Data Assimilation
- A05. 3DVAR and OI -

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**
  - what is the 3D-Var?
  - what is the cost function?
  - maximum likelihood vs. minimum variance
  - how can we get a reasonable B?

- **Training Course**
  - to implement 3DVAR
  - hints to develop KF
  - some tips for KF

Review:
Max. Likelihood Estimation

(復習: 最尤推定)

Maximum Likelihood Estimation

forecast $x_1 = x^{tru} + \varepsilon_1$

observation $x_2 = x^{tru} + \varepsilon_2$

(1) unbias $\langle x_1 \rangle = \langle x_2 \rangle = x^{tru}$

(2) uncorr. $\langle \varepsilon_1 \varepsilon_2 \rangle = 0$

Likelihood Prior (uniform, i.e., no prior info)

$$p(x|x_{1,2}) = \frac{p(x_{1,2}|x)\,p(x)}{p(x_{1,2})}$$

Posterior

constant (since they are given)

Bayesian Estimates

$$maximize \quad p(x|x_{1,2}) \Leftrightarrow maximize \quad p(x_{1,2}|x)$$

$$\Leftrightarrow maximize \quad p(x_1|x) \cdot p(x_2|x)$$

to maximize likelihood

Maximum Likelihood Estimation

forecast $x_1 = x^{tru} + \varepsilon_1$

observation $x_2 = x^{tru} + \varepsilon_2$

(1) unbias $\langle x_1 \rangle = \langle x_2 \rangle = x^{tru}$

(2) uncorr. $\langle \varepsilon_1 \varepsilon_2 \rangle = 0$

$$maximize \quad p(x_1|x) \cdot p(x_2|x)$$

Suppose x_1 & x_2 follow Gaussian PDF $N(x, \sigma)$

$$p(x_i|x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{(x_i - x)^2}{2\sigma_i^2}\right]$$

$$maximize \quad p(x_1|x) \cdot p(x_2|x)$$

$$\Leftrightarrow maximize \quad \frac{1}{\sqrt{2\pi\sigma_1^2}} \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left[-\frac{(x_1 - x)^2}{2\sigma_1^2} - \frac{(x_2 - x)^2}{2\sigma_2^2}\right]$$

$$\Leftrightarrow minimize \quad J(x) = \frac{(x_1 - x)^2}{\sigma_1^2} + \frac{(x_2 - x)^2}{\sigma_2^2}$$

Maximum Likelihood Estimation

Environmental Prediction Science Laboratory

forecast $x_1 = x^{tru} + \varepsilon_1$

observation $x_2 = x^{tru} + \varepsilon_2$

(1) unbias $\langle x_1 \rangle = \langle x_2 \rangle = x^{tru}$

(2) uncorr. $\langle \varepsilon_1 \varepsilon_2 \rangle = 0$

$$minimize \quad J(x) = \frac{(x_1 - x)^2}{\sigma_1^2} + \frac{(x_2 - x)^2}{\sigma_2^2}$$

$$\frac{\partial J}{\partial x} = -2\frac{(x_1 - x)}{\sigma_1^2} - 2\frac{(x_2 - x)}{\sigma_2^2} = 0$$

analysis of maximum likelihood estimates

$$x^a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} x_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} x_2$$

3DVAR

Assumption & Definition

Assumption (1) : unbiased error

$$\mathbf{x}^b = \mathbf{x}^{tru} + \boldsymbol{\varepsilon}^b \qquad \langle \boldsymbol{\varepsilon}^b \rangle = 0$$

$$\mathbf{x}^a = \mathbf{x}^{tru} + \boldsymbol{\varepsilon}^a \qquad \langle \boldsymbol{\varepsilon}^a \rangle = 0$$

$$\mathbf{y}^o = \mathbf{y}^{tru} + \boldsymbol{\varepsilon}^o \qquad \langle \boldsymbol{\varepsilon}^o \rangle = 0$$

$$\mathbf{y}^{tru} = H(\mathbf{x}^{tru})$$

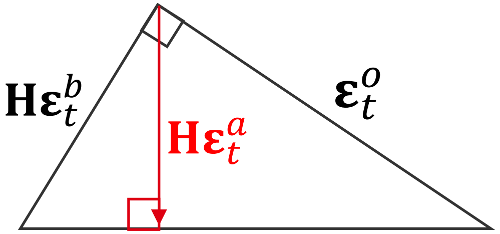
Assumption (2) : uncorrelated error

$$\langle \mathbf{H}\boldsymbol{\varepsilon}_t^b (\boldsymbol{\varepsilon}_t^o)^T \rangle = \langle (\boldsymbol{\varepsilon}_t^o)^T \mathbf{H}\boldsymbol{\varepsilon}_t^b \rangle = 0$$

since background and obs errors are independent

$$\langle \mathbf{H}\boldsymbol{\varepsilon}_t^a (\boldsymbol{\varepsilon}_t^o)^T \rangle \neq 0$$

$$\langle \boldsymbol{\varepsilon}_t^b (\boldsymbol{\varepsilon}_t^a)^T \rangle \neq 0$$



$\mathbf{x}$	model state	$\in \mathbb{R}^n$
$\boldsymbol{\varepsilon}$	error	
$\mathbf{y}$	observation	$\in \mathbb{R}^p$
$M(\,)$	nonlinear model	
$\mathbf{M}$	Jacobian of M	$\in \mathbb{R}^{n \times n}$
$\mathbf{K}$	Kalman gain	$\in \mathbb{R}^{n \times p}$
$H(\,)$	nonlin. obs. operator	
$\mathbf{H}$	Jacobian of H	$\in \mathbb{R}^{p \times n}$
$\mathbf{P}$	model error covariance	$\in \mathbb{R}^{n \times n}$
$\mathbf{R}$	obs. error covariance	$\in \mathbb{R}^{p \times p}$
n	# of model vars.	
p	# of observations	
m	# of ensemble	
tru	truth	
b	background	
a	analysis	
t	time	
o	observation	
< >	expectation	

Multidimensional Extension

Scalar

Suppose x_1 & x_2 follow Gaussian PDF $N(x, \sigma)$

$$p(x_i|x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{(x_i - x)^2}{2\sigma_i^2}\right]$$

➔ $maximize \quad p(x_1|x) \cdot p(x_2|x)$

Multi-dims.

Suppose $\mathbf{x}_t^b$ follow $N(\mathbf{x}, \mathbf{B})$

$$p^b(\mathbf{x}_t^b|\mathbf{x}) \propto \exp[-\frac{1}{2}(\mathbf{x} - \mathbf{x}_t^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_t^b)]$$

Suppose $\mathbf{y}_t^o$ follow $N(H(\mathbf{x}), \mathbf{R})$

$$p^o(\mathbf{y}_t^o|\mathbf{x}) \propto \exp[-\frac{1}{2}(H(\mathbf{x}) - \mathbf{y}_t^o)^T \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}_t^o)]$$

Joint Probability

$$p^b(\mathbf{x}_t^b|\mathbf{x}) \cdot p^o(\mathbf{y}_t^o|\mathbf{x}) \propto \exp[-J(\mathbf{x})]$$

$maximize \ p^b(\mathbf{x}_t^b|\mathbf{x}) \cdot p^o(\mathbf{y}_t^o|\mathbf{x})$
$\Leftrightarrow \ minimize \ J(\mathbf{x})$

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_t^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_t^b) + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y}_t^o)^T \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}_t^o)$$

Variational DA

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_t^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_t^b) + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y}_t^o)^T \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}_t^o)$$

$$\mathbf{x} = \mathbf{x}_t^b + \delta\mathbf{x} \quad \& \quad H(\mathbf{x}_t^b + \delta\mathbf{x}) \approx H(\mathbf{x}_t^b) + \mathbf{H}\delta\mathbf{x}$$

$$J(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^T \mathbf{B}^{-1}(\delta\mathbf{x}) + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}_t^{o-b})^T \mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}_t^{o-b})$$

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

d: innovation, departure

gradient

$$\frac{\partial J(\delta\mathbf{x})}{\partial(\delta\mathbf{x})} = \mathbf{B}^{-1}\delta\mathbf{x} + \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}_t^{o-b}) = \mathbf{0}$$

necessary condition

$$\Leftrightarrow (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\delta\mathbf{x} = \mathbf{H}^T\mathbf{R}^{-1}\mathbf{d}_t^{o-b}$$

$$\Leftrightarrow \delta\mathbf{x} = \mathbf{A}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{d}_t^{o-b}$$

$$\Leftrightarrow \mathbf{x}_t^a - \mathbf{x}_t^b = \delta\mathbf{x} = \mathbf{K}_t\mathbf{d}_t^{o-b}$$

B, $\mathbf{P}^b$: background error covariance
A, $\mathbf{P}^a$: analysis error covariance

Variational DA (cont'd)

Proof of Kalman Gain

$$\mathbf{K}_t = (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}$$

$$= (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{R} + \mathbf{HBH}^T)(\mathbf{R} + \mathbf{HBH}^T)^{-1}$$

$$= (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}(\mathbf{H}^T + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{HBH}^T)(\mathbf{R} + \mathbf{HBH}^T)^{-1}$$

$$= (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\mathbf{BH}^T(\mathbf{R} + \mathbf{HBH}^T)^{-1}$$

$$= \mathbf{BH}^T(\mathbf{R} + \mathbf{HBH}^T)^{-1}$$

Proof of Analysis Error Cov.

$$(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1} = \mathbf{B} - [\mathbf{I} - (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{B}^{-1}]\mathbf{B}$$

$$= \mathbf{B} - (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}[(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}) - \mathbf{B}^{-1}]\mathbf{B}$$

$$= \mathbf{B} - \mathbf{KHB} = (\mathbf{I} - \mathbf{KH})\mathbf{B} = \mathbf{A}$$

Important Equations

Kalman Gain

$$\mathbf{K}_t = \mathbf{B}\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1} = \mathbf{A}\mathbf{H}^T\mathbf{R}^{-1}$$

Analysis Error Covariance

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B} \iff \mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$$

Analysis Update Equation

$$\mathbf{x}_t^a = \mathbf{x}_t^b + \mathbf{K}_t\mathbf{d}_t^{o-b} = \mathbf{A}\left[\mathbf{B}^{-1}\mathbf{x}_t^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_t^o\right]$$

$$\iff \mathbf{A}^{-1}\mathbf{x}_t^a = \mathbf{B}^{-1}\mathbf{x}_t^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_t^o$$

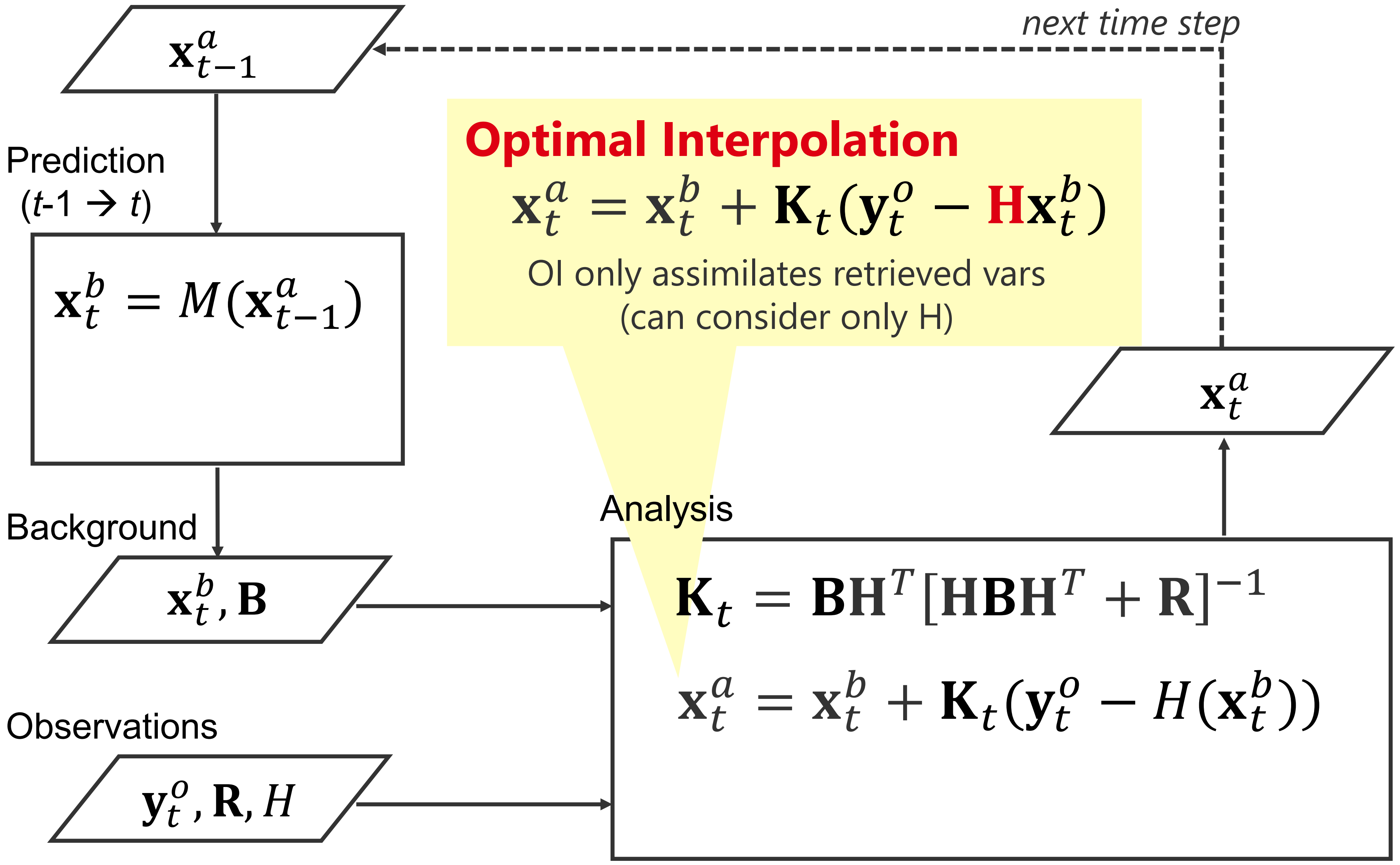
$$\mathbf{A}\left[\mathbf{B}^{-1}\mathbf{x}_t^b + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_t^o\right]$$

$$= \mathbf{A}[\mathbf{A}^{-1} - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}]\mathbf{x}_t^b + \mathbf{A}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_t^o$$

$$= \mathbf{x}_t^b + \mathbf{A}\mathbf{H}^T\mathbf{R}^{-1}\left[\mathbf{y}_t^o - \mathbf{H}\mathbf{x}_t^b\right] = \mathbf{x}_t^a$$

3DVAR

Environmental Prediction Science Laboratory



Deeper Interpretations

How To Construct B?

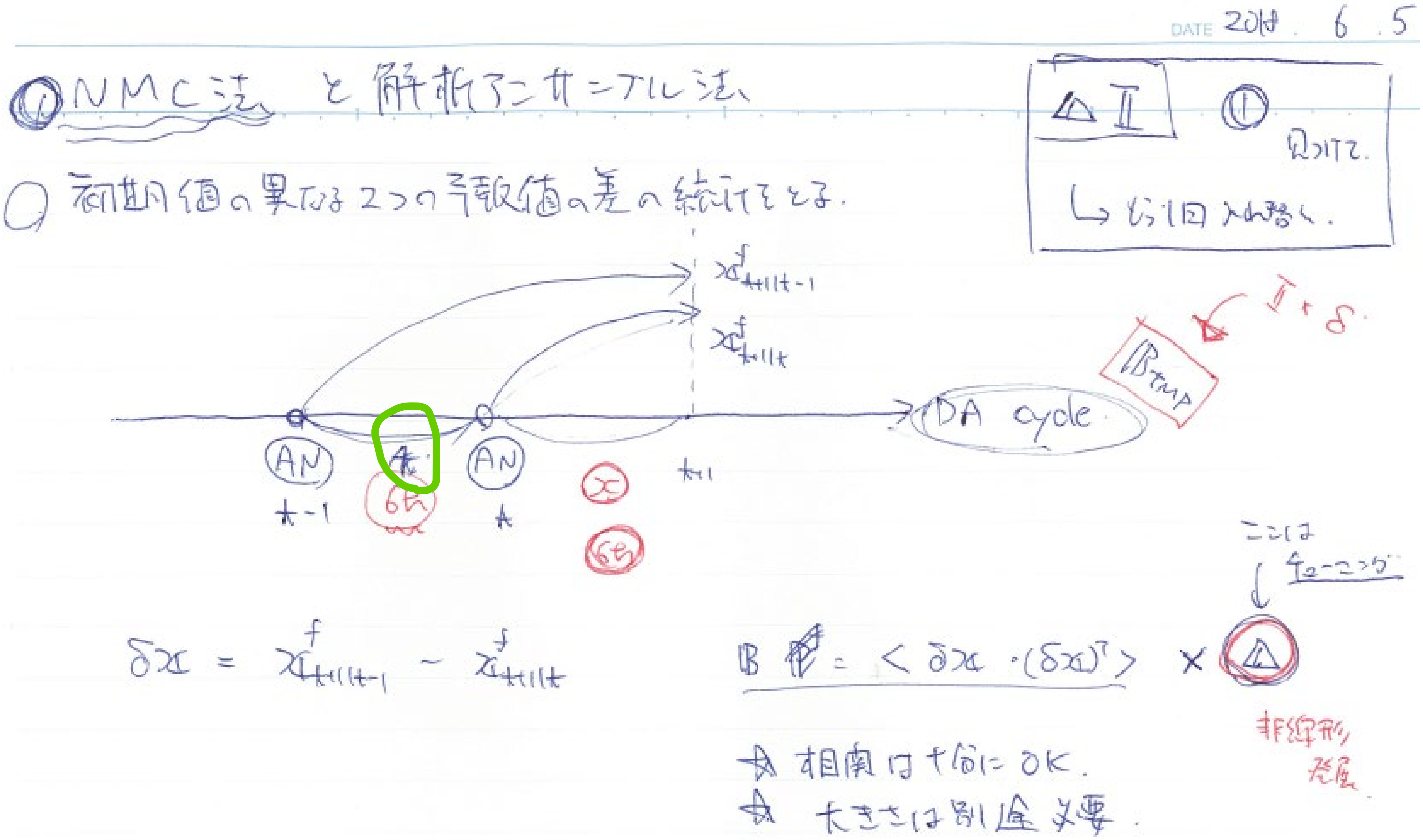
- **(1) A simplistic method**
  - to assume diagonal background error covariance
  - to tune error variance manually

- **(2) NMC Method (NMC: U.S. National Meteorological Center)**
  - Parrish and Derver (1992)
    - taking difference between 24-h and analysis, and rescaled
      - To reflect 6-h forecast error
  - Later, differences between pairs of forecasts valid at the same time (e.g. 48 and 24 h forecasts) were similarly used.
    - one reason for using such a long lag was to mitigate diurnal signals

Parrish, D. F. and Derber, J. C. (1992):
The national meteorological center's spectral statistical-interpolation analysis system
Mon. Wea. Rev., 120, 1747–1763

Errico et al. (2014):
Use of an OSSE to evaluate background-error covariances estimated by the NMC method
QJRMS

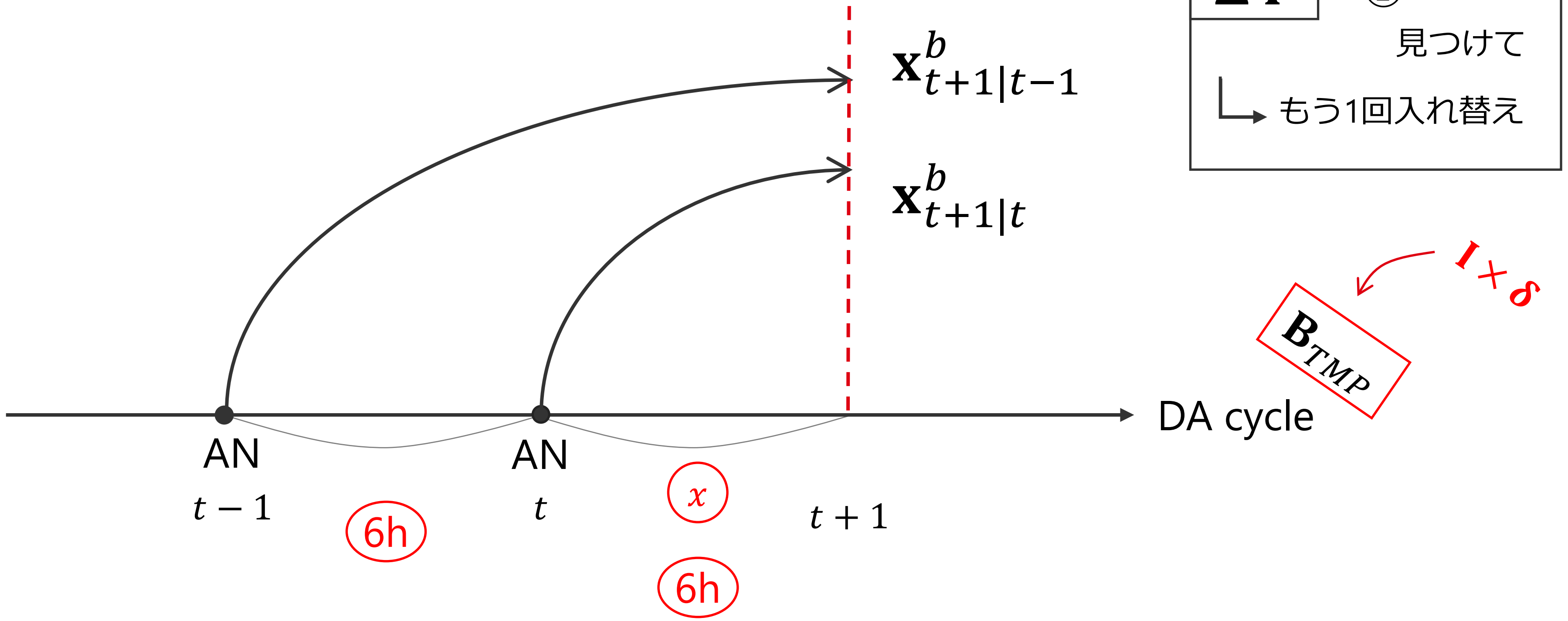
NMC Method


DATE 2018. 6. 5
① NMC法 と 解析アンサンブル法
○ 初期値の異なる2つの予報値の差の統計をとる.
AN
AN
t-1
t
t+1
DA cycle
$x^f_{t+1|t-1}$
$x^f_{t+1|t}$
$\delta x = x^f_{t+1|t-1} - x^f_{t+1|t}$
$B = \langle \delta x \cdot (\delta x)^T \rangle \times \triangle$
ここはチューニング
非線形発展
★ 相関は十分にOK.
★ 大きさは別途必要.

NMC Method

NMC法と解析アンサンブル法

初期値の異なる2つの予報値の差の統計をとる



$$\delta\mathbf{x} = \mathbf{x}^b_{t+1|t-1} - \mathbf{x}^b_{t+1|t}$$

$$\mathbf{B} = \langle \delta\mathbf{x}\cdot(\delta\mathbf{x})^T\rangle \times \boldsymbol{\Delta}$$

ここはチューニング

非線形発展

☆相関は十分にOK
☆大きさは別途必要

Training Course

Environmental
Prediction
Science
Laboratory

DA Study w/ 40-variable Lorenz-96

Lorenz-96 model (Lorenz 1996)

For $j=1,\ldots,N$, $X_j=X_{j+N}$

$$dX_j/dt = \underbrace{(X_{j+1} - X_{j-2})X_{j-1}}_{\text{Advection term}} - \underbrace{X_j}_{\text{Dissipation term}} + \underbrace{F}_{\text{Forcing term}}$$

Advection term　Dissipation term　Forcing term

力学系モデル・データ同化基礎技術の速習コース

Training Course of Dynamical Model and Data Assimilation

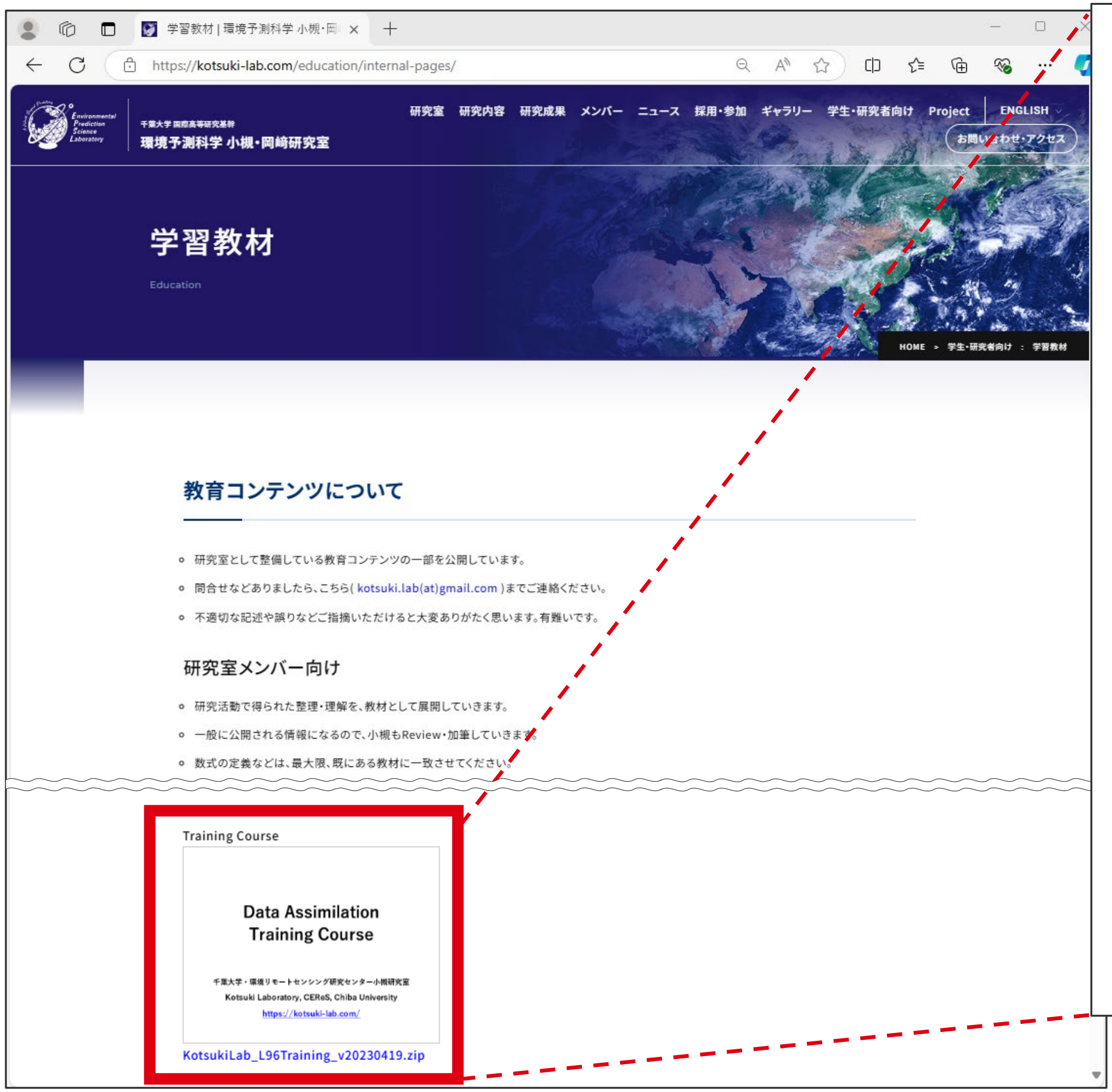
January 31, 2020, Shunji Kotsuki

updated 2020/03/19, 2020/06/29, 2021/07/15

目的： 簡易力学モデル Lorenz の 40 変数モデル（以下 L96; Lorenz 1996）を使って複数のデータ同化手法を自ら実装し、様々な実験を行う。データ同化システムを実際に、0 からコーディングすることで、力学モデリングやデータ同化に関する実践的な「使える」基礎技術を体得する。

Purpose: Using the 40-variable dynamical a.k.a. Lorenz-96 (L96; Lorenz 1996), we are going to perform various experiments with multiple data assimilation (DA) methods. By actually coding a data assimilation system from scratch, you will acquire practically "usable" basic techniques related to mechanical modeling and data assimilation.

① Training Description



pswd: ceres

力学系モデル・データ同化基礎技術の速習コース

Training Course of Dynamical Model and Data Assimilation

January 31, 2020, Shunji Kotsuki
updated 2020/03/19, 2020/06/29, 2021/07/15

目的： 簡易力学モデル Lorenz の 40 変数モデル（以下 L96; Lorenz 1996）を使って複数のデータ同化手法を自ら実装し、様々な実験を行う。データ同化システムを実際に、0 からコーディングすることで、力学モデリングやデータ同化に関する実践的な「使える」基礎技術を体得する。

Purpose: Using the 40-variable dynamical a.k.a. Lorenz-96 (L96; Lorenz 1996), we are going to perform various experiments with multiple data assimilation (DA) methods. By actually coding a data assimilation system from scratch, you will acquire practically "usable" basic techniques related to mechanical modeling and data assimilation.

方法： 以下の課題を自ら実装し、解決していく。使用言語やプラットフォームは問わない。研究室の MTG において、各自が進捗を報告し、問題点を解消していく。質問は MTG の他も、居室で適宜受け付ける。使用言語については、特に拘りがなければ、行列演算の容易な python が扱いやすい。また、単精度ではなく倍精度でコーディングする事。でないと、既往研究と比較して正しく動作しているか確認できない。

Method: Implement and solve the following problems yourself. Any programing languages or platforms can be used in this exercise. At the Kotsuki Lab. mtg, each personnel will report the progress, and try to solve the problems. Questions are accepted during the MTG as well as at the office when necessary. As for the programing language, python, which is easy to perform matrix operations, is recommended unless specific language is preferred. Also, you should code in double precision instead of single precision. Otherwise, confirming whether performing properly or not compared to the previous studies will not be possible.

▸ https://kotsuki-lab.com/internal-pages/

Basic Task 5

Basic Task 5

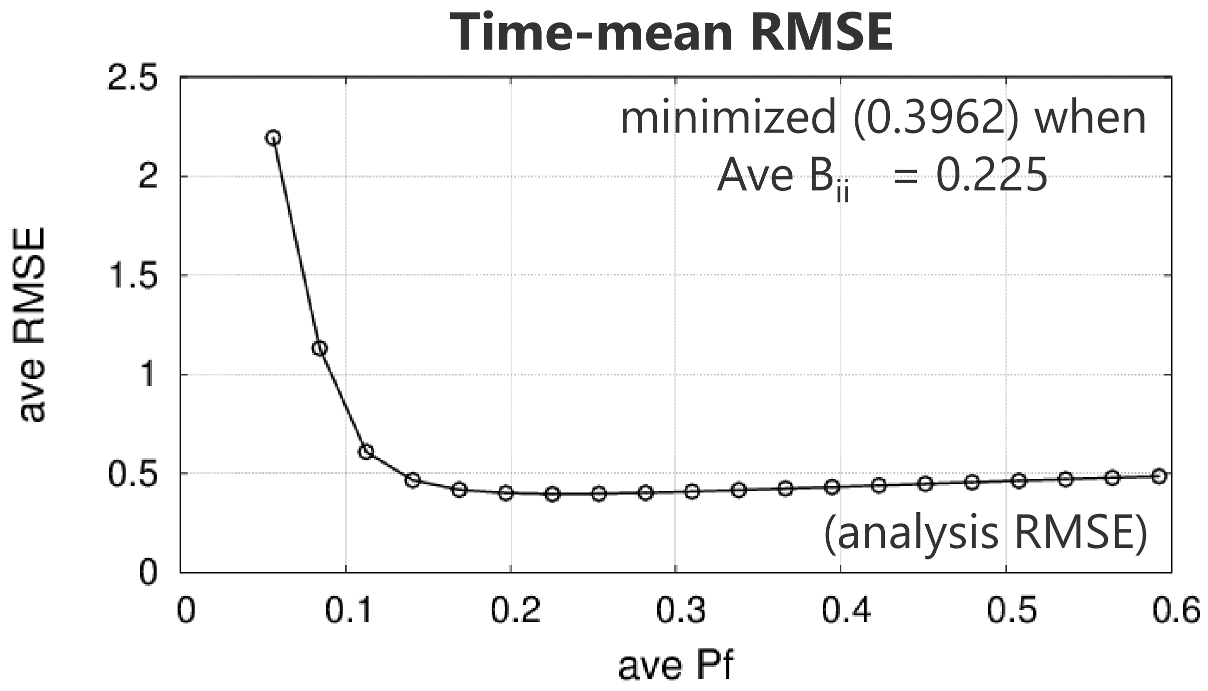
5．３次元変分法と KF の比較実験を行う。この際、観測分布・観測密度への依存性を調べる。

5. Perform a comparative experiment between the 3D variational method and KF. At this time, the dependence on the observation distribution and observation density is investigated.

3DVAR (Full Observations)

when B is diagonal matrix.

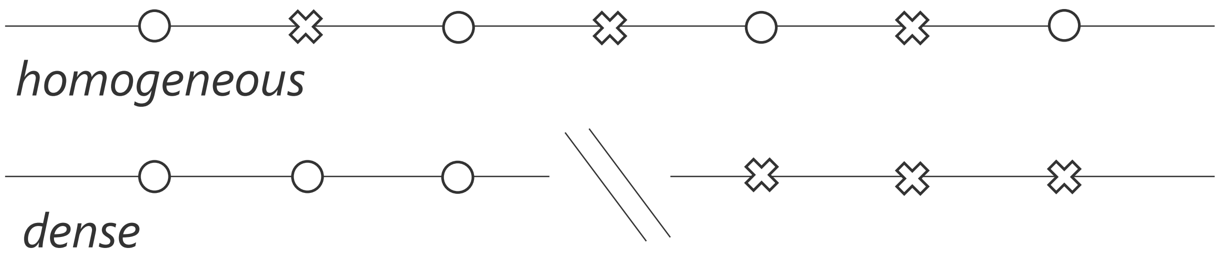
$$B = \begin{pmatrix} b_{11} & & & & 0 \\ & \ddots & & & \\ & & b_{ii} & & \\ & & & \ddots & \\ 0 & & & & b_{nn} \end{pmatrix}$$



Sensitivity to Obs. Network

Case 1: The Num. Obs. = X

Case 2: The Num. Obs. = X



Full Observations

$n=40$, $p=40$

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & & 0 \\ 0 & 1 & 0 & 0 & & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & & 0 \\ & & \vdots & & \ddots & \\ 0 & 0 & 0 & 0 & & 1 \end{pmatrix}$$

***20 Obs* (Case1)**

$n=40$, $p=20$

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 1 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ & & \vdots & & & \ddots \end{pmatrix}$$

***20 Obs* (Case2)**

$n=40$, $p=20$

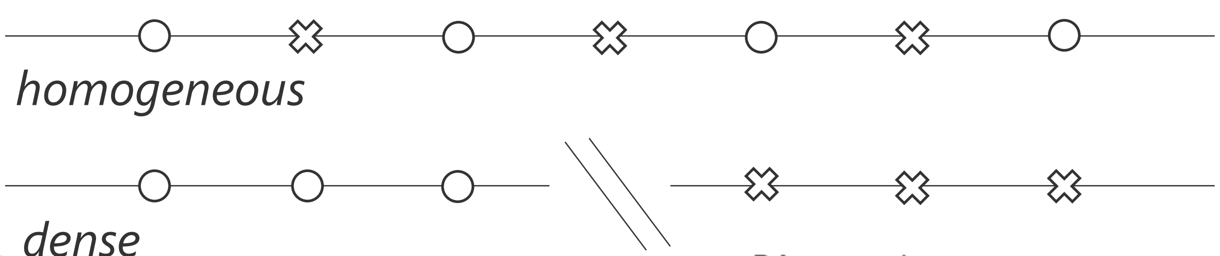
$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 0 & 0 & \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 0 & 1 & \\ & & \vdots & & & \ddots \end{pmatrix}$$

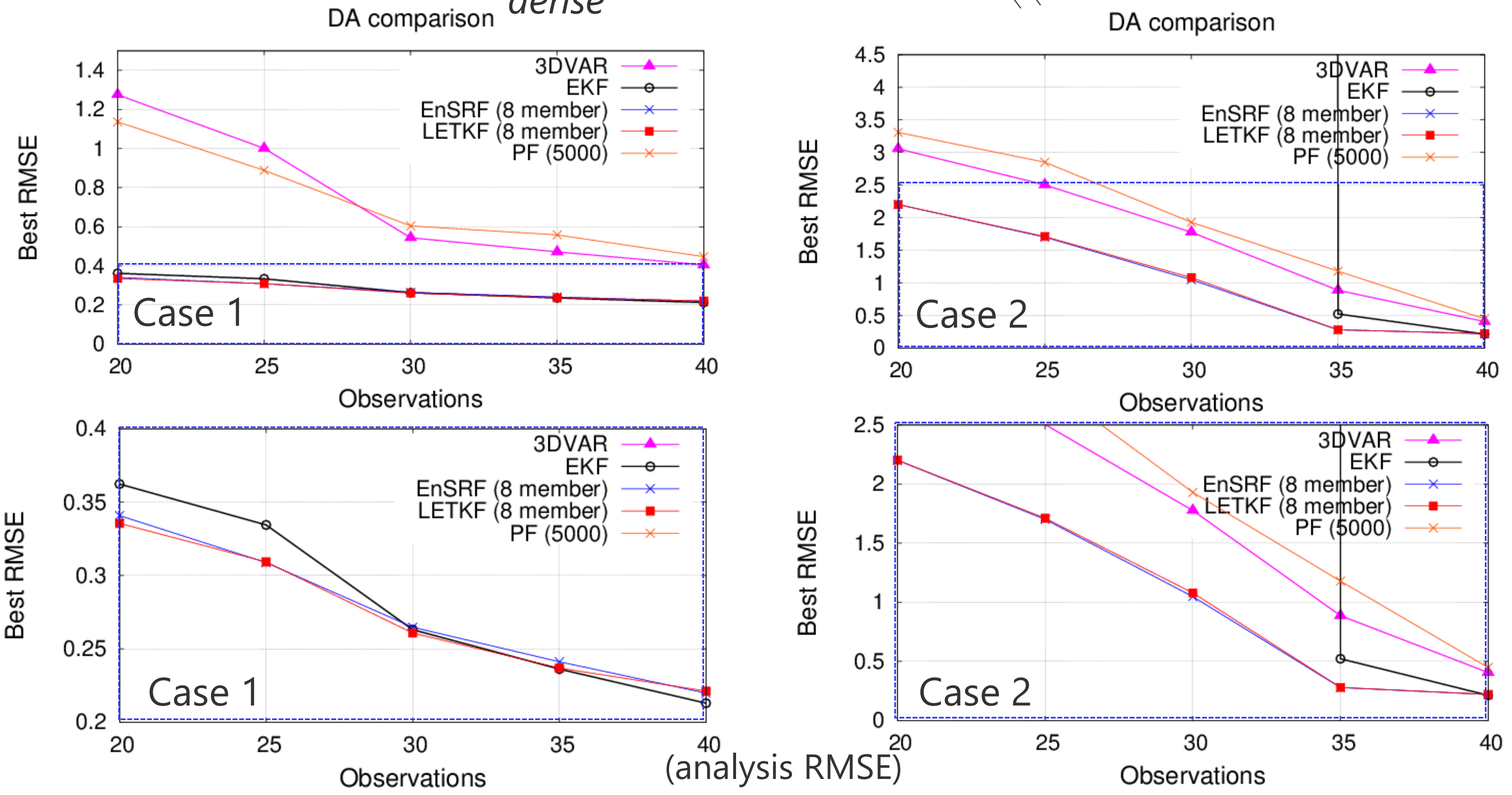
Sensitivity to Obs. Network

Environmental Prediction Science Laboratory

Case 1: The Num. Obs. = X

Case 2: The Num. Obs. = X





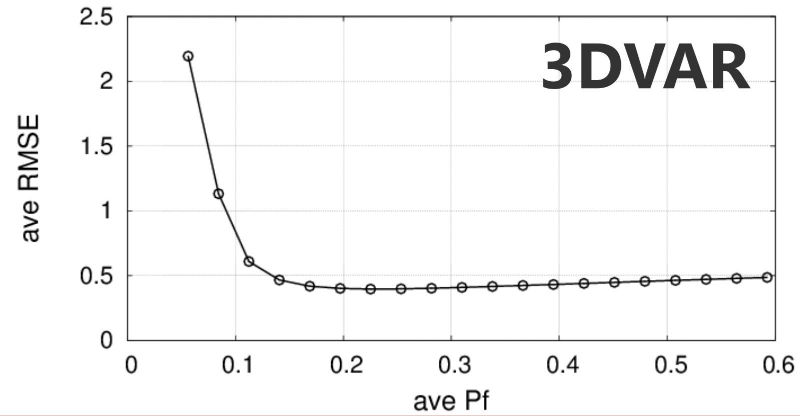
(analysis RMSE)

Hints to Develop KF & 3DVAR

(1) Steps for KF & 3DVAR

	KF	3DVAR
State Prediction	$\mathbf{x}_t^b = M(\mathbf{x}_{t-1}^a)$	
Background Error Cov.	$\mathbf{P}_t^b = \mathbf{M}\mathbf{P}_{t-1}^a\mathbf{M}^T$	$\mathbf{B}$ (static)
Kalman Gain	$\mathbf{K}_t = \mathbf{P}_t^b\mathbf{H}^T\left[\mathbf{H}\mathbf{P}_t^b\mathbf{H}^T + \mathbf{R}\right]^{-1}$	
State Analysis	$\mathbf{x}_t^a = \mathbf{x}_t^b + \mathbf{K}_t(\mathbf{y}_t^o - H(\mathbf{x}_t^b))$	
Analysis Error Cov.	$\mathbf{P}_t^a = [\mathbf{I} - \mathbf{K}_t\mathbf{H}]\mathbf{P}_t^b$	/

Starting with 3DVAR is a good strategy, followed by KF



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