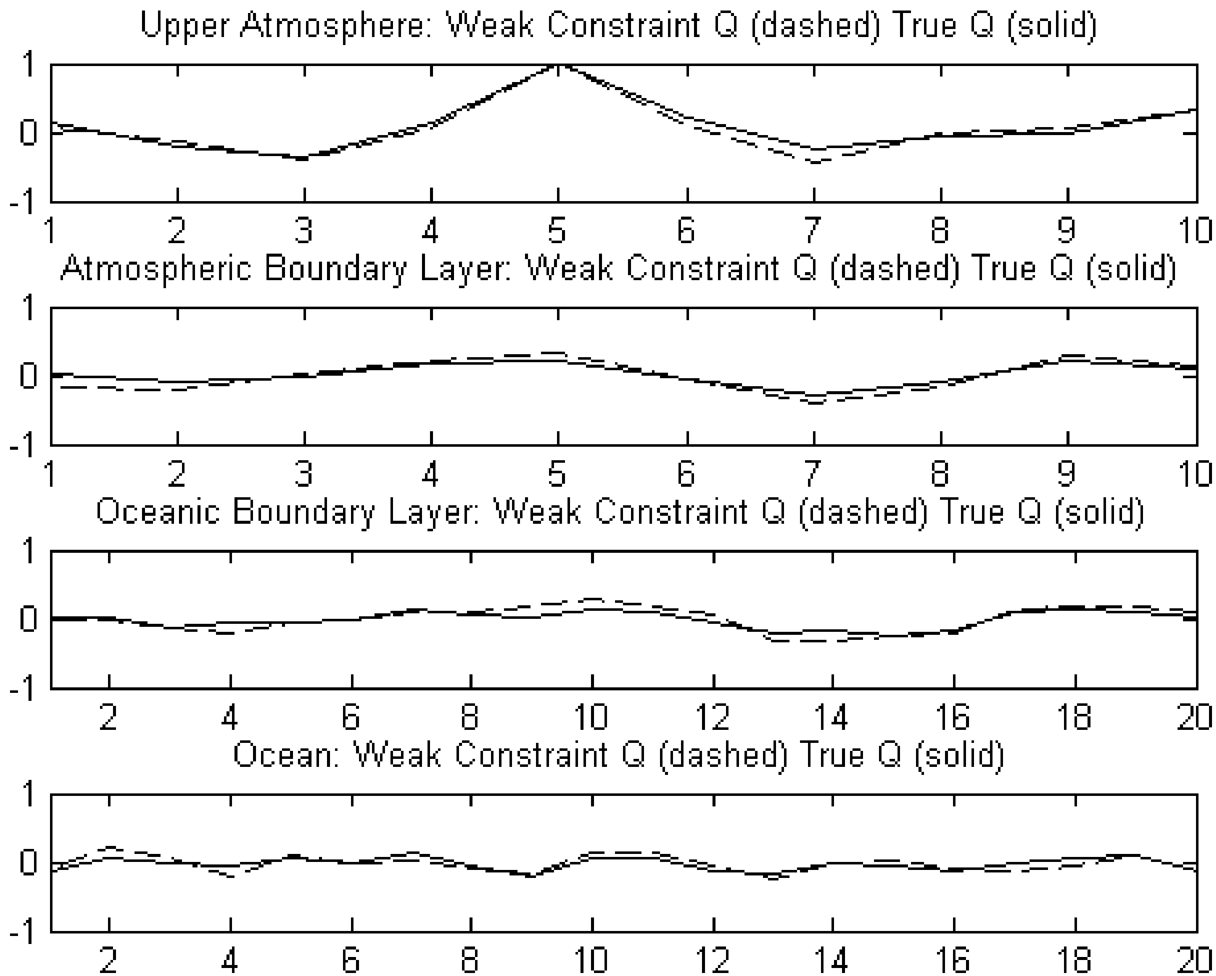
- Q = 0.36, R = 0.36, 60 days, 60x4=240 model error realizations



Weak constraint model error estimates are under-variant!

Correlation of model error proxies from WC 4DVAR with upper atmos

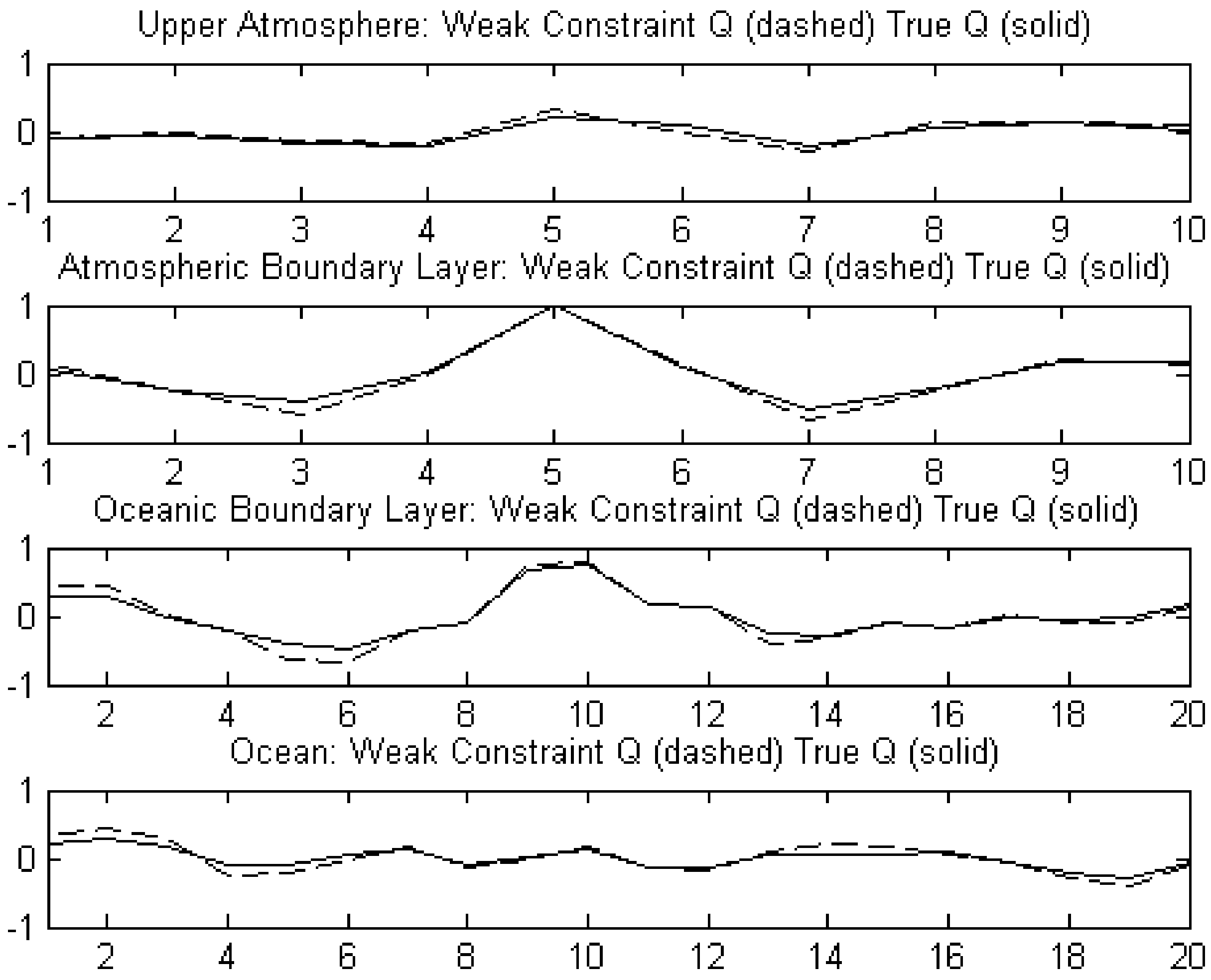
- Q = 0.36, R = 0.36, 60 days, 60x4=240 model error realizations



WC model error estimates have correct correlation function

Correlation of model error proxies from WC 4DVAR with atmos BL

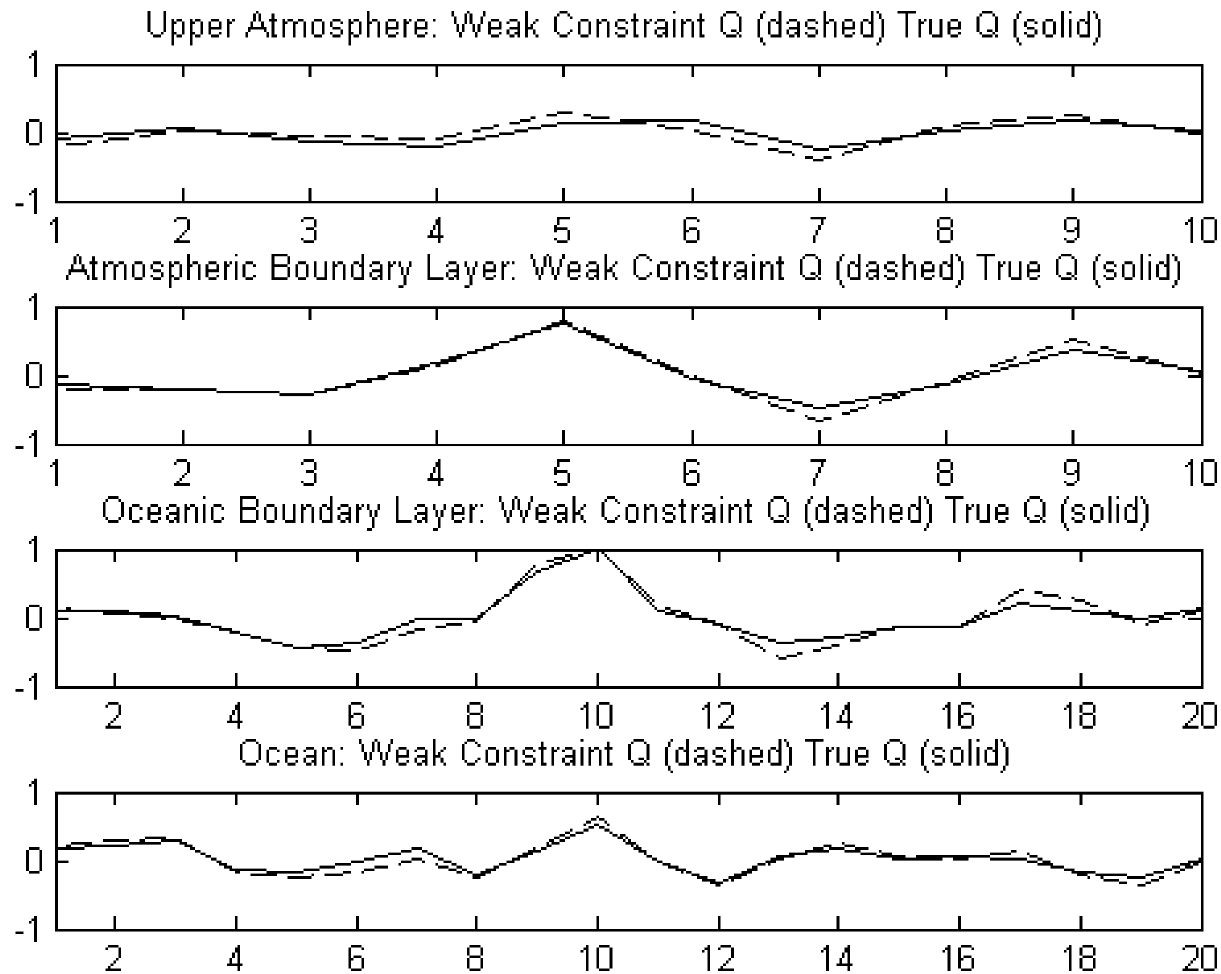
- Q = 0.36, R = 0.36, 60 days, 60x4=240 model error realizations



WC model error estimates have correct correlation function

Correlation of model error proxies from WC 4DVAR with ocean BL

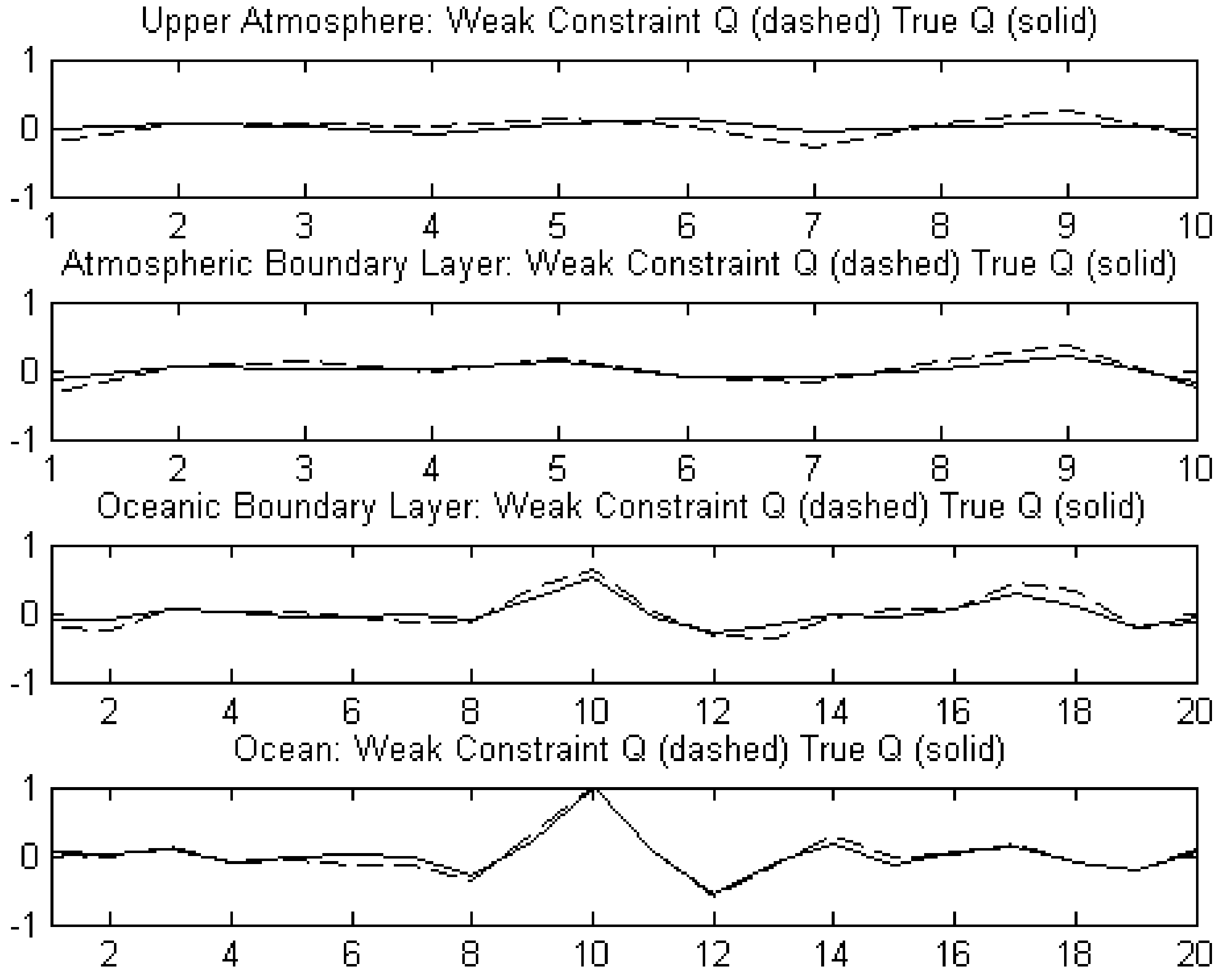
- Q = 0.36, R = 0.36, 60 days, 60x4=240 model error realizations



WC model error estimates have correct correlation function

Correlation of model error proxies from WC 4DVAR with deep ocean

- Q = 0.6, R = 0.6, 60 days, 60x4=240 model error realizations

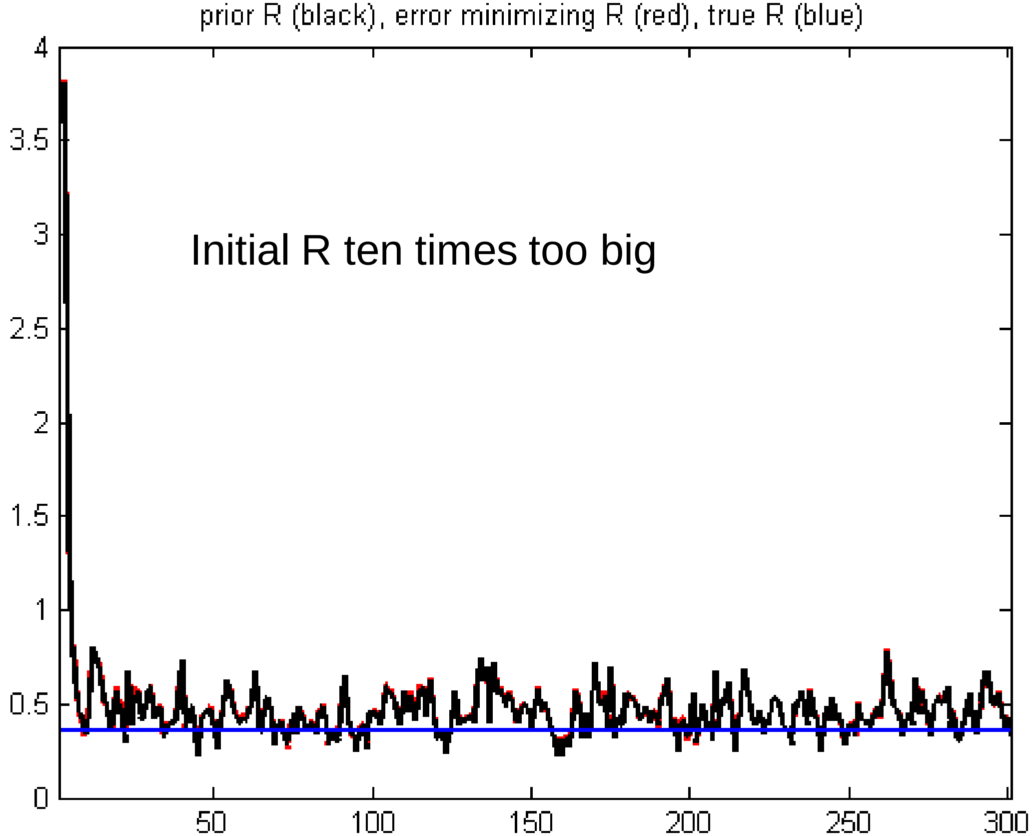


WC model error estimates have correct correlation function

Test of DC method for finding R Q_DA=Q_True

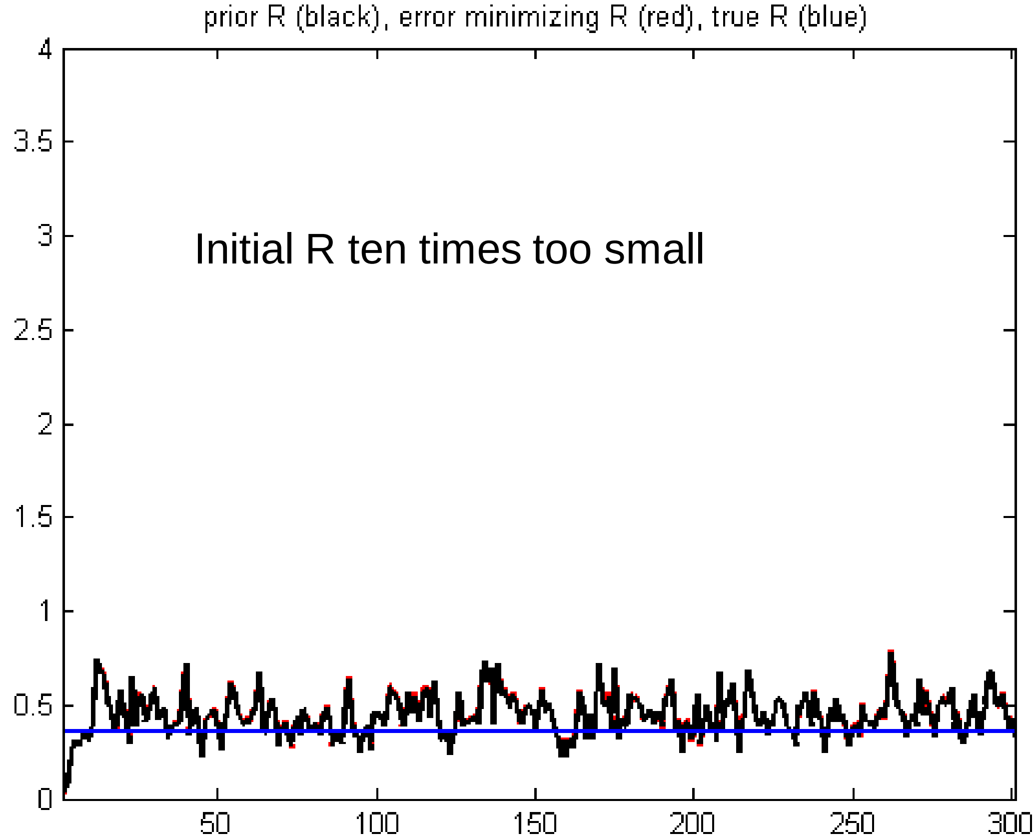
- True model error variance changes in horizontal
  - **Q_true** *variance* ranges from 0.36 on left half of domain to 1.08 on right half.
  - **Q_true** *correlation matrix* equal to the climatological correlation matrix.
- DA has
  - **Q_DA** = **Q_true**
- Actual observations have
  - **R_true**=0.36 **I** (diagonal with variance 0.36)
- DA scheme has
  - **R_DA**=3.6 **I** (10 times bigger than the true error variance)
- Perform Divide and Calibrate **R** estimation for 300 DA cycles.

Test of DC method for finding R
Q_DA=Q_True



R estimate (black) quickly converges to a value close to **R_true** (blue) – but it is slightly higher. This gives smaller corrections at unassimilated observations than if using **R_true** – an error reducing strategy when forecast error correlations are imperfect.

Test of on-line method for finding R Q_DA=Q_True

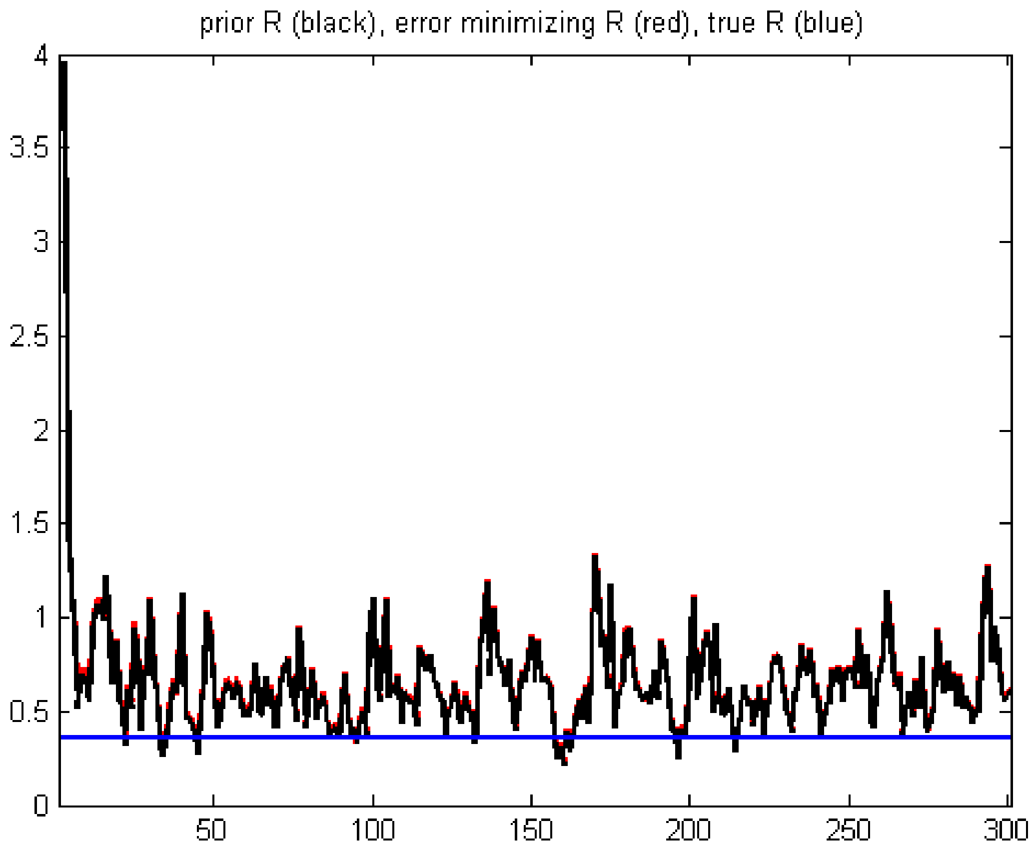


R estimate (black) quickly converges to a value close to **R_true** (blue) – but it is slightly higher. This gives smaller corrections at unassimilated observations than if using **R_true** – an error reducing strategy when forecast error correlations are imperfect.

Test of DC method for finding R
Q_DA not equal Q_True

- True model error variance changes in horizontal
  - **Q_true** *variance* ranges from 0.36 on left half of domain to 1.08 on right half.
  - **Q_true** *correlation matrix* equal to the climatological correlation matrix.
- DA has
  - **Q_DA** *variance* is uniformly equal to 0.72.
  - **Q_DA** *correlation matrix* is the identity matrix (uncorrelated model error)
- Actual observations have
  - **R_true**=0.36 **I** (diagonal with variance 0.36)
- DA scheme has
  - **R_DA**=3.6 **I** (10 times bigger than the true error variance)
- Perform DC **R** estimation for 300 DA cycles.
- Resulting $\mathbf{R}_{\mathbf{DC}}$**(1)** used in Todling equation to get $\mathbf{Q}_{\mathbf{Todling}}$**(1)**
- Repeat DC **R** estimation using $\mathbf{Q}_{\mathbf{Todling}}$**(1)** to obtain $\mathbf{R}_{\mathbf{DC}}$**(2)**
- Use $\mathbf{R}_{\mathbf{DC}}$**(2)** to obtain $\mathbf{Q}_{\mathbf{Todling}}$**(2)**
- **M**easure similarity of $\mathbf{Q}_{\mathbf{Todling}}$**(2)** to **Q_true**

Test of DC method for finding R Q_DA not equal Q_True

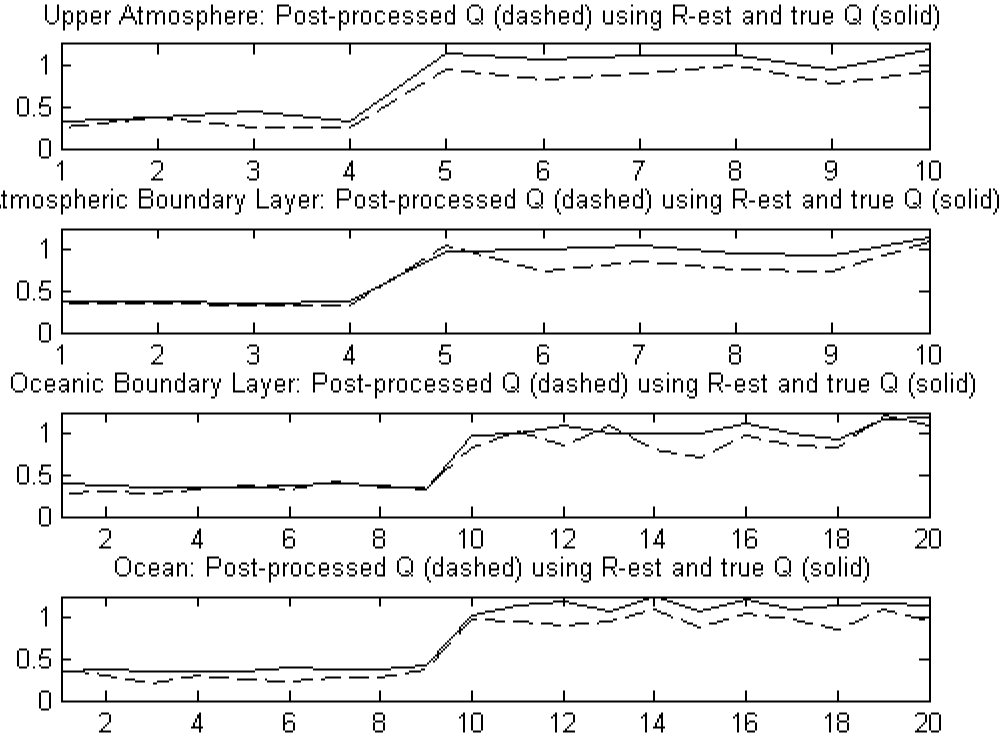


R estimate (black) higher than **R_true** (blue). This gives smaller corrections at unassimilated observations than if using **R_true**. This is an error reducing strategy when **P** is imperfect.

Test of DC method for finding R Q_DA not equal Q_True

- True model error variance changes in horizontal
  - **Q_true** *variance* ranges from 0.36 on left half of domain to 1.08 on right half.
  - **Q_true** *correlation matrix* equal to the climatological correlation matrix.
- DA has
  - **Q_DA** *variance* is uniformly equal to 0.72.
  - **Q_DA** *correlation matrix* is the identity matrix (uncorrelated model error)
- Actual observations have
  - **R_true**=0.36 **I** (diagonal with variance 0.36)
- DA scheme has
  - **R_DA**=3.6 **I** (10 times bigger than the true error variance)
- Perform DC **R** estimation for 300 DA cycles.
- Resulting $\mathbf{R_{DC}(1)}$ used in Todling equation to get $\mathbf{Q_{Todling}(1)}$
- Repeat DC **R** estimation using $\mathbf{Q_{Todling}(1)}$ to obtain $\mathbf{R_{DC}(2)}$
- Use $\mathbf{R_{DC}(2)}$ to obtain $\mathbf{Q_{Todling}(2)}$
- **M**easure similarity of $\mathbf{Q_{Todling}(2)}$ to **Q_true**

$Q_{Todling}$(1) variance
Q_DA not equal Q_True

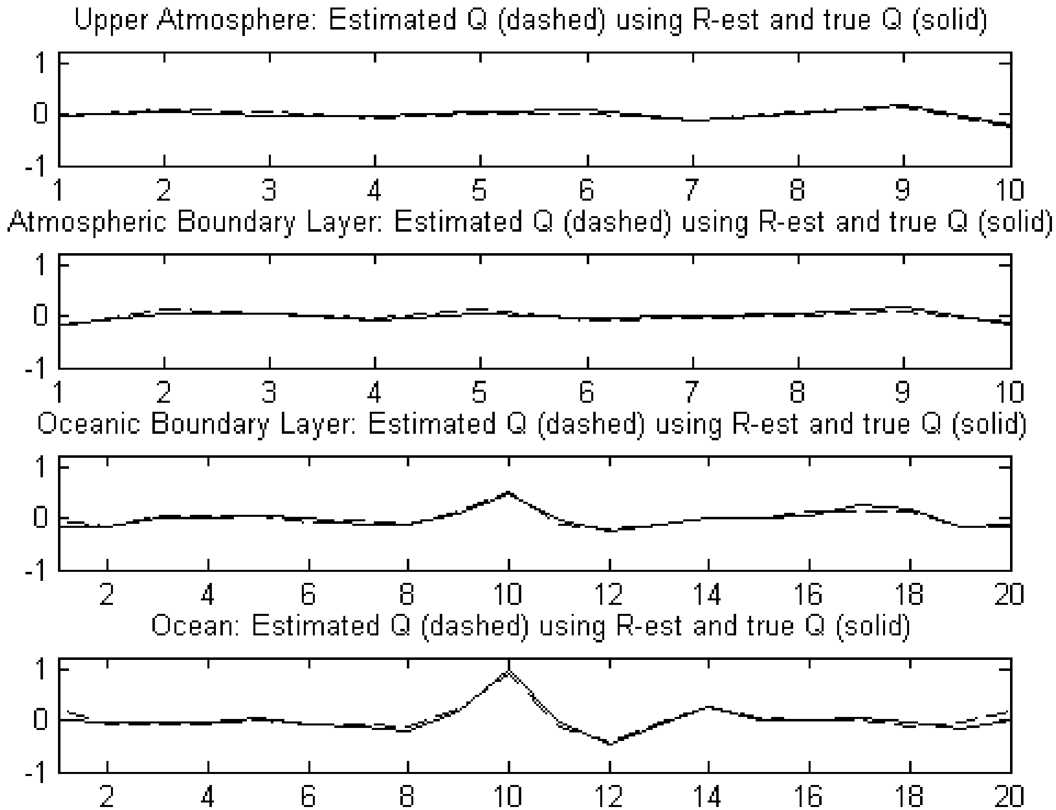


Post-processing involves:

(i) enforcing symmetry by making estimate equal to average of itself and its transpose, and

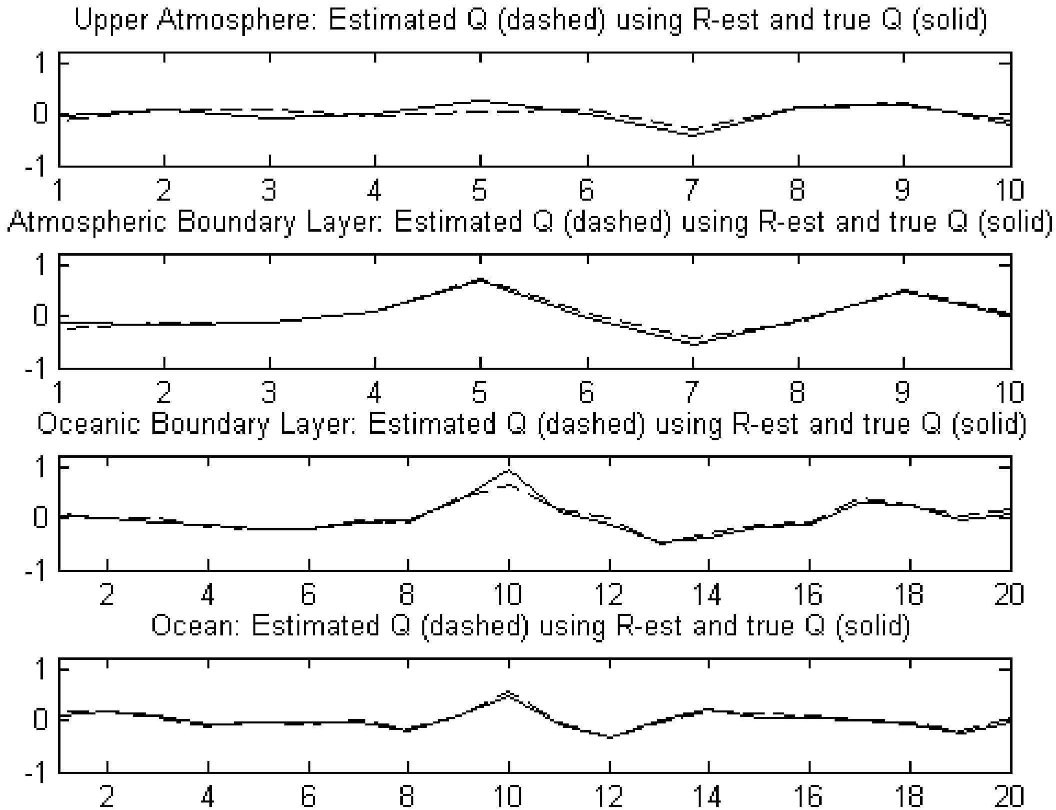
(ii) making negative eigenvalues very small positives

$Q_{Todling}$(1) covariance with deep ocean Q_DA not equal Q_True



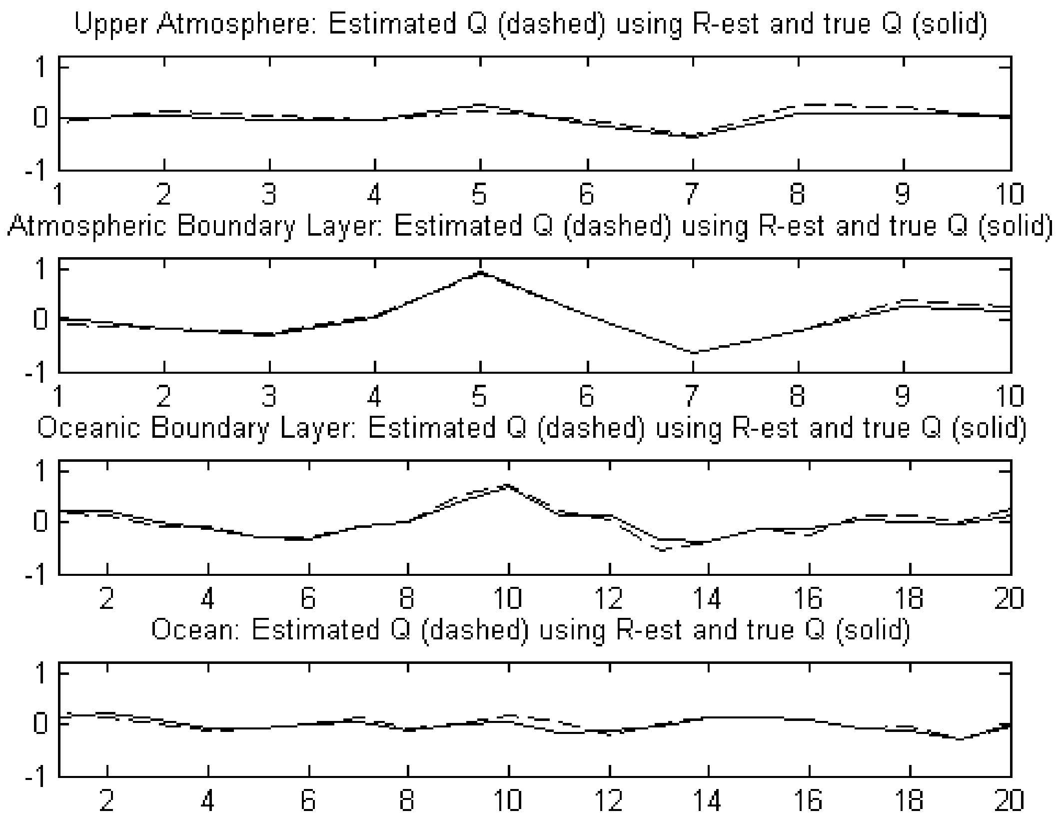
Covariance function retrieved despite diagonal **Q_DA**

$Q_{Todling}$(1) covariance with oceanic BL Q_DA not equal Q_True



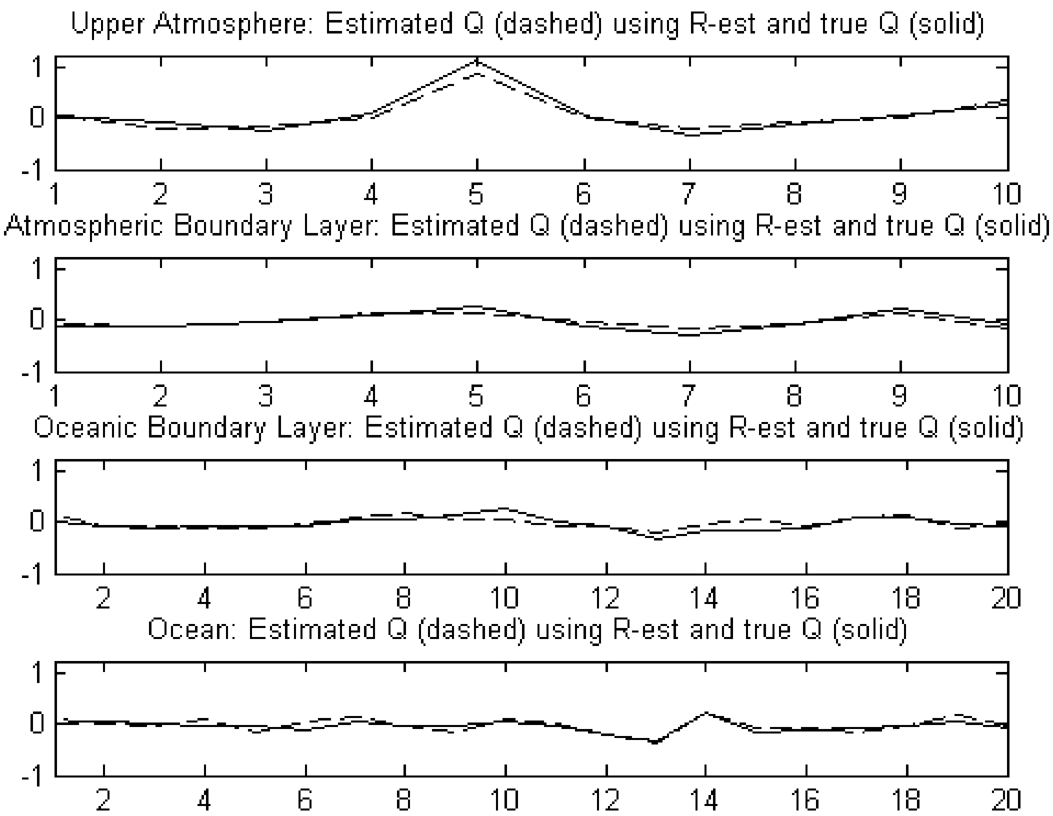
Covariance function retrieved despite diagonal **Q_DA**

$Q_{Todling}$(1) covariance with atmos BL Q_DA not equal Q_True



Covariance function retrieved despite diagonal **Q_DA**

$Q_{Todling}$(1) covariance with upper atmos Q_DA not equal Q_True

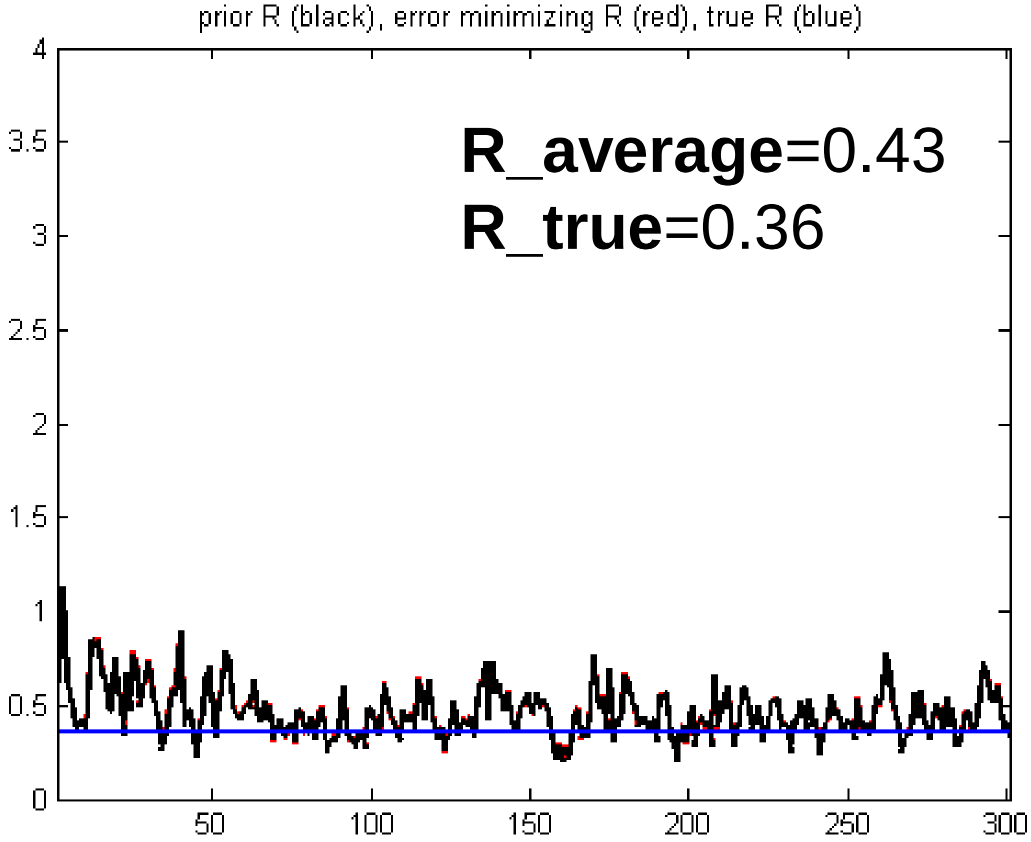


Covariance function retrieved despite diagonal **Q_DA**

Test of DC method for finding R Q_DA not equal Q_True

- True model error variance changes in horizontal
  - **Q_true** *variance* ranges from 0.36 on left half of domain to 1.08 on right half.
  - **Q_true** *correlation matrix* equal to the climatological correlation matrix.
- DA has
  - **Q_DA** *variance* is uniformly equal to 0.72.
  - **Q_DA** *correlation matrix* is the identity matrix (uncorrelated model error)
- Actual observations have
  - **R_true**=0.36 **I** (diagonal with variance 0.36)
- DA scheme has
  - **R_DA**=3.6 **I** (10 times bigger than the true error variance)
- Perform DC **R** estimation for 300 DA cycles.
- Resulting $\mathbf{R_{DC}(1)}$ used in Todling equation to get $\mathbf{Q_{Todling}(1)}$
- Repeat DC **R** estimation using $\mathbf{Q_{Todling}(1)}$ to obtain $\mathbf{R_{DC}(2)}$
- Use $\mathbf{R_{DC}(2)}$ to obtain $\mathbf{Q_{Todling}(2)}$
- **M**easure similarity of $\mathbf{Q_{Todling}(2)}$ to **Q_true**

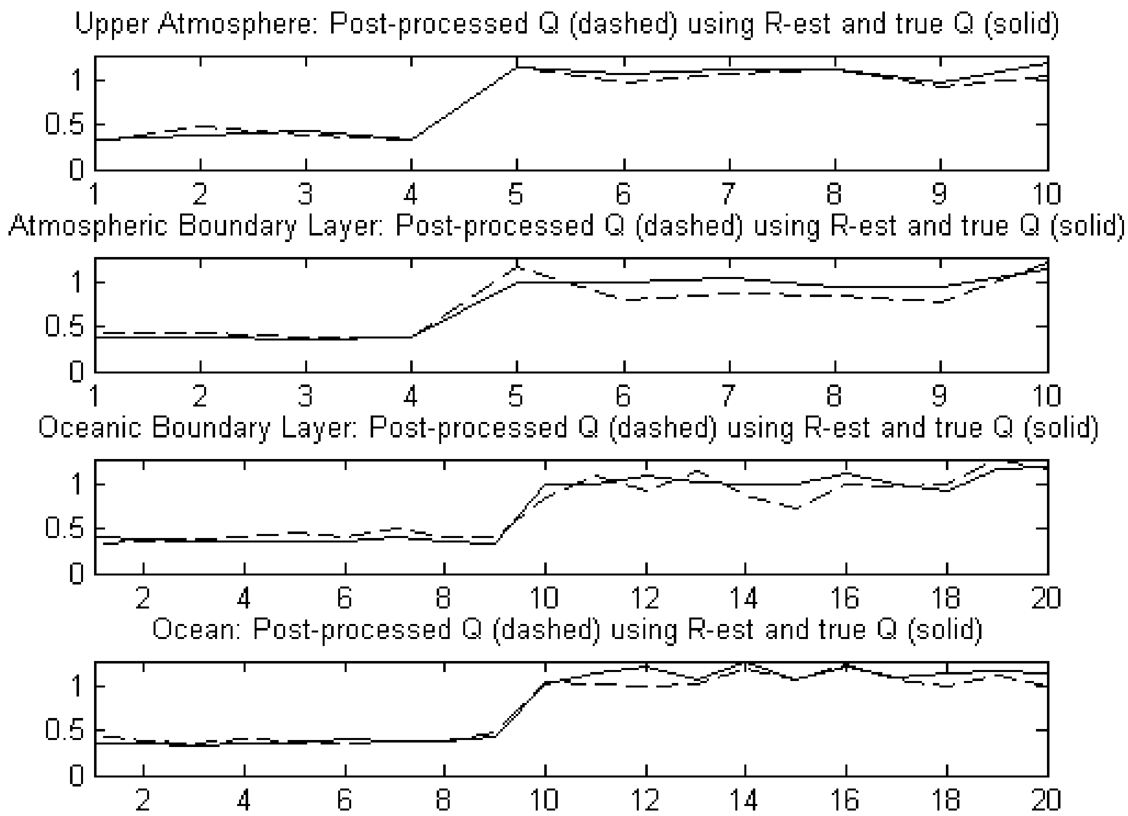
Evolution of R estimate over 300 DA cycles using Q from previous run



R_average is mean of last 240 **R** estimates

R estimate (black) is now much closer to the truth (blue) than in the previous run where **R_average**=0.62.

Estimate of Q variance with improved Q_DA in DA after post-processing



In addition, the recovered model error correlations are almost indistinguishable from the true correlations.

Concluding remarks

i. Broken trajectory weak constraint 4D-VAR simultaneously provides the filter and smoother analyses required by Todling's (2015) method for estimating **Q**

ii. Divide and Calibrate method for estimating **R** requires the gradient of the data assimilation scheme and, possibly, the model.
   i. It provided usefully accurate estimates in the weakly coupled system considered here.
   ii. The more numerous and co-located independent observations, the more accurate this technique is likely to be.

iii. WC-4DVAR, $\mathbf{R}_{\mathbf{DC}}$ + Todling equation recovered true **Q** even with very poor initial guesses of **R, P** and **Q**

iv. Future challenge: Develop methods to inform flow dependent model error representations.

Global measure of accuracy of Q estimates

$$rel\_mse\_diag\left(\mathbf{Q}_{Est}\right)=\frac{\overline{diag\left[\left(\mathbf{Q}_{Est}-\mathbf{Q}_{true}\right)\right].\wedge 2}}{\overline{diag\left(\mathbf{Q}_{true}\right).\wedge 2}}$$

$$rel\_mse\_diag\left[\mathbf{Q}_{Todling}\left(1\right)\right]=0.0286$$

$$rel\_mse\_diag\left[\mathbf{Q}_{Todling}\left(2\right)\right]=0.0132$$