Fast Implementation of Localization for Ensemble DA

B. Wang[1,2] and J.-J. Liu[1]

[1]LASG/IAP, Chinese Academy of Sciences

[2]CESS, Tsinghua University

DAOS-2012，Sep 20, 2012

EnKF

The formula of EnKF without localization

$$\mathbf{x}' = \mathbf{b}\,\mathbf{r}^T(\mathbf{O}+\mathbf{r}\mathbf{r}^T)^{-1}\mathbf{y}'_{obs}$$

Localization:

$$\mathbf{x}' = \boldsymbol{\rho}_{\mathbf{xy}} \circ (\mathbf{b}\,\mathbf{r}^T)(\mathbf{O}+\boldsymbol{\rho}_{\mathbf{yy}} \circ (\mathbf{r}\mathbf{r}^T))^{-1}\mathbf{y}'_{obs}$$

where

$\boldsymbol{\rho}_{\mathbf{xy}}$, $\boldsymbol{\rho}_{\mathbf{yy}}$: Filtering matrices;

$\circ$: Schür product operator

$$\begin{cases} \mathbf{b} = \dfrac{1}{\sqrt{n-1}}(\mathbf{x}_1-\bar{\mathbf{x}}, \mathbf{x}_2-\bar{\mathbf{x}}, \cdots, \mathbf{x}_n-\bar{\mathbf{x}}), \\ \mathbf{r} = \dfrac{1}{\sqrt{n-1}}(\mathbf{y}_1-\bar{\mathbf{y}}, \mathbf{y}_2-\bar{\mathbf{y}}, \cdots, \mathbf{y}_n-\bar{\mathbf{y}}), \end{cases}$$

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i$$

$$\bar{\mathbf{y}} = \frac{1}{n}\sum_{\mathrm{i}=1}^{\mathrm{n}}\mathbf{y}_i$$

The elements of filter matrix ρ_{xy} (similar for ρ_{yy}):

$$\rho_{i,j} = C_0(d^h_{i,j} / d^h_0) \cdot C_0(d^v_{i,j} / d^v_0)$$
$$(i = 1, 2, \cdots, m_y;\ j = 1, 2, \cdots, m_x) \quad ,$$

where the filtering function C_0 is defined as (Gaspari and Cohn, 1999)

$$C_0(r) = \begin{cases} -\frac{1}{4}r^5 + \frac{1}{2}r^4 + \frac{5}{8}r^3 - \frac{5}{3}r^2 + 1, & 0 \le r \le 1 \\ \frac{1}{12}r^5 - \frac{1}{2}r^4 + \frac{5}{8}r^3 + \frac{5}{3}r^2 - 5r + 4 - \frac{2}{3}r^{-1}, & 1 < r \le 2, \\ 0, & 2 < r \end{cases}$$

$d^h_{i,j}$ **and** $d^v_{i,j}$ **respectively represent the horizontal and vertical distances between the *i-th* and *j-th* row vectors;** d^h_0 **and** d^v_0 **are the horizontal and vertical Schür radius.**

Issues caused by Localization

Before localization, it is easy to get the solution:

$$\mathbf{x}' = \mathbf{b}\,\mathbf{r}^T(\mathbf{O}+\mathbf{R})^{-1}\mathbf{y}'_{obs} = \mathbf{b}\,\mathbf{r}^T(\mathbf{O}+\mathbf{r}\,\mathbf{r}^T)^{-1}\mathbf{y}'_{obs}$$

After localization, huge computing time is required for the solution:

$$\mathbf{x}' = \boldsymbol{\rho}_{\mathbf{xy}} \circ (\mathbf{b}\,\mathbf{r}^T)(\mathbf{O}+\boldsymbol{\rho}_{\mathbf{yy}} \circ (\mathbf{r}\,\mathbf{r}^T))^{-1}\mathbf{y}'_{obs}$$

$m_x \sim 10^7$

$m_x \times n \;\; (n \sim 10^2)$

$m_y \sim 10^5$

$m_x \times m_y$

$m_y \times n$

diagonal

$m_y \times m_y$

The extra multiplications due to localization

$$m_x \times m_y \times n + m_y \times m_y \times n \times l \sim 10^{14}$$

Available solution to the issue

1) To split the observation vector into a combination of low-dimension sub-vectors
2) To sequentially assimilate the low-dimension observation sub-vectors

Our new method is to split filter matrix in localization

m_x-dimension filter eigenvector

Filter matrices

$$\boldsymbol{\rho}_{\mathbf{xy}} = \sum_{k=1}^{N} \boldsymbol{\rho}_{\mathbf{x}}^{(k)} \left(\boldsymbol{\rho}_{\mathbf{y}}^{(k)}\right)^T$$

m_y-dimension filter eigenvector

$$\boldsymbol{\rho}_{\mathbf{yy}} = \sum_{k=1}^{N} \boldsymbol{\rho}_{\mathbf{y}}^{(k)} \left(\boldsymbol{\rho}_{\mathbf{y}}^{(k)}\right)^T$$

N is the number of the major modes of the filter matrix

Expansion of filter function in localization

Filter function:

Base function

$$C(x_1, x_2) = \sum_{n=1}^{\infty} a_n e_n(x_1) e_n(x_2)$$

$$\int_0^L w(x)\, e_n(x)\, e_m(x)\, dx = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

Orthogonal

Weighting function

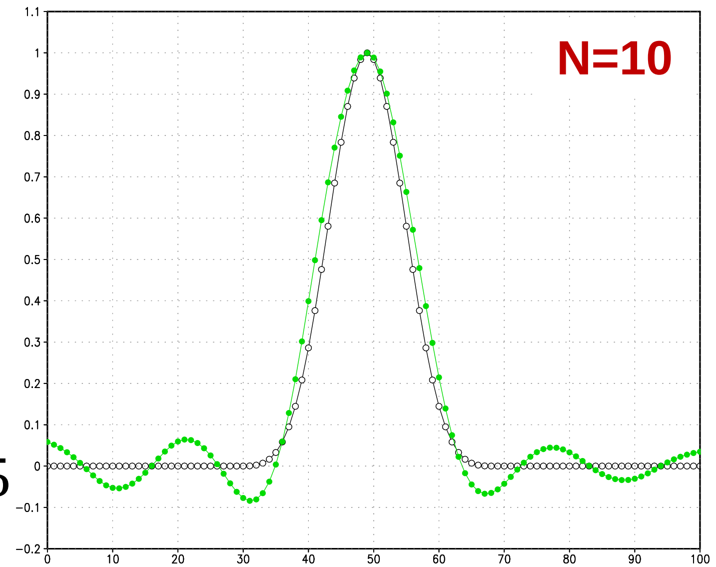
$$a_n = \int_0^L \int_0^L w(x_1) w(x_2) C(x_1, x_2)\, e_n(x_1)\, e_n(x_2)\, dx_1\, dx_2$$

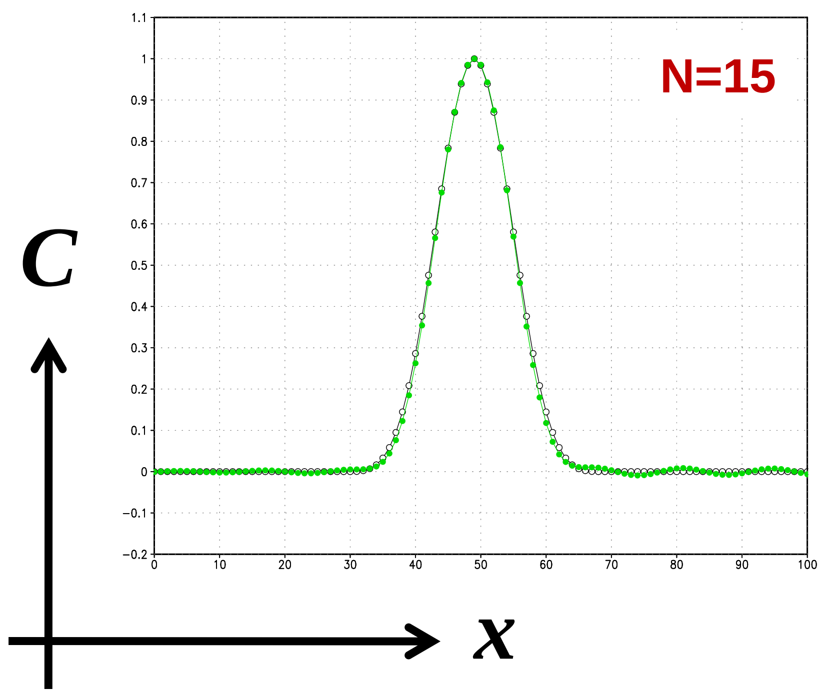
One-Dimension Filter

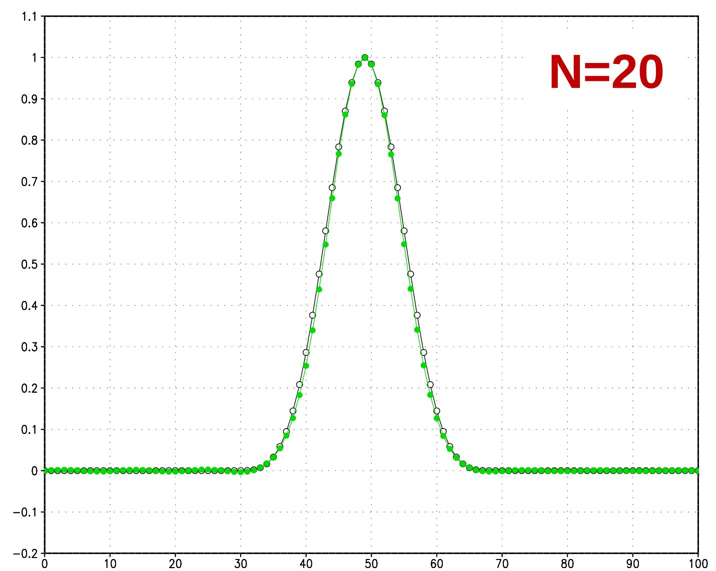
Green: $C_N(x, x_0) = \sum_{n=1}^{N} a_n e_n(x)\, e_n(x_0)$

Black: $C_0(r)$

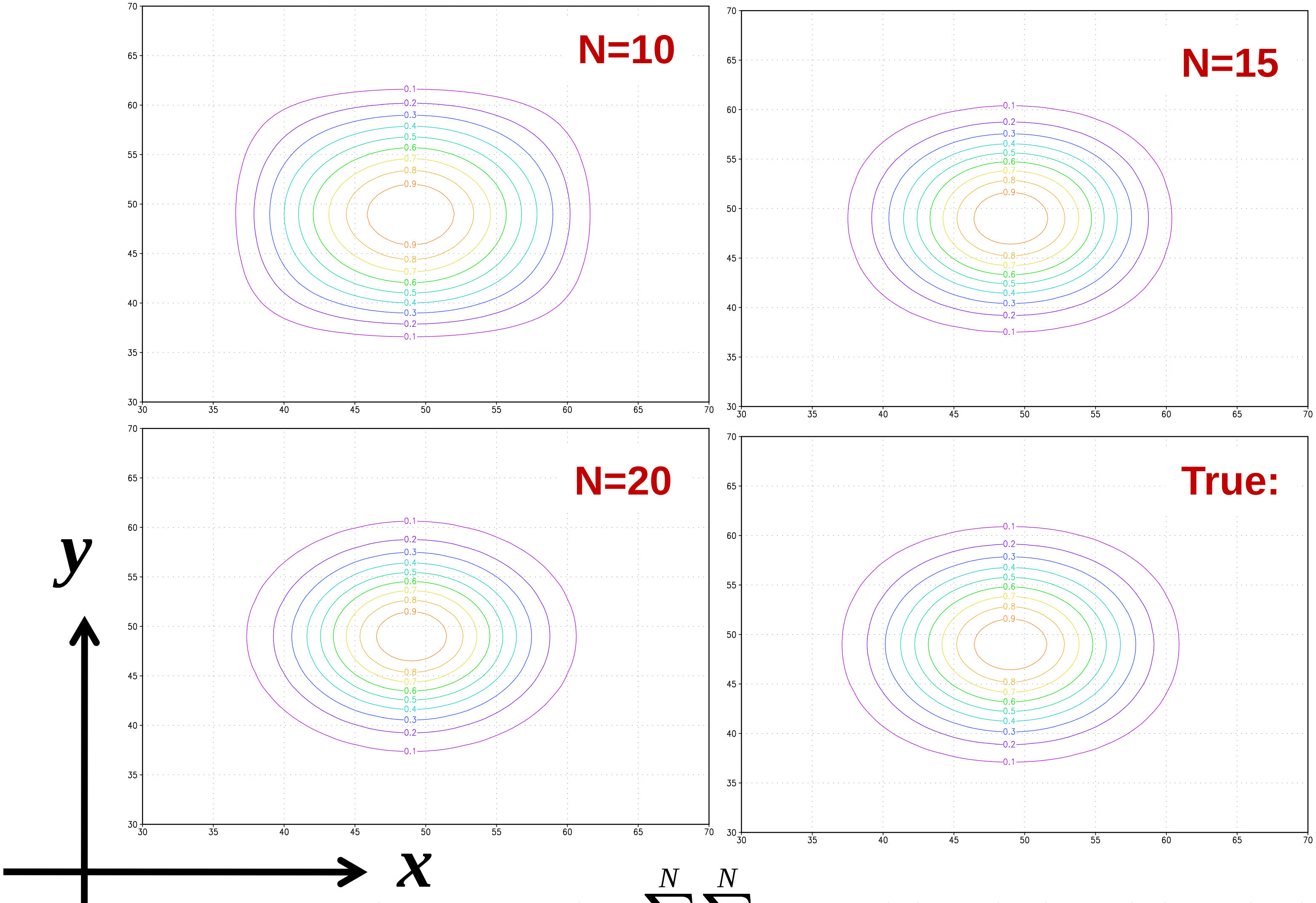
where $r = |x - x_0| / d_0,\ x_0 = 49,\ d_0 = 5$







Two-Dimension Filter



$C_0(r)$

$$C_N(x, x_0; y, y_0) = \sum_{n=1}^{N} \sum_{m=1}^{N} a_{n,m} e_n(x)\, e_n(x_0)\, e_m(y)\, e_m(y_0)$$

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2} / d_0, \;\; x_0 = 49, \;\; y_0 = 49, \;\; d_0 = 5$$

EnKF with extended samples

$$\mathbf{x}' = \tilde{\mathbf{b}}\,(\mathbf{I}_{\tilde{n}\times\tilde{n}} + \tilde{\mathbf{r}}^T \mathbf{O}^{-1}\tilde{\mathbf{r}})^{-1}\tilde{\mathbf{r}}^T \mathbf{O}^{-1}\mathbf{y}'_{obs}$$

where

$$\mathbf{b} = \frac{1}{\sqrt{\tilde{n}-1}}(\boldsymbol{\rho}_{\mathbf{x}}^{(1)} \circ \mathbf{x}_1 - \bar{\mathbf{x}}, \boldsymbol{\rho}_{\mathbf{x}}^{(1)} \circ \mathbf{x}_2 - \bar{\mathbf{x}}, \cdots, \boldsymbol{\rho}_{\mathbf{x}}^{(1)} \circ \mathbf{x}_n - \bar{\mathbf{x}},$$

$$\boldsymbol{\rho}_{\mathbf{x}}^{(2)} \circ \mathbf{x}_1 - \bar{\mathbf{x}}, \boldsymbol{\rho}_{\mathbf{x}}^{(2)} \circ \mathbf{x}_2 - \bar{\mathbf{x}}, \cdots, \boldsymbol{\rho}_{\mathbf{x}}^{(2)} \circ \mathbf{x}_n - \bar{\mathbf{x}},$$

$$\vdots$$

$$\boldsymbol{\rho}_{\mathbf{x}}^{(N)} \circ \mathbf{x}_1 - \bar{\mathbf{x}}, \boldsymbol{\rho}_{\mathbf{x}}^{(N)} \circ \mathbf{x}_2 - \bar{\mathbf{x}}, \cdots, \boldsymbol{\rho}_{\mathbf{x}}^{(N)} \circ \mathbf{x}_n - \bar{\mathbf{x}})$$

$$\bar{\mathbf{x}} = \frac{1}{\tilde{n}}\sum_{k=1}^{N}\sum_{i=1}^{n}\boldsymbol{\rho}_{\mathbf{x}}^{(k)} \circ \mathbf{x}_i \qquad (\tilde{n} = n \times N)$$

$$\mathbf{r} = \frac{1}{\sqrt{\tilde{n}-1}} \Big(\boldsymbol{\rho}_{\mathbf{y}}^{(1)} \circ \mathbf{y}_1 - \overline{\mathbf{y}}, \boldsymbol{\rho}_{\mathbf{y}}^{(1)} \circ \mathbf{y}_2 - \overline{\mathbf{y}}, \cdots, \boldsymbol{\rho}_{\mathbf{y}}^{(1)} \circ \mathbf{y}_n - \overline{\mathbf{y}},$$

$$\boldsymbol{\rho}_{\mathbf{y}}^{(2)} \circ \mathbf{y}_1 - \overline{\mathbf{y}}, \boldsymbol{\rho}_{\mathbf{y}}^{(2)} \circ \mathbf{y}_2 - \overline{\mathbf{y}}, \cdots, \boldsymbol{\rho}_{\mathbf{y}}^{(2)} \circ \mathbf{y}_n - \overline{\mathbf{y}},$$

$$\vdots$$

$$\boldsymbol{\rho}_{\mathbf{y}}^{(N)} \circ \mathbf{y}_1 - \overline{\mathbf{y}}, \boldsymbol{\rho}_{\mathbf{y}}^{(N)} \circ \mathbf{y}_2 - \overline{\mathbf{y}}, \cdots, \boldsymbol{\rho}_{\mathbf{y}}^{(N)} \circ \mathbf{y}_n - \overline{\mathbf{y}} \Big)$$

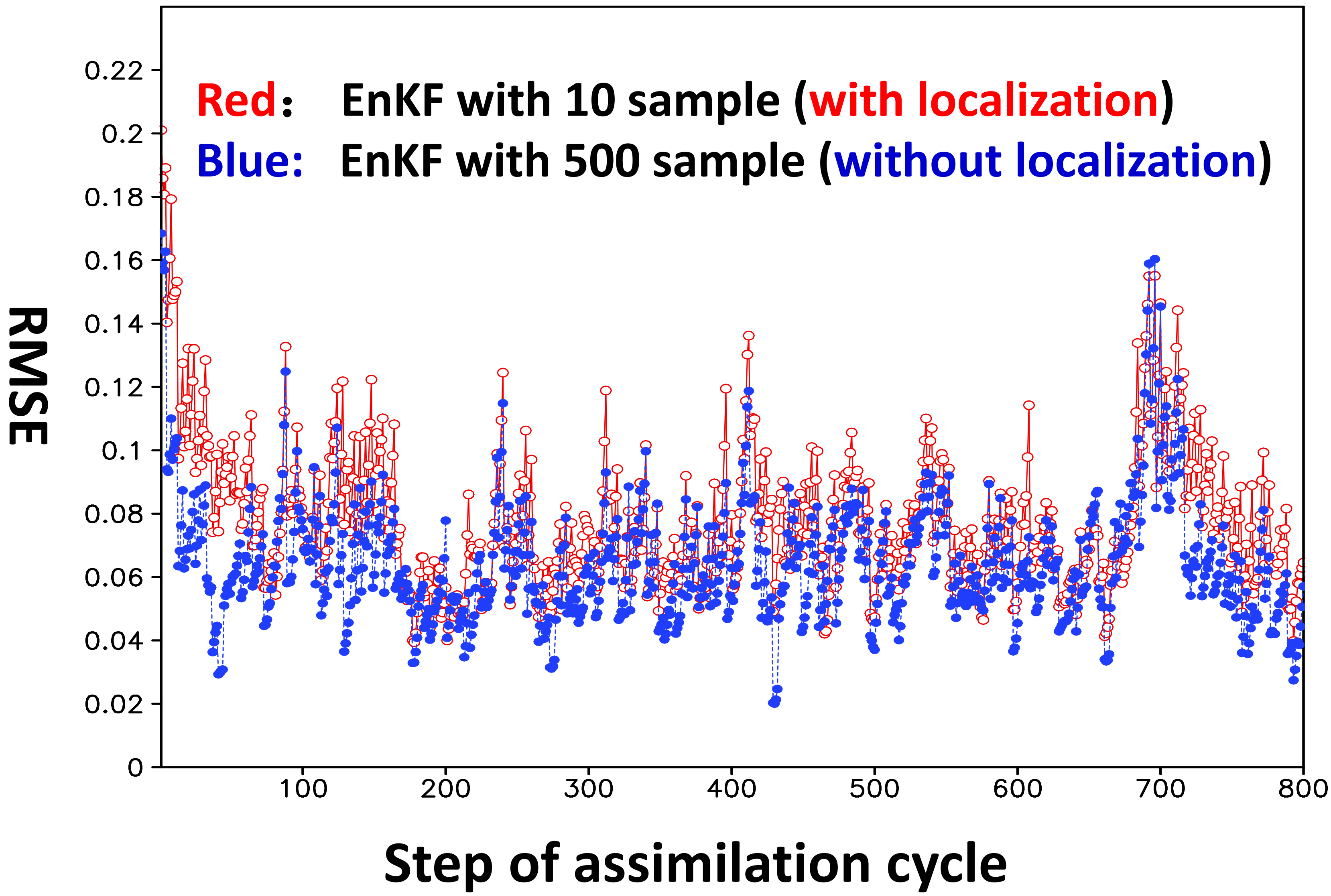
$$\overline{\mathbf{y}} = \frac{1}{\tilde{n}} \sum_{k=1}^{N} \sum_{i=1}^{n} \boldsymbol{\rho}_{\mathbf{y}}^{(k)} \circ \mathbf{y}_i \qquad (\tilde{n} = n \times N)$$

Experiment with simple model

Lorenz-96 model:

$$\frac{dX_j}{dt} = (X_{j+1} - X_{j-2})X_{j-1} - X_j + F$$

- **Spatial coordinate:** $j = 1, \ldots, M \quad (M = 40)$
- **Forcing parameter:** $F = 8$
- **Solution scheme:** 4^{th} -order Runge-Kutta scheme
- **Time stepsize:** 0.05 time unit
- **Periodic boundary conditions:** $X_{j+40} = X_j$
- **Note:** 0.05 time unit ~6h (**Lorenz and Emanuel, 1998**)



Red: EnKF with 10 sample (with localization)
Blue: EnKF with 500 sample (without localization)
RMSE
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
100
200
300
400
500
600
700
800
Step of assimilation cycle

Experiment based on an operational prediction model: AREM

CTRL: No assimilation with prediction as IC

ERA： Using ERA40 as IC, no assimilation

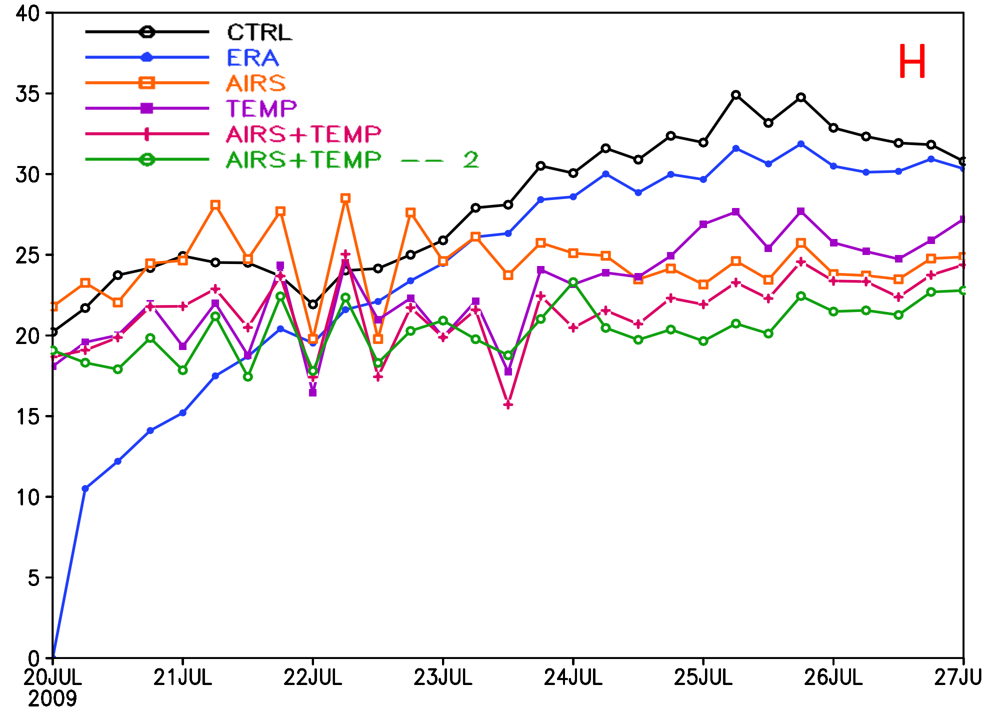
AIRS： Assimilating T & Q of AIRS only

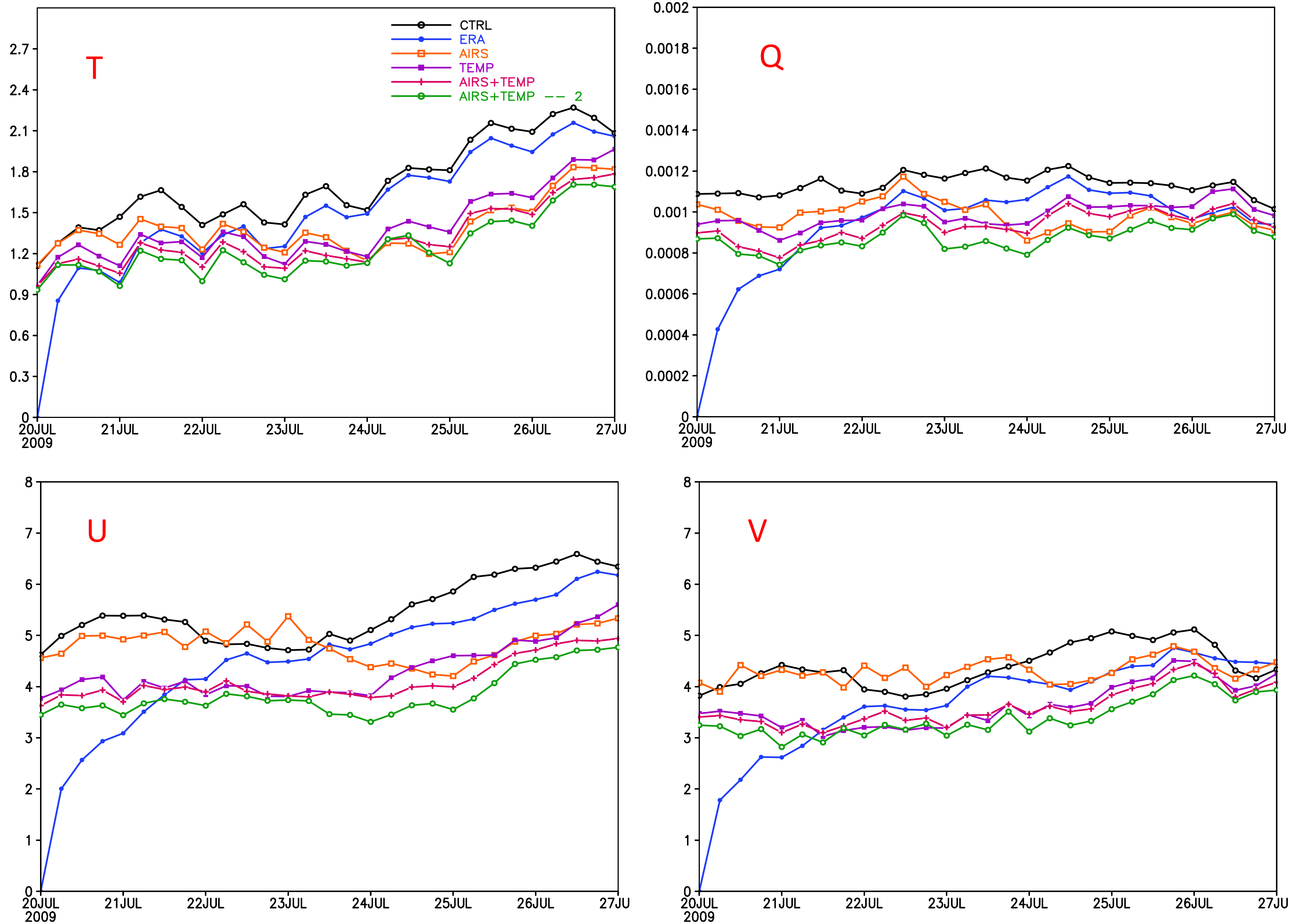
TEMP: Assimilating Radiosondes only

AIRS+TEMP: With random perturbation

AIRS+TEMP—2: No random perturbation

RMSEs averaged from 850hPa to 100hPa




T
Q
U
V
CTRL
ERA
AIRS
TEMP
AIRS+TEMP
AIRS+TEMP -- 2
20JUL
2009
21JUL
22JUL
23JUL
24JUL
25JUL
26JUL
27JU

Summary

- Localization for EnKF can be economically implemented using an orthogonal expansion of filter function;
- The new implementation can be applied to other ensemble-based DA, e.g., En-3DVar and En-4DVar.

Comments, please.

Thank you!

- **“True” states:** simulations during a period of time unit after a long-term integration (e.g., 10^5 model time steps) of the model from an arbitrary IC .
- **Observations** of all model variables: “true” states plus uncorrelated random noise with standard Gaussian distribution (with zero mean and variance of 0.16)
- **Assimilation experiments:** assimilation cycles over a period of 200 days using the **EnKF** with 18-hr assimilation window and, respectively.

500 samples are used for the experiments so that they have good representativeness.