



# Data Assimilation

## - A06. Ensemble Kalman Filter -

**Shunji Kotsuki**

Environmental Prediction Science Laboratory  
Center for Environmental Remote Sensing (CEReS), Chiba University  
([shunji.kotsuki@chiba-u.jp](mailto:shunji.kotsuki@chiba-u.jp))



# 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 Inflation
- ▶ (11) 4D Variational Method (4DVAR)

# Today's Goal

---

- ▶ **Lecture: Ensemble Kalman Filter**
  - ▶ to introduce EnKF
  - ▶ to understand PO method
- ▶ **Training: Lorenz 96**
  - ▶ to implement PO method
  - ▶ to implement localization

# Ensemble Kalman Filter (EnKF)

# Why EnKF?

## Kalman Filter

$$\overset{\longleftarrow n \longrightarrow}{\underset{\updownarrow n}{\mathbf{P}_t^b}}$$

Background error covariance cannot be stored on RAM for high dimensional models such as NWP ( $n \sim O(10^{12} \sim 10^{15})$ )

Ex) if  $n = 10^6 \rightarrow 10^{12} \times 8 \text{ byte} = 8 \text{ TB}$

## Ensemble Kalman Filter

$$\mathbf{P}_t^b \approx \frac{\delta \mathbf{X}_t^b (\delta \mathbf{X}_t^b)^T}{m - 1}$$

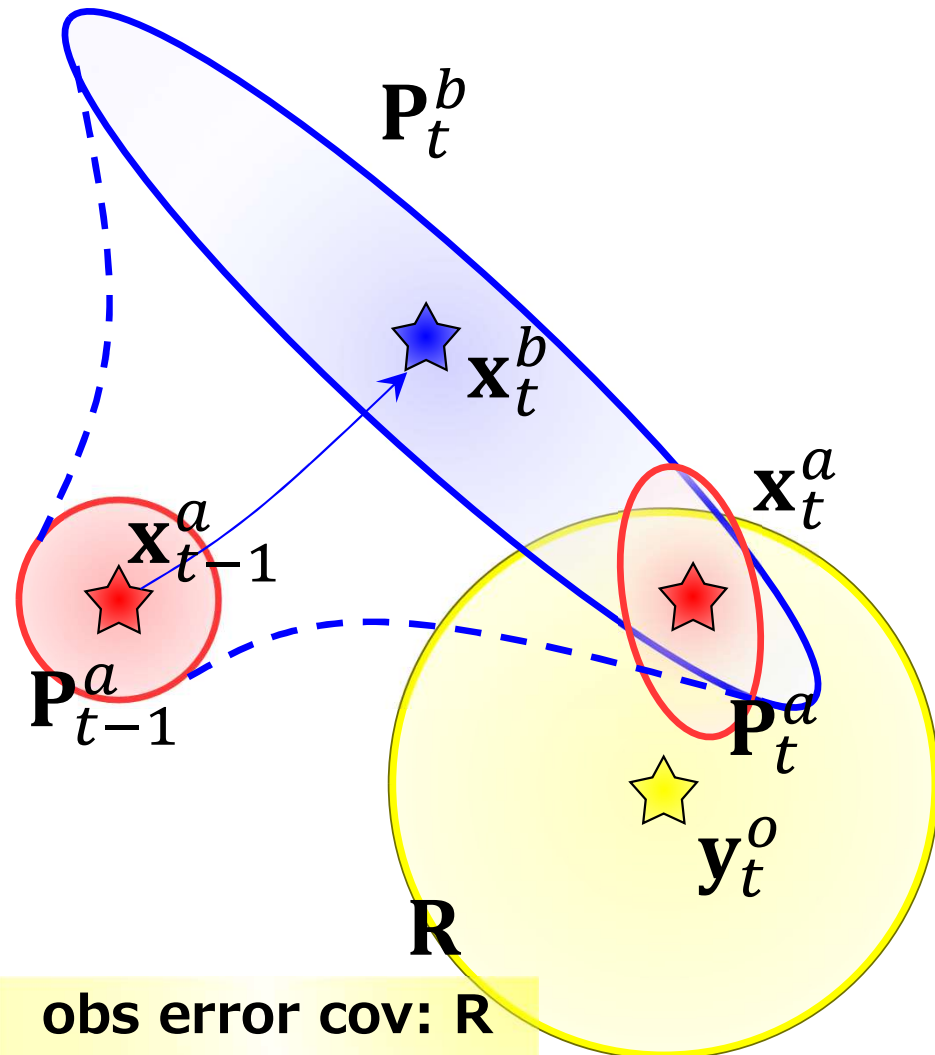
An approximation of error covariance with ensemble perturbation matrix  $\delta \mathbf{X}$ .

$$\overset{\longleftarrow m \longrightarrow}{\underset{\updownarrow n}{\delta \mathbf{X}_t^b}}$$

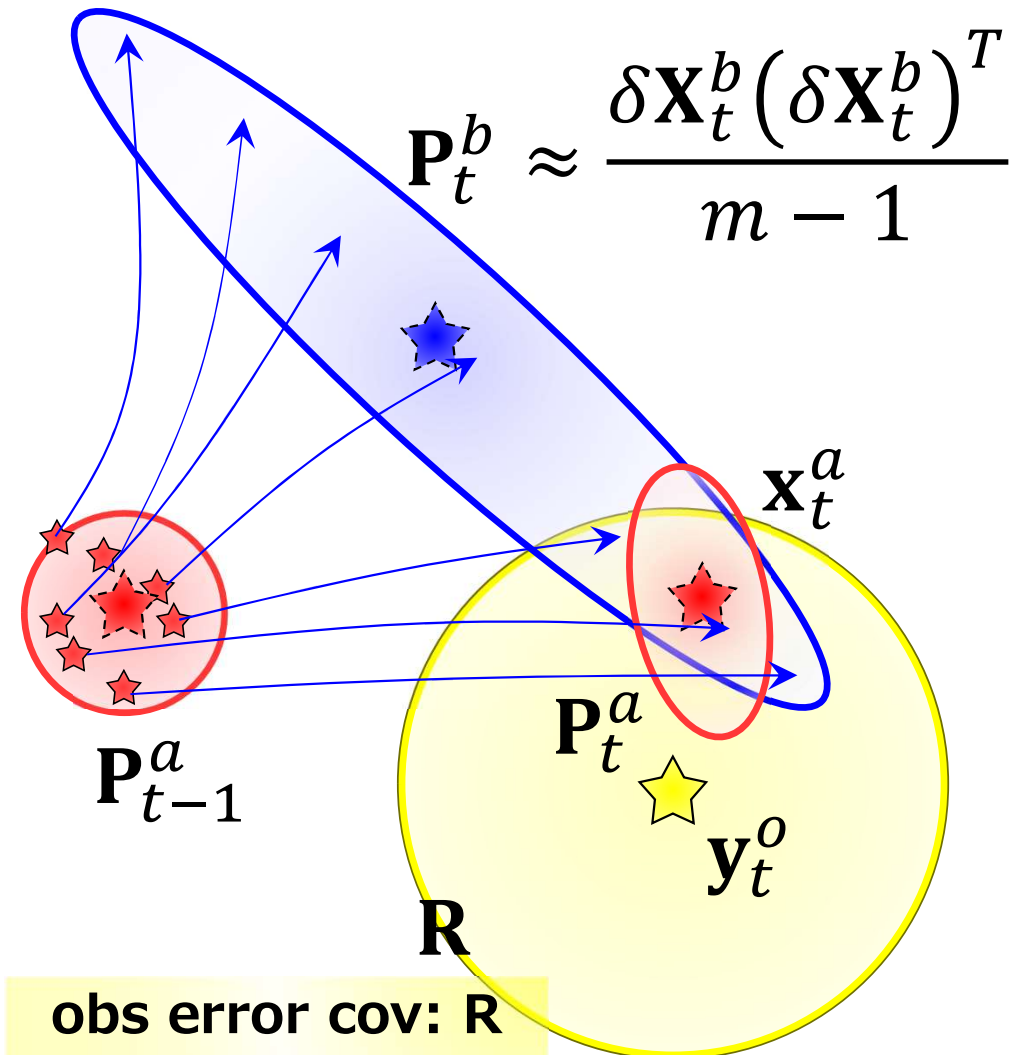
$m$ : ensemble size

# Conceptual Images

## Kalman Filter



## Ensemble Kalman Filter



# Ensemble Forecasts

## Analysis Ensemble

$$\mathbf{X}_{t-1}^a = \left[ \mathbf{x}_{t-1}^{a(1)}, \mathbf{x}_{t-1}^{a(2)}, \dots, \mathbf{x}_{t-1}^{a(m)} \right]$$

## Ensemble Forecasts

$$\mathbf{X}_t^b = \left[ \mathbf{x}_t^{b(1)}, \mathbf{x}_t^{b(2)}, \dots, \mathbf{x}_t^{b(m)} \right] \quad \text{where } \mathbf{x}_t^{b(i)} = M \left( \mathbf{x}_{t-1}^{a(i)} \right)$$

for  $i = 1, \dots, m$

## Ensemble Mean

$$\bar{\mathbf{x}} \equiv \sum_{i=1}^m \mathbf{x}^{(i)} / m$$

## Ensemble Perturbation

$\delta$  represents ensemble perturbation

$$\begin{aligned} \delta \mathbf{X}_t^b &= \left[ \mathbf{x}_t^{b(1)} - \bar{\mathbf{x}}_t^b, \mathbf{x}_t^{b(2)} - \bar{\mathbf{x}}_t^b, \dots, \mathbf{x}_t^{b(m)} - \bar{\mathbf{x}}_t^b \right] \\ &= \left[ \delta \mathbf{x}_t^{b(1)}, \delta \mathbf{x}_t^{b(2)}, \dots, \delta \mathbf{x}_t^{b(m)} \right] \end{aligned} \quad \mathbf{Z}_t^b = \delta \mathbf{X}_t^b / \sqrt{m - 1}$$

# Approximation of $\mathbf{P}^b$

## Error Propagation in Ensemble Forecasts

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

$$\overline{M(\mathbf{x}_{t-1}^a)} = \frac{1}{m} \sum_{i=1}^m M(\mathbf{x}_{t-1}^{a(i)})$$

$$\approx \left[ M(\bar{\mathbf{x}}_{t-1}^a) + \mathbf{M} \delta \mathbf{x}_{t-1}^{a(i)} \right] - \left[ M(\bar{\mathbf{x}}_{t-1}^a) + \langle \mathbf{M} \delta \mathbf{x}_{t-1}^a \rangle \right]$$

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

$$\mathbf{P}_t^b \approx \frac{1}{m-1} \delta \mathbf{X}_t^b (\delta \mathbf{X}_t^b)^T$$

$$= \frac{1}{m-1} \mathbf{M} \delta \mathbf{X}_{t-1}^a (\mathbf{M} \delta \mathbf{X}_{t-1}^a)^T = \mathbf{M} \mathbf{P}_{t-1}^a \mathbf{M}^T$$

**Ensemble forecasts can be used for approximating propagation of error covariance !**

# why $m-1$ ?

---

- ▶ Pls. study discussions on “unbiased variance”

# Ensemble Kalman Filter

## Ensemble Perturbation in Observation Space

$$\mathbf{Y}_t^b \equiv \mathbf{H}\mathbf{Z}_t^b \approx \left[ H(\mathbf{X}_t^b) - \overline{H(\mathbf{X}_t^b)} \cdot \mathbf{1} \right] / \sqrt{m-1} \quad \text{i.e. no } \mathbf{H} \text{ is needed}$$

## Error Covariance Approximation

$$\mathbf{P}_t^b \approx \mathbf{Z}_t^b (\mathbf{Z}_t^b)^T = \delta \mathbf{X}_t^b (\delta \mathbf{X}_t^b)^T / (m-1) \quad \mathbf{H}\mathbf{P}_t^b \mathbf{H}^T \approx \mathbf{Y}_t^b (\mathbf{Y}_t^b)^T$$

## Kalman Gain

$$\begin{aligned} \mathbf{K}_t &= \mathbf{P}_t^b \mathbf{H}^T [\mathbf{H}\mathbf{P}_t^b \mathbf{H}^T + \mathbf{R}]^{-1} \\ &\approx \mathbf{Z}_t^b (\mathbf{Y}_t^b)^T [\mathbf{Y}_t^b (\mathbf{Y}_t^b)^T + \mathbf{R}]^{-1} = \mathbf{Z}_t^b [\mathbf{I} + \underbrace{(\mathbf{Y}_t^b)^T \mathbf{R}^{-1} \mathbf{Y}_t^b}_{m \times m}]^{-1} (\mathbf{Y}_t^b)^T \mathbf{R}^{-1} \end{aligned}$$

$p \times p$   $m \times m$

We can choose inverse computation depending on  $p$  and  $m$ .

✖ usually  $\mathbf{R}$  is diagonal (i.e., no obs error corr.)

$$\begin{aligned} &\mathbf{Z}_t^b (\mathbf{Y}_t^b)^T [\mathbf{Y}_t^b (\mathbf{Y}_t^b)^T + \mathbf{R}]^{-1} \\ &= \mathbf{Z}_t^b [\mathbf{I} + (\mathbf{Y}_t^b)^T \mathbf{R}^{-1} \mathbf{Y}_t^b]^{-1} [\mathbf{I} + (\mathbf{Y}_t^b)^T \mathbf{R}^{-1} \mathbf{Y}_t^b] (\mathbf{Y}_t^b)^T [\mathbf{Y}_t^b (\mathbf{Y}_t^b)^T + \mathbf{R}]^{-1} \\ &= \mathbf{Z}_t^b [\mathbf{I} + (\mathbf{Y}_t^b)^T \mathbf{R}^{-1} \mathbf{Y}_t^b]^{-1} (\mathbf{Y}_t^b)^T [\mathbf{I} + \mathbf{R}^{-1} \mathbf{Y}_t^b (\mathbf{Y}_t^b)^T] [\mathbf{Y}_t^b (\mathbf{Y}_t^b)^T + \mathbf{R}]^{-1} \\ &= \mathbf{Z}_t^b [\mathbf{I} + (\mathbf{Y}_t^b)^T \mathbf{R}^{-1} \mathbf{Y}_t^b]^{-1} (\mathbf{Y}_t^b)^T \mathbf{R}^{-1} [\mathbf{R} + \mathbf{Y}_t^b (\mathbf{Y}_t^b)^T] [\mathbf{Y}_t^b (\mathbf{Y}_t^b)^T + \mathbf{R}]^{-1} \end{aligned}$$

# KF

# EnKF



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

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

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

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

Ensemble Prediction (state)

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

Prediction of Error Covariance (implicitly)

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

Kalman Gain

$$\begin{aligned} \mathbf{K}_t &= \mathbf{Z}_t^b (\mathbf{Y}_t^b)^T [\mathbf{Y}_t^b (\mathbf{Y}_t^b)^T + \mathbf{R}]^{-1} \\ &= \mathbf{Z}_t^b [\mathbf{I} + (\mathbf{Y}_t^b)^T \mathbf{R}^{-1} \mathbf{Y}_t^b]^{-1} (\mathbf{Y}_t^b)^T \mathbf{R}^{-1} \end{aligned}$$

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)

# PO Method (stochastic)

## Analysis of Ensemble

$$\boldsymbol{\varepsilon}_t^o \sim N(0, \mathbf{R})$$

randomly drawn perturbation  
→ perturbed observation

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

## Why do we need perturbation?

if w/o perturbation (to take ave. from both sides)

$$\delta \mathbf{x}_t^{a(i)} \approx \delta \mathbf{x}_t^{b(i)} - \mathbf{K}_t \mathbf{H} \delta \mathbf{x}_t^{b(i)}$$

$$\Leftrightarrow \delta \mathbf{x}_t^a \approx (\mathbf{I} - \mathbf{K}_t \mathbf{H}) \delta \mathbf{x}_t^b$$

$$\mathbf{P}_t^a \approx \delta \mathbf{x}_t^a (\delta \mathbf{x}_t^a)^T / (m - 1)$$

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

$$\mathbf{K}_t \mathbf{R} \mathbf{K}_t^T = \mathbf{K}_t \langle \boldsymbol{\varepsilon}_t^o (\boldsymbol{\varepsilon}_t^o)^T \rangle \mathbf{K}_t^T$$

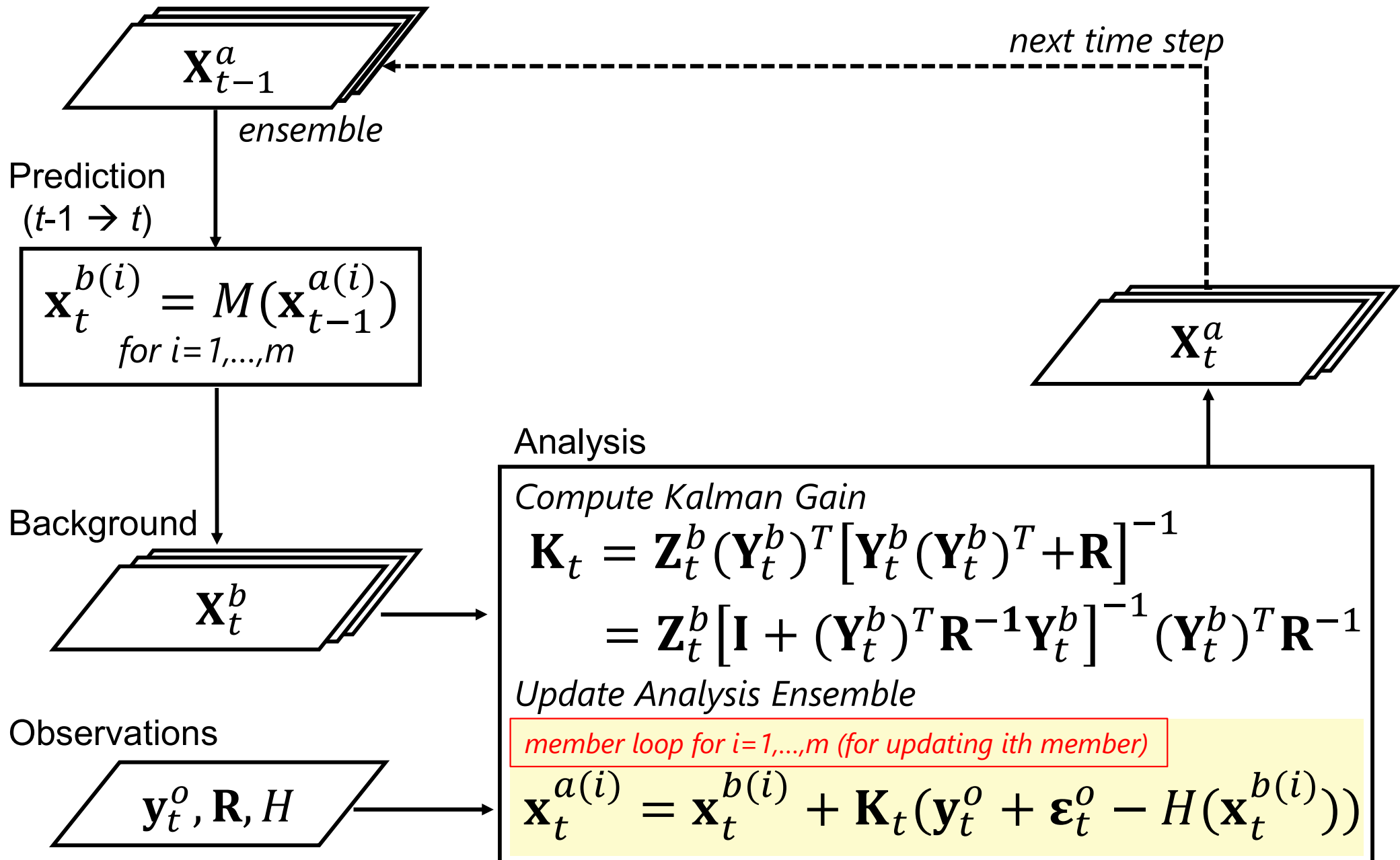
**Analysis error covariance  
is underestimated  
if without perturbation!**

Burgers et al. (1998)

Analysis error covariance  
should be (cf. 4<sup>th</sup> lecture)

$$\mathbf{P}_t^a = (\mathbf{I} - \mathbf{K} \mathbf{H}) \mathbf{P}_t^b (\mathbf{I} - \mathbf{K} \mathbf{H})^T + \mathbf{K} \mathbf{R} \mathbf{K}^T$$

# EnKF (PO) Algorithm



# Ensemble Kalman Filter

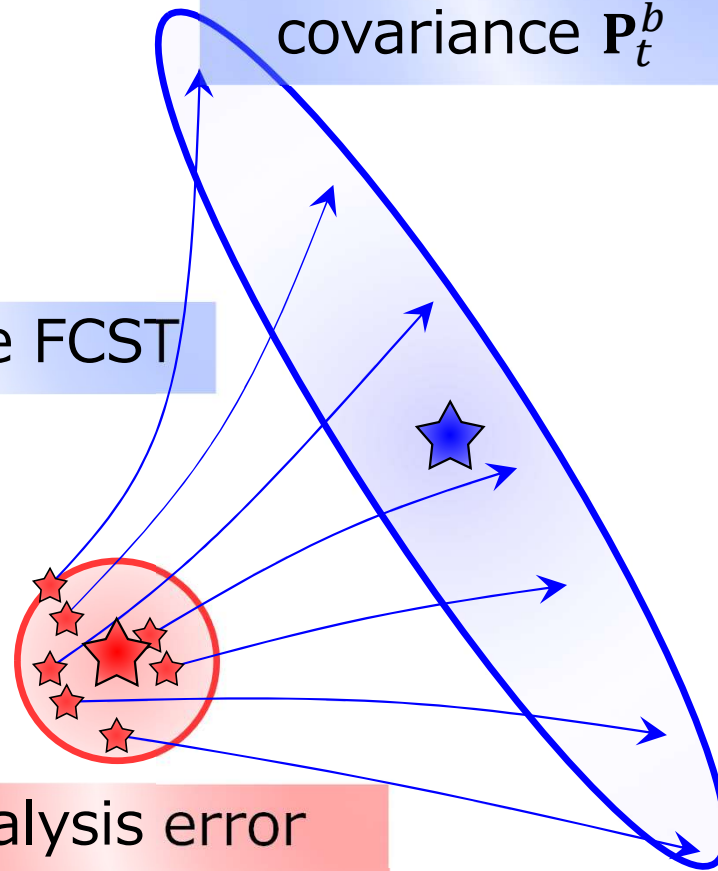
$$\mathbf{X}_t^b = M(\mathbf{X}_{t-1}^a)$$

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

Ensemble FCST

Forecast error  
covariance  $\mathbf{P}_t^b$

Analysis error  
covariance  $\mathbf{P}_{t-1}^a$



$$\mathbf{Z}_t^b = \delta \mathbf{X}_t^b / \sqrt{m-1}$$

# Ensemble Kalman Filter

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

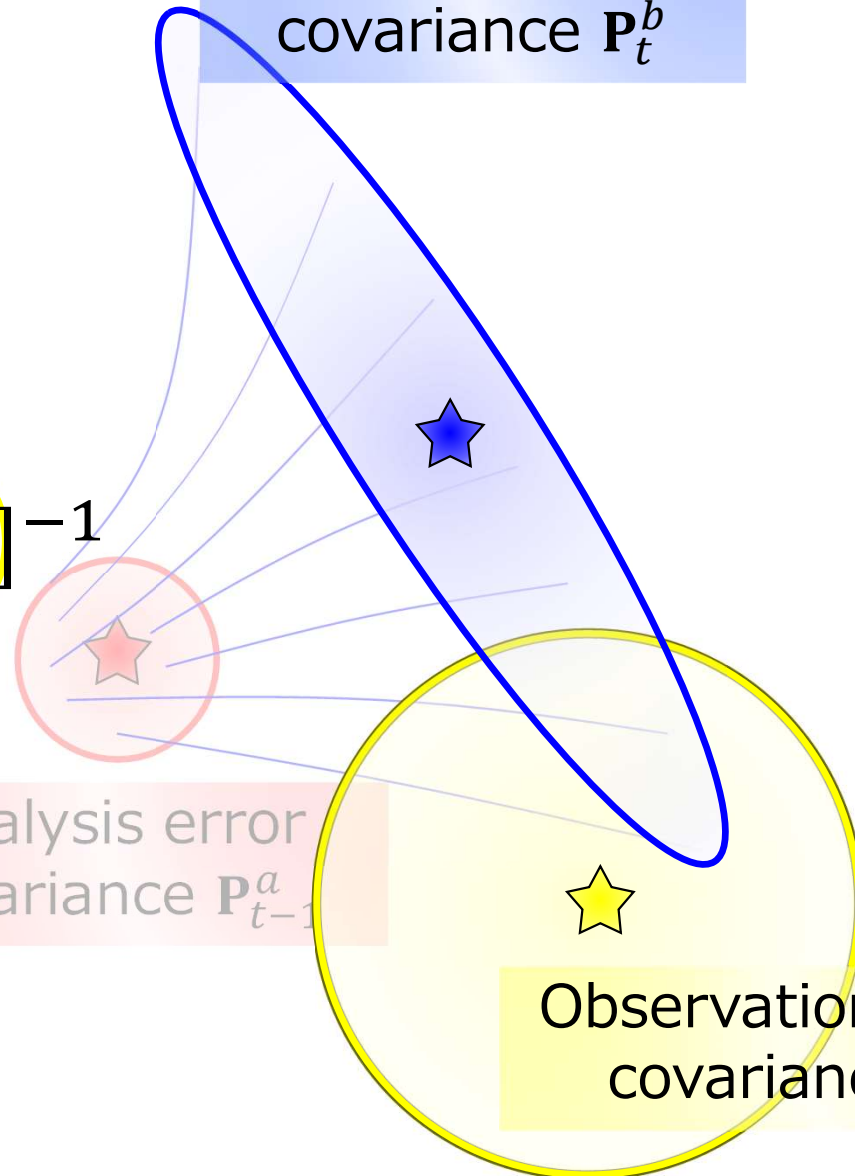
$$\mathbf{P}_t^b = \mathbf{Z}_t^b (\mathbf{Z}_t^b)^T$$

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

Forecast error  
covariance  $\mathbf{P}_t^b$

Analysis error  
covariance  $\mathbf{P}_{t-1}^a$

Observation error  
covariance  $\mathbf{R}$



# Ensemble Kalman Filter

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

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

Forecast error  
covariance  $\mathbf{P}_t^b$

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

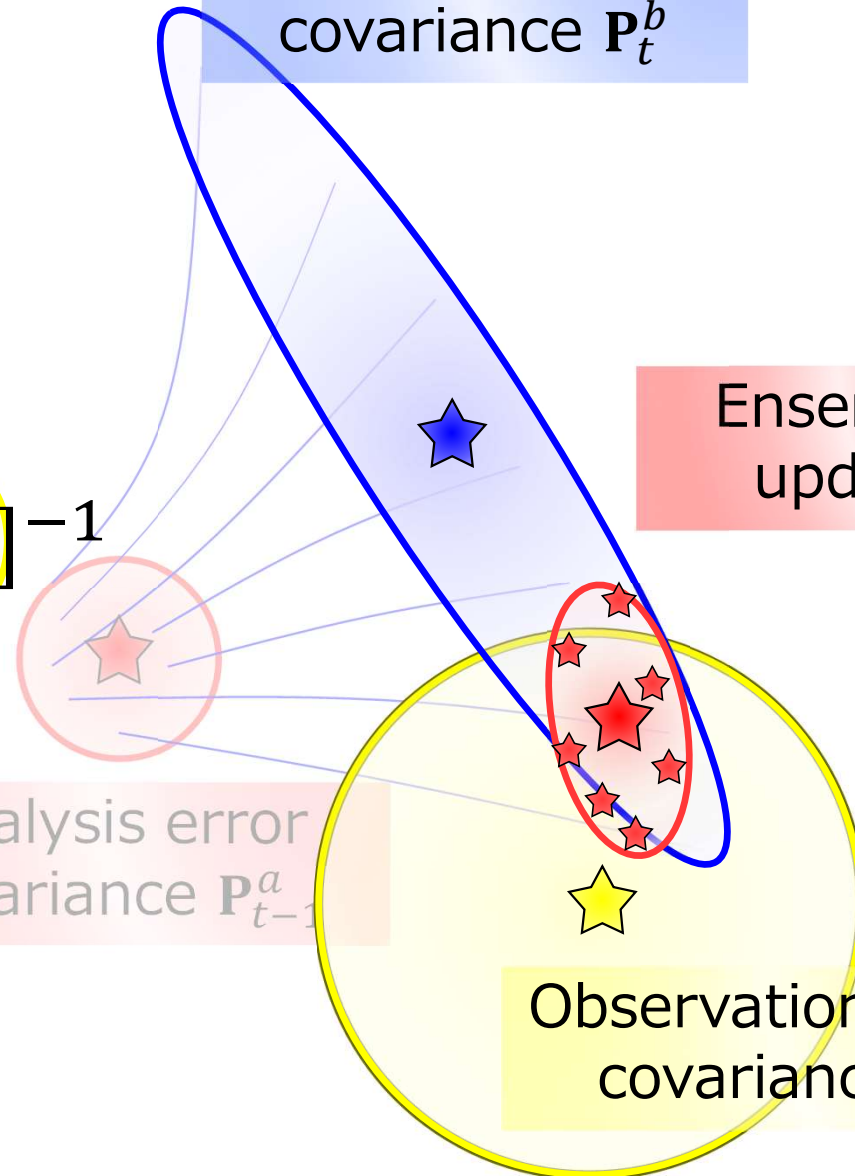
Ensemble  
update

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

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

Analysis error  
variance  $\mathbf{P}_{t-1}^a$

Observation error  
covariance  $\mathbf{R}$

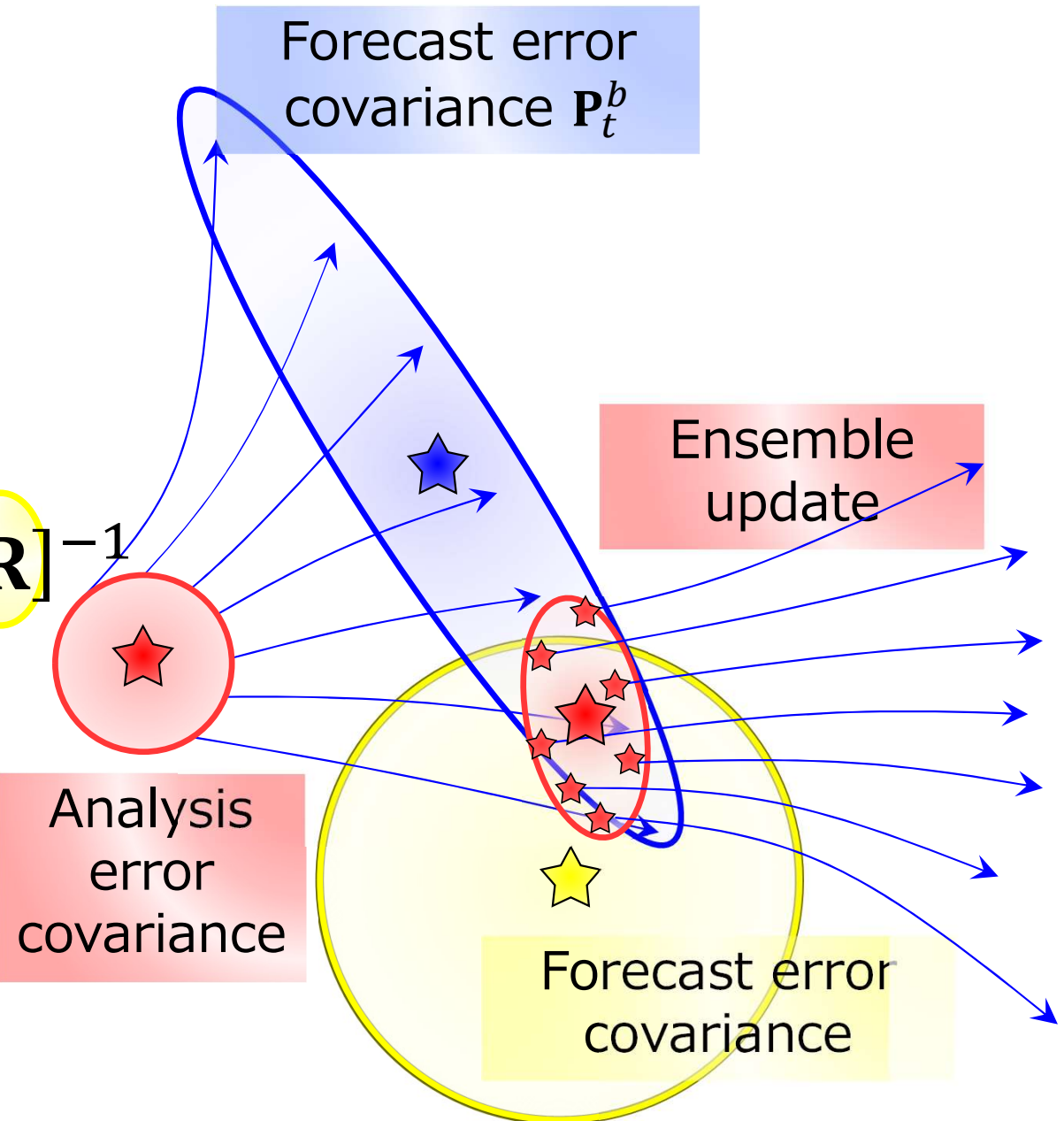


$$\mathbf{P}_t^b = \delta \mathbf{Z}_t^b (\delta \mathbf{Z}_t^b)^T$$

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

$$\mathbf{X}_t^a = \mathbf{X}_t^b + \mathbf{K}(\mathbf{y}_t^o - \mathbf{H}\mathbf{X}_t^b)$$
$$\mathbf{P}_t^a = \mathbf{Z}_t^a (\mathbf{Z}_t^a)^T$$

$$\mathbf{P}_t^a = \mathbf{Z}_t^a (\mathbf{Z}_t^a)^\top$$



# Basic Task 5

# Basic Task 5

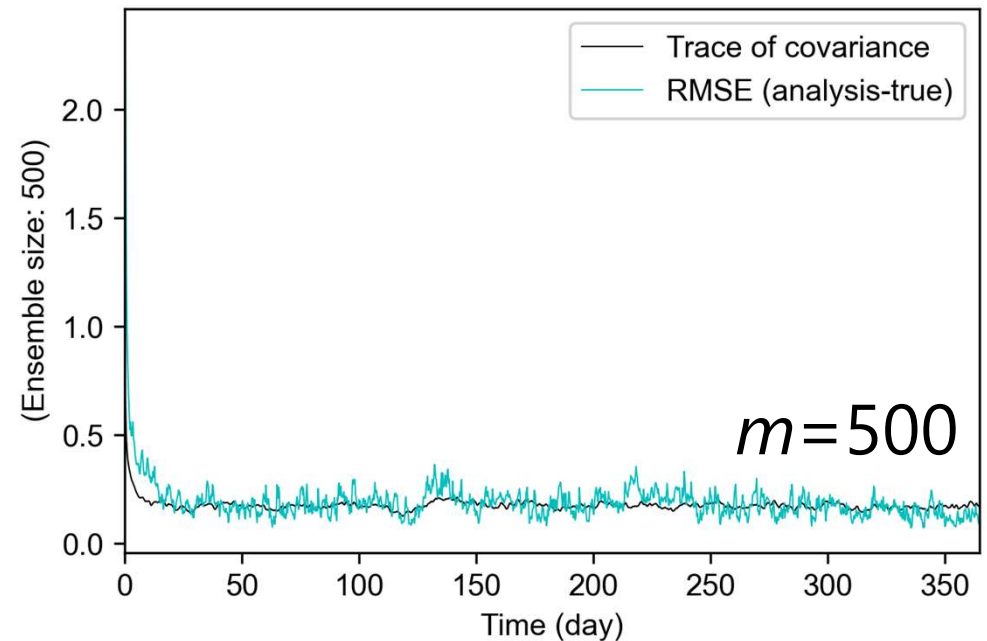
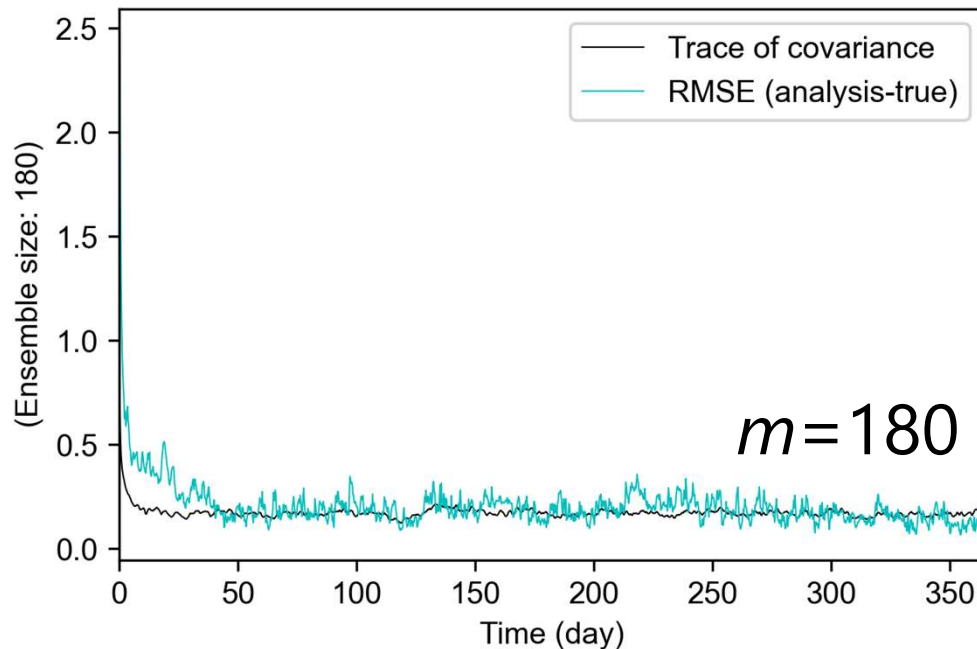
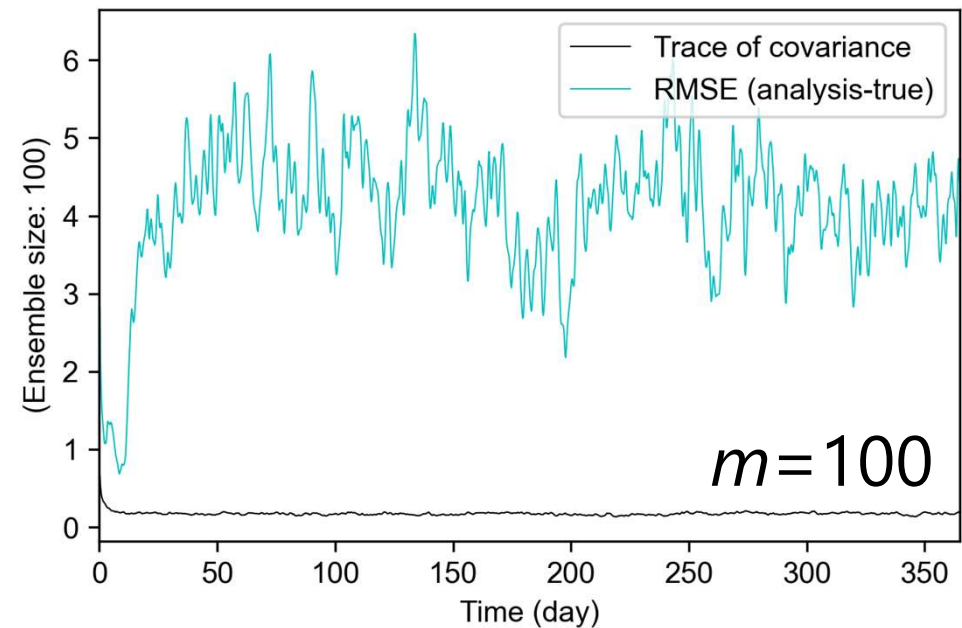
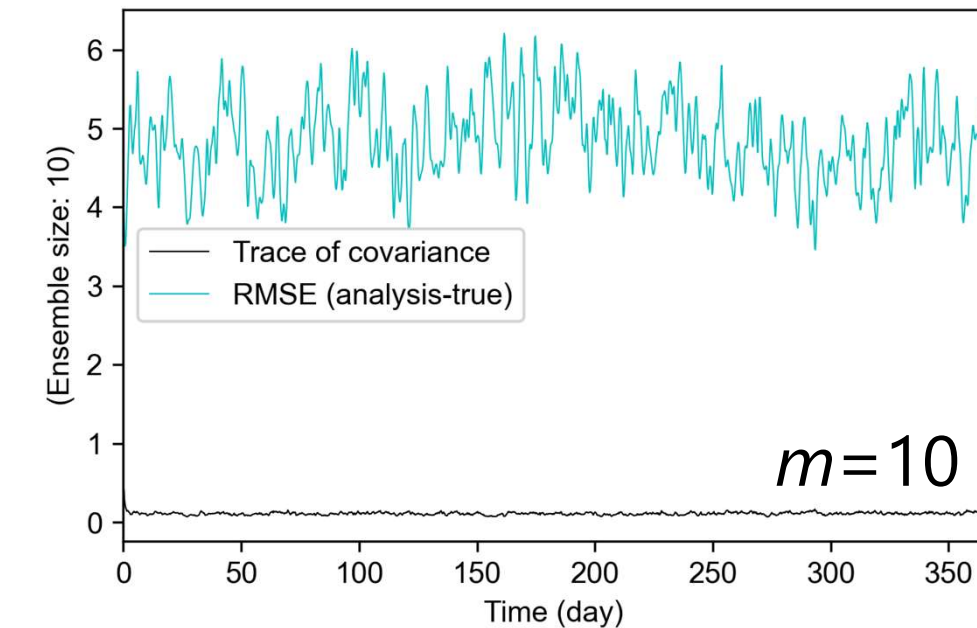
6. EnKF を実装し、KF と比較する。Whitaker and Hamill (2002)による Serial EnSRF, Bishop et al. (2001)による ETKF、Hunt et al. (2007)による LETKF、PO 法などの解法がある。2 つ以上実装すること。

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.

# EnKF (PO) w/o Localization



no inflation is used here

# Treatments

## (1) Perturbed Observations

$$\mathbf{x}_t^{a(i)} = \mathbf{x}_t^{b(i)} + \mathbf{K}_t(\mathbf{y}_t^o + \boldsymbol{\varepsilon}_t^{o(i)} - H(\mathbf{x}_t^{b(i)})) \quad \text{for } i=1,\dots,m$$

Random number should be different  
for each member

this error should be modified so that  $\sum_{i=1}^m \boldsymbol{\varepsilon}_t^{o(i)} = 0$

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

## (2) Variance Inflation

$$\mathbf{P}_{inf}^b = (1 + \delta)^2 \cdot \mathbf{P}^b$$

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

Specifically, generate random numbers  $\boldsymbol{\varepsilon}_t^{o(i)}$  for  $i = 1, \dots, m$   
and compute their average  $\bar{\boldsymbol{\varepsilon}}_t^o = \frac{1}{m} \sum_{i=1}^m \boldsymbol{\varepsilon}_t^{o(i)}$ .

Then, perturb with  $\boldsymbol{\varepsilon}_t^{o(i)} = \boldsymbol{\varepsilon}_t^{o(i)} - \bar{\boldsymbol{\varepsilon}}_t^o$  that satisfies  $\sum_{i=1}^m \boldsymbol{\varepsilon}_t^{o(i)} = 0$

## (3) Localization

- to limit impacts of obs far from analysis grid points  
for erroneous error "co"variance due to sampling errors
- several localizations have been proposed
  - **K** localization (in PO or serial EnKF)
  - **R** localization (in LETKF)
  - **B** localization (usually complex for high-dim modes)

Assimilating surrounding local obs.  
to update centered grid  $\mathbf{x}_1$  (★)

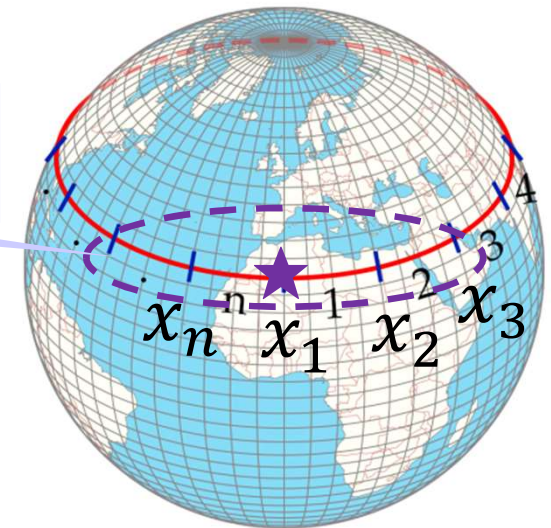


Figure 1.3: Example of a latitude circle of the earth, divided into  $n$  equal sized sectors.  
Kekem and Leendert (2018)

# covariance localization (EnKF)

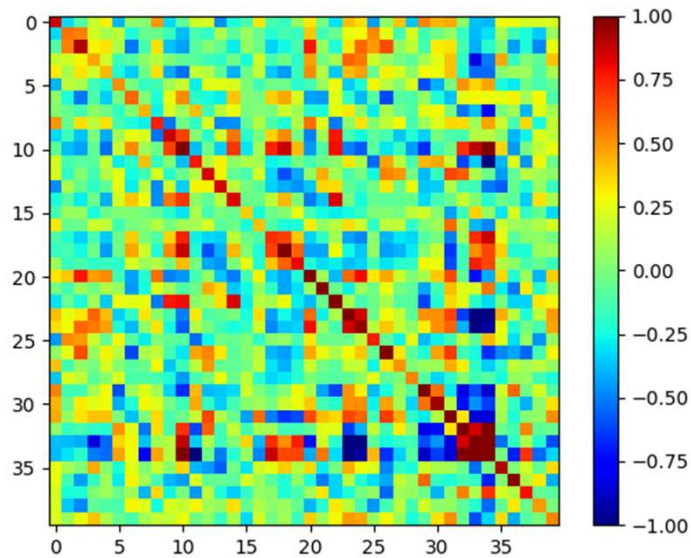
## Empirical treatment for

- (1) reducing sampling noise
- (2) increasing the rank

$$\mathbf{P}^b \rightarrow \mathbf{L} \circ \mathbf{P}^b$$

$\circ$  : Schur product

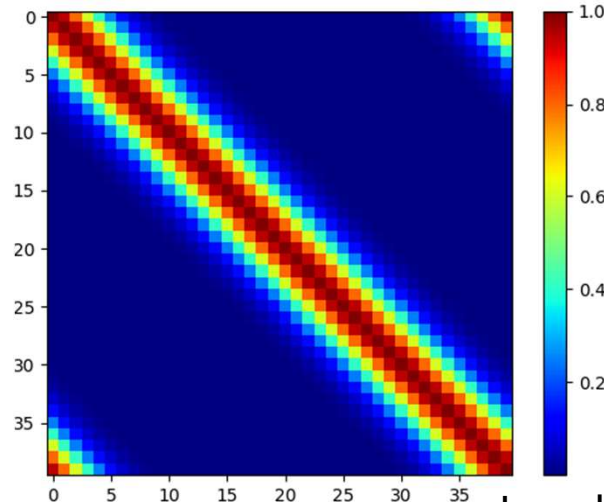
$$\mathbf{P}^b \approx \frac{1}{m-1} \delta \mathbf{X}^b (\delta \mathbf{X}^b)^T$$



Sampled error covariance  
(ensemble approximation)

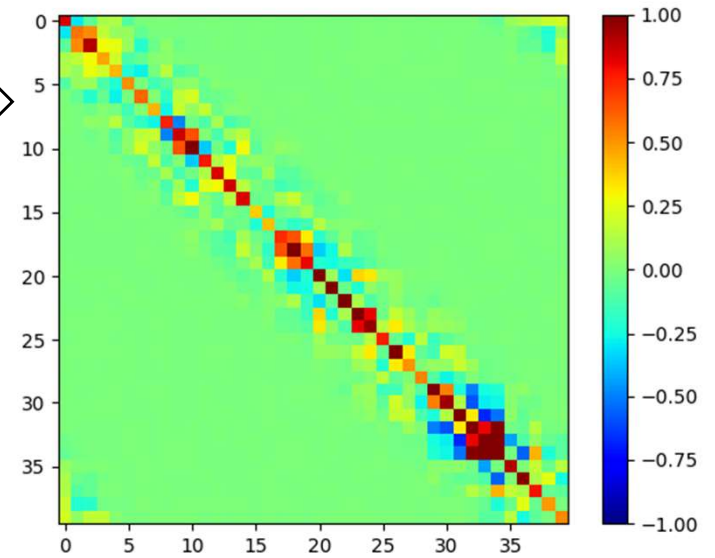
localization

$\mathbf{L}$



Localization Function

$$\mathbf{L} \circ \mathbf{P}^b$$



Error Cov. w/ Localization

# Localization Function

## Gaspari Cohn Function

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

*tuning parameter*

$d$ : distance b/w grids  
 $\sigma$ : localization length scale

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

Gaspari and Cohn (1999)

usually used in PO and serial EnSRF

## Gaussian Function

$$L(d) = \begin{cases} \exp\left(-\frac{d^2}{2\sigma^2}\right) & d < 2\sqrt{10/3}\sigma \\ 0 & \text{else} \end{cases}$$

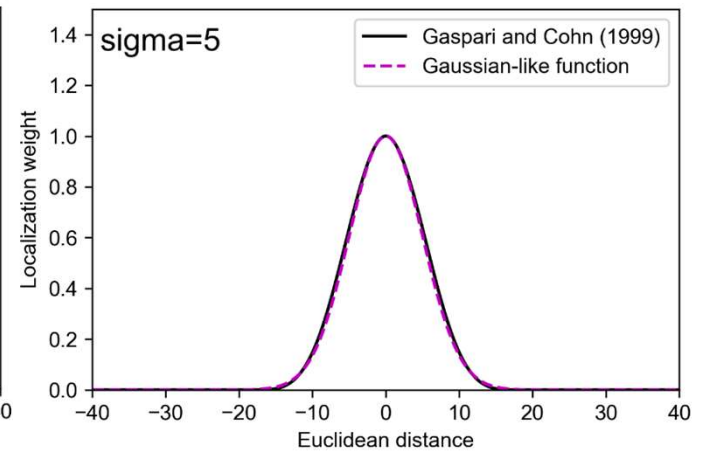
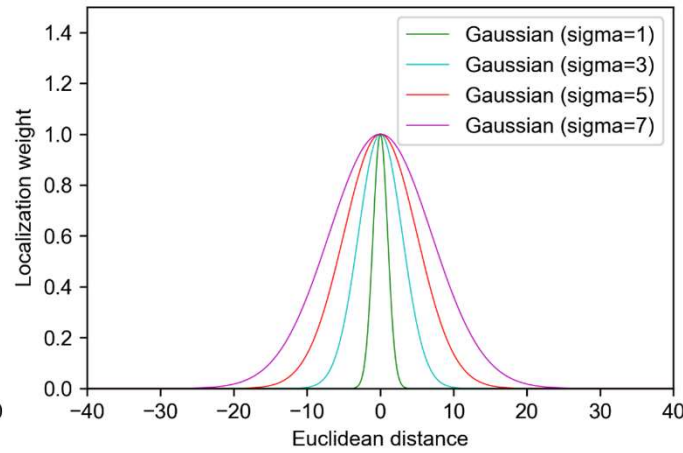
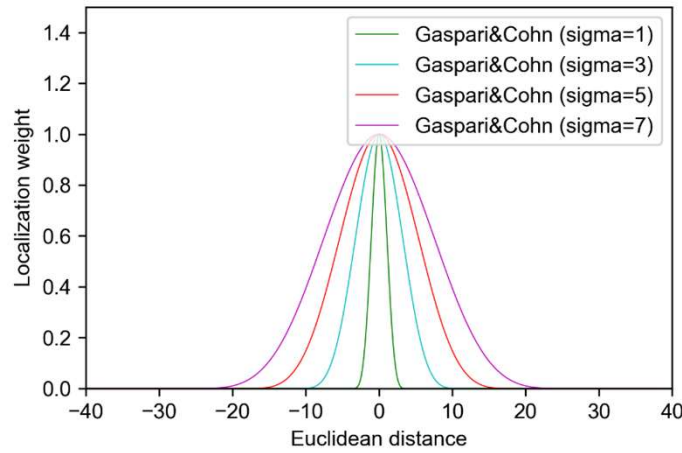
*tuning parameter*

$d$ : distance b/w grids  
 $\sigma$ : localization length scale

usually used in LETKF

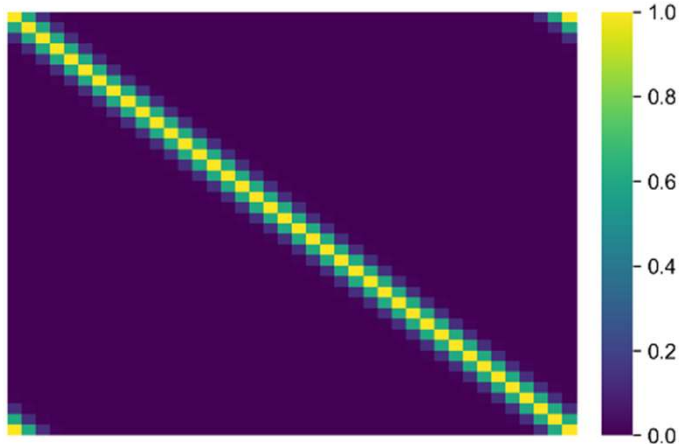
# Localization Function

## Localization Function

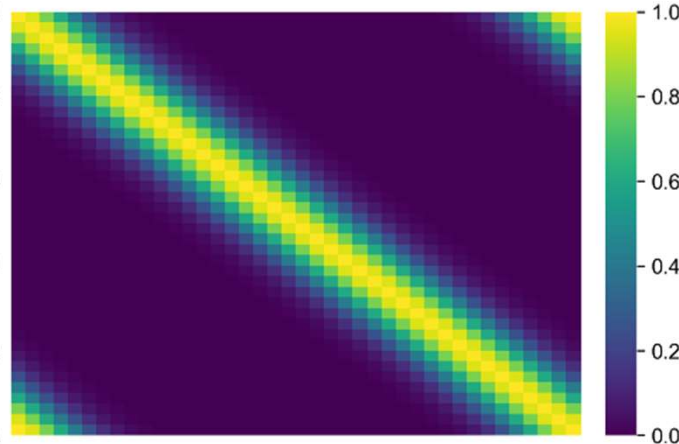


## Localization Matrix (Gaussian Function)

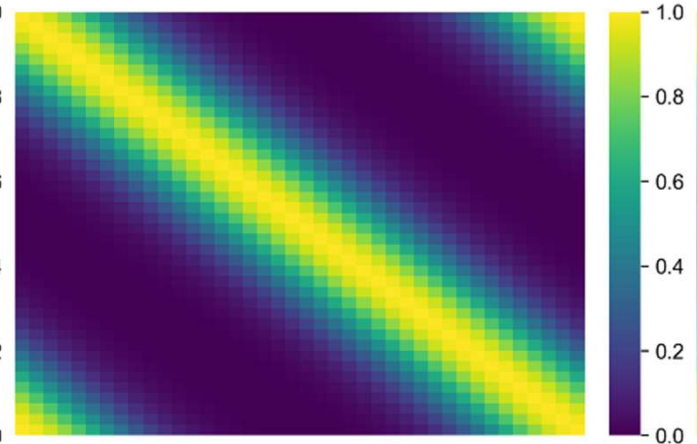
$$\sigma = 1$$



$$\sigma = 3$$



$$\sigma = 5$$



# Localization in PO method

## Kalman Gain

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

$$\mathbf{P}_t^b \mathbf{H}^T = \mathbf{L} \circ \mathbf{Z}_t^b (\mathbf{Y}_t^b)^T$$

$$\mathbf{H} \mathbf{P}_t^b \mathbf{H}^T = \mathbf{L} \circ \mathbf{Y}_t^b (\mathbf{Y}_t^b)^T$$

$\mathbf{L}$  : Localization Matrix

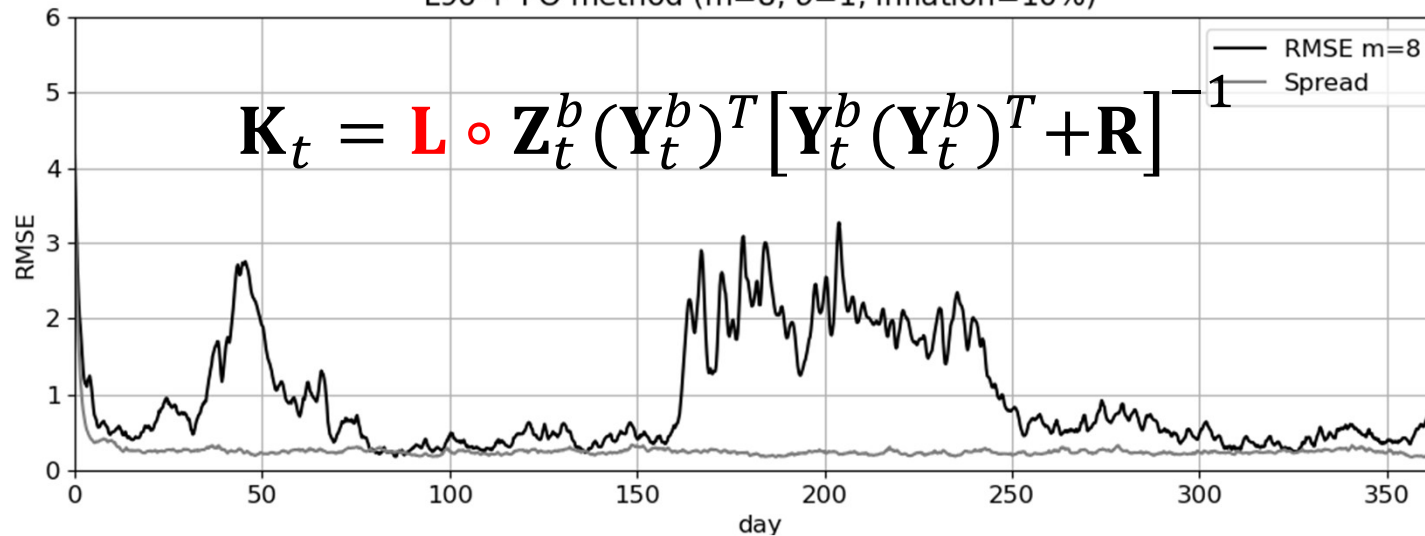
$\circ$  : Shur product  
also known as Hadamard product  
or, element-wise product

# Impacts of Localization (PO)

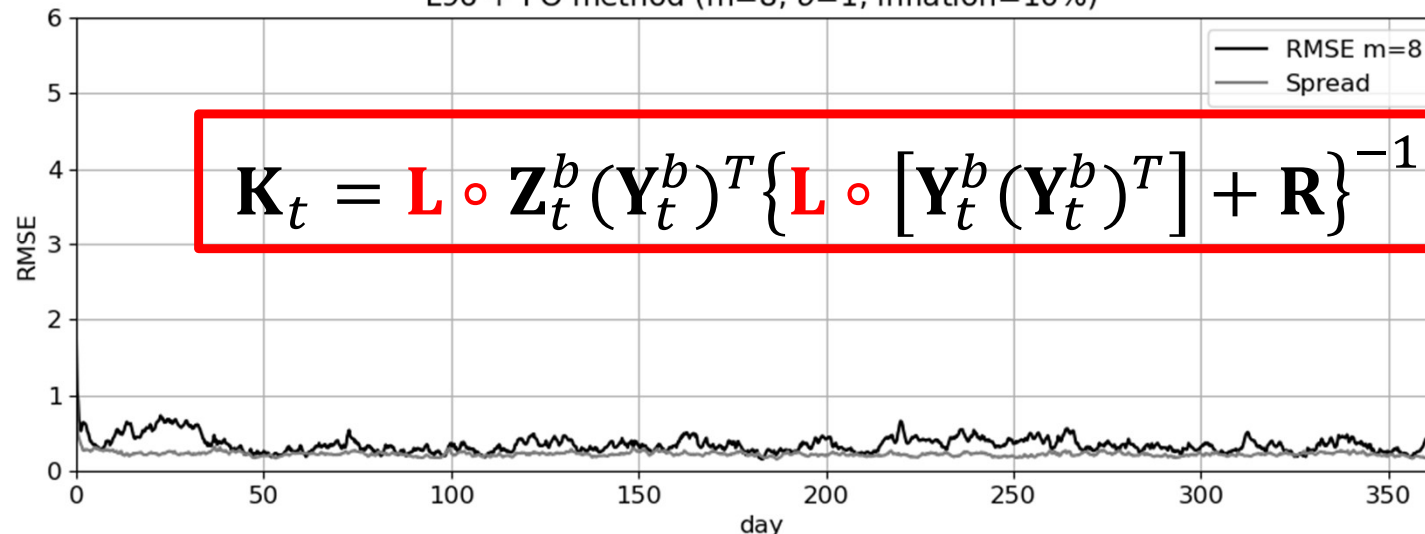
$m=8, \sigma=1, \delta=10\%$

Inflation :  $\sqrt{\delta} \circ \mathbf{Z}_t^b$

L96 + PO method ( $m=8, \sigma=1, \text{inflation}=10\%$ )

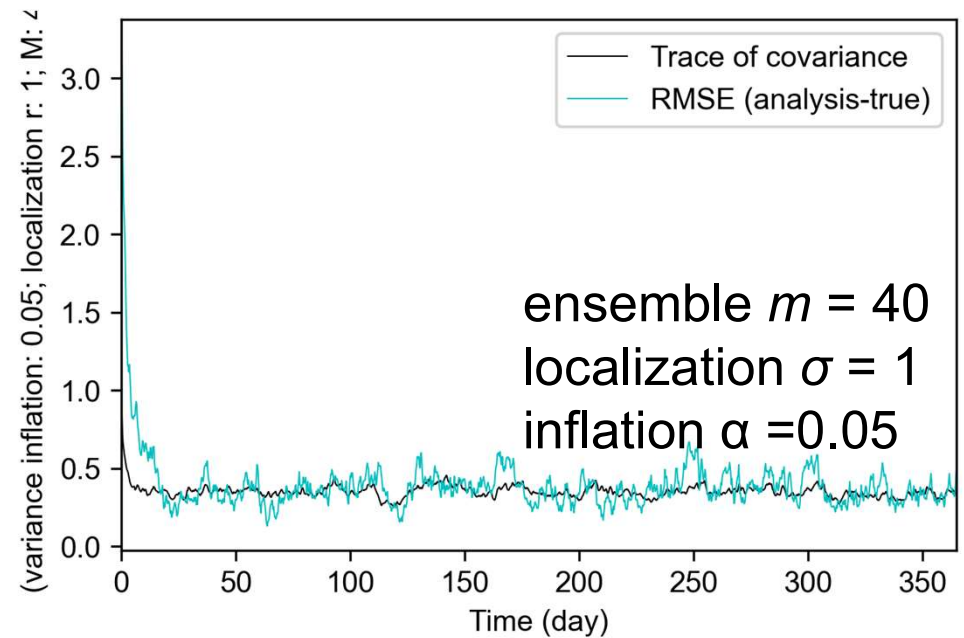
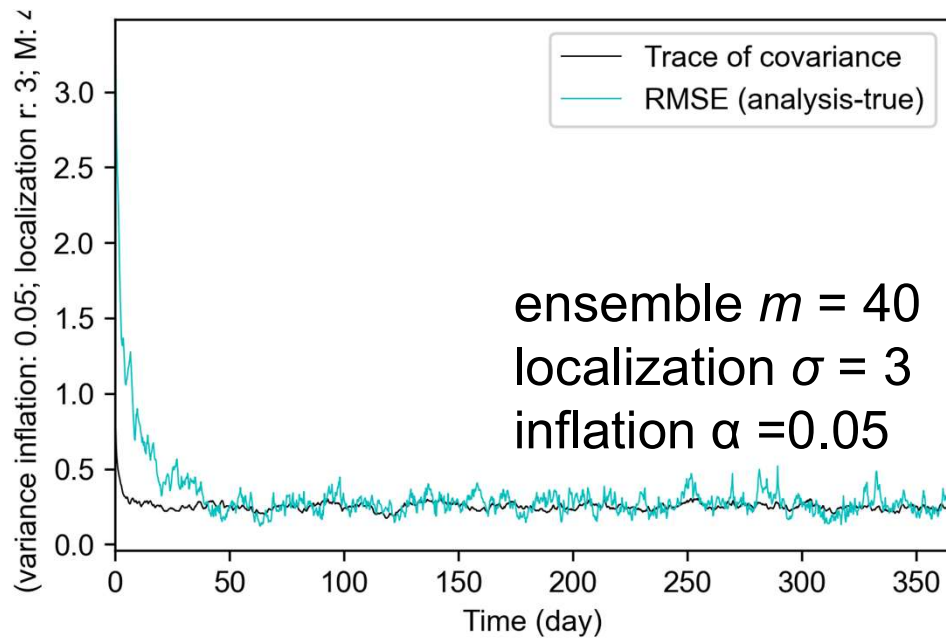
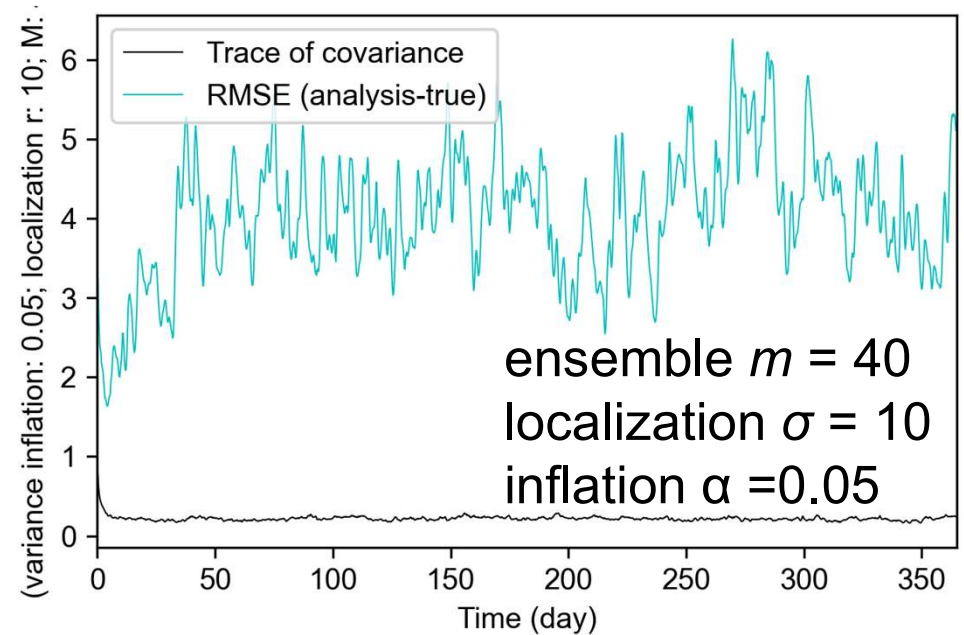
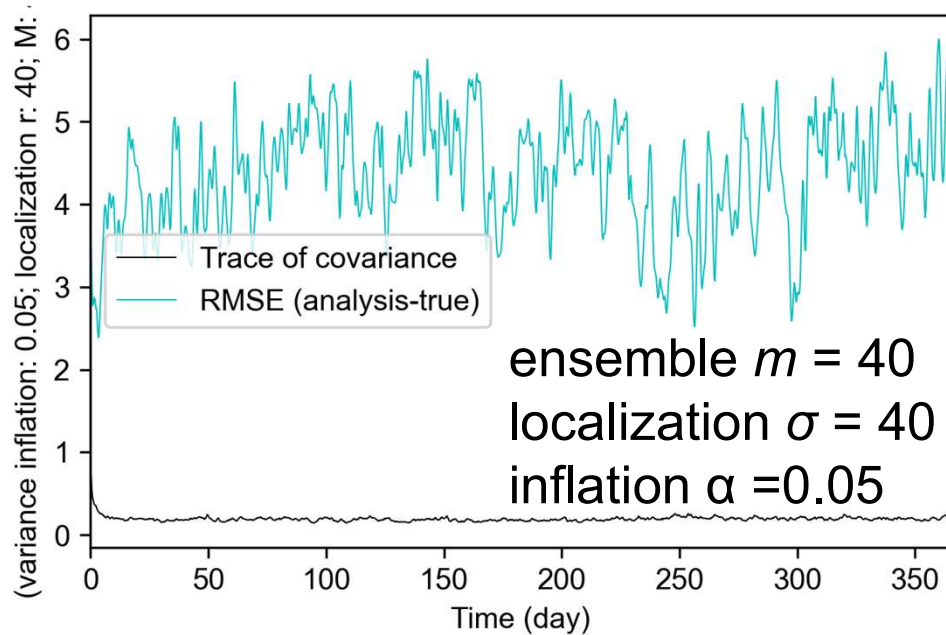


L96 + PO method ( $m=8, \sigma=1, \text{inflation}=10\%$ )



Localization should be applied for all "co"variance

# EnKF (PO) w/ Localization



# **Thank you for your attention!**

**Presented by Shunji Kotsuki**  
([shunji.kotsuki@chiba-u.jp](mailto:shunji.kotsuki@chiba-u.jp))

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

