Data Assimilation
- A07. Serial EnSRF -

Shunji Kotsuki

Center for Environmental Remote Sensing / Institute of Advanced Academic Research

(shunji.kotsuki@chiba-u.jp)

CHIBA UNIVERSITY

Environmental Prediction Science Laboratory

DA Lectures A (Basic Course)

- (1) Introduction and NWP
- (2) Deterministic Chaos and Lorenz-96 model
- (3) A toy model and Bayesian estimation
- (4) Kalman Filter (KF)
- (5) 3D Variational Method (3DVAR)
- (6) Ensemble Kalman Filter (PO method)
- (7) Serial Ens. Square Root Filter (Serial EnSRF)
- (8) Local Ens. Transform Kalman Filter (LETKF)
- (9) Innovation Statistics
- (10) Adaptive Inflations
- (11) 4D Variational Method (4DVAR)

Today's Goal

- **Lecture: serial EnSRF**
  - to introduce serial EnSRF

- **Training Course: Lorenz 96**
  - to implement serial EnSRF into L96

Ensemble Kalman Filter (EnKF)

KF

Prediction (state)

$$\mathbf{x}_t^b = M(\mathbf{x}_{t-1}^a)$$

Prediction of Error Cov. (explicitly)

$$\mathbf{P}_t^b = \mathbf{M}\mathbf{P}_{t-1}^a\mathbf{M}^T(+\mathbf{Q})$$

Kalman Gain

$$\mathbf{K}_t = \mathbf{P}_t^b\mathbf{H}^T\left[\mathbf{H}\mathbf{P}_t^b\mathbf{H}^T + \mathbf{R}\right]^{-1}$$

Analysis (state)

$$\mathbf{x}_t^a = \mathbf{x}_t^b + \mathbf{K}_t(\mathbf{y}_t^o - H(\mathbf{x}_t^b))$$

Analysis Error Covariance

$$\mathbf{P}_t^a = [\mathbf{I} - \mathbf{K}_t\mathbf{H}]\mathbf{P}_t^b$$

EnKF

Ensemble Prediction (state)

$$\mathbf{x}_t^{b(i)} = M\left(\mathbf{x}_{t-1}^{a(i)}\right) \quad \text{for } i = 1,\ldots,m$$

Prediction of Error Covariance (implicitly)

$$\mathbf{P}_t^b \approx \mathbf{Z}_t^b\left(\mathbf{Z}_t^b\right)^T$$

Kalman Gain

$$\mathbf{K}_t = \mathbf{Z}_t^b(\mathbf{Y}_t^b)^T\left[\mathbf{Y}_t^b(\mathbf{Y}_t^b)^T + \mathbf{R}\right]^{-1}$$

$$= \mathbf{Z}_t^b\left[\mathbf{I} + (\mathbf{Y}_t^b)^T\mathbf{R}^{-1}\mathbf{Y}_t^b\right]^{-1}(\mathbf{Y}_t^b)^T\mathbf{R}^{-1}$$

Analysis (state)

$$\mathbf{x}_t^a = \mathbf{x}_t^b + \mathbf{K}_t(\mathbf{y}_t^o - H(\mathbf{x}_t^b))$$

Analysis Error Covariance

(1) Stochastic: PO method

(2) Deterministic: Square Root Filter (SRF) (e.g., serial EnSRF, EAKF, LETKF)

Serial EnSRF

Square Root Filter (SRF)

SRF assumes the following update w/o adding perturbation in obs.

$$\mathbf{Z}_t^a = \mathbf{Z}_t^b \mathbf{W}$$

$\mathbf{W}$ ($\in \mathbb{R}^{m \times m}$): Ensemble Ptb. Transform Matrix

and compute $\mathbf{W}$ that satisfies

$$\mathbf{P}_t^a = \mathbf{Z}_t^b \mathbf{W} (\mathbf{Z}_t^b \mathbf{W})^T$$

$$= [\mathbf{I} - \mathbf{K}_t \mathbf{H}] \mathbf{Z}_t^b (\mathbf{Z}_t^b)^T$$

However, SRF cannot determine $\mathbf{W}$ deterministically.
For example, for $\mathbf{U}$ that satisfies $\mathbf{U}\mathbf{U}^T = \mathbf{I}$,
a new matrix $\mathbf{S} = \mathbf{W}\mathbf{U}$ can be also a ptb. transform matrix since

$$\mathbf{P}_t^a = \mathbf{Z}_t^b \mathbf{W} (\mathbf{Z}_t^b \mathbf{W})^T = \mathbf{Z}_t^b \mathbf{W}\mathbf{U} (\mathbf{Z}_t^b \mathbf{W}\mathbf{U})^T = \mathbf{Z}_t^b \mathbf{S} (\mathbf{Z}_t^b \mathbf{S})^T$$

Question: how can we determine $\mathbf{W}$?

Ensemble SRF (EnSRF)

Environmental Prediction Science Laboratory

EnSRF assumes the following ensemble update

$$\mathbf{Z}_t^a = [\mathbf{I} - \widetilde{\mathbf{K}}\mathbf{H}]\mathbf{Z}_t^b$$

$$\Leftrightarrow [\mathbf{I} - \widetilde{\mathbf{K}}\mathbf{H}]\mathbf{P}_t^b[\mathbf{I} - \widetilde{\mathbf{K}}\mathbf{H}]^T = [\mathbf{I} - \mathbf{K}_t\mathbf{H}]\mathbf{P}_t^b$$

This equation was solved by Andrew (1968) such that

$$\widetilde{\mathbf{K}} = \mathbf{P}_t^b\mathbf{H}^T\left[\underbrace{\left(\mathbf{H}\mathbf{P}_t^b\mathbf{H}^T + \mathbf{R}\right)}_{p\times p \text{ matrix}}{}^{-1/2}\right]^T\left[\underbrace{\left(\mathbf{H}\mathbf{P}_t^b\mathbf{H}^T + \mathbf{R}\right)}_{p\times p \text{ matrix}}{}^{1/2} + \mathbf{R}^{1/2}\right]^{-1}$$

However, solving a *pxp* matrix is computationally expensive...

Serial EnSRF assimilates observation serially (i.e., p=1)

Whitaker & Hamill (2002)

$$\widetilde{\mathbf{K}} = \mathbf{P}_t^b\mathbf{H}^T\left(HP_t^bH^T + R\right)^{-1/2}\left[\left(HP_t^bH^T + R\right)^{1/2} + R^{1/2}\right]^{-1}$$

$$= \mathbf{P}_t^b\mathbf{H}^T\left(HP_t^bH^T + R\right)^{-1}\left(HP_t^bH^T + R\right)^{1/2}\left[\left(HP_t^bH^T + R\right)^{1/2} + R^{1/2}\right]^{-1}$$

$$= \mathbf{K}\left[1 + \underbrace{\sqrt{R/\left(HP_t^bH^T + R\right)}}_{\text{scalar since now assimilating one observation}}\right]^{-1}$$

Serial EnSRF Algorithm

$\mathbf{X}_{t-1}^{a}$

ensemble

Prediction (t-1 → t)

$$\mathbf{x}_t^{b(i)} = M(\mathbf{x}_{t-1}^{a(i)})$$

for i=1,…,m

Background

$\mathbf{X}_t^{b}$

Observations

$\mathbf{y}_t^{o}, \mathbf{R}, H$

next time step

$\mathbf{X}_t^{a}$

Analysis

Update Analysis Ensemble

obs loop for i=1,…,p (for assimilating ith observation)

(1) Compute Kalman Gain by

$$\mathbf{K}_t = \mathbf{Z}_t^b (\mathbf{Y}_t^b)^T \left[\mathbf{Y}_t^b (\mathbf{Y}_t^b)^T + \mathbf{R}\right]^{-1}$$

$$= \mathbf{Z}_t^b \left[\mathbf{I} + (\mathbf{Y}_t^b)^T \mathbf{R}^{-1} \mathbf{Y}_t^b\right]^{-1} (\mathbf{Y}_t^b)^T \mathbf{R}^{-1}$$

(2) Mean update

$$\overline{\mathbf{x}}_t^a = \overline{\mathbf{x}}_t^b + \mathbf{K}_t \left(\mathbf{y}_t^o - H(\overline{\mathbf{x}}_t^b)\right)$$

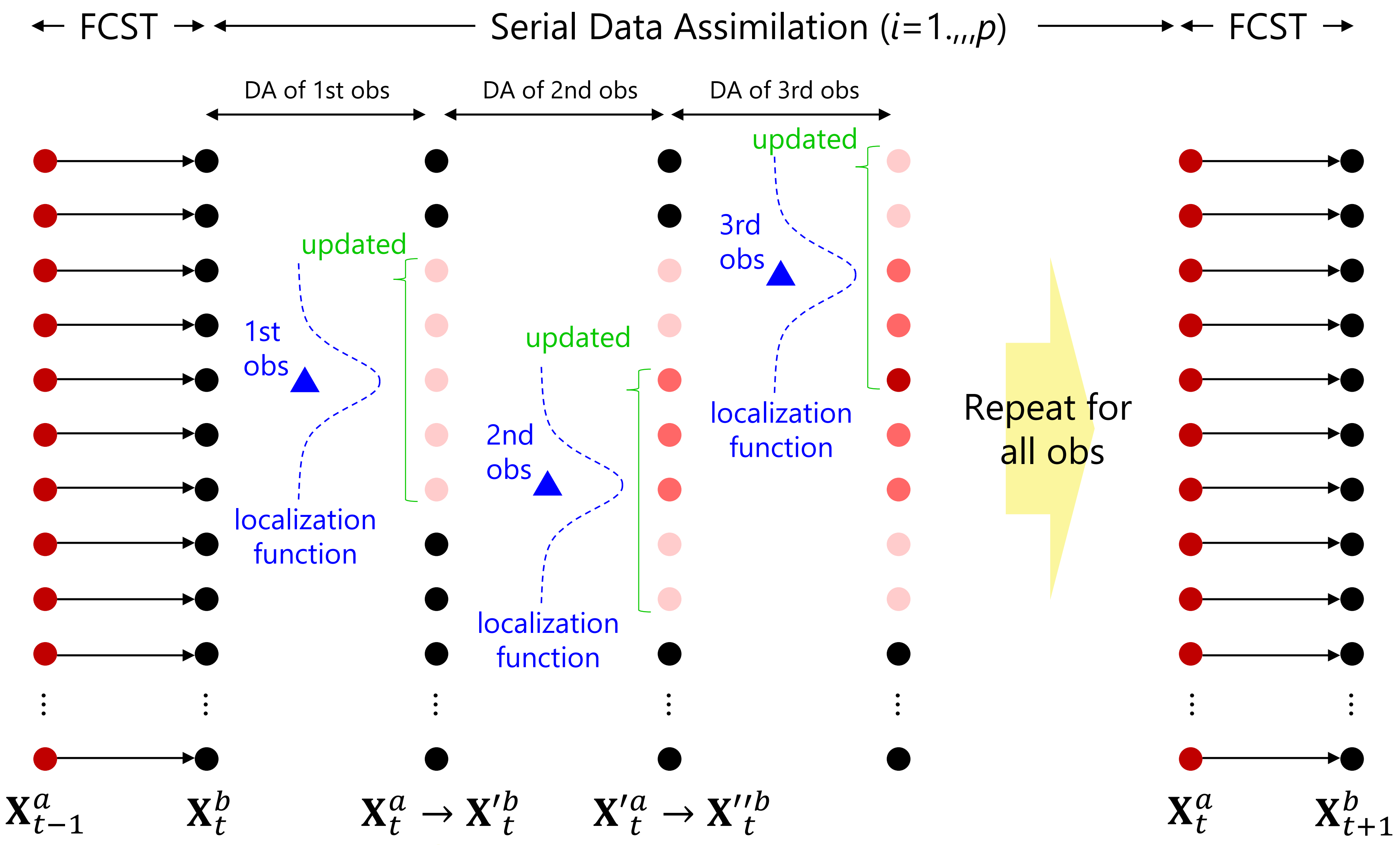
(3) Perturbation update

$$\mathbf{Z}_t^a = [\mathbf{I} - \widetilde{\mathbf{K}}\mathbf{H}]\mathbf{Z}_t^b \quad \textit{where}$$

$$\widetilde{\mathbf{K}} = \mathbf{K}\left[1 + \sqrt{R/(HP_t^b H^T + R)}\right]^{-1}$$

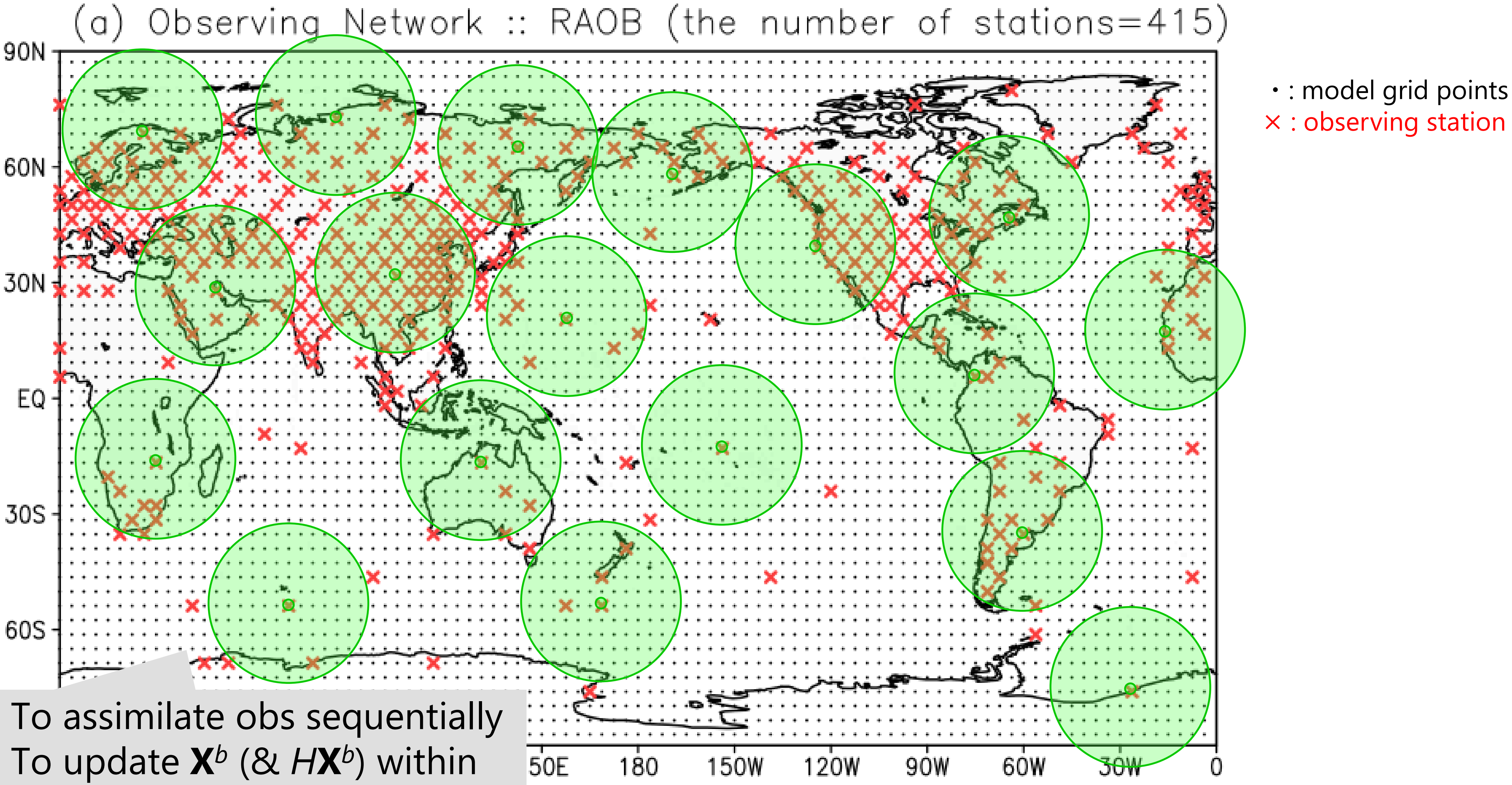
Regarding AN as FG for 2nd-pth obs ($\mathbf{X}_t^{a} \rightarrow \mathbf{X}_t'^{b}$)

Serial EnSRF Algorithm



Point!: Regarding the updated state as the background for assimilating 2nd obs.

Serial Assimilation



- To assimilate obs sequentially
- To update $\mathbf{X}^b$ (& $H\mathbf{X}^b$) within localization radius.

Limitations for parallelization

Basic Task 5

Basic Task 5

6．EnKF を実装し、KF と比較する。Whitaker and Hamill (2002)による Serial EnSRF, Bishop et al. (2001)による ETKF、Hunt et al. (2007)による LETKF、PO 法などの解法がある。2 つ以上実装すること。

ヒント）気象分野の EnKF では、上述の手法が良く用いられている。カナダでは PO 法、米国気象局では Serial EnSRF、ドイツ・日本では LETKF など。小槻研で研究を進める場合、LETKF を用いた研究をしていくことが想定されるため、LETKF の実装には取り組んで欲しい。

6. Implement EnKF and compare with KF. There are solutions such as Serial EnSRF by Whitaker and Hamill (2002), ETKF by Bishop et al. (2001), LETKF and PO method by Hunt et al. (2007). Implement at least two or more.

Hint) The above methods are often used in EnKF in the meteorological field. PO method in Canada, Serial EnSRF in the US Meteorological Bureau, LETKF in Germany and Japan, etc. When proceeding with research at Kotsuki Lab, it is expected that research using LETKF will be carried out, so I would like you to work on the implementation of LETKF at least.

Techniques for serial EnSRF

Environmental Prediction Science Laboratory

Gaspari Cohn Function

$$r = \frac{d}{\sqrt{10/3}\,\sigma}$$

tuning parameter

d: distance b/w grids
σ: localization length scale

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

Gaspari and Cohn (1999)

Localization

Because of assimilating observation serially, analysis equation is

$$\overline{\mathbf{x}}_t^a = \overline{\mathbf{x}}_t^b + \mathbf{k}_t \left(y^{o(i)} - \mathbf{h}\overline{\mathbf{x}}_t^b \right)$$

Kalman Gain ($\in \mathbb{R}^n$) ith obs (scalar) obs ope for ith obs

Localize k depending on distance b/w grid and obs using the GC function.

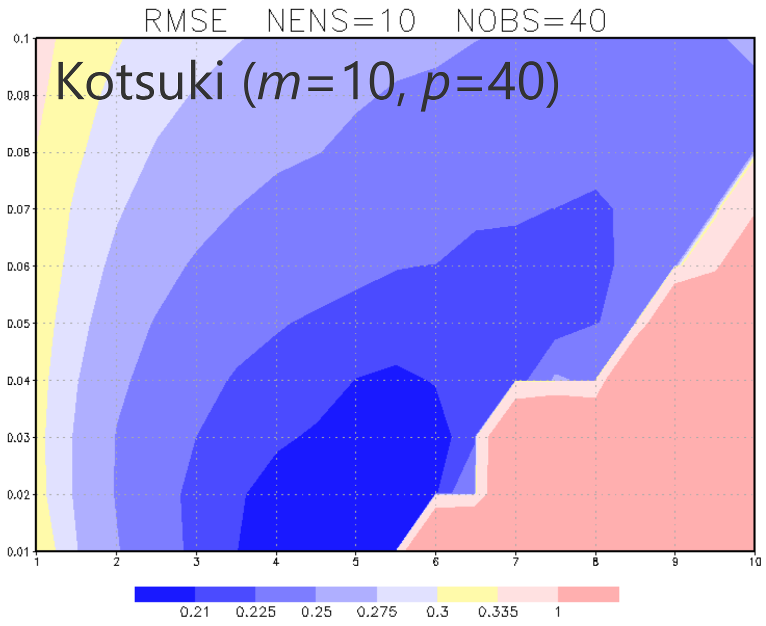
Inflation

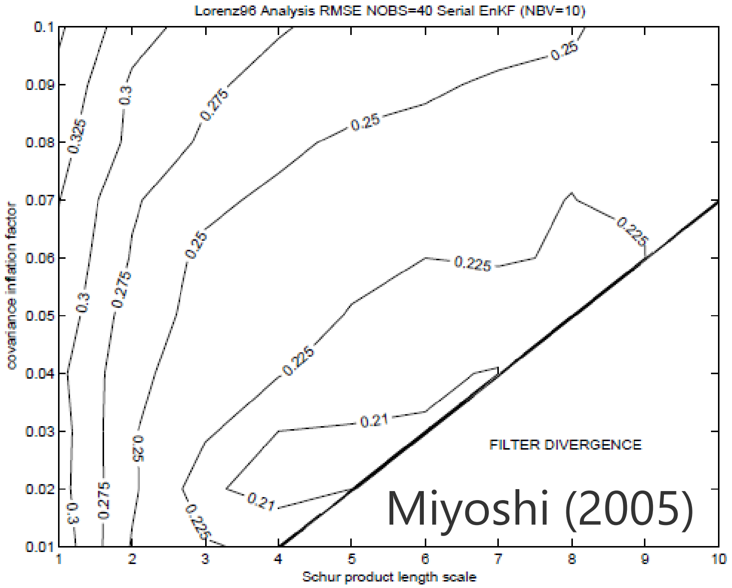
$$\delta \mathbf{X}_{inf}^b = (1 + \delta)\delta \mathbf{X}^b$$

Note: this inflation should be applied only once at assimilating 1st obs!
(i.e. no inflation for assimilating 2nd-pth obs)

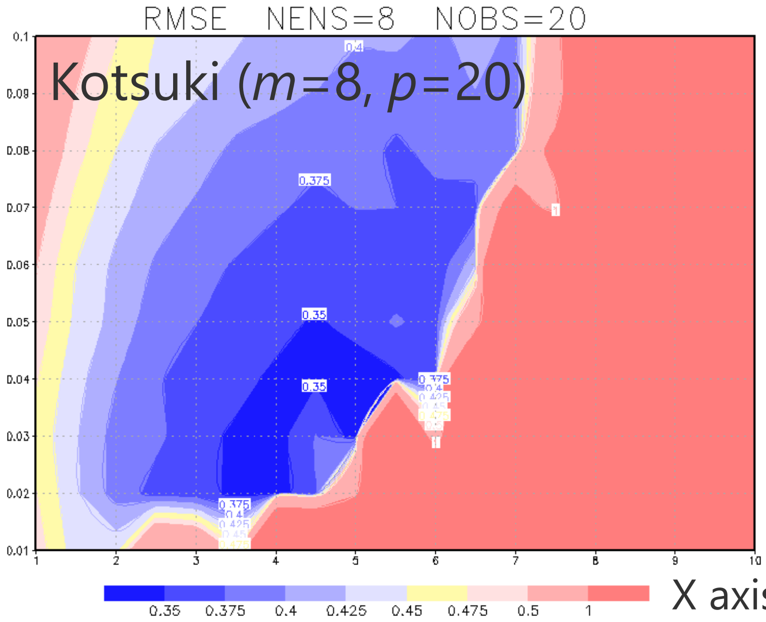
Analysis RMSE (Serial EnSRF)

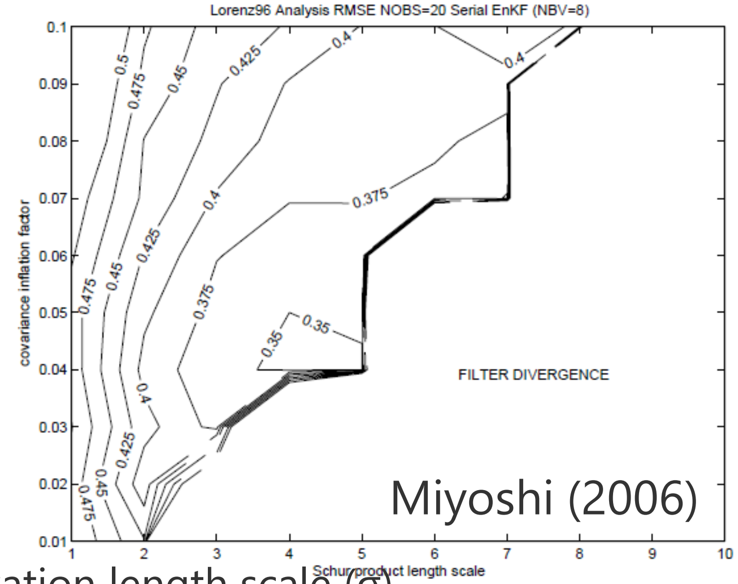
Y axis: inflation factor (δ)





Y axis: inflation factor (δ)





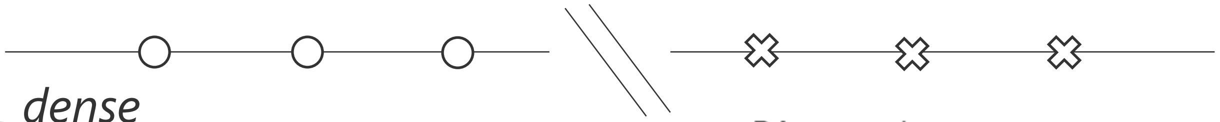
X axis: localization length scale (σ)

Sensitivity to Obs. Network

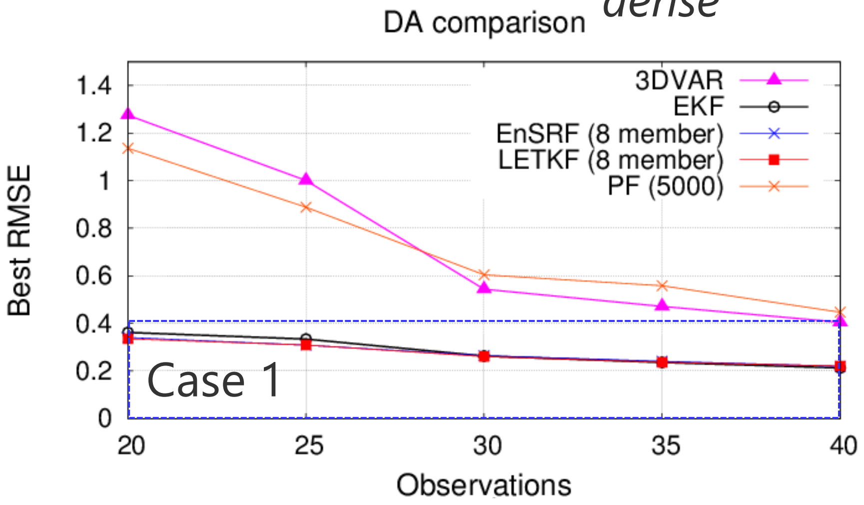
Case 1: The Num. Obs. = X

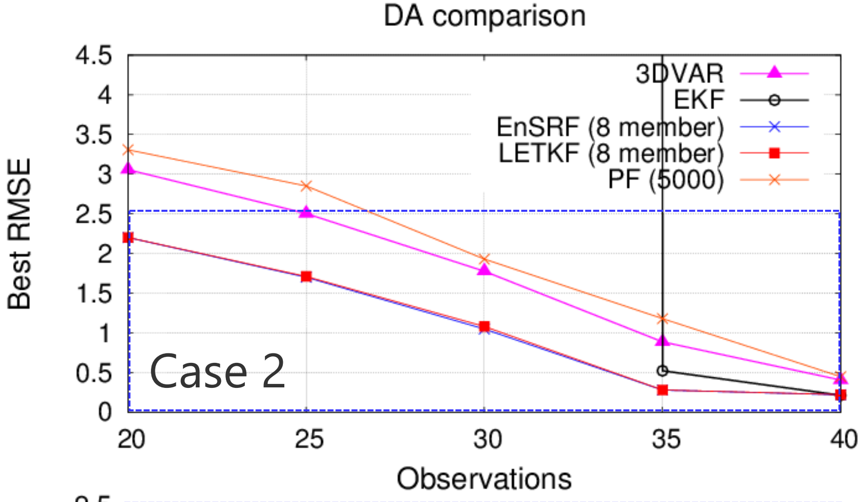
homogeneous

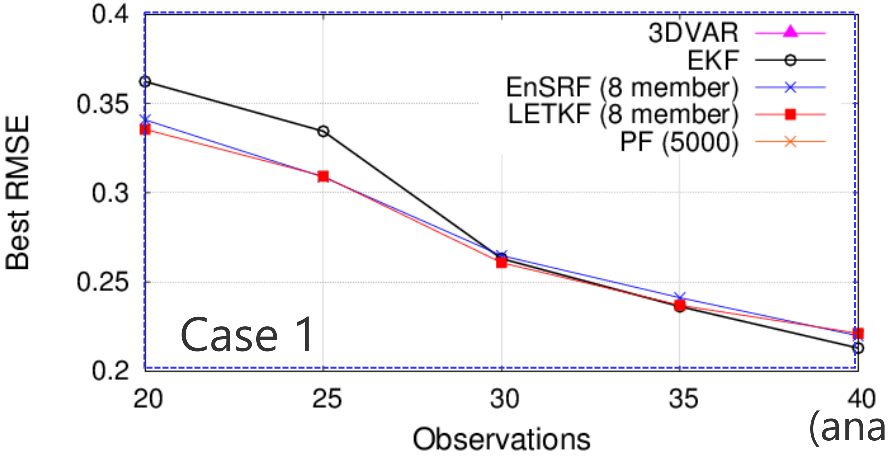
Case 2: The Num. Obs. = X



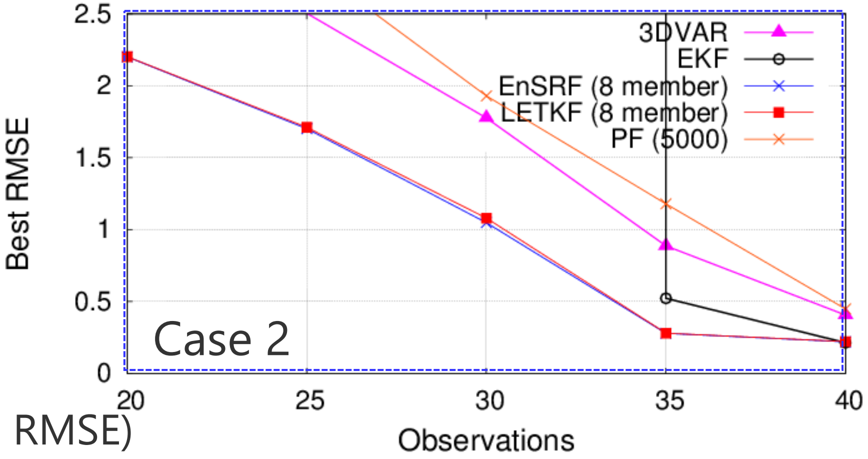
dense







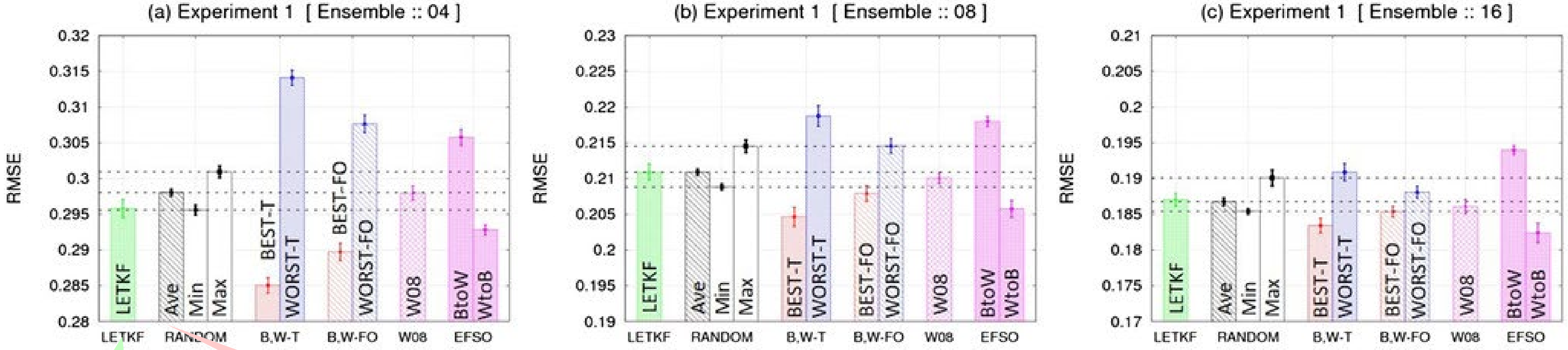
(analysis RMSE)



Kotsuki et al. (2017)

Environmental Prediction Science Laboratory

Analysis RMSE with 40 observations w/ L96 (w/ best loc. scale)



LETKF

Serial EnSRF

Kotsuki, S., Greybush, S., and Miyoshi, T. (2017):
Can we optimize the assimilation order in the serial ensemble Kalman filter? A study with the Lorenz-96 model. *Mon. Wea. Rev.*, 145, 4977-4995.

Thank you for your attention!

Presented by Shunji Kotsuki
(shunji.kotsuki@chiba-u.jp)

Further information is available at
https://kotsuki-lab.com/

Environmental Prediction Science Laboratory